



Université Sidi Mohammed Ben Abdellah
Faculté des Sciences Dhar El Mahraz- Fès
Centre d'Etudes Doctorales
"Sciences et Technologies"

**Formation Doctorale : Sciences et Technologies de
l'Information et de la Communication (STIC)**

Discipline : Informatique

Spécialité : Informatique

Laboratoire d'Informatique et Modélisation (LIM)

THESE DE DOCTORAT

Présentée par

N i s r i n e D a d

**Contributions à la description d'images couleur pour la recherche d'images
par le contenu**

Soutenue le 28/02/2018 devant le jury composé de :

Pr. Mohammed Meknassi	Faculté des Sciences Dhar El Mahraz – Fès	Président
Pr. Hassan Qjidaa	Faculté des Sciences Dhar El Mahraz – Fès	Rapporteur
Pr. Mohammed Rziza	Faculté des Sciences – Rabat	Rapporteur
Pr. Khalid Alaoui Zidani	Faculté des Sciences Dhar El Mahraz – Fès	Rapporteur
Pr. Mohammed Oumsis	Ecole Supérieure de Technologie – Salé	Examineur
Pr. Si Lhoussain Aouragh	Faculté des Sciences Juridiques, Economiques et Sociales – Salé	Examineur
Pr. Saïd Ouatik El Alaoui	Faculté des Sciences Dhar El Mahraz – Fès	Directeur de thèse
Pr. Noureddine En-nahnahi	Faculté des Sciences Dhar El Mahraz – Fès	Co-directeur de thèse

Acknowledgements

I would like to thank a number of people who have accompanied me over the last five years and without whom I would not have been able to complete my thesis work.

First and foremost, I want to thank both of my supervisors Pr. Said El Alaoui Ouatik and Pr. Noureddine En-nahnahi. I am grateful to Pr. Said El Alaoui Ouatik for having faith in me and for having encouraged me to apply for a PhD thesis under his supervision. He had always pulled me forward and awakened the hard-working person in me. I would like to thank him for his insightful remarks and suggestions as well as for helping me with all the administrative procedures, including applying for scholarships and requesting flight-tickets. This work's process provided me with the pleasure of working with Pr. Noureddine En-nahnahi, without whom, I confess, I would have been lost. He oriented me since the beginning of my journey and never failed to offer me his guidance whenever needed. He has several human qualities: Perfectionism, sense of humor, empathy, intuition, creativity, passion, responsibility and kindness. He has been more than a supervisor to me, he was a spiritual teacher and mentor. I am particularly indebted to him for teaching me writing techniques. Language is incapable of expressing my sincere debt and warm thanks to him.

I cannot but record my deep gratitude and appreciation to Pr. Meknassi Mohammed, the chairman of the jury and director of the LIM laboratory. I am also happy to acknowledge my thankfulness to my reading committee members: Pr. Khalid Alaoui Zidani, Pr. Hassan Qjidaa and Pr. Mohammed Rziza for their time, interest, and helpful comments, as well as the two members of my oral defense committee, Pr. Mohammed Oumsis and Pr. Si Lhoussain Aouragh, for their time and invaluable questions.

My thanks goes also to the LIM laboratory which has been a second home for me throughout those years. I hereby acknowledge the valuable work of the committee at the laboratory and their continuous efforts to bring the best out of its members and PhD candidates, especially when it comes to organizing the laboratory days every year. These days allowed me to present the progress of this thesis works and take advantage of the pertinent feed-back of experts in the field. I am also grateful for the flight tickets that I couldn't afford but thanks to which I had been able to communicate and discuss my research works in international conferences held abroad (Qatar and Tunis).

I gratefully acknowledge the CNRST for making my PhD work possible thanks to their funding scholarship.

I would like to thank my colleagues Lamyae, Abdelkader, Afaf, Ahmed, Amal whose encouragement refueled me with energy and love in moments when I most needed them.

I take the opportunity to thank my friends Zahra, Mahmoud, Imane, Andaloussi, Mimi and Fadou, whom I've known during this PhD and who left a good impact in my heart and my life. Lastly, I would like to thank my family for all their love, encouragement and support. My mother (My sweet angel) who worked hard and sacrificed so many things so that I could finish my PhD thesis in good conditions. I thank her for her empathy and her givings without being asked to. I thank my brother Zouhair for being patient with me and for his financial help. My thanks goes also to my sisters Manar and Majda who were always there for me when I mostly needed them.

This thesis is in the world for those who believed in me all along. Too many to name, but not too many to appreciate.

Contributions à la description d'images couleur pour la recherche d'images par le contenu

Résumé: Ce travail de thèse s'inscrit dans le cadre de l'amélioration de la recherche d'images par le contenu en se focalisant sur le processus d'extraction des caractéristiques d'images. Nous nous intéressons particulièrement aux descripteurs d'images basés à la fois sur la forme et la couleur. En effet, en exploitant l'algèbre des quaternions, nous généralisons les moments orthogonaux radiaux permettant de décrire la forme en vue d'intégrer également l'information couleur. Dans un premier temps, nous décrivons l'état de l'art concernant les descripteurs d'images couleur, y compris les moments quaternaires. Ensuite, nous présentons trois contributions concernant la description d'images couleur pour la recherche par le contenu. Dans notre première contribution, nous proposons une adaptation de la transformée en harmoniques sphériques (SHT) pour la description d'images couleur en utilisant à la fois l'algèbre des quaternions et une transformation sphère-disque appropriée. De plus, nous améliorons le temps de calcul de notre descripteur, appelé *Quaternion Disc-Harmonic Moments* (QDHM), en se basant sur la technique d'échantillonnage sphérique *Healpix*. La deuxième contribution consiste à proposer une reformulation sans paramètre des moments orthogonaux quaternaires, appelés *Parameter-free quaternion orthogonal moments* (PFQOM). Ceci est justifié par le fait que le calcul des moments quaternaires existants dépend du choix d'un paramètre quaternion unitaire pur dont la valeur optimale ne peut être fixée qu'empiriquement. De plus, la valeur estimée de ce paramètre est fortement liée au domaine d'application. Dans notre troisième contribution, nous suggérons une méthode de classification d'images couleur basée sur les moments QDHM sans paramètre. L'objectif est de confirmer la pertinence de notre descripteur quaternaire et de le valider dans le cadre d'un système de classification d'images. L'idée de base est de garder la forme quaternaire des coefficients composant le descripteur QDHM au lieu de se restreindre aux modules des moments. Pour maintenir la corrélation entre les quatre parties, réelle et imaginaires, des coefficients du descripteur quaternaire, nous nous sommes basés sur un classificateur de type quaternaire, qui est la machine d'apprentissage extrême quaternaire (QELM). La performance de l'ensemble des descripteurs proposés est évaluée en termes de capacité de reconstruction d'images couleur, de robustesse aux changements photométriques, d'invariance géométrique, de pouvoir discriminatoire et de temps de calcul. Les résultats obtenus montrent l'intérêt de notre propos.

Mots clés: Descripteurs de formes, Moments orthogonaux, Harmoniques sphériques, Image couleur, Algèbre des quaternions, Recherche d'images par le contenu, Classification, Machine d'apprentissage extrême.

Contributions in Color Image Description for Content-Based Image Retrieval

Abstract: This thesis work concerns the enhancement of content-based image retrieval (CBIR) by focusing on the image feature extraction process. We are particularly interested in image description based on both color and shape information. Indeed, by exploiting the algebra of quaternions, we generalize the radial orthogonal moments, meant for describing shapes, in order to also integrate the color information. We first present an extensive state-of-the-art regarding color image descriptors including the quaternion-based moments. Then, we provide three contributions in color image description for CBIR. In our first contribution, we propose an adaptation of the spherical harmonic transform (SHT) for color image description by using both the algebra of quaternions and an appropriate sphere-to-disc transformation. In addition, we improve the computational time of our descriptor, namely quaternion disc-harmonic moments (QDHMs), by employing the Healpix spherical sampling technique. Our second contribution consists in proposing a parameter-free reformulation of the quaternary orthogonal moments, namely the parameter-free quaternary orthogonal moments (PFQOMs). We were motivated by the fact that the calculation of the existing quaternary moments depend on the choice of a unit pure quaternion parameter whose optimal value can be fixed only with the help of experiments. Furthermore, the estimated value of this parameter is strongly dependent on the application domain. In our third contribution, we introduce a color image classification method based on the parameter-free QDHMs. The objective is to confirm the relevance of our quaternary descriptor and validate it as part of an image classification system. The basic idea is to keep the quaternary form of the coefficients composing the QDHMs descriptor instead of being restricted to the moment's modulus. To maintain the correlation between the four parts, real and imaginary, of the quaternary descriptor coefficients, we take advantage of a quaternary classifier that is the quaternary extreme learning machine (QELM). The performance of all the proposed descriptors is evaluated in terms of color image reconstruction capability, robustness to photometric changes, invariance under geometric transformations, discriminative power as well as computational time. The obtained results illustrate the effectiveness of our methods.

Keywords: Shape descriptors, Orthogonal moments, Spherical harmonics, Color image, Quaternion algebra, Content-based image retrieval, Classification, Extreme learning machine.

List of Figures

1	An example of Google search by image.	1
2	Flowchart of a typical content-based image retrieval (CBIR) system.	2
3	Example of the emergence of false colors in the case of the marginal approach (Varian et al., 2013): (a) Original image. (b) Morphological opening result, which shows the emergence of green color (false color). (c) Median filtering result, which shows that the phenomenon of false color appearance is not systematic.	3
1.1	Human eye anatomy and close-up of the retina.	6
1.2	Mechanism of the human perception of colors.	7
1.3	Cones spectrum.	7
1.4	Our taxonomy of color image description techniques.	8
1.5	An example of two different images with similar color histograms (Pass et al., 1996).	9
1.6	A simple example of two different images whose histograms and color coherence vectors are almost similar and color correlograms are different (Huang et al., 1997).	9
1.7	xy chromaticity diagram.	11
1.8	A simple example of two images with same shape and different colors before and after conversion into gray-level. Gray-level image descriptors fail to capture color information, therefore, will consider the two images similar.	12
2.1	(a) Square of the real and imaginary parts of the spherical harmonics of order up to 3. (b) square of the modulus of the spherical harmonics of order up to 3.	27
2.2	Reconstruction of 3D objects by the spherical harmonic transform up to different maximum orders.	28
2.3	Rotational invariance of the amplitude coefficients of the SHT.	29
2.4	The first few associated Legendre ploynomials.	30
2.5	Example of spherical harmonic lighting on a 3D car object (Green, 2003).	31
2.6	Results for different types of queries by the 3D-model search engine developed by Funkhouser et al., 2003.	31
2.7	Method proposed by Funkhouser et al., 2003 for rotation invariant spherical shape representation based on the SHT.	31
2.8	Method proposed by Pu and Ramani, 2006 for representing a 2D drawing as a spherical function.	32
2.9	Method proposed by Sajjanhar et al., 2004; Sajjanhart et al., 2005 for representing a 2D shape in 3D space.	33

2.10	Complex-valued harmonic function up to order $n = 8$ employed by Ennahahi et al., 2010 for representing a 2D shape with the spherical harmonic transform. The top triangle represents the imaginary parts whilst the bottom triangle represents the real parts. The repetition m varies between 0 and n . Green pixels represent regions where the function is negative, blue ones represent where it is positive and the black ones are where the function is zero.	34
2.11	Sphere-to-disc mapping adopted by Ennahahi et al., 2010.	36
2.12	QDHMs reconstruction of image (a) up to order 20, 40, 60 with the corresponding NIRE errors.	41
2.13	Distribution of the pixels on the hemisphere using (a) the new mapping which considers uniform steps of φ and $r = \frac{2 \times \theta}{\pi}$, (b) the old mapping with uniform steps of φ and $r = \cos(\theta)$ (Ennahahi et al., 2010).	42
2.14	Quaternion-valued harmonic function up to order $n = 8$ using (a) the first version [Eq. (2.22)] and (b) the modified version [Eq. (2.37)]. The top triangles represent the first imaginary parts whilst the bottom triangles represent the real parts. The repetition m varies between 0 and n . Green pixels represent regions where the function is negative, blue ones represent where it is positive and the black ones are where the function is zero.	42
2.15	QDHMs reconstruction of image (a) up to order 20,40,60 using (b) the old mapping and (c) the new mapping.	43
2.16	Healpix sampling of the sphere for $N_{side} = 1, 2, 4, 8$ (Górski et al., 2005). . . .	44
2.17	Some images for one object from the COIL-100 data set showing object pose variations.	48
2.18	Some of the challenging object images selected from the COIL-100 data set. . .	48
2.19	Some images for one object from the ALOI-VIEW data set showing object pose variations.	48
2.20	A class of images extracted from the Amsterdam Library of Object Images under illumination color change (ALOI-COL).	48
2.21	A class of images extracted from the Amsterdam Library of Object Images under illumination direction change (ALOI-ILL).	49
2.22	A sample of objects extracted from (a) Columbia Object Image Library (COIL-100) (Nene et al., 1996) and (b) Amsterdam Library of Object Images (ALOI) (Geusebroek et al., 2005).	49
2.23	Comparison of the average NIRE values from Image reconstruction using the QDHMs and the quaternion moments.	50
2.24	A sample of images used in Salt and Pepper (SP) noise robustness test, blur robustness test and rotation invariance test.	51
2.25	The overall average squared errors corresponding to the intensity of the transformation for each of the state-of-the-art moments: a) Salt and Pepper noise effect, b) Gaussian blur effect.	51
2.26	Time consumed in the computation of the feature vector of the image (a) of size 128×128 using the different state-of-the-art moments of number of moments up to 121.	55
2.27	Recall-precision curves for different maximum orders using the QDHMs. . . .	57
2.28	Recall-precision curves for different similarity measures using the QDHMs. . .	58
2.29	Recall-precision curves for different color spaces using the QDHMs.	59

2.30	Content-based color image retrieval by QDHMs (a) using the HSV color space, (b) using the RGB color space. The first row in both (a) and (b) represents the query images set. Each column is the result of the corresponding query.	60
2.31	Comparative analysis of the discriminative power (recall-precision rates) of (a) the DHMs and QDHMs using the COIL-100, (b) the DHMs and QDHMs using the ALOI-VIEW, (c) the quaternion and Clifford based moments using the COIL-100 and (d) the quaternion and Clifford based moments using the ALOI-VIEW.	61
2.32	Recall-Precision curves based on Chi-squared distance for the QDHMs and the QDHMs-Healpix using (a) COIL-100 and (b) ALOI-VIEW	62
3.1	Parameter-free harmonic function up to order $n = 8$. At left, the top triangle represents the first imaginary parts whilst the bottom triangle represents the real parts. At right, the top triangle represents the third imaginary parts whilst the bottom triangle represents the second imaginary parts. The repetition m varies between 0 and n . Green pixels represent regions where the function is negative, blue ones represent where it is positive and the black ones are where the function is zero.	68
3.2	Some head variations of one individual from Faces96 collection.	69
3.3	Six images for one individual from Faces95 collection, showing some major variation of head scale.	70
3.4	Reconstruction errors (NIRE) of the PFQDHMs up to orders from 10 to 80 with a step of 10 for the standard Lena and Barbara, a face image from MPIK and an object image from COIL-100, all of size 128×128	70
3.5	Example of reconstructed images of size 128×128 from the PFQDHMs. The orders are 20,40,60 and 80.	71
3.6	Mean squared error (MSE) of the different transformed images up to different orders.	73
3.7	Average precision using different color image moments when processing the COIL-100 database.	74
3.8	Average precision using different color image moments when processing the Faces96 collection.	74
3.9	ROC curves of different color image moments when processing COIL-100 database.	75
3.10	ROC curves of different color image moments when processing Faces96 collection.	76
3.11	Classification accuracies in % of different color image moments using an ELM with different numbers of hidden layer nodes when processing (a) COIL-100 database and (b) Faces96 collection.	77
3.12	Classification accuracies in % of different color image moments up to different maximum orders $nMax$ using the ELM when processing (a) COIL-100 database and (b) Faces96 collection.	78
3.13	Feature vector size for each of the color image moments up to different maximum orders $nMax$. The size of the PFQDHMs, PFQBFMs, QDHMs and QBFMs is $(nMax + 1)^2$ moments, of the PFQZMs and QZMs is $(nMax + 1)(nMax + 2)/2$ moments, of LCDM is $nMax(nMax + 3)/2$ moments, for CM is $(nMax + 1)(nMax + 2)$ moments and for poFMD is $2nMax^2$ moments.	79

4.1	Structure of a typical neuron.	82
4.2	Single neuron proposed by McCulloch and Pitts, 1943.	83
4.3	Structure of a three-layered neural network.	84
4.4	Classification using the back-propagation neural networks (BPNN): (a) training process to find the optimal parameters, i.e., weights and thresholds of BPNN; (b) the classification process using the trained BPNN.	85
4.5	Example of local minima of an error function.	86
4.6	The excellent generalization performance of the least squares support vector machine (LS-SVM) with RBF kernel on a two-spiral benchmark problem (Suykens and Vandewalle, 1999).	88
4.7	Network architecture proposed by LeCun et al., 1989 for recognizing handwritten zip codes.	89
4.8	The different rotation invariants from quaternary moments and the different classification approaches to deal with them.	94
4.9	Marginal approach for color image classification, which consists of feeding a real-valued classifier with the four parts of the quaternion-valued moment invariants separately (the classifiers' input is four times larger).	94
4.10	Quaternary approach for color image classification, which consists of using a quaternary classifier fed with the quaternion-valued moment invariants as a whole.	95
4.11	Flowchart of the proposed method for color image classification.	96
4.12	Classification accuracy in % of the proposed method and other quaternary methods using the quaternion-valued invariants of the quaternary orthogonal moments with quaternary classifiers, processing the noise-free Faces95 collection.	97
4.13	Classification accuracy in % for different methods using the modulus of the quaternary orthogonal moments with real-valued classifiers, processing the noise-free Faces95 collection.	98
4.14	Ten images for one individual from faces95 after adding Gaussian blur and salt and pepper noise.	99
4.15	Classification accuracy in % of the proposed method and other quaternary methods using the quaternion-valued invariants of the quaternary orthogonal moments with quaternary classifiers, processing corrupted Faces95 collection.	99
4.16	Classification accuracies of the QELM and ELM using different numbers of hidden layer nodes. (a) processing the Faces95 collection without corrupted images. (b) Processing Faces95 with corrupted images.	100
4.17	Classification accuracy in % for different methods using the modulus of the quaternary orthogonal moments with real-valued classifiers, processing corrupted Faces95 collection.	100

List of Tables

2.1	The computational complexity of the quaternion disc-harmonic moments. . . .	39
2.2	The σ/μ values for the classes of the ALOI-COL using each order of the QDHMs up to 8.	52
2.3	The σ/μ values for the classes of the ALOI-ILL using each order of the QDHMs up to 8.	53
2.4	The σ/μ values for the rotated images using each order of the QDHMs up to 8.	54
2.5	The overall average values of σ/μ for the different state-of-the-art moments. The first row is for illumination color change (RGB), the second row for illumination color change (HSV), the third row for illumination direction change and the fourth row for rotation.	54
2.6	Number of the associated Legendre polynomials computed using both the conventional and Healpix method with different values of $nMax$ for 128×128 size image and $N_{side} = 32$	55
2.7	Average time for the computation of QDHMs-Healpix and some of the state-of-the-art quaternion moments (the image size is 128×128)	56
2.8	Average precision based on the Chi-squared distance comparing the QDHMs-Healpix and some well-known quaternion moments using both the COIL-100 and ALOI-VIEW databases	62
3.1	Real-valued radial polynomials R needed for the computation of some orthogonal moments.	67
3.2	Parameter-free quaternary orthogonal basis function $B_{n,m}^q$ up to order $n = 3$ in case $ m \leq n$	68
3.3	Modulus-based PFQZMs invariants up to $nMax = 2$ for rooster image (Fei-Fei et al., 2004) of size 300×300 under various photometric and geometric transformations. The last two columns represent the corresponding MSE for $nMax = 2$ and $nMax = 8$, respectively.	72
3.4	Precision in % using the first nearest neighbor (1-NN) for the different color image moments when processing the COIL-100 and the Faces96 collections. Bold characters refer to the results obtained by our moments.	74
3.5	Area under curve (AUC) in % for the different color image moments when processing the COIL-100 and the Faces96 collections. Bold characters refer to the results obtained by our moments.	76
3.6	Classification accuracies in % for the different color image moments using the ELM with 600 hidden layer nodes when processing the COIL-100 and with 280 hidden layer nodes when processing Faces96 collection. Bold characters refer to the results obtained by our moments.	77
3.7	Elapsed CPU times in seconds for the computation of different color image moments of an image of size 128×128 up to different maximum orders $nMax$.	78

4.1	Number of hidden layer nodes M for each method processing faces95 database with no corrupted images. For QBPNN and BPNN, M is set to $2(L + N)/3$, whilst for QELM and ELM, M is set empirically according to Figure 4.16 (a). .	97
4.2	Number of hidden layer nodes M for each method processing faces95 database with corrupted images. For QBPNN and BPNN, M is set to $2(L + N)/3$, whilst for QELM and ELM, M is concluded from Figure 4.16 (b).	99
4.3	Comparison between the methods based on QELM, ELM, QBPNN and BPNN, including our method, in terms of classification accuracy (in %) processing ALOI-COL.	100
4.4	Comparison between the methods based on QELM, ELM, QBPNN and BPNN, including our method, in terms of classification accuracy (in %) processing ALOI-ILL.	101
4.5	Comparison between the methods based on QELM, ELM, QBPNN and BPNN, including our method, in terms of training time.	102
1	Temps moyen en seconds pour le calcul de QDHMs-Healpix et différents moments quaternaires (La taille des images est 128×128)	R 14

List of Abbreviations

ACMIs	A ffine C olor M oment I nvariants
ALOI	A msterdam L ibrary of O bject I mages
ARTD	A ngular- R adial T ransform D escriptor
AUC	A rea U nder C urve
BFMs	B essel F ourier M oments
BP	B ack- P ropagation learning algorithm
BPNN	B ack- P ropagation N eural N etworks
CBIR	C ontent- B ased I mage R etrieval
CCV	C olor C oherence V ectors
CFMD	C olor F ourier- M ellin D escriptors
CFMDi	C olor F ourier- M ellin D escriptors invariant under the choice of bivector B
CFT	C lifford- F ourier T ransform
CIE	C omité I nternationale de l' E clairage
CIFAR	C anadian I nstitute for A dvanced R esearch
CKELM	C onvolutional K ernel-based E xtrme L earning M achine
CM	C romaticity M oments
CMB	C osmic M icrowave B ackground
CNN	C onvolutional N eural N etworks
COIL	C olumbia O bject I mage L ibrary
CPU	C entral P rocessing U nit
CSOMs	C ircularly S emi- O rthogonal M oments
CSSD	C urvature S cale S pace D escriptor
DBM	D eep B oltzmann M achines
DBN	D eep B elief N etwork
DFT	D iscrete F ourier T ransform
DHMs	D isc- H armonic M oments
DNN	D eep N eural N etworks
EFM	E xponent F ourier M oments
ELM	E xtrme L earning M achine
EMs	E xponent M oments
FCH	F uzzy C olor H istograms
FCM	F uzzy C -means clustering
FDHT	F ast D isc- H armonic T ransform
FFNN	F eed F orward N eural N etworks
FFT	F ast F ourier T ransform
FMMs	F ourier- M ellin M oments
FSHT	F ast S pherical H armonic T ransform
FT	F ourier T ransform
GB	G aussian B lur
GCFD	G eneralized C olor F ourier D escriptors

GPDIs	G eometric and P hotometric D eformations I nvariants
GPU	G raphic P rocessing U nit
HEALPIX	H ierarchical E qual A rea i so L atitude P ixelization
HSV	H ue S aturation V alue
ILSVRC	I mage N et L arge S cale V isual R ecognition C hallenge
IULPM	I mproved U niform L og- P olar M apping
JPEG	J oint P hotographic E xperts G roup
JPL	J et P ropulsion L aboratory
KELM	K ernel E xtrême L earning M achine
k-NN	k - N earest N eighbor
LBP	L ocal B inary P atterns
LCC	L ocal C oordinate C oding
LCDM	L egendre C romaticity D istribution M oments
LDA	L inear D iscriminant A nalysis
LFW	L abeled F aces in the W ild
LM	L evenberg- M arquardt training algorithm
LS-SVM	L east S quares S upport V ector M achine
MNIST	M odified N ational I nstitute of S tandards and T echnology
MPEG	M oving P icture E xperts G roup
MSE	M ean S quare E rror
NIRE	N ormalized I mage R econstruction E rror
NIST	N ational I nstitute of S tandards and T echnology
OCR	O ptical C haracter R ecognition
OFMMs	O rthogonal F ourier- M ellin M oments
PCANet	P rincipal C omponent A nalysis N etwork
PCET	P olar C omplex E xponential T ransforms
PCT	P olar C osine T ransform
PHT	P olar H armonic T ransform
POC	P hase- O nly C orrelation
poFMD	P arallel- O rthogonal F ourier- M ellin D escriptors
PST	P olar S ine T ransform
PZMs	P seudo- Z ernike M oments
QDA	Q uadratic D iscriminant A nalysis
QDBN	Q uaternion-based D eep B elief N etworks
QLBP	Q uaternionic L ocal B inary P attern
QLRBP	Q uaternionic L ocal R anking B inary P attern
QM	Q uaternary M oments
QMD	Q uaternion- M ichelson D escriptor
QP	Q uadratic P rogramming
QPCANet	Q uaternion P rincipal C omponent A nalysis N etwork
QWLD	Q uaternionic W eber L ocal D escriptor
RBF	R adial B asis F unction
RCNN	R ecurrent C onvolutional N eural N etworks
RELM	R egularized E xtrême L earning M achine
RGB	R ed G reen B lue
RHFM	R adial H armonic F ourier M oments
RMSE	R oot M ean S quare E rror

ROC	R eciever O perating C haracteristic
RST	R otation, S cale and T ranslation
SHT	S pherical H armonic T ransform
SLFNs	S ingle hidden- L ayer F eedforward N eural networks
SP	S alt and P epper
SVM	S upport V ector M achine
TSVM	T win S upport V ector M achines
UCI	U niversity of C alifornia, I rvine
YTF	Y ou T ube F aces
ZMs	Z ernike M oments

Contents

Acknowledgements	i
Abstract/Résumé	i
List of Figures	iv
List of Tables	viii
List of Abbreviations	x
General Introduction	1
1 State-Of-the-Art in Color Image Description	6
1.1 Introduction	6
1.2 Color Histogram-based Methods	7
1.2.1 Color Space based Methods	8
1.2.2 Chromaticity Space based Methods	10
1.3 Marginal Methods	12
1.3.1 Geometric Moments	12
1.3.2 Orthogonal Moments	13
1.3.2.1 Complex Zernike Moments	14
1.3.2.2 Complex Orthogonal Fourier-Mellin Moments	14
1.3.2.3 Complex Bessel-Fourier Moments	15
1.3.2.4 Complex Disc-Harmonic Moments	16
1.3.2.5 Separable Orthogonal Moments	16
1.4 Algebraic Methods	16
1.4.1 Quaternion-based Descriptors	16
1.4.1.1 Quaternion Fourier Transforms	16
1.4.1.2 Quaternion Fourier-Mellin Moments	18
1.4.1.3 Quaternion Zernike and Pseudo-Zernike Moments	18
1.4.1.4 Quaternion Bessel-Fourier Moments	19
1.4.1.5 Quaternion Radial Harmonic Fourier Moments	20
1.4.1.6 Quaternion Polar Harmonic Transform	21
1.4.1.7 Quaternion Exponent Moments	21
1.4.1.8 Quaternion Circularly Semi-Orthogonal Moments	22
1.4.2 Clifford-based Descriptors	22
1.5 Conclusion	23
2 Proposed Quaternary Moments for Color Image Description	25
2.1 Introduction	25

2.2	Preliminaries	26
2.2.1	3-Dimensional Spherical Harmonics	26
2.2.2	Spherical Harmonic Transform Definition	28
2.2.3	Associated Legendre Polynomials	29
2.2.4	Applications of the Spherical Harmonic Transform for 3D Models	30
2.2.5	Adaptation of the Spherical Harmonic Transform to 2D Shape Representation	32
2.2.5.1	2.5D Spherical Harmonic Transform	32
2.2.5.2	Connectivity	32
2.2.5.3	Disc-Harmonic Moments	33
2.3	Set of Quaternary Orthogonal Moments	34
2.3.1	Operations on Quaternions	35
2.3.2	Quaternion Disc-Harmonic Moments	35
2.3.2.1	Sphere-Disc Mapping	35
2.3.2.2	Basis Functions	36
2.3.2.3	QDHMs Definition	37
2.3.2.4	Relation between QDHMs and DHMs	38
2.3.2.5	Some Properties	40
2.3.3	Modified Quaternion Disc-Harmonic Moments	41
2.3.4	QDHMs based on Healpix Sampling	43
2.3.5	QDHMs-based Image Descriptor	45
2.3.5.1	Translation and Scale Invariance	46
2.3.5.2	Rotation Invariance	46
2.3.5.3	Feature Vector Construction	46
2.4	Experimental Results	47
2.4.1	Color Image Sets	47
2.4.2	Color Image Reconstruction	49
2.4.3	Robustness and Invariance	49
2.4.3.1	Salt and Pepper Noise	50
2.4.3.2	Blur	51
2.4.3.3	Illumination Color Change	52
2.4.3.4	Illumination Direction Change	52
2.4.3.5	Rotation	53
2.4.4	Computational Time	53
2.4.5	Color Image Retrieval	56
2.4.5.1	K-Nearest-Neighbor Search	56
2.4.5.2	Effect of the Maximum Order	56
2.4.5.3	Effect of the Similarity Measure	57
2.4.5.4	Effect of the Color Space	58
2.4.5.5	Retrieval Comparison	59
2.5	Conclusion	62
3	Proposed Parameter-Free Reformulation of the Quaternary Moments	64
3.1	Introduction	64
3.2	Parameter-Free Quaternary Moments	65
3.2.1	Quaternion-Valued Spherical Harmonics	65
3.2.2	Quaternary Orthogonal Moments μ -based Computation	66

3.2.3	Parameter-Free Quaternion Orthogonal Moments	67
3.3	Experimental Results	69
3.3.1	Color Image Sets	69
3.3.2	Color Image Reconstruction	70
3.3.3	Robustness and Invariance	71
3.3.4	Color Image Retrieval	73
3.3.5	Color Image Classification	76
3.3.6	Computational Time	78
3.4	Conclusion	80
4	Proposed Color Image Classification Method	81
4.1	Introduction	81
4.2	Review of Machine-Learning Algorithms	82
4.2.1	Real-valued Machine Learning Algorithms	82
4.2.1.1	Feed-Forward Neural Networks	82
4.2.1.2	Back-Propagation Neural Networks	85
4.2.1.3	Extreme Learning Machine	87
4.2.1.4	Least Squares Support Vector Machines	88
4.2.1.5	Deep Learning Algorithms	88
4.2.2	Quaternion-Valued Machine Learning Algorithms	90
4.2.2.1	Quaternion Back-Propagation Neural Networks	90
4.2.2.2	Quaternion Extreme Learning Machine	91
4.3	Image Classification Method based on the Quaternary Approach	92
4.4	Experimental Results	95
4.4.1	Classification using Faces95 Collection	96
4.4.2	Classification using Corrupted Faces95 Collection	98
4.4.3	Classification using ALOI-COL Collection	99
4.4.4	Classification using ALOI-ILL Collection	101
4.4.5	Training Time	101
4.5	Conclusion	102
	Conclusion and Perspectives	103
	Publications of the Author	106
	French Summary/ Résumé détaillé en Français	R 1

General Introduction

In our daily lives, we receive, process and analyze information that is mostly of a visual nature such as images and videos. Particularly, *'an image can encapsulate a very important amount of information just like a dozen pages of text do'*. For instance, we nowadays witness a huge quantity of digital images that are shared everyday through social networks for communication purposes. Images are also used in a wide range of application areas such as forensics, robot vision, astronomy, industrial automation, medicine, environment, crafts, military and road traffic, to name but a few.

This excessive quantity of images that are applied in numerous fields and need to be managed requires efficient image search systems. The latter aim to retrieve relevant images from a large image database based on either a textual query (metadata) or an example query, e.g., image or sketch (visual content).

For this purpose, content-based image retrieval (CBIR) systems were introduced and widely applied in a broad range of domains including crime prevention, architectural design, fashion design, medical diagnosis, remote sensing systems, cultural heritage and web searching.

In fact, the term *CBIR* relates to the process of retrieving similar images from large collections based on the visual characteristics of the image query rather than the metadata associated with it, such as keywords and tags. Figure 1 shows an example of Google search by image, where the system receives an image (rather than text) as a query and returns the visually similar images.

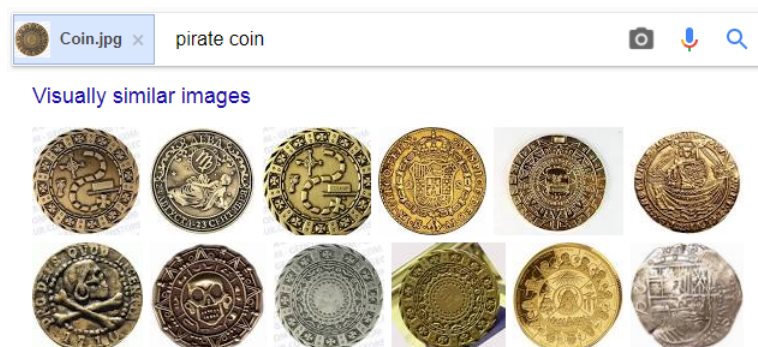


Figure 1 An example of Google search by image.

Figure 2 illustrates the flowchart of a typical content-based image retrieval (CBIR) system. The latter consists of the following stages: (1) An offline stage where the visual contents of the images are extracted through feature vectors to construct the descriptors database. (2) An online stage during which a user provides the system with a sample image. The system transforms this query image into a descriptor vector. Then, it measures the dissimilarity between the query image and those stored in the database by evaluating the distance between their feature vectors. As a result, the system returns the top-ranked images based on this similarity measure. Several factors can impact the performance of a CBIR system including the employed descriptor, the similarity measure, the indexing structure, the search technique, etc.

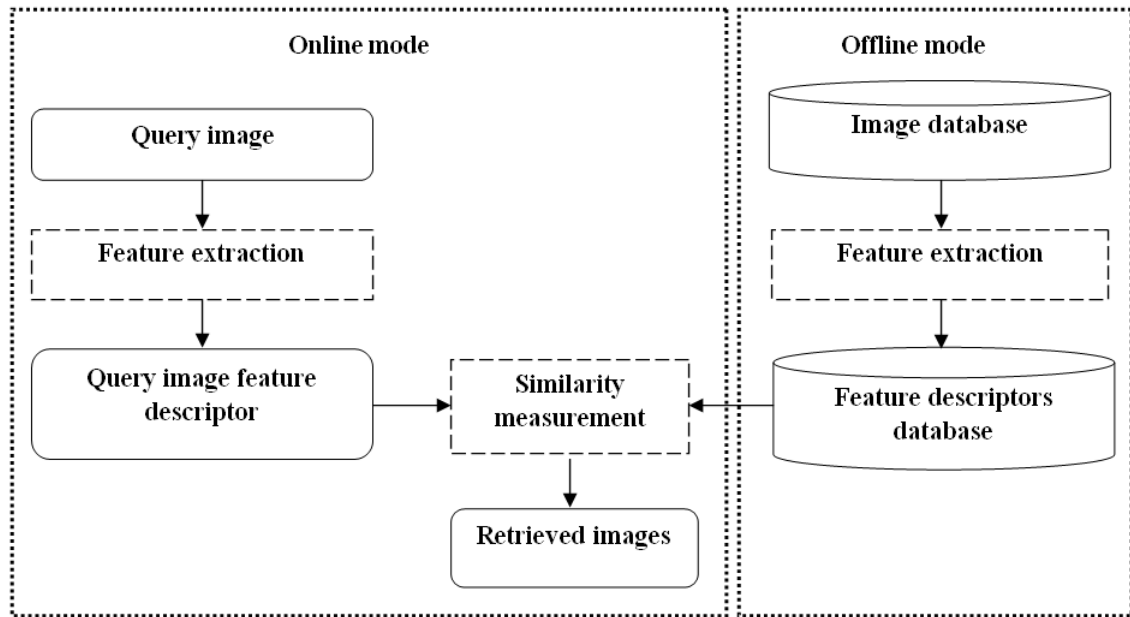


Figure 2 Flowchart of a typical content-based image retrieval (CBIR) system.

From Figure 2, it is obvious that image feature extraction is one of the most important issues in CBIR. In fact, three low-level criteria are generally utilized for image feature extraction: texture, shape, and color.

- A texture representation is employed to define the visual patterns in images based on neighborhood. It is useful in describing the contents of different real-world images such as fingerprints, bricks, trees and fabrics. Among texture descriptors intended for CBIR, we cite the works by Wang et al., 2014b and Park et al., 2002.
- A shape representation could be employed to identify the object contained in an image. In this case, robustness to photometric transformations (blur, noise, illumination color, illumination direction, and so on) and invariance under geometric transformations (translation, scaling, rotation, and so on) are required for a better recognition of the object. According to the review of shape description techniques by Zhang and Lu, 2004, there are two classes of methods for shape description:
 - (1) the methods that use only the contour information to describe the shape, namely the contour-based methods, and
 - (2) the region-based methods, which exploit the entire shape, i.e., all the pixels contained in the shape.

Under each class, there are two approaches: the structural approach, where the shape is represented by segments, and the global approach, where the shape is represented as a whole.

It is generally accepted that the global region-based shape descriptors are the most useful as they can cope well with shape of significant defection, i.e., less likely affected by noise and variations, which poses a problem for contour-based methods. The region-based methods include the orthogonal moments, which have some very important properties yielding to preserving the maximum amount of information on the shape, as a result they have been widely used for CBIR.

- Color, which can be perceived by a combination of three components, is a very important attribute for describing an image. Several works have been investigating color image

description, we cite among them the works by Swain and Ballard, 1991; Pass et al., 1996; Albuz et al., 2001; Paschos et al., 2003.

Many works have been addressing the construction of novel image descriptors, the improvement of the computation accuracy and time, the invariance properties, etc. In this work, we are specifically interested in image descriptors based on both shape and color information. Color can be encoded using different approaches and mathematical formalisms, e.g., marginal and vector approaches.

In fact, many color image processing works involve adapting the gray-level image processing algorithms to deal with color images. This approach, often called marginal, consists of applying a processing on each of the three color components separately then combining the results to produce the final output. Therefore, the color information is not handled globally, but rather split into the three color components independently. This raises problems such as the loss of correlation between the color channels and in most of applications, it leads to the emergence of false colors. The example in Figure 3 (b) illustrates how a marginal morphological opening transformation results in the appearance of green color which is not present in the original image (false color). This undesirable color impedes the visual perception of the user and may distort subsequent operations' results such as a color histogram. However, the phenomenon of false color appearance is not systematic as shown in the result of the marginal median filtering [Figure 3 (c)].

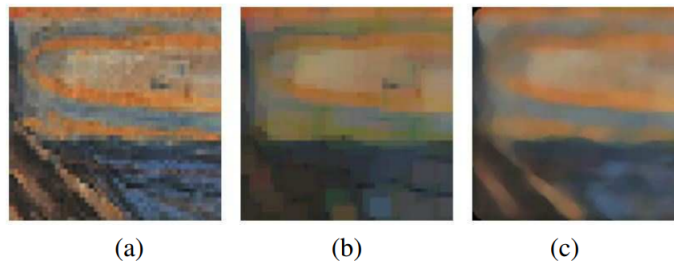


Figure 3 Example of the emergence of false colors in the case of the marginal approach (Varian et al., 2013): (a) Original image. (b) Morphological opening result, which shows the emergence of green color (false color). (c) Median filtering result, which shows that the phenomenon of false color appearance is not systematic.

An alternative to the marginal approach is to consider color as a vector information. In this case, a color pixel is seen as a three-dimensional vector referring to a coordinate within a color space. Although the color representation by means of vectors deals with color information globally, some problems raise. For instance, it is difficult to define order relations between vectors in the three-dimensional space. This is needed in applications such as the median filtering, which is based on pixel ordering. The median filtering is largely employed with gray-level images, but its extension to color images using the vector approach is quite hard. This is because there exists no order between vectors like it exists between scalars.

Another issue of the vector approach is that it is necessary to redefine vector operations to deal with three-dimensional color vectors, e.g., the convolution product, the differentiability (and hence the color gradient) or even more generally all image analysis approaches. Note also the absence of a frequency analysis specific to using the notion of vector colors. Generally speaking, although the vector approach deals with the color information globally, it is disadvantageous in the way it induces more complex treatments compared to the marginal approach.

In this context we focus on the opportunity to encode the color information of images using algebraic representations. These representations use the information of a color pixel as a number of the algebra used. This will allow us to take advantage of the standard operations provided by the algebra, e.g., cross product, inner product, summation, instead of the vector operations, which are very complex to implement. For example, the algebra of complex numbers is widely used in different signal and image processing applications. Precisely, the disc-harmonic moments (DHMs), which requires the use of complex numbers, is a region-based shape descriptor that proves its performances for CBIR. In fact, the DHMs have been proposed by our research team (Ennahnahi et al., 2010) who have been inspired by the well-known spherical harmonic transform (SHT) that has the properties of being complete and orthogonal on the unit sphere.

The initial objective of our work is to propose an extension of these conventional DHMs to deal with color images. For this purpose, we are interested to use the quaternion formalism associated to color images (Sangwine, 1996). In fact, quaternion or hyper-complex numbers, since they are a generalization of complex numbers, can be employed to represent the color information holistically. In addition, they are manipulated by means of simple operations defined in the quaternion algebra. As a quaternion number consists of one real part and three imaginary parts, a color pixel is expressed by a pure quaternion where the real part is set to zero and the three imaginary parts are the three color components.

In this thesis work, we propose different quaternary moment invariants for content-based color image retrieval. The performance of the proposed descriptors is evaluated in terms of color image reconstruction capability, robustness to photometric changes, invariance under geometric transformations, discriminative power as well as computational time.

Contributions of the Thesis Work

The goal of this thesis work is to construct image descriptors based on a combination of both shape and color for a content-based image retrieval purpose. Fast computation, accuracy, numerical stability as well as invariance embodiment are to be considered. This thesis work mainly revolves around the following contributions:

- Establishing a state-of-the-art concerning color image descriptors.
- Improving both the quality and time performances of color image retrieval by proposing fast and invariant quaternary moments for describing color shape images.
- Introducing a parameter-free reformulation of the existing quaternary orthogonal moments.
- Adapting conventional machine-learning algorithms to accept quaternion-valued moment invariants as input for a color image classification purpose.

Outline of the Thesis

This thesis is divided into 4 chapters in addition to the general introduction, conclusion and perspectives. It also contains a detailed french summary.

General Introduction: It highlights the importance of color image description in the context of content-based image retrieval. Then, it describes the motivation and objectives of this thesis work.

Chapter 1: State-Of-the-Art in Color Image Description: It is devoted to a review concerning the existing color image descriptors and their applications especially for CBIR. It also points out the advantages and limitations of each descriptors family.

Chapter 2: Proposed Quaternary Moments for Color Image Description: It presents our first contribution, which consists in proposing a quaternary descriptor, namely the quaternion disc-harmonic moments (QDHMs), inspired by the spherical harmonics (SHs) for color image retrieval. This chapter first gives some preliminaries on the SHs and their various applications dealing with both 3D and 2D models. Then, it details our invariant quaternary descriptor. At the end, this chapter provides the results obtained by the various evaluations. The tests include color image reconstruction capability, invariance to geometric transformations, robustness to photometric changes, computational time as well as discriminative power. The k-nearest-neighbor search involves studying the effect of the similarity measure, color space as well as the maximum order.

Chapter 3: Proposed Parameter-Free Reformulation of the Quaternary Moments: It describes our second contribution which is to introduce a parameter-free reformulation of the quaternary orthogonal moments, namely parameter-free quaternary orthogonal moments (PFQOMs), whose modulus-based invariants are evaluated for color image retrieval and classification. The experimental results show the effectiveness of the proposed descriptors.

Chapter 4: Proposed Color Image Classification Method: It details our third contribution which consists in introducing a color image classification method based on a combination of our quaternion disc-harmonic moments and a quaternion-type classifier, that is the quaternion extreme learning machine (QELM). Our method is evaluated using different challenging color image collections. The obtained results show the ability of our method to guarantee good classification accuracies in a reasonable time.

Conclusion and Perspectives: It summarizes the contributions made in this thesis work and opens the door to their perspectives.

Chapter 1

State-Of-the-Art in Color Image Description

1.1 Introduction

For decades, researchers have been inspired by the human perception of color in order to make computers discriminate objects. Considering the human eye as an example, the light reflected off an object from the outside world is focused by the cornea and lens. Within the eye, the retina is the part that senses the focused light. In fact, the retina, illustrated in Figure 1.1, contains three layers of nerve cells: Ganglion cells on the top, Bipolar cells in the middle and photo-receptors, located in the bottom cell layer. These latter come in two kinds: Cones and Rods, so-named because of their shapes. Their role is to convert the light into electrochemical signals, which are delivered to the visual center in the back of the brain. This allows us to see the color image as it is. Figure 1.2 illustrates the mechanism of the human perception of colors.

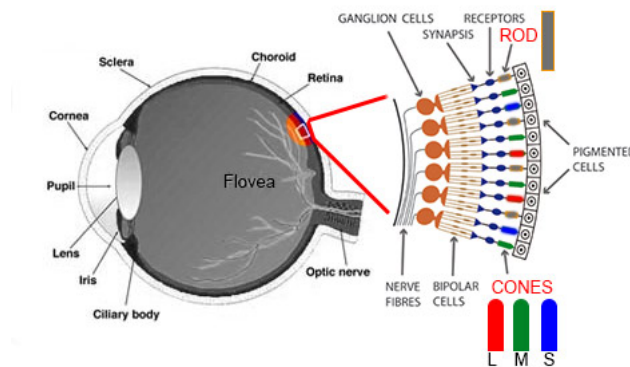


Figure 1.1 Human eye anatomy and close-up of the retina.

The light, which is represented by an electromagnetic spectrum with wavelengths in the region of 380nm to 780nm, pass through the cones and rods. With regard to cones, they function in bright light and give us our color vision, unlike the rods, which function in dim light and do not contribute to color vision in day light. The total number of colors we can distinguish range into the millions. Young, 1802 and Hermann Von Helmholtz realized that the eye can not possibly hold a receptor for each color. They theorized that all colors could be perceived by a combination of just three receptors: Red, green and blue, which are identified as the cones.

Blue has a short wavelength (S), its peak sensitivity is 437 nm. The green sensor cone operates in the middle wavelength (M) and its peak sensitivity is 533 nm. The red cone comes with a

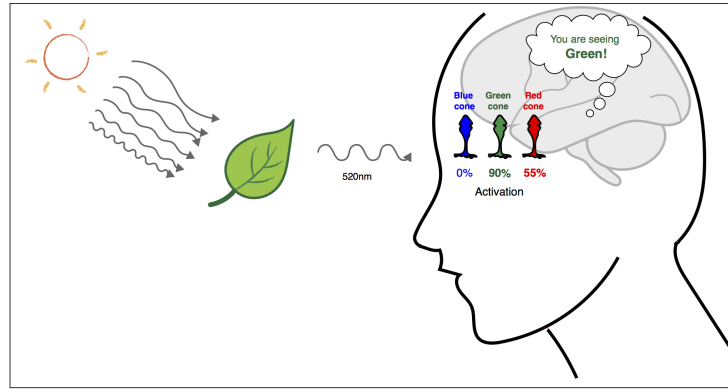


Figure 1.2 Mechanism of the human perception of colors.

longer wavelength (L) with a peak sensitivity of 564 *nm*. Figure 1.3 shows the spectrum of the three cones.

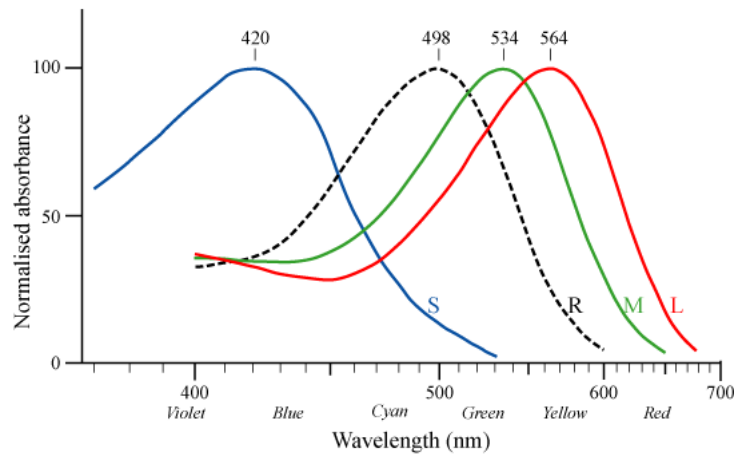


Figure 1.3 Cones spectrum.

Just like the human vision, a computer encodes a color pixel as three numerical components. There exist many ways for defining these three components leading to several color spaces or models including RGB, HSV, HSL.

Generally speaking, color information is a very important attribute of an image. On the other hand, shape information allows to recognize the object in the image. In this chapter, we will give the state-of-the-art concerning the color image descriptors and their numerous applications as well as the advantages and limitations of each descriptor family. In this work, we classify the color descriptors into three categories, i.e., histogram-based methods, marginal approach based methods and algebraic approach based methods. Figure 1.4 shows this classification including the different descriptors under each category.

1.2 Color Histogram-based Methods

Color histogram-based methods for image description can be divided into two categories: (1) Methods that are directly performed on color spaces by using adequate conversions between

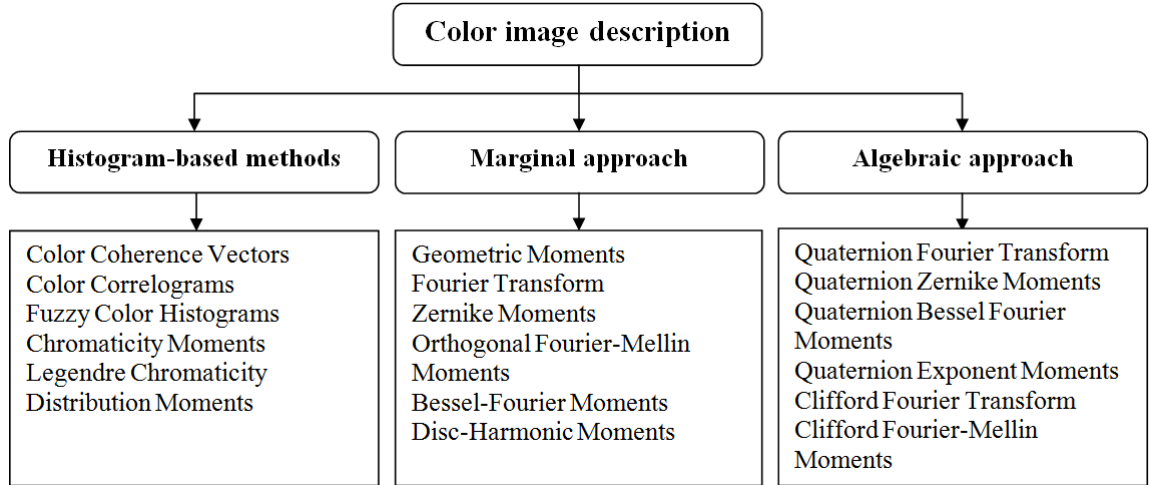


Figure 1.4 Our taxonomy of color image description techniques.

spaces. (2) Methods that are performed on chromaticity spaces in order to eliminate the impact of the brightness component.

1.2.1 Color Space based Methods

The most well-known method for color image description is the color histogram which is employed in several content-based image retrieval systems like QBIC (Flickner et al., 1995) and Chabot (Ogle and Stonebraker, 1995).

The color histogram is obtained from the color space by dividing it into uniform regions according to a quantization level Q and the color pixel counts are assigned to the related histogram bins. Many color spaces, such as RGB, HSV and XYZ, can be employed. Swain and Ballard, 1991 proposed to extract color histograms from the opponent axis color space (Ballard and Brown, 1982) for image retrieval. This latter was based on histogram intersection as similarity measure. The histogram intersection is expressed as follows:

$$HI(I, M) = \sum_{i=1}^n \min(I_i, M_i) / \sum_{i=1}^n M_i, \quad (1.1)$$

where I (input image) and M (model) is a pair of histograms each with n bins.

In other words, the intersection of the model histogram M with the input image histogram I is the number of pixels from M that have corresponding pixels of the same color in I normalized by the number of pixels in M .

Although color histograms succeed to extract the color aspect of the image and are robust to small transformations, they lack spatial information. This is disadvantageous in the way that totally different images, for example those illustrated in Figure 1.5, can have similar histograms. In order to overcome this limitation, Pass et al., 1996 introduced color coherence vectors (CCV) as a color histogram-refinement method by incorporating spatial information. In fact the coherence vector partitions histogram bins by the spatial coherence of pixels, i.e., it classifies each color in the image as either coherent or incoherent. A color's coherence represents the degree to which that color pixels are a part of some sizable contiguous region. The aim of the color coherence vectors is to prevent coherent pixels in one image from matching incoherent pixels in



Figure 1.5 An example of two different images with similar color histograms (Pass et al., 1996).

another, which allows to refine the results obtained by color histograms. An enhanced extension of these color coherence vectors, namely CCV with successive refinement (CCV/C), where the image center is used as additional feature, has been introduced by the same authors (Pass and Zabih, 1996).

In turn, Han and Ma, 2002 introduced the notion of fuzzy color histograms (FCH) to alleviate the problem of histograms sensitivity to illumination changes and quantization errors. In fact, fuzzy color histograms consist of spreading each color pixel to all the histogram bins via a fuzzy membership function (fuzzy-means (FCM) clustering algorithm), therefore, allows to consider the color similarity information. It has been shown that the fuzzy color histogram with *CIELAB* color space is less sensitive to lighting intensity changes and quantization errors compared to the conventional color histograms.

On the other hand, color correlograms have been proposed by Huang et al., 1997 for image indexing/retrieval based on the *RGB* color space. The color correlogram of an image is defined as a table indexed by color pairs, where the entry (i, j) specifies the probability of finding a pixel of color j at a distance k from a pixel of color i in the image. Unlike histograms and color coherent vectors (CCV), color correlograms involve the local spatial correlation of colors and allow the description of the global distribution of this correlation. For instance, color histograms and color coherent vectors would fail to distinguish the two images in Figure 1.6 but color correlograms are capable of distinguishing them.



Figure 1.6 A simple example of two different images whose histograms and color coherence vectors are almost similar and color correlograms are different (Huang et al., 1997).

1.2.2 Chromaticity Space based Methods

Chromaticity designates the quality of a color regardless of its luminance or brightness. In fact, the concept of color can be divided into two parts: brightness and chromaticity. For example, the color white is a bright color, while the color grey is considered to be a less bright version of that same white. In other words, the chromaticity of white and grey are the same while their brightness differs.

In image description, the most common representation of a color pixel is in the RGB color space. However, the direct utilization of the (R, G, B) triplet is unreliable due to its susceptibility to change of brightness. Hence, the triplet can be mapped to brightness independent chromaticities prior to indexing. In particular, each pixel in a given image yields a pair (x, y) of chromaticity values, thus forming a set of chromaticities for the entire image. This set of (x, y) pairs is called the chromaticity set of the specific image, while the corresponding xy Euclidean space is called the chromaticity space. Some color spaces separate the three dimensions of color into one luminance dimension and a pair of chromaticity dimensions. We cite, for example, xyY , $YCbCr$, YIQ , YUV , YCC and the opponent color space.

The xyY color space, for instance, is an extension of the CIE XYZ space. The latter is defined by the following transformation of the RGB color space (Billmeyer, 1983):

$$\begin{aligned} X &= 0.607R + 0.174G + 0.200B; \\ Y &= 0.299R + 0.587G + 0.114B; \\ Z &= 0.066G + 1.111B. \end{aligned} \tag{1.2}$$

The (x, y) chromaticities are derived within XYZ as follows:

$$\begin{aligned} x &= \frac{X}{X + Y + Z}; \\ y &= \frac{Y}{X + Y + Z}. \end{aligned} \tag{1.3}$$

The xy chromaticity diagram is illustrated in Figure 1.7.

Apart from the color image descriptors presented above, many others based on the chromaticity space have been proposed including chromaticity moments (Paschos et al., 2003) and Legendre chromaticity distribution moments (Yap and Paramesran, 2006).

The color histogram variations presented above (Swain and Ballard, 1991; Pass et al., 1996) are consuming in terms of both space and retrieval time requirements. Paschos et al., 2003 thought of representing the color contents of an image with the regular moments of both the trace and the distribution of the xy chromaticity space. In fact, for an image I of size (L_x, L_y) , the trace of its chromaticity as well as the distribution (histogram) are respectively given by:

$$T(x, y) = \begin{cases} 1 & \text{if } \exists(i, j) : I(i, j) \text{ yields } (x, y) \\ 0 & \text{otherwise} \end{cases} \tag{1.4}$$

with $0 \leq i \leq L_x - 1$ and $0 \leq j \leq L_y - 1$,

$$D(x, y) = \# \text{pixels yielding } (x, y). \tag{1.5}$$

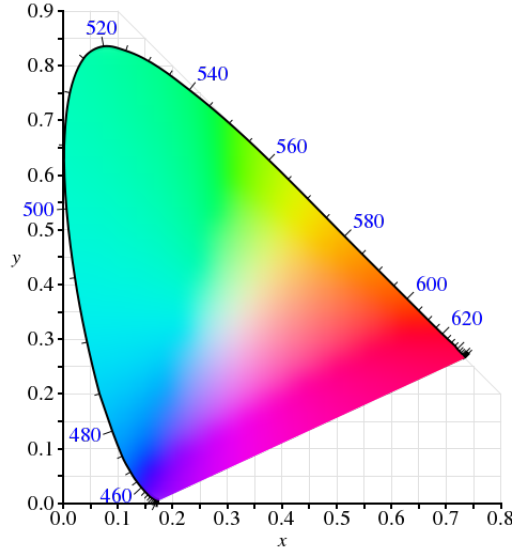


Figure 1.7 xy chromaticity diagram.

The trace T and distribution D can be characterized by a set of moments with $n + m$ being the order defined, respectively, as follows:

$$M_T(n, m) = \sum_x^{X_s-1} \sum_y^{Y_s-1} x^n y^m T(x, y), \quad (1.6)$$

$$M_D(n, m) = \sum_x^{X_s-1} \sum_y^{Y_s-1} x^n y^m D(x, y), \quad (1.7)$$

where X_s, Y_s are the dimensions of the xy space.

The obtained moments, namely the chromaticity moments (CM), were formed by concatenating both the trace and distribution moments and then applied for content-based image retrieval based on the L1 distance. Compared to color histograms (Swain and Ballard, 1991) only few chromaticity moments are needed for comparably effective retrieval performance.

As for the Legendre chromaticity distribution moments (LCDMs) (Yap and Paramesran, 2006), they are computed directly from the opponent chromaticity space derived from the opponent color space (i.e., consists of three axis white-black (wb), yellow-blue (yb) and red-green (rg)) without the need to construct the chromaticity histogram. The LCDMs of order $n + m$ of an image of size $N \times M$ are given by:

$$LCDM_{n,m} = \frac{(2n+1)(2m+1)}{MN} \sum_{l=0}^{MN-1} P_n(rg(l)) P_m(yb(l)), \quad (1.8)$$

where P_n is the Legendre polynomial and rg and yb correspond respectively to the red-green and yellow-blue values in pixel l .

The Legendre chromaticity distribution moments were applied for content-based image retrieval based on the L1 metric for similarity measurement.

1.3 Marginal Methods

For describing a color image, one can intuitively convert that image into gray-scale, then compute its conventional descriptor. However, this direction may result in a significant color information loss in the sense that the same shapes of different colors are deemed similar. An example of this drawback is illustrated in Figure 1.8.

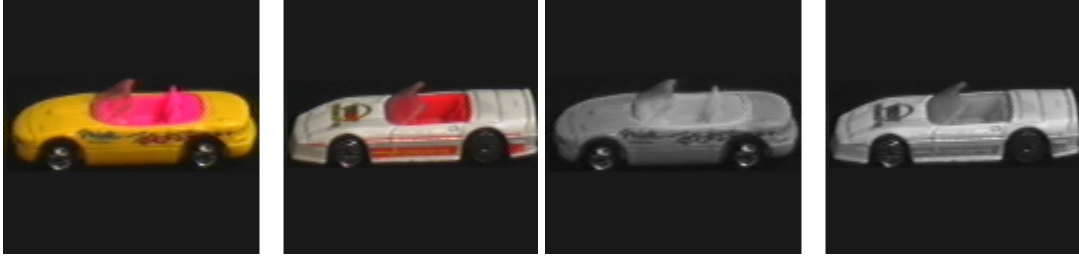


Figure 1.8 A simple example of two images with same shape and different colors before and after conversion into gray-level. Gray-level image descriptors fail to capture color information, therefore, will consider the two images similar.

An alternative consists in calculating the conventional descriptor of each color channel of the image separately, then combining the individual results. This direction, called marginal, may introduce a loss of correlation between the color channels.

Following, we will describe two well-known conventional image descriptor families that are the geometric moments and the orthogonal moments.

1.3.1 Geometric Moments

Geometric moments were first introduced by Hu, 1962 for two-dimensional pattern recognition applications. The geometric moments of order n and repetition m are given by:

$$m_{n,m} = \int_{-1}^1 \int_{-1}^1 x^n y^m f(x, y) dx dy, \quad (1.9)$$

where the image function $f(x, y)$ is defined on a unit square $[-1, 1] \times [-1, 1]$. The lowest-order geometric moments have been widely used in many applications (Hu, 1962; Teh and Chin, 1988; Taubin and Cooper, 1991; Belkasim et al., 1991; Prokop and Reeves, 1992; Flusser and Suk, 1993; Mehtre et al., 1997) owe to their interesting properties. For example, the zero-order moment $m_{0,0}$ represents the total image power, the first-order moments $m_{1,0}$, $m_{0,1}$ locate the centroid of the image (x_c, y_c) , i.e.,

$$x_c = m_{1,0}/m_{0,0} \quad \text{and} \quad y_c = m_{0,1}/m_{0,0}, \quad (1.10)$$

and the second-order moments $m_{2,0}$, $m_{1,1}$ and $m_{0,2}$ characterize the size and orientation of the image. They are expressed as follows:

$$m_{0,0} = \int f(x, y) dx dy, \quad (1.11)$$

$$m_{1,0} = \int x f(x, y) dx dy, \quad m_{0,1} = \int y f(x, y) dx dy, \quad (1.12)$$

$$m_{2,0} = \int x^2 f(x, y) dx dy, \quad m_{1,1} = \int xy f(x, y) dx dy, \quad m_{0,2} = \int y^2 f(x, y) dx dy. \quad (1.13)$$

Mindru et al., 2004 proposed a set of color image invariants to geometric transformation and photometric changes (GPDIs), whose calculation is based on the combination of the geometric moments for one or more color channels. Also Suk and Flusser, 2009 introduced the notion of common centroid for deriving the central moments in order to achieve the translation invariance.

1.3.2 Orthogonal Moments

Although the image representation by means of the geometric moments is complete, i.e., the image function can be exactly reconstructed from the set of its finite number of moments, Teague, 1980 showed that the computation of high-order geometric moments is very difficult as it relies on standard powers which increase rapidly with the increase of the order. It was also demonstrated that the image reconstruction from these high-order moments is not faithful to the original image, which proves there is a considerable amount of information redundancy. On the other hand, Zhang and Lu, 2004 proved that the low-order geometric moment invariants are suitable for describing simple noise-free shapes (MPEG-7); therefore are not sufficient to accurately describe a real-world object.

These drawbacks were the motivation behind the orthogonal moments. In fact Teague, 1980 thought of replacing the conventional transform kernel $x^n y^m$ in Eq. (1.9) with complete sets of an orthogonal basis $OB_{n,m}(x, y)$, verifying the following:

$$\int \int_D w(x, y) OB_{p,q}(x, y) OB_{n,m}(x, y) dx dy = 0, \quad (1.14)$$

for all $(p, q) \neq (n, m)$, where $w(x, y)$ is a weight function and D is the area of orthogonality.

Thanks to this orthogonality property, the information redundancy is eliminated, thus, the orthogonal moments possess better reconstruction power than geometrical moments. On the other hand, the orthogonal basis $OB_{n,m}(x, y)$ is expressed as:

$$OB_{n,m}(r, \varphi) = OP_{n,m}(r) e^{jm\varphi} \quad (1.15)$$

where $OP_{n,m}(r)$ is orthogonal polynomials computed through recurrences, $e^{jm\varphi}$ is a complex kernel and (r, φ) are polar coordinates. This makes the computation of the orthogonal moments less demanding than the geometric moments as they are based on recurrent relations rather than standard powers.

Inspired by Teague, 1980's work, many orthogonal moments have been introduced in the literature for gray-level and binary image analysis. We cite, among them, the orthogonal Fourier-Mellin moments (Sheng and Shen, 1994), the Bessel-Fourier moments (Xiao et al., 2010), the disc-harmonic moments (Ennahahi et al., 2010), the separable two-dimensional discrete orthogonal moments (Abdeslam Hmimid, 2014; Hmimid et al., 2015; Sayyouri et al., 2016), and so one. The theory of moments has been applied in a wide range of applications, including line fitting in noisy images (Qjidaa and Radouane, 1999), lapped block image analysis (El Fadili et al., 2003), watermarking (Kim and Lee, 2003; Gao and Jiang, 2015), pattern recognition (Khotanzad and Hong, 1990; Abdallah et al., 1997), etc. Besides, many researchers proposed fast and accurate computation methods of orthogonal moments (Hosny et al., 2011; Singh and Upneja, 2012b; Mesbah et al., 2016; Deng et al., 2016). Following, we will shed light on some of these orthogonal moments and their various applications in image processing and computer vision.

1.3.2.1 Complex Zernike Moments

As aforementioned, Teague, 1980 was the first to introduce the notion of orthogonal moments into image analysis, particularly Zernike moments (ZMs) and Legendre moments. He provided the Zernike moments invariants up to the eighth order, which can be expressed in terms of the usual geometric moments. Teague also took into consideration that the image moments have to be independent of the object's size, lateral displacement, and orientation. Both scale and translation invariance have been taken care of by the use of normalized, central moments, whilst the invariants under rotation and reflection were mathematically constructed by retaining the modulus of the complex-valued moments. This was to make sure that the constructed invariants are real valued rather than complex valued.

The Zernike moments of order n and repetition m for an image $f(r, \varphi)$, defined in polar coordinates, are given by:

$$ZM_{n,m} = \frac{n+1}{\pi} \int_0^{2\pi} \int_0^1 R_{n,m}(r) e^{-jm\varphi} f(r, \varphi) r dr d\varphi, \quad n - |m| \text{ being even}, \quad (1.16)$$

where $R_{n,m}$ are the Zernike radial polynomials, expressed as:

$$R_{n,m}(r) = \sum_{k=0}^{(n-|m|)/2} \frac{(-1)^k (n-k)!}{k! \left(\frac{n+|m|}{2} - k\right)! \left(\frac{n-|m|}{2} - k\right)!} r^{n-2k} \quad (1.17)$$

The Zernike moments are complete over the unit disc and satisfy the following orthogonality relation:

$$\int_0^{2\pi} \int_0^1 [R_{n,p}(r) e^{jp\varphi}]^* R_{m,q}(r) e^{jq\varphi} r dr d\varphi = \frac{\pi}{n+1} \delta_{nm} \delta_{pq}, \quad (1.18)$$

where δ is the Kronecker symbol.

Later, the Zernike moments were used as features in several computer vision and image processing applications, including but not limited to image watermarking (Kim and Lee, 2003), object recognition (Khotanzad and Hong, 1990; Abdallah et al., 1997), object classification (Nagarajan and Devendran, 2012) and template matching (Beijing et al., 2011).

1.3.2.2 Complex Orthogonal Fourier-Mellin Moments

The orthogonal Fourier-Mellin moments (OFMMs) were introduced by Sheng and Shen, 1994 for invariant pattern recognition. It was shown that these moments outperform the Zernike moments. These latter seem to have a difficulty in describing images of small size, when it comes to scale-invariant pattern recognition. In fact, for small objects the orthogonal Fourier-Mellin moments have better performance than the Zernike moments in terms of noise-free and noisy image reconstruction as well as for pattern recognition. In other words, the OFMMs have higher discriminative power and are more robust to noise compared to the Zernike moments.

The orthogonal Fourier-Mellin moments of order n and repetition m for a gray-scale image $f(r, \varphi)$ are defined as:

$$OFMM_{n,m} = \frac{n+1}{\pi} \int_0^{2\pi} \int_0^1 Q_n(r) e^{-jm\varphi} f(r, \varphi) r dr d\varphi, \quad (1.19)$$

where Q_n are radial polynomials expressed by:

$$Q_n(r) = \sum_{s=0}^n (-1)^{n+s} \frac{(n+s+1)!}{(n-s)!s!(s+1)!} r^s. \quad (1.20)$$

These moments obey the following orthogonality relation:

$$\int_0^{2\pi} \int_0^1 [Q_n(r)e^{jp\varphi}]^* Q_m(r)e^{jq\varphi} r dr d\varphi = \frac{\pi}{n+1} \delta_{nm} \delta_{pq}, \quad (1.21)$$

Several works have been investigated later for fast (Hosny et al., 2011) and accurate (Singh and Upneja, 2012b) computation of the orthogonal Fourier-Mellin moments. Also, Zhang et al., 2010 assessed their performances for invariant pattern recognition dealing with both the gray-scale images of the Columbia database (Nene et al., 1996), which contains images of real-world objects under change of view point, as well as a set of binary images of alphanumeric characters.

1.3.2.3 Complex Bessel-Fourier Moments

In turn, the Bessel-Fourier moments (BFMs) were proposed by Xiao et al., 2010 based on the orthogonal Bessel function of the first kind. The authors assessed their performances in terms of noise-free and noisy image reconstruction, invariant recognition accuracy as well as computational time compared to the Zernike moments and the orthogonal Fourier-Mellin moments.

The Bessel-Fourier moments of order n and repetition m of an image $f(r, \varphi)$ are defined as follows:

$$BFM_{n,m} = \frac{1}{2\pi a_n} \int_0^{2\pi} \int_0^1 J_\nu(\lambda_n r) e^{-jm\varphi} f(r, \varphi) r dr d\varphi, \quad (1.22)$$

where λ_n is the n -th zero of the first-kind Bessel polynomials J_ν of order ν set to be 1. This latter can be expressed in function of the Gamma function Γ , which generalizes the factorial function with non-integer values:

$$J_\nu(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k! \Gamma(\nu + k + 1)} \left(\frac{x}{2}\right)^{\nu+2k}, \quad (1.23)$$

and a_n is the normalization constant:

$$a_n = \frac{[J_{\nu+1}(\lambda_n)]^2}{2}, \quad (1.24)$$

These moments are orthogonal over the unit disc according to the following formula:

$$\int_0^{2\pi} \int_0^1 [J_\nu(\lambda_n r) e^{-jp\varphi}]^* J_\nu(\lambda_m r) e^{-jq\varphi} r dr d\varphi = 2\pi a_n \delta_{nm} \delta_{pq}, \quad (1.25)$$

where δ is the Kronecker symbol.

The Bessel-Fourier moments have been and still used in several computer vision and image processing applications, such as content-based image retrieval (CBIR) and classification (Ma et al., 2011), image reconstruction (Toxqui-Quitl et al., 2011) and image watermarking (Gao and Jiang, 2015).

1.3.2.4 Complex Disc-Harmonic Moments

The disc-harmonic moments (DHMs) were introduced by Ennahnahi et al., 2010 for fast and robust shape description employed for image retrieval. The authors were inspired by the well-known 3D spherical harmonic transform (SHT) that has the properties of being complete and orthogonal on the unit sphere. A detailed description of the disc-harmonic moments will be given in subsection 2.2.5.3.

1.3.2.5 Separable Orthogonal Moments

Koornwinder, 1975 gave a general method for generating bivariate orthogonal polynomials from univariate ones. Generally, bivariate orthogonal polynomials are obtained as a tensor product of two different orthogonal polynomials in one variable. Since then, these bivariate polynomials have attracted considerable research interest (Lewanowicz and Woźny, 2010; Fernández et al., 2011; Fernández et al., 2012; Morales et al., 2009; Fernández et al., 2010; Dunkl and Xu, 2014). Both continuous and discrete orthogonal moments can be defined using bivariate orthogonal polynomials. Many researchers have introduced several continuous and discrete orthogonal moments with bivariate orthogonal polynomials as basis functions for image analysis. Among these separable orthogonal moments, we cite the works by Zhu, 2012; Tsougenis et al., 2013; Tsougenis et al., 2015; Abdeslam Hmimid, 2014; Hmimid et al., 2015; Sayyouri et al., 2016.

1.4 Algebraic Methods

The quaternion algebra (Hamilton, 1844) and the Clifford algebra (Hestenes and Sobczyk, 1987) are deemed as the trend for representing the color information of the image. Hamilton, 1844 introduced the algebra of quaternions, which are a generalization of complex numbers. In fact, a quaternion consists of one real part and three imaginary parts. Note that the Clifford algebra is more general than the quaternion one.

In the literature, the conventional moments dealing with binary and gray-level images described in the previous section were generalized from complex domain to multi-dimensional domain for color image description. This technique allows to encode the three color channels in a single feature vector without passing through the gray-level, which is also advantageous in the way the correlation between the three channels is preserved.

1.4.1 Quaternion-based Descriptors

In this subsection, we will shed light on the pioneer work that shows how a color image can be represented by quaternions as well as the existing works which were inspired by the same technique.

1.4.1.1 Quaternion Fourier Transforms

Being inspired by the works on quaternion versions of Fourier transforms established by Ell, 1993 and Ueda and Takahashi, 1996, Sangwine, 1996 was the first to introduce the color image analysis by the quaternion algebra. He proposed the quaternion discrete/fast Fourier transform (QDFT/QFFT), which makes possible the handling of color images in the frequency domain as a whole, without separate handling of the color components. The author thought of expressing a

color image pixel by a pure quaternion where the real part is set to zero and the three imaginary parts are the three color components. For example, an RGB (red, green, blue) color pixel $f(x, y)$ can be represented as follows:

$$f(x, y) = f_R(x, y)i + f_G(x, y)j + f_B(x, y)k, \quad (1.26)$$

where $f_R(x, y)$, $f_G(x, y)$ and $f_B(x, y)$ represent respectively the red, green and blue channel of the color image, which can obviously be substituted by the three color components of other color spaces such as HSV (hue, saturation, value). This is analogous to the conventional techniques dealing with the monochromatic aspect, except that the pixel intensity is placed in the real part of complex numbers rather than in the imaginary part. As for the inverse transformed image, the real part should be zero; therefore, ignored, just like the imaginary part is ignored in the case of inverse transform of a complex.

The performances of the quaternion discrete/fast Fourier transform were assessed in terms of computational requirements (complexity and memory). It has been shown that these quaternary transforms require less real multiplications compared to the three complex discrete/fast Fourier transform (DFT/FFTs) (Blahut, 1985) (the color image is split to three binary images corresponding to each color channel). Hence, the quaternion version of this transform is more efficient and requires less memory.

Since then, the quaternion Fourier transform have been used in several applications dealing with color images, we can cite:

(1) Color image registration: Feng et al., 2008 extended the conventional phase-only correlation (POC) technique to the quaternion field (QPOC) and proposed a sub-pixel color image registration algorithm based on this proposed quaternion phase-only correlation. This latter is established on the basis of the quaternion Fourier Shift Theorem. In conclusion, the authors drew up that their proposed quaternion correlation makes full use of the chrominance information in color images and also can directly estimate the sub-pixel translational shift from the 2D data array.

Pan et al., 2010 proposed an invariant method combining the quaternion Fourier transform and parametric template method for color image registration.

(2) Color image watermarking: Bas et al., 2003 introduced a color image watermarking scheme based on the quaternion Fourier transform and the quantization index modulation. The authors assessed the performances of their method for different color image filtering processes, e.g., JPEG (Joint Photographic Experts Group) and blur. They found out that the quaternion Fourier transform embedding allows robustness to luminance filtering technique. Wang et al., 2013 proposed a robust blind color image watermarking based on quaternion fast Fourier transform and least squares support vector machine (LS-SVM) (Gestel et al., 2004). The watermark is embedded into original color image blocks by modulating the real quaternion Fourier transform coefficients of these blocks, whilst the LS-SVM correction with pseudo-Zernike moments was employed for watermark decoding. This technique proves to be robust against common image processing operations such as filtering, JPEG compression, histogram equalization, image blurring and also against the geometrical distortions. Chen et al., 2014b introduced a watermark scheme trying to avoid watermark energy loss at the embedding stage when modulating only the real coefficients of the quaternion discrete Fourier transform (QDFT). They established mathematically a way to spread the watermark by fully modulating the four coefficients of the QDFT. In the same line, Ouyang et al., 2015 proposed a blind color image watermarking scheme based on quaternion discrete Fourier transform and on an improved uniform log-polar

mapping (IULPM). Unlike the previous method, the watermark embedding employs the image as a whole, rather than in a block pattern. It has been shown that the use of the quaternion fast Fourier transform makes the scheme resistant to geometric attacks and the IULPM allows good accuracy when estimating the rotation angle of a geometric attack.

(3) Grigoryan et al., 2015 used the quaternion discrete Fourier transform along with the alpha-rooting concept for color image measurement and enhancement. The authors show that their proposed algorithm are fast and simple in implementation and design and it outperforms the traditional Fourier transform-based alpha-rooting and the multi-scale Retinex algorithms in terms of color image enhancement.

(4) As a summary, Ell et al., 2014 recapitulated the use of quaternion Fourier transforms for signal and image processing.

(5) The quaternion discrete Fourier transform was used for multi-modal color medical image fusion Nawaz et al., 2016.

1.4.1.2 Quaternion Fourier-Mellin Moments

Inspired by the one-dimensional quaternion Fourier transform (QFT) and the Mellin transform, Guo and Zhu, 2011 proposed the quaternion Fourier-Mellin moments (QFMMs) for describing color images. They also derived the QFMMs invariants to rotation, scale and translation (RST) transformations and investigated them for the application of color image registration. This latter's purpose is to match two or more color images of the same scene taken at different times, from different viewpoints or by different sensors. The authors were motivated by the fact that the conventional image registration methods, including but not limited to correlation methods (Pratt, 1974; Althof et al., 1997), feature-based methods (Goshtasby et al., 1986; Chang et al., 1997; Wang and Chen, 1997; Alhichri and Kamel, 2003) and frequency domain methods (Chen et al., 1994; Reddy and Chatterji, 1996; Lucchese et al., 2002), deal only with gray-scale images.

1.4.1.3 Quaternion Zernike and Pseudo-Zernike Moments

Chen et al., 2010 introduced the notion of quaternion Zernike moments (QZMs), based on the orthogonal Zernike polynomial. They showed that these moments can be obtained from the conventional Zernike moments of each channel. The authors also constructed a set of combined invariants against rotation and translation (RT) transformation. The Translation invariance was obtained by placing the origin of the coordinate system at the common centroid and the modulus of the QZMs was retained for rotation invariance. The authors tested the performances of their quaternion Zernike moments compared with two existing color moments: the affine geometric deformations and diagonal photometric transformations invariants (GPDIs) (Mindru et al., 2004) and the affine color moment invariants (ACMIs) (Suk and Flusser, 2009), in terms of color image reconstruction, invariance with respect to rotation and translation as well as template matching for the real indoor scene. It was concluded that the quaternion Zernike moments have better discriminative power compared to the existing two methods. This has been explained by the fact that (1) the GPDIs and the ACMIs are nothing but a generalization of the geometric moments for multi-channel images, which means they are based on the product of the powers of the pixel coordinates and those of the intensities for one or more color channels. This is disadvantageous in the way they are not orthogonal and suffer from a high level of information redundancy, contrary to the quaternion Zernike moments, which are based on

orthogonal polynomials. (2) Using the quaternion color image representation proves to be a useful technique for color image analysis.

In another work, Chen et al., 2012 mathematically derived a set of quaternion Zernike moment invariants against rotation, scale and translation (RST) transformation based, this time, on the phase of the quaternary moments rather than their magnitudes as done in their previous work (Chen et al., 2010) and also in the work by Guo and Zhu, 2011. In the color image reconstruction test, the authors used the standard Lena and Pepper images and it has been shown how the reconstructed images from the quaternion Zernike moments are very close to the true image. The results also confirmed that the proposed quaternion descriptor is invariant under rotation, scale and translation (RST) transformations using the Amsterdam Library of Object Images (ALOI) (Geusebroek et al., 2005). The quaternion Zernike moments were also applied for color object recognition and template matching. As for color object recognition, the QZMs invariants were used along with the minimum-Euclidean distance classifier and were compared with (1) the Zernike moment invariants based on RGB decomposition and graying, (2) the modulus-based QZMs rotation invariants previously presented (Chen et al., 2010) and with (3) the QFMMs (Guo and Zhu, 2011), the GPDIs (Mindru et al., 2004) and the ACMIs (Suk and Flusser, 2009). The Columbia Object Image Library (COIL-100) (Nene et al., 1996) database was used in this experiment. The results showed that the phase-based QZMs rotation invariants have better recognition accuracy compared to the other descriptors even with corrupted images. With regard to template matching, an experiment was conducted to verify the performance of the QZMIs for real outdoor scene images after RST transformation. It has been shown that the QZMs invariants can match correctly for more templates with a reasonable error than the other moments.

Overall, the authors drawn up the conclusion that their quaternion-valued QZMs invariant descriptor is more robust to noise and has more discriminative power compared to the existing methods.

Likewise, Bei-Jing et al., 2012 derived a set of phase-based quaternion Pseudo-Zernike moments (QPZMs) invariants with respect to RST transformations for color face recognition. This latter was performed by combining these invariants with quaternion back propagation neural network (QBPNN). The results showed the efficiency of utilizing the QBPNN with the quaternion-valued QPZMs invariants in terms of robustness to the facial expression and illumination and position changes.

1.4.1.4 Quaternion Bessel-Fourier Moments

As previously mentioned, Chen et al., 2010 managed to express the quaternion Zernike moments in function of their conventional version. This can be the case of all quaternary moments, i.e., quaternary moments can be computed via their conventional versions. Shao et al., 2014 introduced the quaternion Bessel-Fourier moments (QBFMs), which can be expressed via the complex Bessel-Fourier moments (BFMs), for object reconstruction and retrieval. The authors provided a complexity study to show the added-value of calculating the QBFMs through the BFMs over calculating them directly. They also highlighted the significance of phase information versus magnitude only through a color image reconstruction example. This is reflected in the fact that using the phase coefficients of the quaternion Bessel-Fourier moments not only embody the contour but also the color components. This conclusion was their motivation for deriving quaternion-valued QBFMs RST invariants, where the scaling invariance is obtained by

using a normalization process. The authors also described an accurate estimation method for rotation angle.

The quaternion Bessel-Fourier moment invariants efficiency was compared to several quaternion orthogonal moments such as the quaternion Zernike moments (Chen et al., 2010), the quaternion Pseudo-Zernike moments (Bei-Jing et al., 2012) and the quaternion version of the orthogonal Fourier-Mellin moments, in terms of color image reconstruction and object recognition under geometric and photometric transformations. It has been shown that the reconstructed images from the quaternion Bessel-Fourier moments are the closest to the original images. Besides, the QBFMs invariants yield the best results for object recognition regardless of the number of moments or the amount of added noise.

1.4.1.5 Quaternion Radial Harmonic Fourier Moments

According to Xiang-yang et al., 2015a, the existing quaternion-based moments (quaternion Zernike moments, quaternion pseudo-Zernike moments, quaternion orthogonal Fourier-Mellin moments and so on) and their conventional versions suffer from many drawbacks recapitulated in three points:

- (1) They cannot effectively capture the image edges, which can be seen from any image reconstruction example.
- (2) These moments' computation is based on factorial terms, which causes a numerical stability problem as the number of moments increases.
- (3) They suffer from high computational complexity, especially when computing high-order moments because many factorial terms are involved.

These drawbacks incited the authors to introduce the quaternion radial harmonic Fourier moments (QRHFM) and their RST invariants for color image retrieval. In fact, the QRHFM is a generalization of the radial harmonic Fourier moments (RHFMs) (Ren et al., 2003) dealing with gray-scale images. It has been proven that the RHFMs are the most performing in terms of image reconstruction, noise robustness and rotation invariance. Besides, they are based on a triangular function as the radial function which makes them free from numerical instability issues and of a lower computational complexity, even with high-order moments, compared to other orthogonal moments.

Xiang-yang et al., 2015a showed a gray image reconstruction example comparing the Zernike moments and the radial harmonic Fourier moments, which proves that (i) for low-order moments, the image reconstruction of RHFMs is better as it well-defines the edges, and (ii) also the reconstructed images from high-order RHFMs are getting closer to the original image, unlike the Zernike moments which crack at high orders leading to numerical instability.

On the other hand, the authors provided the relationship between the QRHFM and the conventional RHFMs of each color channel, not only for moments computation but also for color image reconstruction from these moments. As for RST invariants, the translation invariance was achieved by putting the origin of coordinates at the color image centroid, the rotation invariance was achieved by taking the modulus and the scale invariance was obtained in the color image mapping to a unit-disc domain.

The authors presented a content-based color image retrieval (CBIR) scheme based on the quaternion radial harmonic Fourier moments modulus coefficients. First, the moments of non-zero orders and positive repetitions are selected as features for CBIR. Second, an overall modulus-based distance to measure the similarity between the query color image and each target color image in the database is calculated, according to which the retrieval results are sorted.

This content-based color image retrieval approach was compared to three existing content-based color image retrieval systems for real outdoor color scene images: (1) The work by Swain and Ballard, 1991, in which color histograms of multicolored objects are used for indexing into a large database of models. (2) Using running sub-blocks with different similarity weights for object-based image retrieval as proposed by Li, 2003. (3) The paper by Yap and Paramesran, 2006 where a set of Legendre chromaticity distribution moments (LCDM) are used for color image retrieval. Results showed that the quaternion radial harmonic Fourier moments modulus-based retrieval scheme performs significantly better than the other methods.

1.4.1.6 Quaternion Polar Harmonic Transform

Li, 2013 introduced the quaternion polar harmonic transform (QPHT) as a generalization of the polar harmonic transform (PHT) (Yap et al., 2010) for color images. The PHT is based on kernels that are harmonic in nature, i.e., they are basic waves. This polar transform involves three different transforms, namely, polar complex exponential transform (PCET), polar cosine transform (PCT), and polar sine transform (PST), that were extended by quaternions.

Xiang-yang et al., 2015b studied thoroughly the quaternion polar complex exponential transform (QPCET) for invariant color image description in a similar way as they did with the quaternion radial harmonic transform (Xiang-yang et al., 2015a). In fact, they adopted similar methods for establishing RST invariants as well as for content-based image retrieval. Even in their experiments, the authors used the same real outdoor scene image databases and compared the QPCET with the works by Li, 2003, Yap and Paramesran, 2006 and Wang et al., 2014b. The results showed that the QPCET modulus perform significantly better than the other methods. This conclusion incited researchers to apply the QPCET modulus for color image watermarking (Yang et al., 2015; Wang et al., 2016b). Besides, Hosny and Darwish, 2017 proposed an accurate computation of the quaternion polar complex exponential transform for color images in Cartesian and polar coordinate systems. As for Cartesian coordinates, the authors applied the Gaussian quadrature method to compute higher-order moments of QPCET as done by Singh and Upneja, 2012a, whilst for polar coordinates, an exact kernel-based method was employed to compute the higher-order moments of QPCET (Hosny and Darwish, 2016). Results showed the superiority of the QPCET in polar coordinates against the QPCET in Cartesian coordinates, which was visually assessed in color image reconstruction examples at high-order moments.

1.4.1.7 Quaternion Exponent Moments

Enlightened by the relation between the exponential function and triangular function, Jing, 2011 extended the radial harmonic Fourier moments (RHFMs) (Ren et al., 2003) and introduced the exponent moments (EMs) for invariant pattern recognition. Wang et al., 2014a and Yang et al., 2014 introduced simultaneously the quaternion exponent moments (QEMs), as an extension of the exponent moments, for color image watermarking. Through a gray image reconstruction example, the authors showed the efficiency of the exponent moments against the Zernike moments. The color images reconstructed from the QEMs also seem to be very close to the original image at high-order moments. As for watermark embedding scheme, the authors used the QEMs modulus. Later, the quaternion exponent moments and their RST invariants were exhaustively studied by Wang et al., 2016a for color image zero-watermarking. It has been proven that the magnitudes of the quaternion exponent moments remain almost invariable under various attacks, which confirms the stability of these moments. Also Wang et al., 2016c applied the QEMs modulus together with the twin support vector machines (TSVM) (Jayadeva et al.,

2007) for pixel classification based color image segmentation, and Wang et al., 2016d applied them for copy-move forgery detection.

1.4.1.8 Quaternion Circularly Semi-Orthogonal Moments

Hu et al., 2014 proposed a set of exponent Fourier moments (EFMs) for gray-level images. These moments are similar to polar complex exponential transforms (Yap et al., 2010) and radial harmonic fourier moments (Ren et al., 2003) as they are computationally inexpensive. However, the exponent Fourier moments seem to be imprecise and suffer from numerical instabilities. To avoid this problem, an improvement of these moments have been proposed by Xiao et al., 2015. But later on, Hu et al., 2016 found that these improved exponent Fourier moments are with crucial mistakes in their definition, i.e. incorrect orthogonality condition was employed to find the normalization coefficient, which brought the authors to define the corrected definition. According to Wang et al., 2016e, the circular moments above suffer from numerical and geometric errors, which can be minimized using their proposed circularly semi-orthogonal moments (CSOMs) for gray-level images. Ananth Raj, 2017 proposed the quaternary CSOMs (QSOMs) as an extension of the CSOMs to color images using the quaternion algebra. The advantages of the CSOMs, which are: (1) Higher-order moments are numerically stable than the lower-order moments. (2) The zeroth order approximation is more robust to numerical errors compared with other approximations. (3) They are computationally inexpensive as their radial function does not involve factorial terms, were the motivation of the authors to study the CSOMs quaternion version. The image rotation, translation and flipping invariants were mathematically derived. Note that the outer unit disc mapping transformation, which places the image pixels inside the unit circle was employed. Through a color image reconstruction example, it has been shown that high frequency information like edges is well preserved. Also, the quaternion circularly semi-orthogonal moments seem to be robust against geometric transformations because they are almost equal before and after rotation/translation/flipping image changes. No comparisons were made in this conference paper.

1.4.2 Clifford-based Descriptors

Since the quaternion algebra is contained in the Clifford algebra, this latter can be another direction for dealing with color images. Batard et al., 2010 introduced the Clifford-Fourier transform (CFT) with respect to a unit bivector B for color images and proved that the quaternion Fourier transform by Sangwine, 1996 is a particular case of the CFT. This latter has motivated Mennesson et al., 2010 to propose different generalized color Fourier descriptors (GCFD1 and GCFD2). On the other hand, the same authors (Mennesson et al., 2014) constructed different Clifford-based extensions of the Fourier-Mellin moments for color image description by using the Clifford-Fourier transform, namely parallel-orthogonal Fourier-Mellin descriptors (poFMD), color Fourier-Mellin descriptors (CFMD) and color Fourier-Mellin descriptors invariant under the choice of bivector B (CFMDi). In fact, Fourier-Mellin is based on the polar transformation of the image, followed by a Fourier transform then a Mellin transform.

The parallel-orthogonal Fourier-Mellin descriptors (poFMD) is obtained by concatenating the Fourier-Mellin moments computed on parallel and orthogonal parts of the color image with respect to a bivector B . Then rotation invariance is achieved by taking the modulus of these parallel-orthogonal Fourier-Mellin moments while the scale invariance is obtained by dividing by the first coefficient of each row ($poFMM_{f,B}(m, 0)$). This normalization process is explained by Sheng and Arsenault, 1986.

On the other hand, the authors were inspired by the original definition of the Fourier-Mellin moments, i.e., Fourier-Mellin is based on the polar transformation of the image, followed by a Fourier transform then a Mellin transform, to derive the color Fourier-Mellin descriptors (CFMD). This approach has also been employed by Guo and Zhu, 2011 for defining the quaternion Fourier-Mellin moments (QFMMs) presented in subsection 1.4.1.2). In fact, the color Fourier-Mellin descriptors is based on the Clifford-Fourier transform instead of the classical Fourier transform, followed by a Mellin transform. The rotation and scale invariants are obtained in the same way as for the poFMD. For the translation invariance, the center of the image is defined to be the origin of coordinates. (Mennesson et al., 2014) explained that the color Fourier-Mellin descriptors (CFMD) are preferable for images where a precise color is privileged since they depend on the choice of the bivector B . Otherwise, the CFMDi which is insensitive to the choice of the bivector B is required.

The different Clifford-based extensions of the Fourier-Mellin moments have been evaluated for image recognition using support vector machine (SVM), linear discriminant analysis (LDA) and first nearest neighbor (1-NN) classifiers. The experiments generally showed the superiority of the three proposed extensions compared to other image descriptors and particularly the parallel-orthogonal Fourier-Mellin descriptors (poFMD).

1.5 Conclusion

This review of the literature on color image descriptors suggests that this field of research is widely investigated and can be tackled from different points, i.e., the construction of novel image descriptors, the improvement of the computation accuracy, the development of fast computation algorithms, the embodiment of invariant properties to the image descriptors, etc. In fact, it is always useful to know each family's performances under a thorough examination in a variety of different problems. This leads us to study each of the methods advantages and drawbacks.

For instance, color histograms are very popular among the computer vision researchers thanks to their advantages such as efficiency, and insensitivity to small changes in camera viewpoint. However, they suffer from lacking spatial information since they rely on statistics. Besides, they are highly sensitive to illumination changes and quantization errors.

Furthermore, one can opt for the marginal approach, which consists of computing the conventional descriptors for each color channel separately, then concatenating them to produce a color image descriptor. This solution results in losing the correlation between the three channels. To overcome this limitation, one can represent the vector color pixel by means of a quaternion number or more generally a Clifford number. This technique is easy to implement as it takes advantage of the operations provided by the algebras. It also allows to preserve the correlation between the three color channels with no loss of color information.

Amongst the quaternion descriptors, the quaternion Fourier transform (QFT) and the quaternion Fourier-Mellin moments (QFMMs) are the most popular and have been used in several applications dealing with color images. However, they are not based on orthogonal functions; thus they suffer from a high level of information redundancy. On the other hand, the quaternary orthogonal moments, e.g., the quaternion Zernike moments (QZMs), quaternion pseudo-Zernike moments (QPZMs), quaternion Bessel-Fourier moments (QBFMs), quaternion radial harmonic Fourier moments (QRHFM), quaternion polar harmonic transform (QPHT), quaternion exponent moments (QEMs) and quaternion circular semi-orthogonal moments (QCSOMs), are free

from any shape information redundancy. As a result, one can use them to reconstruct the color image, which can provide an insight into their properties and limitations and allows us to decide about the required maximum order.

As for the quaternion Zernike moments (QZMs), quaternion pseudo-Zernike moments (QPZMs) and quaternion orthogonal Fourier-Mellin moments (QOFMMs), they suffer from the incapability of capturing the image edges, numerical instability and high computational complexity as the number of moments increases. These limitations can be explained by the fact these moments' computation is based on factorial terms.

It was also highlighted that the quaternary moments have been extensively used as color image descriptors for several applications such as watermarking, content-based image retrieval, object recognition and so on. Therefore, it is imperative to wisely choose the adequate descriptor that gives the best results.

In the next chapters, we will present our proposed quaternary moments with orthogonal basis functions for efficient color image retrieval and classification.

Chapter 2

Proposed Quaternary Moments for Color Image Description

2.1 Introduction

In the previous chapter, we detailed the most well-known state-of-the-art color image descriptors. It was drawn up that the quaternary moments of orthogonal basis are the most performing as they eliminate information redundancy, which explains their remarkable performances in several applications such as color image registration (Chen et al., 1994; Reddy and Chatterji, 1996; Lucchese et al., 2002), color image watermarking (Yang et al., 2015; Wang et al., 2016b; Wang et al., 2014a; Yang et al., 2014), color image retrieval (Chen et al., 2012; Bei-Jing et al., 2012; Shao et al., 2014; Xiang-yang et al., 2015a; Xiang-yang et al., 2015b), and so on.

Some of these quaternary orthogonal moments, however, e.g., the quaternion Zernike moments (QZMs) and the quaternion pseudo-Zernike moments (QPZMs), suffer from a high computational complexity especially for higher-order moments, which consequently influences their computational time. In our first contribution, we make use of the 3-dimensional spherical harmonic transform (SHT) and its 2-dimensional version namely the disc-harmonic moments (DHMs) (Ennahnahi et al., 2010), which deal with binary and gray-level object images, to construct quaternary orthogonal moment invariants intended for 2-dimensional color image description.

In this chapter, we first shed light on the numerous advantages and properties of the 3D spherical harmonic transform (SHT), e.g., rotation invariance, orthogonality as well as rapidity; thanks to which, it has been employed in a wide range of applications in different disciplines including but not limited to physics, chemistry, computer vision and has been an inspiration for many researchers aiming at efficient 3D and 2D shapes description. In particular, Ennahnahi et al., 2010 introduced the disc-harmonic moments (DHMs) as an adaptation of the spherical harmonic transform for 2D shape representation using a suitable transformation from the 3D space to the 2D space.

Second, we detail the proposed quaternary moment invariants including their computation formulas. Practically, three descriptor versions are proposed: (1) We propose a set of moment invariants, namely the quaternion disc-harmonic moments (QDHMs), by employing the same mapping sphere-to-disc as the one used by the disc-harmonic moments. (2) Then, we propose a modified version of the QDHMs based on a more adequate mapping sphere-to-disc. (3) We further enhance the computational time of the proposed moments by adapting the internal mechanism of the Hierarchical Equal Area isoLatitude Pixelization (Healpix) technique (Górski et al., 2005).

Finally, we present the experimental results that highlight the effectiveness of the proposed

moments in terms of color image reconstruction, robustness and invariance, computational time as well as content-based color image retrieval based on the k-nearest neighbor. In this latter, we also study the effects of the maximum order, the color space and the similarity measure on the retrieval performance of the proposed moments. The performances of our moments are compared with some well-known state-of-the-art quaternary and Clifford moments.

2.2 Preliminaries

This section is devoted to preliminaries regarding the spherical harmonic transform (SHT). It includes the formulas for their computation as well as some of their applications in several fields dealing with 3D and 2D objects.

2.2.1 3-Dimensional Spherical Harmonics

The spherical harmonics (SHs) are special functions that satisfy Laplace's equation. As their name indicates, the spherical harmonics are defined on the surface of the unit sphere and are separable in spherical coordinates. The SHs are applied in a wide range of fields such quantum chemistry and atomic physics, e.g., electron configuration, quantum mechanics, e.g., atomic orbital, physics, e.g., gravitational field, 3D computer graphics, e.g., indirect lighting and modeling of 3D objects, and so on.

The spherical harmonics of order n and repetition m , denoted $Y_{n,m}(\theta, \varphi)$ where $0 \leq \theta \leq \pi$ is the colatitude and $0 \leq \varphi \leq 2\pi$ is the longitude, are expressed as follows:

$$Y_{n,m}(\theta, \varphi) = N_{n,m} P_{n,m}(\cos(\theta)) e^{jm\varphi}, \text{ such as } |m| \leq n, \quad (2.1)$$

where $N_{n,m}$ is the normalization coefficient defined as

$$N_{n,m} = \sqrt{\frac{(2n+1)(n-m)!}{4\pi(n+m)!}}, \quad (2.2)$$

$P_{n,m}$ is the real-valued associated Legendre polynomial and $e^{jm\varphi}$ is the complex kernel.

The spherical harmonics are indexed according to two integer constants, the order n and the repetition m , satisfying the constraint that $0 \leq n \leq nMax$ and $-n \leq m \leq n$; thus, there are $2n+1$ functions of order n and in total $(nMax+1)^2$ functions, where $nMax$ is the maximum order.

Eq. (2.1) leads us to the real and imaginary parts of the complex SHs given by:

$$\begin{aligned} \Re(Y_{n,m}(\theta, \varphi)) &= N_{n,m} P_{n,m}(\cos(\theta)) \cos(m\varphi), \\ \Im(Y_{n,m}(\theta, \varphi)) &= N_{n,m} P_{n,m}(\cos(\theta)) \sin(m\varphi). \end{aligned} \quad (2.3)$$

The spherical harmonics verify the orthonormality property according to the following equation:

$$\int_{\theta=0}^{\pi} \int_{\varphi=0}^{2\pi} Y_{n,m} Y_{n',m'} \sin \theta d\theta d\varphi = \delta_{n,n'} \delta_{m,m'}, \quad (2.4)$$

where $\delta_{n,n'}$ and $\delta_{m,m'}$ are Kronecker deltas.

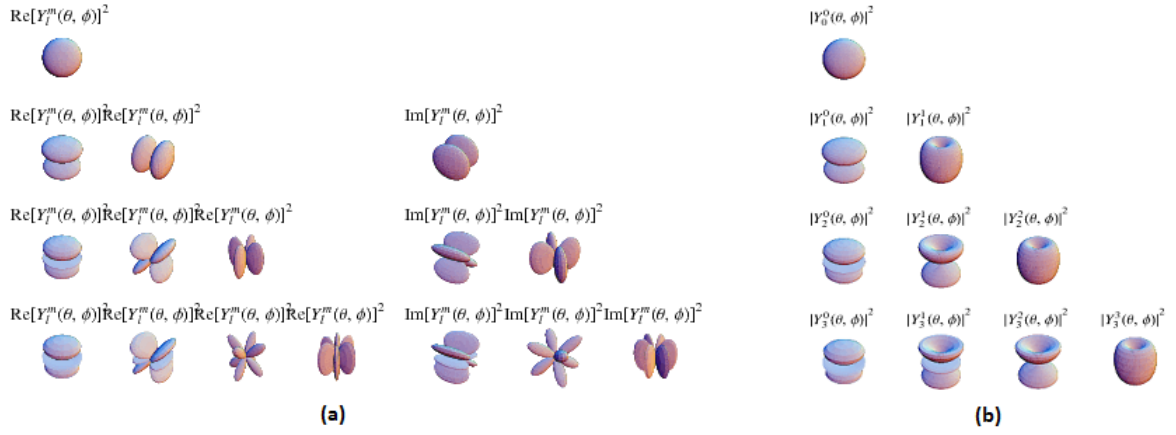


Figure 2.1 (a) Square of the real and imaginary parts of the spherical harmonics of order up to 3. (b) square of the modulus of the spherical harmonics of order up to 3.

Figure 2.1 shows the $\Re^2(Y_{n,m}(\theta, \varphi))$, $\Im^2(Y_{n,m}(\theta, \varphi))$ and $\|Y_{n,m}(\theta, \varphi)\|^2$ of order n up to 3, some of which are given in spherical coordinates as follows:

$$\begin{aligned}
 n = 0 & \left\{ \begin{aligned} Y_{0,0}(\theta, \varphi) &= \frac{1}{2} \frac{1}{\sqrt{\pi}} \end{aligned} \right. \\
 n = 1 & \left\{ \begin{aligned} Y_{1,-1}(\theta, \varphi) &= \frac{1}{2} \sqrt{\frac{3}{2\pi}} \sin \theta e^{-i\varphi} \\ Y_{1,0}(\theta, \varphi) &= \frac{1}{2} \sqrt{\frac{3}{\pi}} \cos \theta \\ Y_{1,1}(\theta, \varphi) &= -\frac{1}{2} \sqrt{\frac{3}{2\pi}} \sin \theta e^{i\varphi} \end{aligned} \right. \\
 n = 2 & \left\{ \begin{aligned} Y_{2,-2}(\theta, \varphi) &= \frac{1}{4} \sqrt{\frac{15}{2\pi}} \sin^2 \theta e^{-2i\varphi} \\ Y_{2,-1}(\theta, \varphi) &= \frac{1}{2} \sqrt{\frac{15}{2\pi}} \sin \theta \cos \theta e^{-i\varphi} \\ Y_{2,0}(\theta, \varphi) &= \frac{1}{4} \sqrt{\frac{5}{\pi}} (3 \cos^2 \theta - 1) \\ Y_{2,1}(\theta, \varphi) &= -\frac{1}{2} \sqrt{\frac{15}{2\pi}} \sin \theta \cos \theta e^{i\varphi} \\ Y_{2,2}(\theta, \varphi) &= \frac{1}{4} \sqrt{\frac{15}{2\pi}} \sin^2 \theta e^{2i\varphi} \end{aligned} \right. \quad (2.5)
 \end{aligned}$$

From these spherical harmonics values, we can see that there is an important relation connecting $Y_{n,-m}$ with $Y_{n,m}$ that is:

$$Y_{n,-m} = (-1)^m Y_{n,m}^*, \quad (2.6)$$

where X^* is the conjugate of the complex number X .

2.2.2 Spherical Harmonic Transform Definition

In mathematics, a transform maps a function from a domain to another. In image processing, transforms are generally related to mapping an image from spatial domain to frequency domain. In fact, one can express the integrable original function, i.e., image, in terms of the sum of its projections onto a set of orthonormal basis functions via inner product. The spherical harmonics are basis functions that are the essence for the spherical harmonic Transform where the original function is integrable on the unit sphere.

The spherical harmonic transform (SHT) of a complex-valued function $f^O(\theta, \varphi)$ defined on the unit sphere is expressed using the complex inner product as follows (Abramowitz, 1974):

$$SHT_{n,m}(f) = \int_0^\pi \int_0^{2\pi} Y_{n,m}^*(\theta, \varphi) f^O(\theta, \varphi) \sin(\theta) d\theta d\varphi, \quad (2.7)$$

where θ is the colatitude, φ is the longitude, X^* denotes the conjugate of the complex X and $Y_{n,m}$ is the complex-valued spherical harmonics.

On the other hand, one can perform the inverse transform, also called reconstruction, which is simply achieved by summing up all the spherical harmonic transform components multiplied by their corresponding basis function as follows:

$$f^R(\theta, \varphi) = \sum_{n=0}^{\infty} \sum_{m=-n}^n Y_{n,m} SHT_{n,m}(f). \quad (2.8)$$

Note that the inverse transform uses the conjugate of the spherical harmonics, which prevents a sign change occurring in the reconstruction process. The reconstructed function $f^R(\theta, \varphi)$ tends to the exact original image $f^O(\theta, \varphi)$ as n tends to infinity. Practically, the reconstruction is limited to a finite order n denoted $nMax$, which occurs in truncation errors. This can also be concluded from Figure 2.2 which gives 3D objects reconstruction from the SHT.

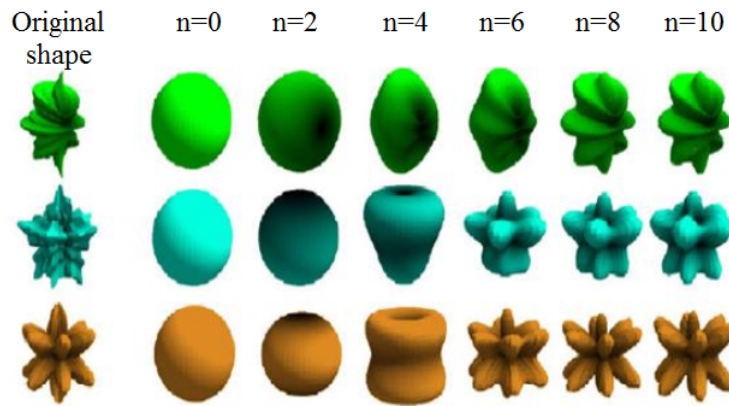


Figure 2.2 Reconstruction of 3D objects by the spherical harmonic transform up to different maximum orders.

The amplitude information of the spherical harmonic transform are rotationally invariant as shown in Figure 2.3; thus, in many applications, the phase information is discarded. Figure 2.3 shows that an object and its rotated version have the same magnitude-based SHT descriptor.

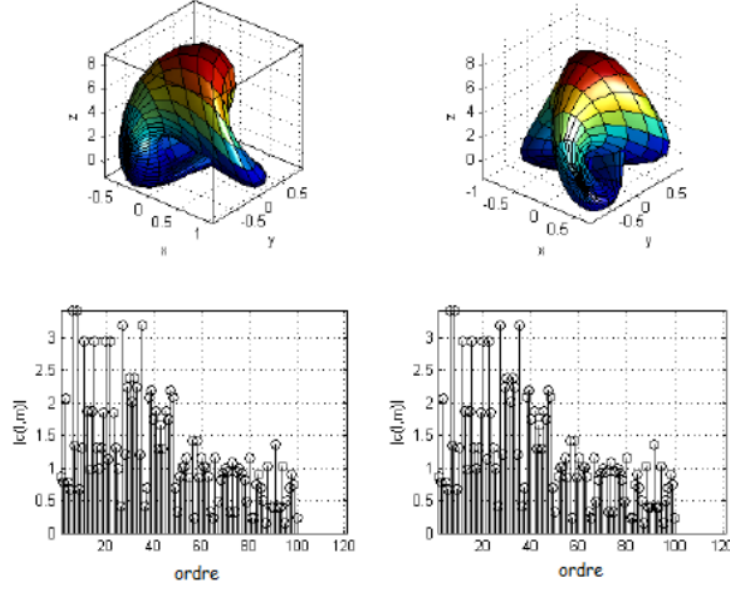


Figure 2.3 Rotational invariance of the amplitude coefficients of the SHT.

2.2.3 Associated Legendre Polynomials

The associated Legendre polynomials occur when solving Laplace's equation in spherical coordinates; thus, they play an important role in the definition of the spherical harmonics, which in turn form a complete orthonormal set of functions in the sense of Fourier series.

These polynomials, denoted $P_{n,m}(x)$, are real valued where $x \in [-1, 1]$ and accept two integer parameters, i.e., the order n and repetition m such as $0 \leq m \leq n$, or with trivially equivalent negative values.

There are many ways to evaluate the associated Legendre polynomials numerically (Abramowitz, 1974; Press et al., 2007). The computational method based on recurrences appears to be the most robust, stable and rapid. These recurrence relations are as follows:

$$(n - m)P_{n,m}(x) = x(2n - 1)P_{n-1,m}(x) - (n + m - 1)P_{n-2,m}(x). \quad (2.9)$$

The starting value is

$$P_{m,m}(x) = (-1)^m (2m - 1)!! (1 - x^2)^{\frac{m}{2}}, \quad (2.10)$$

(The notation $n!!$ denotes the product of all odd integers less than or equal to n .) With $n = m + 1$, and setting $P_{m-1,m}(x) = 0$, we find

$$P_{m+1,m}(x) = x(2m + 1)P_{m,m}(x). \quad (2.11)$$

Equations (2.10) and (2.11) provide the two starting values required for (2.9) for general n and positive m . On the other hand, the polynomials for negative values of m are proportional to those of positive m as follows:

$$P_{n,-m}(x) = (-1)^m \frac{(n - m)!}{(n + m)!} P_{n,m}(x). \quad (2.12)$$

Note that it is from this relation between the associated Legendre polynomials of positive and negative m , that the symmetry of the spherical harmonics, given in Eq. (2.6) arose.

Figure 2.4 shows a plot of the first few associated Legendre polynomials.

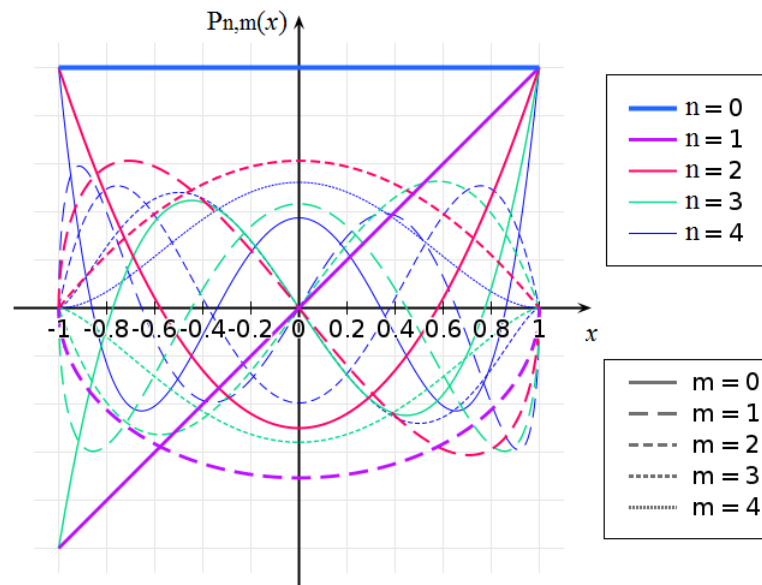


Figure 2.4 The first few associated Legendre polynomials.

2.2.4 Applications of the Spherical Harmonic Transform for 3D Models

The spherical harmonic transform is an essential tool for solving physical and computer vision problems related to 3D models, e.g., multipole electrostatic and electromagnetic fields, atomic orbital electrons, planetary bodies and stars, cosmic microwave background radiation, 3D objects and so on.

One of the several applications of the spherical harmonic transform for 3D models is called spherical harmonic lighting (SH lighting), which was first introduced by Sloan et al., 2002. This is a technique for calculating the lighting on 3D models from area light sources that allows us to capture, relight and display global illumination style images in real time. Later, Green, 2003 provided gritty details about SH lighting, including the background, which was lacked in the first paper (Sloan et al., 2002), and he explained how the spherical harmonic lighting can be added to video games. An example of SH lighting is given in Figure 2.5.

Another widely investigated application of the spherical harmonic transform is 3D shape matching, whose principal challenge is to consider to be the same two models that differ by a rotation. For example, Kazhdan et al., 2003 proposed an SHT-based mathematical tool for 3D models rotation invariant representation, which allowed improving the matching performances of many descriptors as well as reducing the dimensionality of the descriptor, thus reducing the time and space requirements.

In order to obtain a spherical rotation invariant descriptor, the authors adopted a method which consists of (1) decomposing the spherical function into its harmonics, (2) summing the harmonics within each frequency, and (3) computing the modulus of each frequency component.

They also showed that this method is not only reserved to spherical shape descriptors, but can be generalized to the case of voxel descriptors.



Figure 2.5 Example of spherical harmonic lighting on a 3D car object (Green, 2003).

For 3D matching, one can simply compute the Euclidean distance between the descriptors, then for a given query, the k nearest neighbors are returned.

On the other hand, Funkhouser et al., 2003 established a search engine for 3D models based on a shape query, 3D and 2D sketch query, text query as well as multimodal query. Figure 2.6 shows results for different types of queries.

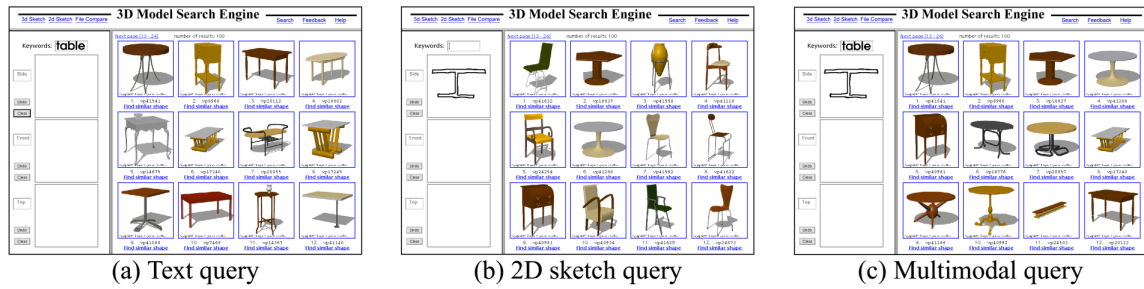


Figure 2.6 Results for different types of queries by the 3D-model search engine developed by Funkhouser et al., 2003.

For 3D shape queries, the authors' method for 3D rotation invariant shape representation is quite similar to the one presented by Kazhdan et al., 2003 except that polygons within the 3D model are first rasterized into a voxel grid. This is shown in Figure 2.7.

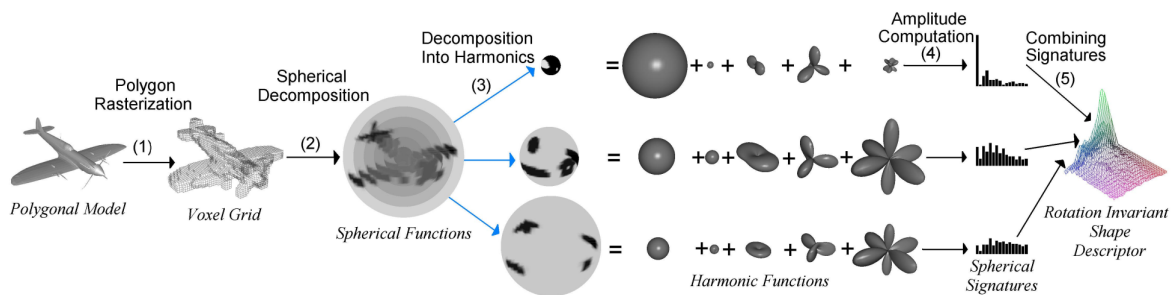


Figure 2.7 Method proposed by Funkhouser et al., 2003 for rotation invariant spherical shape representation based on the SHT.

2.2.5 Adaptation of the Spherical Harmonic Transform to 2D Shape Representation

2.2.5.1 2.5D Spherical Harmonic Transform

Pu and Ramani, 2006 proposed a method, namely 2.5D spherical harmonics transformation, for retrieving 2D drawings by measuring their shape similarity. This is established by transforming the 2D drawing from a 2D space into a 3D space, hence the name 2.5D comes from, then employing the fast spherical harmonic transform to get a rotation invariant descriptor.

In fact, the 2D drawing is represented as a spherical function as shown in Figure 2.8 by following these steps:

1. Determining a sphere that satisfies the following conditions:
 - Its center is the center of the bounding box of this drawing.
 - Its radius is equal to half the longer diagonal length of bounding box. This is to guarantee a complete enclosing of the drawing in the sphere.
 - The 2D drawing lies in the equator plane of the sphere.
2. Intersection of the 2D drawing with the sphere by :
 - Generating a set of rays uniformly in plane xy where the 2D drawing lies. These rays start from the center of the sphere.
 - Computing the intersections between these rays and the 2D drawing.
 - Transforming all intersection points in plane xy into a spherical function in plane xyz by introducing a new variable, which is the longitude φ .

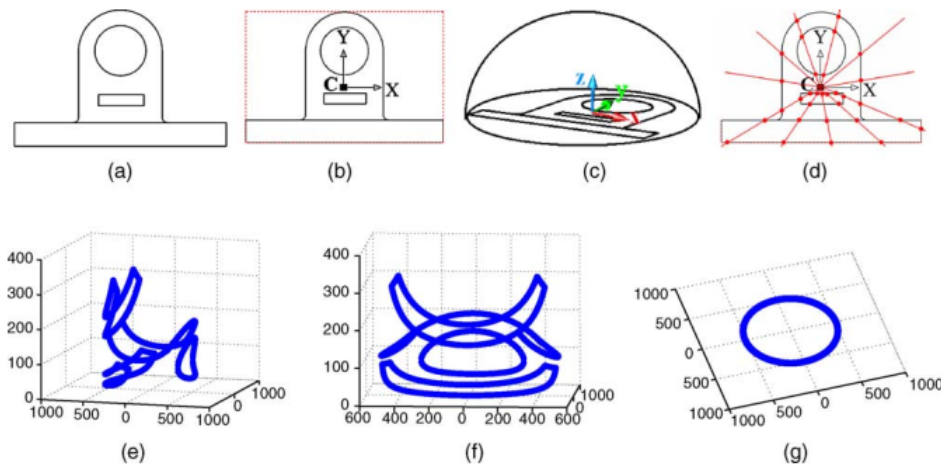


Figure 2.8 Method proposed by Pu and Ramani, 2006 for representing a 2D drawing as a spherical function.

After having represented the 2D drawing as a spherical function, one can perform a normalization step to obtain scale invariance then easily apply the fast spherical harmonic transform as proposed by Healy et al., 2003. The modulus coefficients at each frequency are retained in order to guarantee rotation invariance and the Euclidean distance is used as similarity measure.

2.2.5.2 Connectivity

Sajanhar et al., 2004 introduced the concept of Connectivity for transforming a 2D shape into 3D space. Later on, the same authors (Sajjanhar et al., 2005) used this concept for 2D shape

retrieval based on the spherical harmonic transform just like Pu and Ramani, 2006 did.

The Connectivity captures the pixel neighborhood information and it is calculated to be a number between 0 and 8 corresponding to the number of the nearest 8-neighbors deemed to be dark, i.e., has intensity below a predefined threshold. Connectivity information is represented in the axis z in Cartesian coordinates.

Figure 2.9 (b) shows the Connectivity information added to the 2D original shape in Figure 2.9 (a) from different perspectives, then triangulated into the polygon mesh in Figure 2.9 (c).

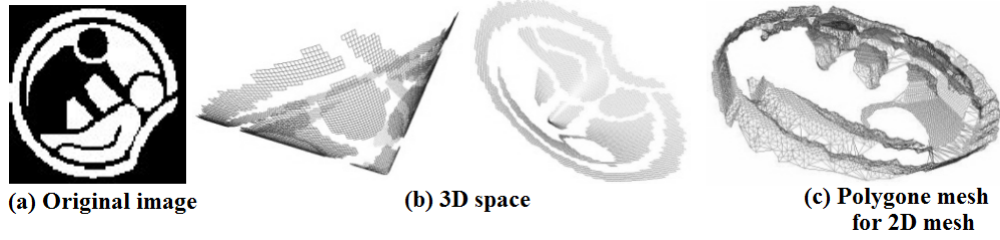


Figure 2.9 Method proposed by Sajjanhar et al., 2004; Sajjanhart et al., 2005 for representing a 2D shape in 3D space.

2.2.5.3 Disc-Harmonic Moments

The disc-harmonic moments (DHMs) were introduced by Ennahnahi et al., 2010 as an adaptation of the spherical harmonic transform for a faster and more robust 2D shape description for a retrieval purpose. Unlike 2.5DSHT and Connectivity which try to adapt the 2D shape to the sphere, the disc-harmonic moments do not require a 3D representation in order to apply the spherical harmonic transform (SHT). In fact, inspired by the SHT, the DHMs were proposed as a set of orthogonal moments defined on the unit disc, which can be put under orthogonal moments' category. Practically, the authors adapted the spherical harmonics in order to conceive a transformation defined on the unit disc. The real and imaginary parts of the complex-valued harmonic function employed by the disc-harmonic moments up to order $n = 8$ are shown in Figure 2.10 according to the following equation:

$$\begin{aligned} H_{n,m}(r, \varphi) &= N_{n,m} P_{n,m}(r) e^{jm\varphi}, \\ &= N_{n,m} P_{n,m}(r) \cos(m\varphi) + j(N_{n,m} P_{n,m}(r) \sin(m\varphi)). \end{aligned} \quad (2.13)$$

where (r, φ) are polar coordinates.

The complex-valued disc-harmonic moments (DHMs) of order n and repetition m are expressed for a gray-scale image $f_C(r, \varphi)$ as follows:

$$CDHM_{n,m}(f_C) = \int_0^{2\pi} \int_0^1 f_C(r, \varphi) H_{n,m}^*(r, \varphi) r dr d\varphi, |m| \leq n, \quad (2.14)$$

where X^* is the conjugate of the complex number X .

Evidently, the disc-harmonic moments inherit the properties of the spherical harmonic transform including completeness, orthonormality, rotation invariance and so on. The translation and scaling invariance were obtained using a pre-processing step, which moves the shape to

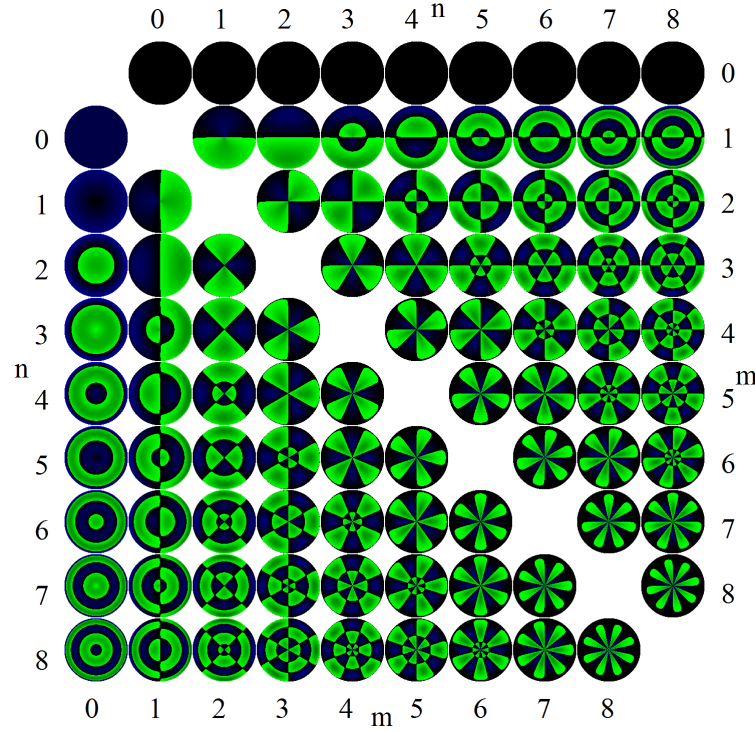


Figure 2.10 Complex-valued harmonic function up to order $n = 8$ employed by Ennahnahi et al., 2010 for representing a 2D shape with the spherical harmonic transform. The top triangle represents the imaginary parts whilst the bottom triangle represents the real parts. The repetition m varies between 0 and n . Green pixels represent regions where the function is negative, blue ones represent where it is positive and the black ones are where the function is zero.

the center and normalizes all the images to a standard scale. With regard to rotation invariance, the modulus of the moments is employed. It was experimentally shown that the disc-harmonic moments outperform some contour-based descriptors, e.g., Fourier descriptor (Zhang and Lu, 2002; Zhang and Lu, 2001) and curvature scale space descriptor (CSSD) (Mokhtarian et al., 1996; Abbasi et al., 1999; Abbasi et al., 2000; Mokhtarian and Mackworth, 1992), some region-based descriptors, e.g., Zernike moments descriptor (Teague, 1980; Chong et al., 2003), angular-radial transform descriptor (ARTD) (Lee and Kim, 2012) and geometric moments descriptor (Hu, 1962; Teh and Chin, 1988; Taubin and Cooper, 1991), as well as some spherical harmonic transform-based methods, e.g., Connectivity (Sajjanahart et al., 2005), 2DSHT (Funkhouser et al., 2003) and 2.5DSHT (Pu and Ramani, 2006) in terms of image retrieval.

2.3 Set of Quaternary Orthogonal Moments

In the previous section, we saw that although the spherical harmonic transform (SHT) is defined over the sphere, it can be used for 2D shape representation. According to Ennahnahi et al., 2010, the spherical harmonic transform's 2D version, namely the disc-harmonic moments (DHMs), allows a faster and more robust 2D shape description compared to the other SHT-based methods, i.e., 2.5DSHT and Connectivity.

In fact, the disc-harmonic moment invariants can capture the significant image features while remaining stable under image transformations like translation, scale and rotation. They are also

less likely affected by noise and variations. Furthermore, similarly to the spherical harmonic transform, the disc-harmonic moments are based on the spherical harmonic orthogonal basis; thus they eliminate information redundancy and allow image reconstruction.

Despite all the advantages of the disc-harmonic moments, these latter can only capture the monochromatic aspect of the image as they are complex valued. In order to incorporate color information, we propose a set of region-based 2D color shape descriptors inspired by the SHT and the algebra of quaternions (Hamilton, 1844).

2.3.1 Operations on Quaternions

A quaternion Hamilton, 1844 is a generalization of a complex number. It consists of one real part and three imaginary parts as $q = q^a + q^b i + q^c j + q^d k$.

The imaginary units i , j and k follow the combinations:

$$\begin{aligned} i^2 = j^2 = k^2 &= -1, \\ ij = k, jk &= i, ki = j, ji = -k, kj = -i, ik = -j. \end{aligned} \quad (2.15)$$

The conjugate and the modulus of a quaternion number q are respectively defined by

$$q^* = q^a - q^b i - q^c j - q^d k, \quad (2.16)$$

$$||q|| = \sqrt{(q^a)^2 + (q^b)^2 + (q^c)^2 + (q^d)^2}. \quad (2.17)$$

The dot product of two quaternions q_1 and q_2 is defined as:

$$q_1 \bullet q_2 = q_1^a q_2^a + q_1^b q_2^b i + q_1^c q_2^c j + q_1^d q_2^d k, \quad (2.18)$$

and the cross product as:

$$\begin{aligned} q_1 \otimes q_2 &= (q_1^a q_2^a - q_1^b q_2^b - q_1^c q_2^c - q_1^d q_2^d) + (q_1^a q_2^b + q_1^b q_2^a + q_1^c q_2^d - q_1^d q_2^c) i \\ &\quad + (q_1^a q_2^c - q_1^b q_2^d + q_1^c q_2^a + q_1^d q_2^b) j + (q_1^a q_2^d + q_1^b q_2^c - q_1^c q_2^b + q_1^d q_2^a) k. \end{aligned} \quad (2.19)$$

The quaternions space, or also called Hamilton space, is denoted $\mathbb{H}$.

2.3.2 Quaternion Disc-Harmonic Moments

2.3.2.1 Sphere-Disc Mapping

The key idea behind our first contribution is to generalize the spherical harmonic transform from complex domain to quaternion domain in order to take into consideration the color information contained in the image. These proposed set of moments are called quaternion disc-harmonic moments (QDHMs), which inherit the numerous properties of the spherical harmonic transform, i.e., discrimination power, orthonormality of the basis function, image reconstruction capability and so on. On the other hand, our QDHMs are destined to 2D image representation; thus a reformulation of the spherical harmonics so as to accept the unit disc as domain of definition is required. This is established by means of a mapping from spherical coordinate system to polar

coordinates. In this work, we adopt the same mapping sphere-disc as by Ennahnahi et al., 2010 as following:

$$\sin(\theta) = \sqrt{1 - r^2} \quad , \quad \cos(\theta) = r \quad , \quad \varphi_{sphere} = \varphi_{disc}, \quad (2.20)$$

where $(\theta, \varphi_{sphere})$ is spherical coordinate and (r, φ_{disc}) is polar coordinate, such as $0 \leq \varphi \leq 2\pi$, $0 \leq \theta \leq \pi/2$ and $0 \leq r \leq 1$. Over the sphere, θ varies between 0 and π , nevertheless, as r must be positive, the mapping to the unit disc was restricted only to the northern hemisphere ($0 \leq \theta \leq \pi/2$) as shown in Figure 2.11.

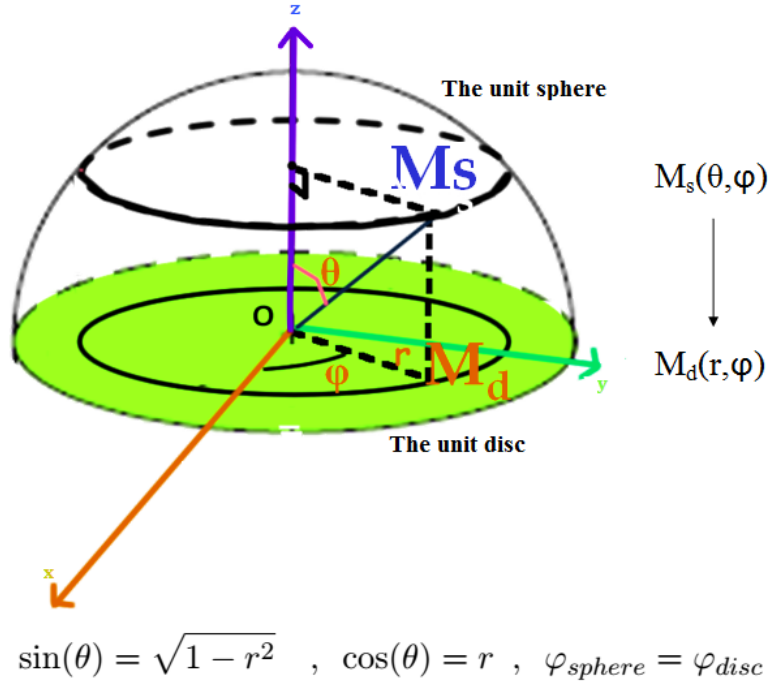


Figure 2.11 Sphere-to-disc mapping adopted by Ennahnahi et al., 2010.

2.3.2.2 Basis Functions

The quaternion-valued harmonic functions defined on the unit disc are derived from the spherical harmonics given in Eq. (2.1), based on the mapping sphere-disc used by Ennahnahi et al., 2010 and by replacing the kernel $e^{jm\varphi}$ with $e^{\mu m\varphi}$, as follows:

$$H_{n,m}(r, \varphi) = N_{n,m} P_{n,m}(r) e^{\mu m\varphi}. \quad (2.21)$$

Here μ is a unit pure quaternion:

- Unit: This property is needed to provide rotation invariance.
- Pure: Each imaginary part corresponds to a color space channel.
- μ is, generally, chosen to be the gray line in the RGB space or the luminance axis, i.e., $\mu = (i + j + k)/\sqrt{3}$. The three imaginary parts are weighted with the same factor $1/\sqrt{3}$ in order to provide the same privileges to the three color space channels.

Recall that $N_{n,m}$ is the normalization coefficient previously given in Eq. (2.2) and $P_{n,m}$ are the associated Legendre polynomials detailed in subsection 2.2.3.

Thus, the basis functions in Eq. (2.21) can be written in their quaternion form as follows:

$$\begin{aligned}
 H_{n,m}(r, \varphi) &= N_{n,m} P_{n,m}(r) e^{\mu m \varphi}, \\
 &= N_{n,m} P_{n,m}(r) (\cos(m\varphi) + \mu \sin(m\varphi)), \\
 &= N_{n,m} P_{n,m}(r) \cos(m\varphi) + \frac{N_{n,m} P_{n,m}(r) \sin(m\varphi)}{\sqrt{3}} i \\
 &\quad + \frac{N_{n,m} P_{n,m}(r) \sin(m\varphi)}{\sqrt{3}} j + \frac{N_{n,m} P_{n,m}(r) \sin(m\varphi)}{\sqrt{3}} k.
 \end{aligned} \tag{2.22}$$

Visual representations of the real and first imaginary parts of the quaternion-valued harmonic functions given in Eq. (2.22) of order up to $n = 8$ and $0 \leq m \leq n$ are illustrated in Figure 2.14 (a). It is obvious that the imaginary parts of the harmonic functions of repetition $m = 0$ are zero. Also, the imaginary parts can be obtained by performing an adequate rotation over the real parts of these harmonic functions.

2.3.2.3 QDHMs Definition

In this subsection, we show how we derive an efficient color region-based shape descriptor, namely the quaternion disc-harmonic moments (QDHMs), from the quaternion-valued disc-harmonic functions described in the previous subsection. The proposed set of moments arise from the employment of the disc-harmonic transform analogous to what has been explained in subsection 2.2.5.3.

Let $f(r, \varphi)$ be a color image defined on the unit disc where (r, φ) are polar coordinates. According to Sangwine, 1996, this color image can be represented as a pure quaternion, where the real part is set to zero and the three imaginary parts are three color channels, as follows:

$$f(r, \varphi) = f_R(r, \varphi)i + f_G(r, \varphi)j + f_B(r, \varphi)k, \tag{2.23}$$

where $f_R(r, \varphi)$, $f_G(r, \varphi)$ and $f_B(r, \varphi)$ represent respectively the red, green and blue channel of the color image.

In other color spaces such as HSV and XYZ, the components f_R, f_G, f_B are simply replaced by the first, the second and the third channel of the corresponding color space.

The right-sided QDHMs of order n and repetition m are given by:

$$QDHM_{n,m}^R(f) = \int_0^{2\pi} \int_0^1 f(r, \varphi) \otimes H_{n,m}^*(r, \varphi) r dr d\varphi, \text{ with } |m| \leq n, \tag{2.24}$$

where $H_{n,m}^*(r, \varphi)$ is the conjugate of the quaternion-valued disc-harmonic functions and $\otimes$ is the quaternary cross product, given respectively in Eq. (2.16) and Eq. (2.19).

As the multiplication of quaternions is not commutative, we also define the left-sided QDHMs by moving the $H_{n,m}^*(r, \varphi)$ part in Eq. (2.24) to the left-side of $f(r, \varphi)$ as follows:

$$QDHM_{n,m}^L(f) = \int_0^{2\pi} \int_0^1 H_{n,m}^*(r, \varphi) \otimes f(r, \varphi) r dr d\varphi, \text{ with } |m| \leq n. \tag{2.25}$$

In this work, we will use the right-sided QDHMs denoted by the abbreviation 'QDHM _{n,m} '.

In order to numerically approximate the QDHMs, the simplest way is to perform a zero-order approximation which consists of replacing the double integral in Eq. (2.24) with a double summation on the surface of the disc as follows:

$$QDHM_{n,m}(f) \approx \frac{1}{N^2} \sum_0^1 \sum_0^{2\pi} f(r, \varphi) \otimes H_{n,m}^*(r, \varphi), \quad \text{with } |m| \leq n. \quad (2.26)$$

Since the basis function $H_{n,m}$ is independent of the image content, it can be pre-computed and stored. Note that:

- The multiplication of a pure quaternion ($f(r, \varphi)$) and a quaternion number ($H_{n,m}$) requires 12 multiplications and 8 additions.
- The sum of two quaternions requires 4 additions.

Therefore, computing the QDHMs using their direct formula given in Eq. (2.24) necessitates $(12N^2)$ multiplications and $(8N^2 + 4(N^2 - 1) = 12N^2 - 4)$ additions.

2.3.2.4 Relation between QDHMs and DHMs

Starting from the definition of the QDHMs given in Eq. (2.24) and substituting $f(r, \varphi)$ by Eq. (2.23) and $H_{n,m}$ by Eq. (2.22) while using the operations on quaternions given in equations (2.15), (2.16), (2.19), we can obtain the relationship between the QDHMs and the complex-valued DHMs of each color channel as follows:

$$QDHM_{n,m}(f) = A_{n,m} + iB_{n,m} + jC_{n,m} + kD_{n,m}, \quad (2.27)$$

where

$$\begin{aligned} A_{n,m} &= -\frac{1}{\sqrt{3}} [\Im(CDH M_{n,m}(f_R)) + \Im(CDH M_{n,m}(f_G)) + \Im(CDH M_{n,m}(f_B))], \\ B_{n,m} &= \Re(CDH M_{n,m}(f_R)) + \frac{1}{\sqrt{3}} [\Im(CDH M_{n,m}(f_G)) - \Im(CDH M_{n,m}(f_B))], \\ C_{n,m} &= \Re(CDH M_{n,m}(f_G)) + \frac{1}{\sqrt{3}} [\Im(CDH M_{n,m}(f_B)) - \Im(CDH M_{n,m}(f_R))], \\ D_{n,m} &= \Re(CDH M_{n,m}(f_B)) + \frac{1}{\sqrt{3}} [\Im(CDH M_{n,m}(f_R)) - \Im(CDH M_{n,m}(f_G))]. \end{aligned} \quad (2.28)$$

Here $CDH M_{n,m}(f_R)$, $CDH M_{n,m}(f_G)$ and $CDH M_{n,m}(f_B)$ represent respectively the complex-valued DHMs [Eq. (2.13)] for the red, green and blue channel. $\Re(X)$ and $\Im(X)$ are respectively the real part and the imaginary part of the complex number X .

Computing the QDHMs via the complex-valued DHMs as in Eq. (2.28) requires to numerically approximate the DHMs. These latter can be computed from Eq. (2.14) using the zero-order approximation as follows:

$$CDH M_{n,m}(f_C) \approx \frac{1}{N^2} \sum_0^1 \sum_0^{2\pi} f_C(r, \varphi) H_{n,m}^*(r, \varphi), \quad \text{with } |m| \leq n. \quad (2.29)$$

where $f_C(r, \varphi)$ is the real-valued binary image of size $N \times N$ and $H_{n,m}$ is the complex-valued disc-harmonic function.

The $CDHM_{n,m}$ involves $(2N^2)$ multiplications and $2(N^2 - 1) = (2N^2 - 2)$ additions. And since the computation of the QDHMs using Eq. (2.28) requires the computation of three DHMs for each color channel, 8 additions and 4 multiplications, the total necessary additions and multiplications are respectively $(6N^2 + 2)$ and $(6N^2 + 4)$.

Table 2.1 summarizes the number of operations consumed in both the direct computation of the QDHMs and the computation using the DHMs of each color channel for positive values of m .

Table 2.1 The computational complexity of the quaternion disc-harmonic moments.

	Computational complexity	
	Additions	Multiplications
Direct computation of the QDHMs	$12N^2 - 4$	$12N^2$
Computation of the QDHMs via the DHMs	$6N^2 + 2$	$6N^2 + 4$

It can be seen that the computation of the QDHMs via the DHMs reduces considerably the number of operations for positive values of m .

On the other hand, the QDHMs for negative values of m can be expressed in function of the DHMs of negative values of m by exploiting the following symmetry property of the DHMs, which is analogous to the symmetry of SHT defined in Eq. (2.6):

$$\begin{aligned}\Re(CDH M_{n,m}(r, \theta)) &= (-1)^m \Re(CDH M_{n,-m}(r, \varphi)), \\ \Im(CDH M_{n,m}(r, \theta)) &= (-1)^{m+1} \Im(CDH M_{n,-m}(r, \varphi)).\end{aligned}\tag{2.30}$$

The QDHMs for negative values of m are:

$$\begin{aligned}A_{n,-m} &= (-1)^{m+1} A_{n,m}, \\ B_{n,-m} &= (-1)^{m+1} [-2\Re(CDH M_{n,m}(f_R)) + B_{n,m}], \\ C_{n,-m} &= (-1)^{m+1} [-2\Re(CDH M_{n,m}(f_G)) + C_{n,m}], \\ D_{n,-m} &= (-1)^{m+1} [-2\Re(CDH M_{n,m}(f_B)) + D_{n,m}].\end{aligned}\tag{2.31}$$

Since the QDHMs and the DHMs for positive values of m are already computed, they will not be considered. Therefore, the QDHMs for negative values of m need 7 multiplications and 3 additions.

In this work, we compute the DHMs using the S2kit tool (Kostelec and Rockmore, 2004), which is known for performing a fast computation of the spherical harmonic transform. In fact, the real function $f_C(r_i, \varphi_j)$, where C is the corresponding color channel, will be sampled on a uniform $N \times N$ grid such as $r_i = (2i + 1)/2N$ and $\varphi_j = 2\pi j/N$ where $i = 0, \dots, N - 1$ and $j = 0, \dots, N - 1$.

Thus the computational time of the DHMs can be accelerated using the following:

$$\begin{aligned} CDHM_{n,m}(f_C) &= \frac{\sqrt{2\pi}N_{n,m}}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} w_i P_{n,m}(r_i) f_C(r_i, \varphi_j) e^{-jm\varphi_j}, \\ &= \frac{\sqrt{2\pi}N_{n,m}}{N} \left(\sum_{i=0}^{N-1} w_i P_{n,m}(r_i) \left(\sum_{j=0}^{N-1} f_C(r_i, \varphi_j) e^{-jm\varphi_j} \right) \right). \end{aligned} \quad (2.32)$$

where w_i are the Chebyshev-Gauss quadrature weights (Abramowitz, 1974).

In Eq. (2.32) we perform the *Separation of variables* technique. In fact, the DHMs are computed by doing a bunch of discrete Legendre transforms, followed by a bunch of “regular” Fast Fourier Transforms (FFTs) (Dilts, 1985). This does not require the computation of the associated Legendre polynomials at each point of the unit disc but only at a limited number of r values. Therefore, the DHMs’ computation is faster.

In the work by Driscoll and Healy, 1994, the authors announce that the computational complexity of this algorithm is $O(n(\log n)^2)$ which is asymptotically better than the naive algorithm $O(n^2)$, where n is the number of sampled values of the function f_C .

2.3.2.5 Some Properties

The proposed quaternion disc-harmonic moments (QDHMs) inherit the property of completeness and orthogonality on the unit disc according to the following formula:

$$\int_0^{2\pi} \int_0^1 [H_{n,m}]^* \otimes H_{n',m'} r dr d\varphi = \delta_{nn'} \delta_{mm'}, \quad (2.33)$$

where δ is the Kronecker symbol.

This property makes the proposed moments free of shape information redundancy. Thus the RGB image reconstruction from them is possible by summing up all the QDHMs components multiplied by their corresponding basis function $H_{n,m}$ via inner product as follows:

$$f^r(r, \varphi) = \sum_{n=0}^{n=Max} \sum_{m=-n}^{m=n} QDHM_{n,m} \otimes H_{n,m}. \quad (2.34)$$

where $f^r(r, \varphi)$ is the reconstructed image up to the maximum order $nMax$.

The reconstructed function $f^r(r, \varphi)$ tends to the exact original image as n tends to infinity. Practically, limiting the reconstruction to a finite order $nMax$, occurs in truncation errors, called Normalized Image Reconstruction Error (NIRE) which measures the difference between the reconstructed image and the original image as follows:

$$NIRE = \frac{\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} |f(x, y) - f^r(x, y)|^2}{\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} |f(x, y)|^2}, \quad (2.35)$$

where $f(x, y)$ is the original image with size $(N \times N)$ and $f^r(x, y)$ is the reconstructed image. Figure 2.12 shows an example of reconstructed color images from our QDHMs and their NIRE values up to different orders.

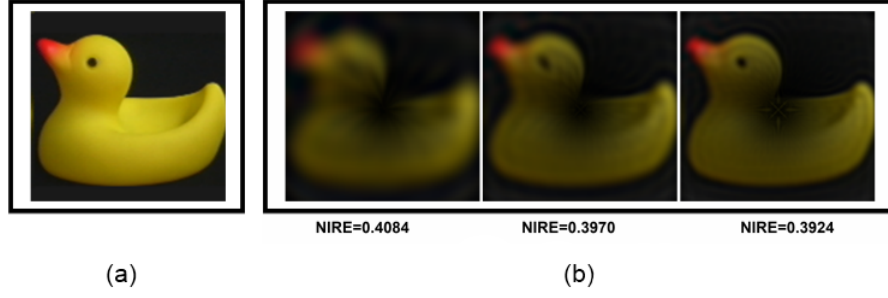


Figure 2.12 QDHMs reconstruction of image (a) up to order 20, 40, 60 with the corresponding NIRE errors.

This color image reconstruction example confirms that the reconstructed images tend to the original image as the order n increases. However, one can clearly see that the reconstructed images are somehow dark at the center resulting in large NIRE values. This limitation of the QDHMs will be overcome in their modified version detailed in the next subsection.

2.3.3 Modified Quaternion Disc-Harmonic Moments

The adaptation of the 3D spherical harmonic transform to deal with 2D images requires mapping the pixels between the spherical and the polar coordinate systems. In our first version of the quaternion disc-harmonic moments, we used the same mapping sphere-disc as done by Ennahahi et al., 2010. This mapping was previously given in subsection 2.3.2.1.

The issue with this mapping is that we run into the problem of the nonuniform variation of θ with a uniform step of r , as shown in Figure 2.13 (b). To overcome this issue, we propose a more adequate mapping sphere-disc, which is defined as:

$$\theta = \frac{\pi}{2}r, \quad \varphi_{\text{sphere}} = \varphi_{\text{disc}}, \quad (2.36)$$

where $0 \leq \theta \leq \pi/2$ and $0 \leq r \leq 1$.

Most of the implementations of the spherical harmonics require an angular distribution of the pixels over the sphere uniform under θ and φ . Different representations of sampled spherical pixels needed in the computation of the disc-harmonic moments are possible. Both representations using the old mapping and the new one are shown in Figure 2.13.

One can see clearly from Figure 2.13 that the distribution of pixels under θ and φ is uniform in (a) and not uniform in (b).

Based on the new mapping, the quaternion-valued harmonic function defined on the unit disc will be derived as follows:

$$H_{n,m}(r, \varphi) = N_{n,m} P_{n,m}(\cos(\frac{\pi}{2}r)) e^{\mu m \varphi}. \quad (2.37)$$

Visual representations of the real and first imaginary parts of the quaternion-valued harmonic function given in Eq. (2.37) of order up to $n = 8$ and $0 \leq m \leq n$ are illustrated in Figure 2.14 (b). It is obvious that the imaginary parts of the harmonic functions of repetition $m = 0$ are zero. Also, the imaginary parts can be obtained by performing an adequate rotation over the real parts of these harmonic functions.

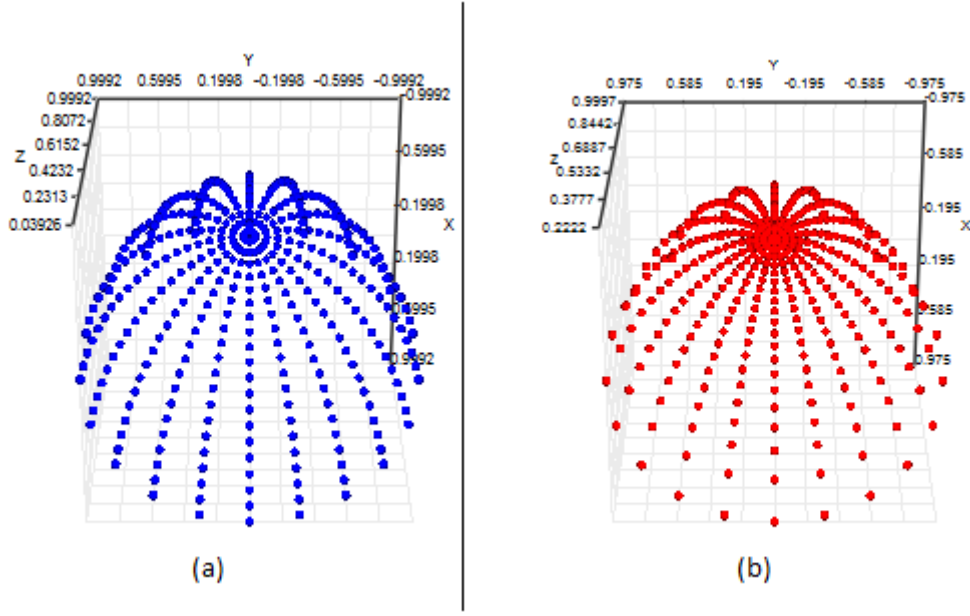


Figure 2.13 Distribution of the pixels on the hemisphere using (a) the new mapping which considers uniform steps of φ and $r = \frac{2 \times \theta}{\pi}$, (b) the old mapping with uniform steps of φ and $r = \cos(\theta)$ (Ennahnahi et al., 2010).

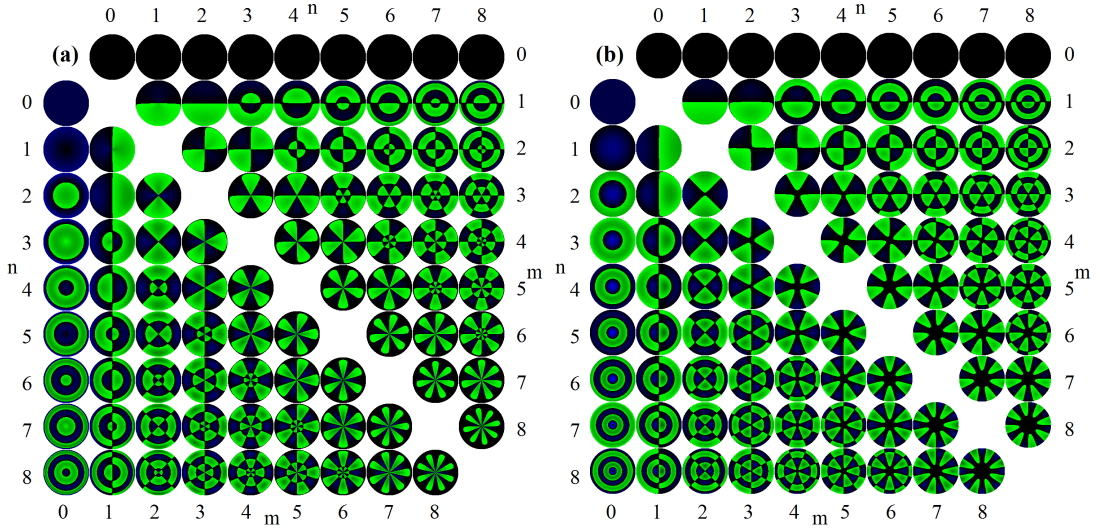


Figure 2.14 Quaternion-valued harmonic function up to order $n = 8$ using (a) the first version [Eq. (2.22)] and (b) the modified version [Eq. (2.37)]. The top triangles represent the first imaginary parts whilst the bottom triangles represent the real parts. The repetition m varies between 0 and n . Green pixels represent regions where the function is negative, blue ones represent where it is positive and the black ones are where the function is zero.

Similarly to the first version of the QDHMs, the modified QDHMs are defined by substituting $H_{n,m}$ in Eq. (2.24) by Eq. (2.37) and can be computed by considering the relation between them and the DHMs, given in subsection 2.3.2.4, based on the new mapping.

We perform color image reconstruction comparison between the QDHMs based on the old

mapping sphere-disc and those based on the proposed mapping. The maximum orders are: 20, 40 and 60. Results are illustrated in Figure 2.15.

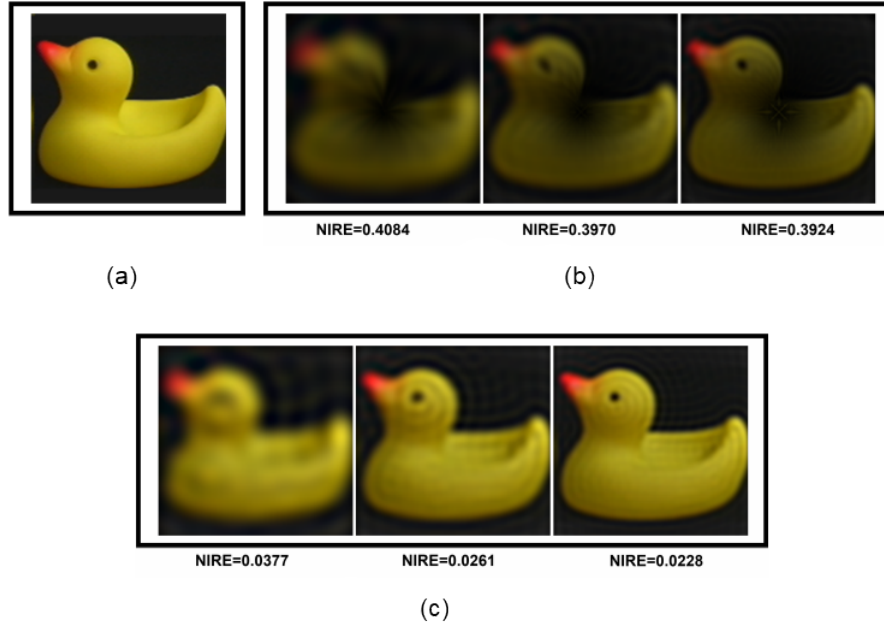


Figure 2.15 QDHMs reconstruction of image (a) up to order 20,40,60 using (b) the old mapping and (c) the new mapping.

From this example, it is seen clearly that the proposed mapping performs better than the old one based on NIRE values [Eq. (2.35)]. The NIRE value for $nMax = 60$ decreases from 0.39 using the old mapping (Ennahahi et al., 2010) to 0.02 using our mapping.

2.3.4 QDHMs based on Healpix Sampling

Given that the quality of any image descriptor isn't only about the accuracy of the results but also about its computational time, we present in this subsection a major improvement in the computational time consumed by the quaternion disc-harmonic moments (QDHMs) without any loss in quality and discriminative power. As aforementioned, the QDHMs are computed via the complex disc-harmonic moments (DHMs) of each color component, which are inspired from the 3D spherical harmonics (Dilts, 1985) using a mapping sphere-disc. Hereafter, we analyze the sampling method used by the disc-harmonic moments with the aim of enhancing the time consumed in the computation of the quaternion disc-harmonic moments.

The limitation of computing the disc-harmonic moments using the simple formula given in Eq. (2.29) lies in the necessity of calculating the associated Legendre polynomials as well as the complex angular part at each point of the unit disc and for each order n and m . In order to reduce the computational time, we pre-calculate and store the polynomials of each order n and m at each point of the unit disc. Although this solution reduces the feature vector extraction time, it proves to be consuming in terms of memory resources. A detailed discussion will be presented in subsection 2.4.4.

On the other hand, the spherical harmonics computation can be made faster by operating an adequate spherical sampling, which is not trivial. In fact, an iso-latitudinal location of samples is essential for fast discrete transform into spherical harmonics (FSHT). The icosahedral partitions

provide a uniform pixelation but not aligned on iso-latitudinal rings, which prevents applying the fast spherical harmonic transform. In the work by Healy et al., 2003, the calculation of the conventional FSHT is based on an equiangular grid. This iso-longitudinal samples layout suffers from the presence of an affinity between the pixels at the polar regions of the sphere. In addition, although this partition offers the possibility of applying sampling theorems and exact quadrature for numerical integration, its pixels are of widely varied size. In this conventional partition, a set of radial segments, starting at the center of the unit disc, are generated with a regular angular step. The sampled points are placed on the intersections between these segments and the concentric circles of radii ($r = \sin \theta$) by varying the angle θ also with a fixed step. Another drawback of this method is that the star-shaped radial segments may not take into account some details localized in areas between consecutive segments. In this case, all sample points could be nothing but a representation more or less unfaithful to the original shape.

To overcome this situation, we will operate the internal mechanism of the Hierarchical Equal Area isoLatitude Pixelization (Healpix) technique (Górski et al., 2005) for achieving rapid transformation into harmonic unit disc. Healpix technique was developed, originally, by the Jet Propulsion Laboratory of Nasa to meet the increased need for processing and analyzing large amounts of data from satellite missions to measure the cosmic microwave background (CMB) anisotropy. It allows an equal area and iso-latitude sphere pixelation and provides rapid spherical harmonics analysis even at very high resolutions.

Healpix is based on a first division of the sphere into 12 wide pixels that can be further subdivided into the desired resolution, generating a map of $N_{pix} = 12 \times N_{side}^2$ pixels of resolution N_{side} (Figure 2.16).

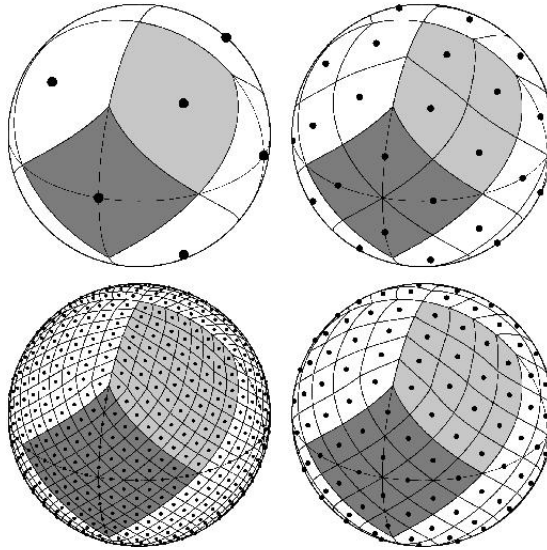


Figure 2.16 Healpix sampling of the sphere for $N_{side} = 1, 2, 4, 8$ (Górski et al., 2005).

A Healpix map is carefully constructed in order to have an equal area pixelization and the pixels are also aligned on iso-latitudinal rings. In fact, the pixel property of having equal area allows overcoming the problem of polar pixels affinity on the sphere as well as giving an equal weight to pixels.

The positions on the sphere are defined by $(\cos \theta, \varphi)$ with $\theta \in [0, \pi]$ is the co-latitude in radians measured from the northern hemisphere and $\varphi \in [0, 2\pi]$ represents the longitude. For a resolution N_{side} , the pixels are distributed on $(4 \times N_{side} - 1)$ iso-latitudinal rings, and can be ranked

by the index $p \in [0, N_{pix}]$ that travels the rings from the northern hemisphere to the southern hemisphere.

On the northern hemisphere, the positions of the spherical pixels are given by the following equations:

For pixels of the northern hemisphere cap ($z > 2/3$):

Given $p_h = (p + 1)/2$, ($1 \leq i \leq N_{side}$) the ring index and $1 \leq j \leq 4i$ the pixel index in the ring i , we have:

$$\begin{aligned} i &= I(\sqrt{p_h - \sqrt{I(p_h)}}) + 1, \\ j &= p + 1 - 2i(i - 1), \\ z &= 1 - \frac{i^2}{3N_{side}^2}, \\ \varphi &= \frac{\pi}{2i}(j - \frac{s}{2}), \\ s &= 1. \end{aligned} \tag{2.38}$$

For pixels of the northern equatorial belt ($0 \leq z \leq 2/3$):

Given $p' = p - 2N_{side}(N_{side} - 1)$, ($N_{side} \leq i \leq 2N_{side}$) and ($1 \leq j \leq 4N_{side}$):

$$\begin{aligned} i &= I(\frac{p'}{4N_{side}}) + N_{side}, \\ j &= (p' \bmod 4N_{side}) + 1, \\ z &= \frac{4}{3} - \frac{2i}{3N_{side}^2}, \\ \varphi &= \frac{\pi}{2N_{side}}(j - \frac{s}{2}), \\ s &= (i - N_{side} + 1) \bmod 2. \end{aligned} \tag{2.39}$$

where the auxiliary index s is the longitudinal phase on the arches and $I(x)$ is the greatest integer number less than x .

The southern hemisphere pixels can be obtained by a glass symmetry regarding the equatorial plan.

For computing our Healpix-based quaternion disc-harmonic moments, we follow these steps:

- (1) We sample the unit disc containing the image using Eq. (2.38), Eq. (2.39) and Eq. (2.36),
- (2) we perform a rapid discrete disc-harmonic transform (FDHT) on the three color components separately as in Eq. (2.32).
- (3) Finally, we compute the QDHMs feature vector by substituting Eq. (2.32) into Eq. (2.27).

A detailed complexity analysis will be discussed in subsection 2.4.4.

2.3.5 QDHMs-based Image Descriptor

In this subsection, we derive the Rotation/Translation/Scale (RST) invariants of the proposed 2D shape descriptors. Then we show how to obtain our image descriptor from these invariants.

2.3.5.1 Translation and Scale Invariance

In order to compute the quaternion disc-harmonic moments, defined on the unit disc, it is first necessary to do a mapping transformation from the Cartesian coordinates (x, y) to the polar coordinates (r, φ) (Flusser et al., 2009). This is achieved as follows:

$$r = \frac{\sqrt{(x-x_c)^2+(y-y_c)^2}}{r_{max}}, \text{ where } r_{max} = \frac{\sqrt{m_{00}}}{2} \sqrt{\frac{M}{N} + \frac{N}{M}}, \text{ and}$$

$$\varphi = \arctan \left(\frac{y - y_c}{x - x_c} \right), \quad (2.40)$$

where

$$(x_c, y_c) = ([m_{1,0}(R) + m_{1,0}(G) + m_{1,0}(B)]/m_{0,0}, \\ [m_{0,1}(R) + m_{0,1}(G) + m_{0,1}(B)]/m_{0,0}), \\ \text{and } m_{0,0} = m_{0,0}(R) + m_{0,0}(G) + m_{0,0}(B),$$

with $m_{1,0}(X)$, $m_{0,1}(X)$, $m_{0,0}(X)$ as the first-order and zero-order geometric moments for the X channel, and $(N \times M)$ is the size of the RGB image.

The quaternion disc-harmonic moments invariance to translation is achieved in this mapping, which locates the origin of the coordinate system at the common centroid of the three channels (x_c, y_c) (Suk and Flusser, 2009).

Note that the scale invariance is implicitly obtained by normalizing the image on the unit disc.

2.3.5.2 Rotation Invariance

The rotation invariance can be easily achieved by considering the modulus of the quaternion disc-harmonic moments according to the following:

$$\begin{aligned} QDHM'_{n,m} &= N_{n,m} \int_0^{2\pi} \int_0^1 P_{n,m}(\cos(\frac{\pi}{2}r)) f(r, \varphi + \varphi') \otimes e^{-\mu m \varphi} r dr d\varphi, \\ &= N_{n,m} \int_0^{2\pi} \int_0^1 P_{n,m}(\cos(\frac{\pi}{2}r)) f(r, \varphi) \otimes e^{-\mu m(\varphi - \varphi')} r dr d\varphi, \\ &= N_{n,m} \int_0^{2\pi} \int_0^1 P_{n,m}(\cos(\frac{\pi}{2}r)) f(r, \varphi) \otimes e^{-\mu m \varphi} \otimes e^{\mu m \varphi'} r dr d\varphi, \\ &= QDHM_{n,m} \otimes e^{\mu m \varphi'}, \end{aligned} \quad (2.41)$$

where $QDHM_{n,m}$ and $QDHM'_{n,m}$ are, respectively, the quaternion disc-harmonic moments of the color image $f(r, \varphi)$ and its rotated version $f(r, \varphi + \varphi')$.

According to which, we have:

$$\|QDHM'_{n,m}\| = \|QDHM_{n,m}\|. \quad (2.42)$$

where $\|q\|$ is the modulus of the quaternion q given in Eq. (2.17).

2.3.5.3 Feature Vector Construction

Our QDHMs-based descriptor vector (QDHMD) is build by concatenating the quaternion disc-harmonic moment modulus-based invariants of orders n up to a maximum order $nMax$ while satisfying the condition $-n \leq m \leq n$ as follows:

$$\begin{bmatrix} ||M_{0,0}|| \\ ||M_{1,-1}|| & ||M_{1,0}|| & ||M_{1,1}|| \\ \vdots & \vdots & \vdots \\ ||M_{nMax,-nMax}|| & \dots & ||M_{nMax,-1}|| & ||M_{nMax,0}|| & ||M_{nMax,1}|| & \dots & ||M_{nMax,nMax}|| \end{bmatrix} \quad (2.43)$$

As each order n of the moments provides $n \times 2 + 1$ real-valued coefficients based on the modulus of the moments, the total size of our QDHMs feature vector depends on the maximum order as below:

$$\text{size(QDHMD)} = (nMax + 1)^2. \quad (2.44)$$

Using the modulus to obtain the rotation invariance of our moments is advantageous in the way the feature's size is reduced to a quarter (Real valued instead of quaternion-valued). This is helpful in reducing response time of content-based image retrieval systems. In fact, during the online stage (Figure 2), the feature vector of the query image is compared to the feature vectors of the database images according to a similarity distance, e.g., Euclidean distance. Therefore, considering a small feature size will definitely reduce the retrieval time.

2.4 Experimental Results

In this section, we undertake several experiments to evaluate the performance of our quaternion disc-harmonic moments descriptor (QDHMs) in terms of color image reconstruction, robustness to geometric and photometric transformations, computational time as well as color image retrieval. The performance of the QDHMs is compared with the complex disc-harmonic moments (DHMs), with four orthogonal quaternion-based moments, the quaternion Zernike moments (QZMs), the quaternion pseudo-Zernike moments (QPZMs), the quaternion Bessel-Fourier moments (QBFMs) and the named quaternion orthogonal Fourier-Mellin moments (QOFMMs) derived hereafter which are an extension of the orthogonal Fourier-Mellin moments (OFMMs) (Sheng and Shen, 1994) as well as with two Clifford-based moments, namely the parallel-orthogonal Fourier-Mellin descriptor (poFMD) and the color Fourier-Mellin descriptor invariant to the choice of bivector B (CFMDi).

In order to consider the same number of moments as the QDHMs, the constraint $-n \leq m \leq n$ was imposed for both the QBFMs and the QOFMMs. Note that the DHMs' descriptor is computed for images after conversion from color to gray-level. Besides, "QDHMs" denotes the quaternion disc-harmonic moments in their modified version.

2.4.1 Color Image Sets

In these experiments, we use four color image collections:

- The Columbia object Image Library (COIL-100) (Nene et al., 1996), which contains 72,000 images with size 128×128 categorized into 100 objects. This collection is challenging as it presents object pose variations, as shown in Figure 2.17, and also some objects with the same shape and different colors, as shown in Figure 2.18.
- The Amsterdam Library of Object Images Viewing Direction (ALOI-VIEW) (Geusebroek et al., 2005), which is partitioned into 1000 classes where each class contains the same object captured under 72 different viewing directions, as shown in Figure 2.19. The images are with size 192×144 .

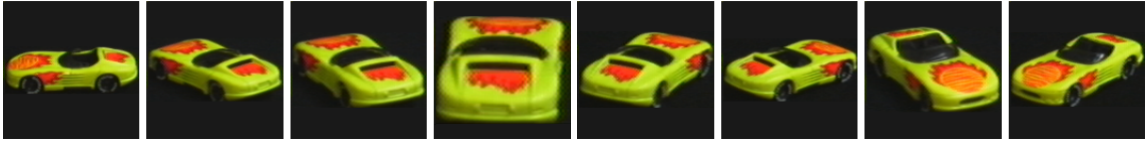


Figure 2.17 Some images for one object from the COIL-100 data set showing object pose variations.

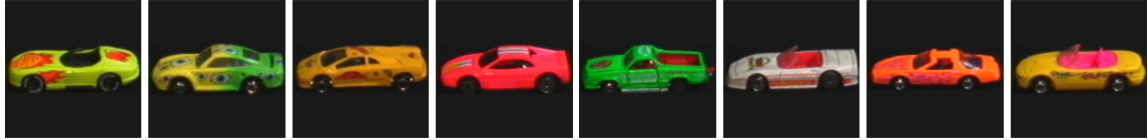


Figure 2.18 Some of the challenging object images selected from the COIL-100 data set.



Figure 2.19 Some images for one object from the ALOI-VIEW data set showing object pose variations.

- The Amsterdam Library of Object Images Illumination Color (ALOI-COL) (Geusebroek et al., 2005), which consists of 1000 classes varied in 12 configurations by changing the illumination color temperature. The structure of the ALOI-COL images is illustrated in Figure 2.20.



Figure 2.20 A class of images extracted from the Amsterdam Library of Object Images under illumination color change (ALOI-COL).

- The Amsterdam Library of Object Images Illumination Direction (ALOI-ILL) (Geusebroek et al., 2005), which is partitioned into 1000 classes with 24 configurations taken under different illumination angles. The structure of the ALOI-ILL images is illustrated in Figure 2.21.

A sample of objects extracted from COIL-100 and ALOI is shown in Figs. 2.22(a) and 2.22 (b), respectively.



Figure 2.21 A class of images extracted from the Amsterdam Library of Object Images under illumination direction change (ALOI-ILL).

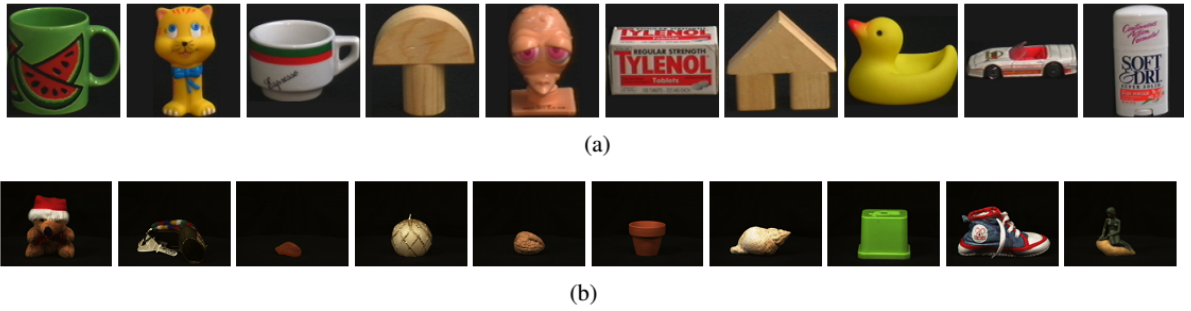


Figure 2.22 A sample of objects extracted from (a) Columbia Object Image Library (COIL-100) (Nene et al., 1996) and (b) Amsterdam Library of Object Images (ALOI) (Geusebroek et al., 2005).

2.4.2 Color Image Reconstruction

In this subsection, we evaluate the color image reconstruction capability of our QDHMs descriptor compared to the QZMs, QBFMs, QOFMMs and QPZMs. For this purpose, we consider the 10 images extracted from the COIL-100 shown in Figure 2.22 (a).

The proposed QDHMs and the other feature vectors of each image up to orders varying between 4 and 40 with a step of 4 are computed. The average reconstruction errors [Eq. (2.35)] as well as an example of reconstructed images by the QDHMs and the QZMs are shown in Figure 2.23. The example in Figure 2.23 shows clearly that even for low order QDHMs, the reconstructed image is close to the original image. It is also illustrated that color image reconstruction from the QDHMs is better than reconstruction from the QZMs, QOFMMs and QPZMs. One can see that the image reconstruction from the QDHMs and the QBFMs are close.

2.4.3 Robustness and Invariance

The aim of this subsection is to test the robustness of our descriptor (the modified QDHMs) to both photometric and geometric transformations separately. For this test, we consider the following transformations: Salt and Pepper noise, blur, illumination color change, illumination direction change and rotation.

For Salt and Pepper noise and blur tests, the process for the evaluation consists of computing the squared error between the original image feature vector and its transformed versions, then reporting the mean squared errors. The more the squared errors are of a small value, the better results are.

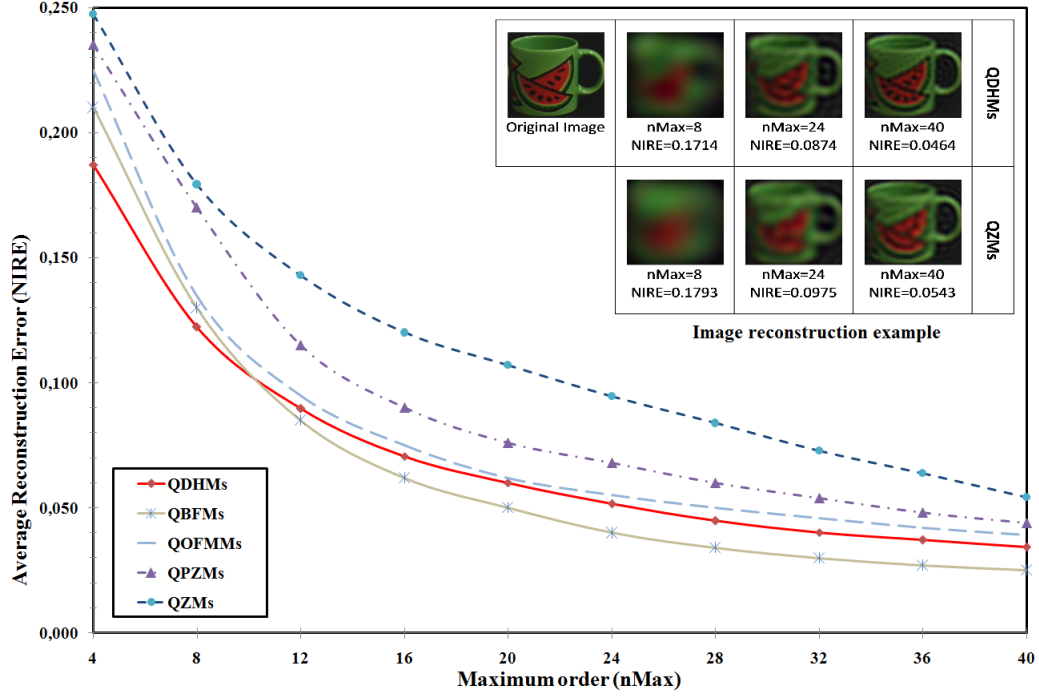


Figure 2.23 Comparison of the average NIRE values from Image reconstruction using the QDHMs and the quaternion moments.

For illumination color change, illumination direction change and rotation tests, the evaluation consists of computing the descriptor for each image of the data set. Then, we compute, for each class of the testing set, the rate σ/μ where σ represents the standard deviation of each order of the descriptor and μ the respective mean. Next, we calculate the mean of σ/μ for each order. Finally, we report the overall mean of σ/μ . The smaller this latter is, the better results are.

In these experiments, we take into consideration the following:

- The maximum order $nMax$ of the image descriptor when computed using the QDHMs, the QPZMs, the QBFMs, the QOFMMs and the DHMs is 8 (81 coefficients).
- The maximum order $nMax$ of the image descriptor is set to 11 when computed using the QZMs (78 coefficients).
- The maximum order $nMax$ of the image descriptor is set to 6 when computed using the poFMD (72 coefficients).
- The maximum order $nMax$ of the image descriptor is set to 9 when computed using the CFMDi (81 coefficients).

We recall that the CFMDi is independent from the choice of the color direction (i.e., the bivector) unlike the poFMD in which the bivector was set to B_{red} in the following experiments.

2.4.3.1 Salt and Pepper Noise

The 10 images of the COIL-100 database shown in Figure 2.22 (a) are used in this test. To each of these 10 images, we add different intensities of Salt and Pepper noise, varying between 5% and 20% with 5% increment. Therefore, the testing set is composed of $50 = 10 \times 5$ images. Figure 2.24 depicts a sample of this testing set while Figure 2.25.a) shows the mean squared errors.

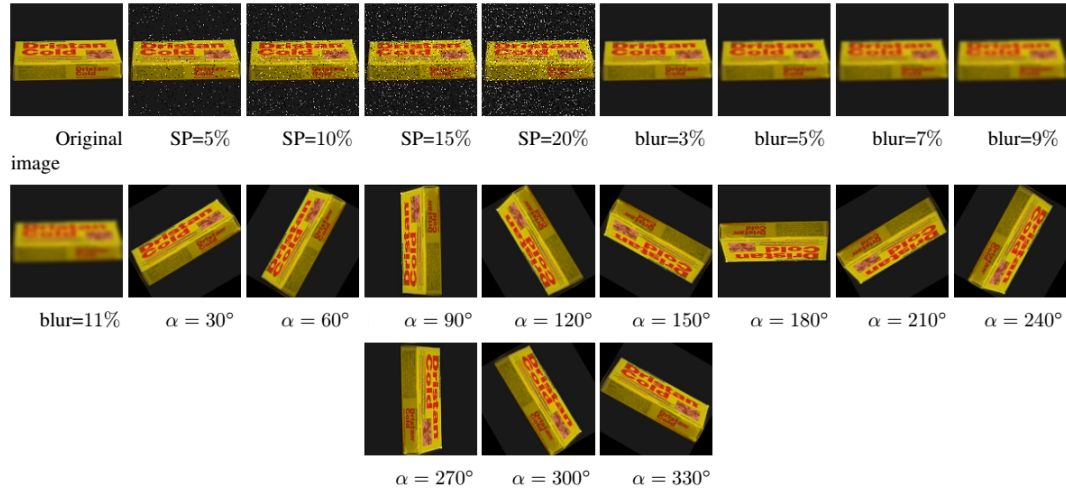


Figure 2.24 A sample of images used in Salt and Pepper (SP) noise robustness test, blur robustness test and rotation invariance test.

From these results, it is seen that the QDHMs have a better noise robustness (The squared error ≤ 0.4404) compared to the other quaternion-based moments. The Clifford-based descriptors are the most robust to noise.

2.4.3.2 Blur

We consider the 10 objects of the COIL-100 shown in Figure 2.22 (a). Each object undergoes different amounts of Gaussian blur between 3% and 11% with a step of 2%. The testing set contains $60 = 10 \times 6$ images. A sample is shown in Figure 2.24. Figure 2.25.b) illustrates the respective mean squared errors.

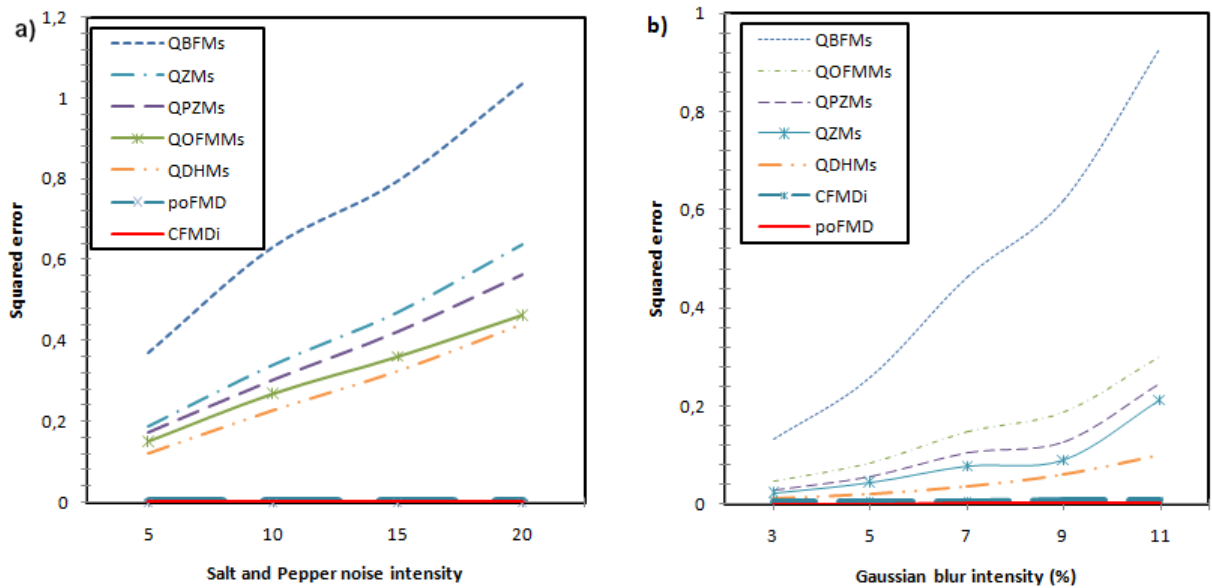


Figure 2.25 The overall average squared errors corresponding to the intensity of the transformation for each of the state-of-the-art moments: a) Salt and Pepper noise effect, b) Gaussian blur effect.

The results show that the QDHMs' feature vector remains almost unchanged under different blur intensities and most of the squared errors are less than 0.1014. A better blur robustness compared to the other quaternion- moments is also seen. However, the Clifford-based descriptors are the most robust to blur.

2.4.3.3 Illumination Color Change

Ten classes extracted from the Amsterdam Library of Object Images Illumination Color (ALOI-COL) (Geusebroek et al., 2005) are used hereafter. The testing set consists of a total of $120 = 10 \times 12$ images.

Table 2.2 shows the values of σ/μ for each class of the testing set using the QDHMs of order up to 8. Besides, the σ/μ values were also computed for the quaternion-based moments. The overall mean σ/μ values for the different moments can be found in the first row of Table 2.5.

As experimental results show, among all the moments used in this test, the QDHMs are the most influenced by this change. Therefore, the QDHMs could be useful in applications where it is important to distinguish same objects of different illumination colors.

On the other hand, we have undertaken the same experiment after converting the testing images to the HSV color space. Results, which are illustrated in the second row of Table 2.5, show clearly that the HSV color space gives more efficient results than the RGB color space as far as the QDHMs' robustness to the illumination color change is concerned.

Table 2.2 The σ/μ values for the classes of the ALOI-COL using each order of the QDHMs up to 8.

Order	1	2	3	4	5	6	7	8	9	10	Mean of σ/μ
0	0.0710	0.0638	0.0484	0.0681	0.0639	0.0806	0.0463	0.0719	0.0543	0.0572	0.0626
1	0.0831	0.0753	0.0583	0.0796	0.0751	0.0897	0.0546	0.0784	0.0669	0.0677	0.0729
2	0.0893	0.0869	0.0793	0.0860	0.0889	0.1072	0.0558	0.0838	0.0696	0.0777	0.0825
3	0.1089	0.0672	0.0648	0.0947	0.0845	0.1269	0.0503	0.0933	0.0775	0.0644	0.0833
4	0.0920	0.0803	0.0742	0.0983	0.0772	0.1089	0.0843	0.0944	0.0739	0.0737	0.0858
5	0.0931	0.0925	0.0972	0.0865	0.0904	0.1187	0.1401	0.0803	0.0747	0.0848	0.0959
6	0.0913	0.1267	0.0863	0.1668	0.1069	0.1280	0.0872	0.0890	0.0704	0.0956	0.1049
7	0.0940	0.1035	0.0929	0.1616	0.1144	0.1195	0.0721	0.0920	0.0777	0.0974	0.1025
8	0.0972	0.0821	0.1009	0.1245	0.1124	0.1371	0.0602	0.0943	0.0764	0.0948	0.0980
Overall mean											0.0876

2.4.3.4 Illumination Direction Change

For this test, 10 classes from the Amsterdam Library of Object Images Illumination Direction (ALOI-ILL) (Geusebroek et al., 2005) are used. This results in a testing set of a total of $240 = 10 \times 24$ images. The results for the QDHMs are shown in Table 2.3 while the general results are depicted in the third row of Table 2.5.

As one can see from these results, among all the moments used in this test, the QZMs the poFMD and the QPZMs are the less sensitive to the illumination direction change. The QDHMs

are more robust to the illumination direction change (the overall average $\sigma/\mu = 0.2801$) compared to the QOFMMs (the overall mean $\sigma/\mu = 0.3608$), the QBFMs (the overall mean $\sigma/\mu = 0.3396$) and the CFMDi (the overall mean $\sigma/\mu = 0.03098$).

In general, the overall mean σ/μ for the illumination direction change is significant (less than 0.3608 for all the moments) compared to the illumination color change (less than 0.0876 for all the moments). This is explained by the fact that, in some cases, parts of the object become invisible for the illumination direction change (Non-linear transformation) unlike the illumination color change which doesn't affect the shape of the object (Linear transformation).

Table 2.3 The σ/μ values for the classes of the ALOI-ILL using each order of the QDHMs up to 8.

Order	1	2	3	4	5	6	7	8	9	10	Mean of σ/μ
0	0.2285	0.2467	0.1032	0.1551	0.1479	0.2068	0.2134	0.2598	0.2220	0.1808	0.1964
1	0.2558	0.2384	0.1004	0.1472	0.1431	0.2095	0.2072	0.2634	0.2283	0.1738	0.1967
2	0.3488	0.3123	0.1053	0.1590	0.1676	0.2616	0.2160	0.2987	0.2498	0.2188	0.2338
3	0.4513	0.2784	0.1149	0.1841	0.1866	0.3053	0.2346	0.3608	0.3042	0.2436	0.2664
4	0.3760	0.2756	0.1178	0.2398	0.1802	0.2811	0.3286	0.3607	0.2932	0.1870	0.2640
5	0.4854	0.3719	0.1363	0.3145	0.2160	0.3125	0.4104	0.3165	0.3267	0.2253	0.3116
6	0.4787	0.4331	0.1491	0.3977	0.2524	0.3712	0.4016	0.3367	0.3856	0.2730	0.3479
7	0.5444	0.4418	0.1635	0.3507	0.2587	0.4244	0.3286	0.3692	0.3839	0.2896	0.3555
8	0.4882	0.4699	0.2007	0.2783	0.2422	0.4431	0.2871	0.3935	0.4117	0.2776	0.3492
Overall mean											0.2801

2.4.3.5 Rotation

The 10 images of the COIL-100 dataset shown in Figure 2.22 (a) are used in this test. We rotate each of these 10 images by an angle varying between 30° and 330° with 30° increment forming a testing set of $120 = 10 \times 12$ images. Figure 2.24 depicts a sample of this testing set. The average σ/μ values for the QDHMs and the overall average σ/μ values for the different moments are given in Table 2.4 and the fourth row of Table 2.5, respectively.

According to these results, among all the moments used in this test, the QOFMMs are the most unchanged under rotation. The invariance of the QDHMs to rotation is also acceptable (the overall average $\sigma/\mu = 0.1101$).

Although there are some black parts added to the image when it is rotated (Figure 2.24), the overall average σ/μ value for rotation is, generally, less than 0.1963.

2.4.4 Computational Time

The program for the moments' computation was implemented on a personal computer i3-2310M 2.10 GHz, 4 Go RAM. We consider the image in Figure 2.26.(a) of size 128×128 , which is extracted from the COIL-100 database, as a test image. The number of moments varies from 1 to 121. Note that we have computed the Clifford-based descriptors using their C++ implementation which is available online.

Table 2.4 The σ/μ values for the rotated images using each order of the QDHMs up to 8.

Order	1	2	3	4	5	6	7	8	9	10	Mean of σ/μ
0	0.0298	0.0173	0.0320	0.0221	0.0346	0.0403	0.0347	0.0423	0.0211	0.0278	0.0302
1	0.1604	0.0602	0.1112	0.0525	0.0696	0.2082	0.0606	0.1554	0.0404	0.0487	0.0967
2	0.1564	0.0898	0.1240	0.1085	0.0799	0.1394	0.0692	0.1402	0.0444	0.0864	0.1038
3	0.1280	0.1143	0.1190	0.1300	0.0484	0.1532	0.0812	0.1364	0.1205	0.0617	0.1093
4	0.1447	0.0692	0.1143	0.1095	0.0770	0.1630	0.0916	0.1366	0.0616	0.0548	0.1022
5	0.1508	0.1100	0.1233	0.1495	0.0962	0.1427	0.1383	0.1611	0.1018	0.0586	0.1232
6	0.1120	0.1525	0.1400	0.1763	0.0807	0.1392	0.1189	0.1300	0.1600	0.0839	0.1294
7	0.1723	0.1467	0.1574	0.1850	0.1295	0.1711	0.1207	0.1830	0.1162	0.0887	0.1471
8	0.2068	0.1284	0.1570	0.1741	0.1276	0.1572	0.1554	0.1435	0.1416	0.1009	0.1492
Overall mean											0.1101

Table 2.5 The overall average values of σ/μ for the different state-of-the-art moments. The first row is for illumination color change (RGB), the second row for illumination color change (HSV), the third row for illumination direction change and the fourth row for rotation.

	QDHMs	QZMs	QBFMs	QPZMs	QOFMMs	CFMDi	poFMD
Illumination color change (RGB)	0.0876	0.0668	0.0838	0.0730	0.0858	0.0375	0.0456
Illumination color change (HSV)	0.0793	0.0902	0.0711	0.0826	0.0854	0.1689	0.1086
Illumination direction change	0.2801	0.2297	0.3396	0.2664	0.3608	0.3098	0.2367
Rotation	0.1101	0.1378	0.1021	0.1016	0.0925	0.1294	0.1963

Figure 2.26 shows the time in seconds consumed for the computation of the feature vector depending on the number of moments and using the QDHMs and its competitors.

From these results, one can remark that the time consumed in the computation of the poFMD and the CFMDi is not proportional to the order, and it is far less than the quaternion-based moments computation times. On the other hand, the time needed for the computation of the QDHMs is less than the QOFMMs, QZMs, QPZMs, and it is comparable to QBFMs. Thus, the QDHMs could be useful as feature vector in real-time color image retrieval systems.

Second, we discuss the computational time of the proposed QDHMs-Healpix in comparison with the basic QDHMs by presenting a complexity analysis to highlight the perceived gain in memory and computation time of the Legendre polynomials. On an image of size 128×128 pixels, we count $128 \times 128 = 16,384$ pixels. For each of these points, 81 moments ($n_{\text{Max}}=8$) are calculated, but only 45 associated Legendre Polynomials. In total, we have to compute $16,384 \times 45 = 737,280$ Legendre polynomials. Whereas with Healpix, if $N_{\text{side}} = 32$ is the basic resolution, we will get a total of $12 \times 32^2 = 12,288$ pixels corresponding to $(4 \times 32 - 1) = 127$ iso-latitudinal rings. But, we will compute the Legendre functions only for the $(2 \times 32 - 1) = 63$ iso-latitudinal rings situated on the northern hemisphere. The result is therefore that the number of Legendre polynomials to be calculated does not exceed $63 \times 45 = 2835$ in total. The gain in memory and complexity is remarkably persuasive (Table. 2.6).

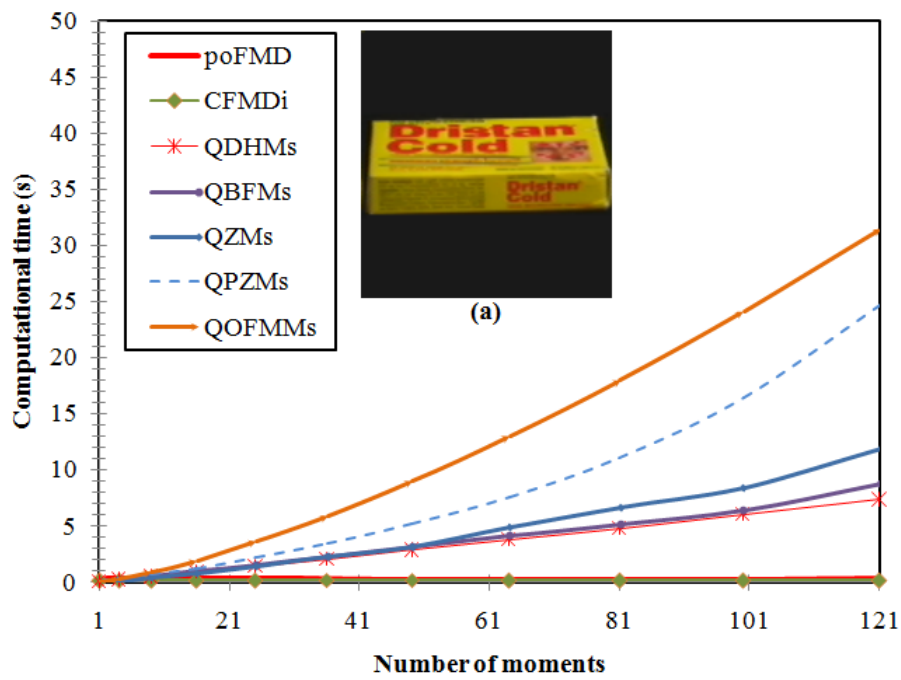


Figure 2.26 Time consumed in the computation of the feature vector of the image (a) of size 128×128 using the different state-of-the-art moments of number of moments up to 121.

Table 2.6 Number of the associated Legendre polynomials computed using both the conventional and Healpix method with different values of $nMax$ for 128×128 size image and $N_{side} = 32$

nMax	8	12	16	20
Basic QDHMs	737,280	1,490,944	2,506,752	3,784,704
QDHMs-Healpix	2835	5733	9639	14,553

Górski et al., 2005 proposed a graphical complexity comparison between different types of partitions and realized the convenience of Healpix for rapid calculation of Spherical Harmonics. They also stressed that the extra time of computation dissipated on non-optimal generation of associated Legendre functions typically results in a complexity of order $O(N_{pix}^2)$, whereas this complexity is reduced to $O(\sqrt{N_{pix}}^3)$ if one opts for Healpix pixelation. This improvement is of purely geometric origin: The Legendre functions members of the spherical harmonics, which must be generated via iterations and recurrences, are evaluated just once for each ring.

We also emphasize another heavy improvement which is the employment of the FFT when computing the DHMs with Eq. (2.32).

The discussion above is illustrated in Table. 2.26, which includes the time for the extraction of the QDHMs and QDHMs-Healpix descriptors without pre-calculating Legendre polynomials. Regarding the computation of the quaternion moments, which necessitates the computation of the conventional moments three times (for R , G , and B channels), we mention that we adopt a parallel implementation (based on OpenMP, for shared memory architectures) in order to compute the three moments of each channel in the same time.

Table 2.7 Average time for the computation of QDHMs-Healpix and some of the state-of-the-art quaternion moments (the image size is 128×128)

QDHMs-Healpix	QDHMs	QBFMs	QZMs	QPZMs
0.0937s	1.1278s	1.3125s	1.6852s	2.7856s

Time in Table. 2.26 is expressed in seconds and represents the respective average time calculated on the totality of the COIL-100 data set (7200 images of size 128×128). The program for the moments' computation was implemented on a 64-bits ubuntu uni-processor, 2.5Ghz, 1 Go RAM, running on a virtual machine.

According to Table.2.26, the contribution of the QDHMs-Healpix is remarkable since the calculation is roughly 12 times faster than the basic method.

2.4.5 Color Image Retrieval

We first fixed some parameters that allow the highest recall-precision rates using our QDHMs, such as the maximum order $nMax$, similarity measure and color space. Then, a color image retrieval comparison of the proposed descriptors and their competitors considering these parameters was undertaken.

2.4.5.1 K-Nearest-Neighbor Search

In the following experiments, we will opt for the K-Nearest-Neighbor search (k-NN). In fact, this latter aims at finding the top k points in a given set that are the closest to a given point. Typically, this point is taken to be a multidimensional vector and closeness is taken to be the dissimilarity measure, i.e., distance metric, which is of larger values when the objects are less similar.

In this work, we perform image retrieval based on the k-nearest-neighbor search. First, we extract the feature vectors from the color images in the database up to a maximum order $nMax$. Then, we compute the dissimilarity between the query image descriptor (each image in the database) and the database color images' descriptors. Finally, we compute the recall-precision at each value of k ($1 \leq k \leq M$ where M is the number of classes in the database).

In fact the recall and precision measures are defined by

$$\text{precision} = \frac{\text{No. of relevant retrieved images}}{\text{total number of retrieved images}}, \quad (2.45)$$

$$\text{recall} = \frac{\text{No. of relevant retrieved images}}{\text{total number of relevant images}}. \quad (2.46)$$

2.4.5.2 Effect of the Maximum Order

The goal of this experiment is to determine the optimal maximum order $nMax$. We extract the descriptor of each image of the test set using the QDHMs up to different maximum orders (from 4 to 14). The respective recall-precision curves based on the Euclidean distance for similarity measurement are given in Figure 2.27.

From this figure, we can clearly see that the discriminative power of the descriptor increases with the increase of the order (in other words, the size of the descriptor), this is explained by

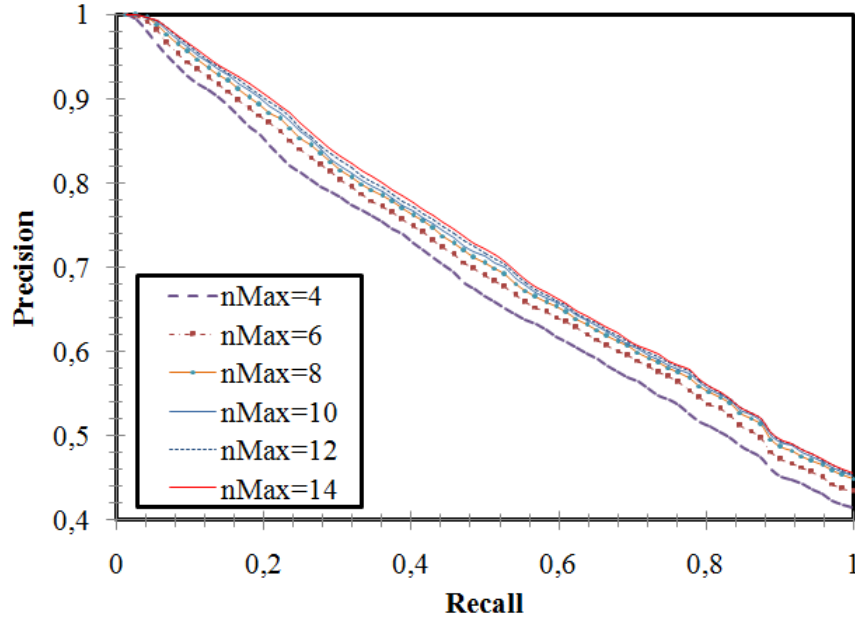


Figure 2.27 Recall-precision curves for different maximum orders using the QDHMs.

the fact that high dimensional feature vectors contain more information about the color image. However, when the maximum order is up to 8 (Which corresponds to a size of the QDHMs descriptor of 81), the performance of the QDHMs remains almost stable. And since the dimensionality of the feature vector influences considerably the time required for the similarity measurement phase, the order $nMax = 8$ of the QDHMs is retained for the rest of the experiments.

2.4.5.3 Effect of the Similarity Measure

This experiment aims to study the effect of the similarity measure on the retrieval performance of our descriptor. To this end, different distances [Eqs. (2.47)-(2.51)] including Euclidean distance, Canberra distance, Chi-squared distance, Squared-chord distance and Cosine distance (Cha, 2007) are used to measure the similarity between two different feature vectors P and Q of size d in the k-nearest-neighbor search.

$$\text{Euclidean distance: } d_{euc} = \sqrt{\sum_{i=1}^d |P_i - Q_i|^2}, \quad (2.47)$$

$$\text{Canberra distance: } d_{can} = \sum_{i=1}^d \frac{|P_i - Q_i|}{P_i + Q_i}, \quad (2.48)$$

$$\text{Chi-squared distance: } d_{chisq} = \sum_{i=1}^d \frac{(P_i - Q_i)^2}{P_i + Q_i}, \quad (2.49)$$

$$\text{Squared-chord distance: } d_{sqc} = \sum_{i=1}^d (\sqrt{P_i} - \sqrt{Q_i})^2, \quad (2.50)$$

$$\text{Cosine distance: } d_{cos} = \sqrt{\sum_{i=1}^d P_i^2} \sqrt{\sum_{i=1}^d Q_i^2}. \quad (2.51)$$

We plot the recall-precision curves based on these different distances ($nMax$ is set to 8). Results are depicted in (Figure 2.28).

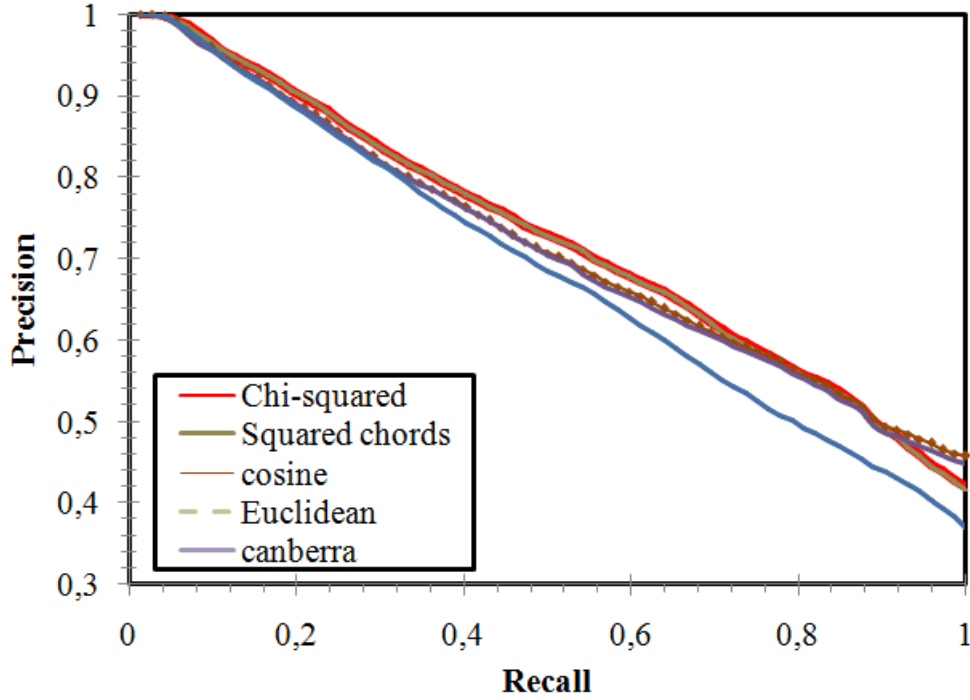


Figure 2.28 Recall-precision curves for different similarity measures using the QDHMs.

The results indicate that our descriptor shows better performance using the Chi-squared distance as a similarity measure for content-based color image retrieval.

2.4.5.4 Effect of the Color Space

This experiment aims to study the effect of the color space on the retrieval performance of our descriptor. To this end, the testing images were converted into other color spaces such as HSV, XYZ, LAB, LUV, YCrCB. The corresponding recall-precision curves based on the Chi-squared similarity measure are displayed in (Figure 2.29) ($nMax$ is set to 8).

From the results, one can conclude that the HSV color space is suitable for color image retrieval using the QDHMs descriptor. Note that for each RGB pixel represented by a quaternion vector f , HSV (hue, saturation, value) color space coordinates can be defined using some operations on quaternions. The value component is the norm of the color's orthogonal projection vector $(f \cdot \mu_{gray})\mu_{gray}$ on the gray axis μ_{gray} ($\mu_{gray} = (i + j + k)/\sqrt{3}$). The saturation component is the distance between the color vector f and the gray axis μ_{gray} . The hue component is the angle between the color vector f and a color vector ν taken anywhere on the plane orthogonal to μ_{gray} and which sets the zero hue reference angle (Denis et al., 2007).

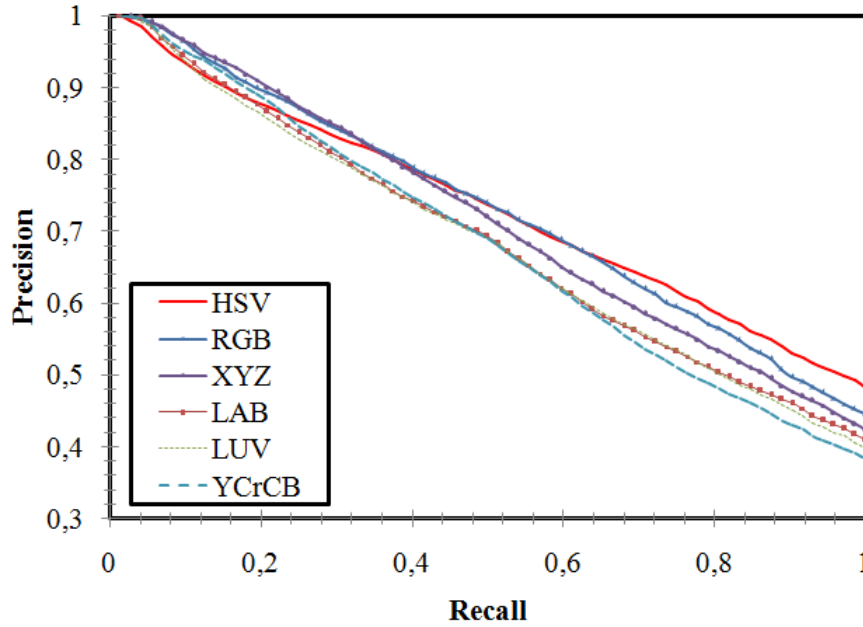


Figure 2.29 Recall-precision curves for different color spaces using the QDHMs.

In order to visually assess the benefit of the use of the HSV color space over the RGB color space in content-based color image retrieval, we colored an apple shape using different shades (Forming a set of 21 image each of size 256×256). Then, we used the k-nearest neighbor classifier (where $k=7$) based on Chi-square distance to measure the correct classification percentage (CCP) which is nothing but the *precision* given in Eq. (2.45).

As the hue is specified by an angle in degrees between 0 and 360, and both saturation and value are specified by values between 0 and 1, it is obvious that we have used a normalization so as the H, S and V are all between 0 and 255.

Figure 2.30 illustrates the results. The precision for the HSV color space is 78.91 % whereas that for the RGB space is 76.19%. Though a small difference is obtained, it is more likely to discriminate two objects with the same shape and different combinations of colors using the HSV color space and the QDHMs.

2.4.5.5 Retrieval Comparison

In the previous subsections, we studied the effect of the maximum order, similarity measure and color space on the retrieval performance of our QDHMs. The optimal parameters are recapitulated in the following:

- The maximum order is $nMax = 8$ for the DHMs, the QDHMs, the QBFMs, the QPZMs, the QOFMMs, $nMax = 11$ for the QZMs, $nMax = 6$ for the poFMD and $nMax = 9$ for the CFMDi.
- The similarity measure used for the k-NN retrieval is the Chi-squared distance.

The retrieval performance of the QDHMs is compared to the classical DHMs in order to show the improvement brought by integrating the color information. Figure 2.31 (a) and Figure 2.31 (b) show the corresponding recall-precision curves using the COIL-100 database images and the ALOI-VIEW database images, respectively.

From these results, one can remark that the QDHMs give more efficient results than the DHMs. This is explained by the fact that the transformation of the color images into gray-level ones

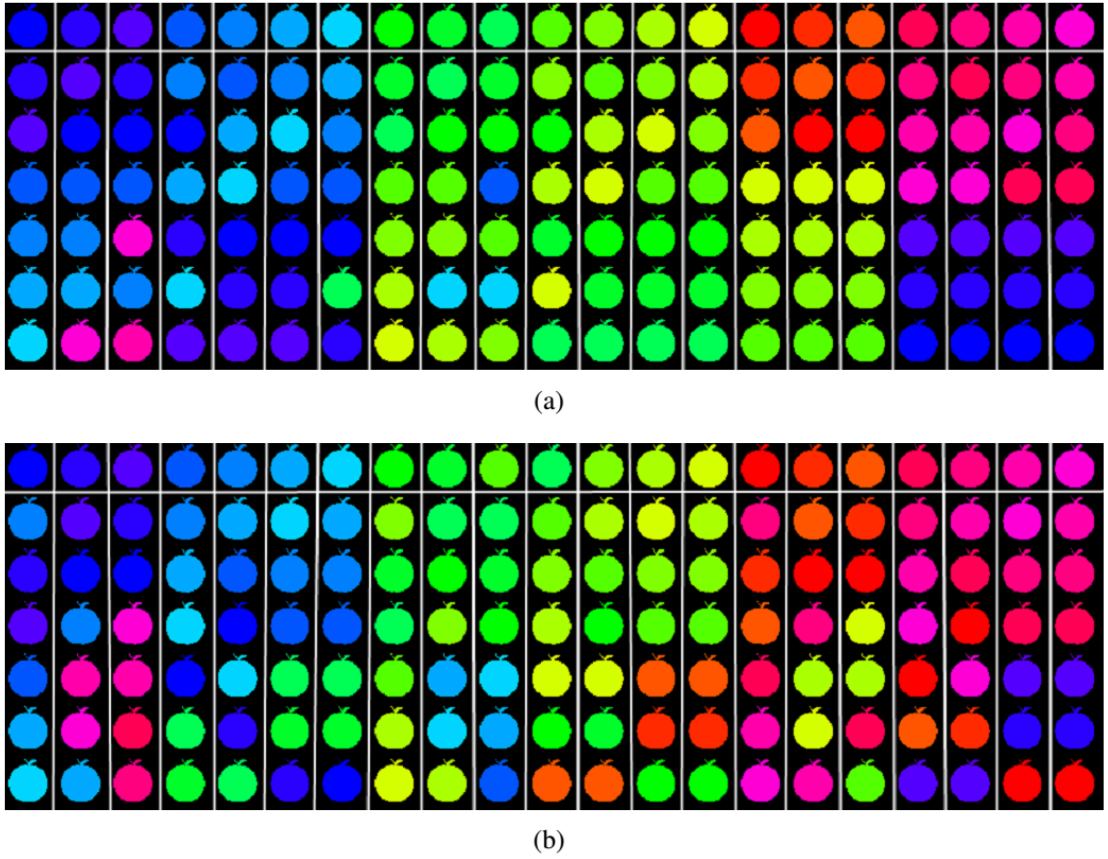


Figure 2.30 Content-based color image retrieval by QDHMs (a) using the HSV color space, (b) using the RGB color space. The first row in both (a) and (b) represents the query images set. Each column is the result of the corresponding query.

results in a significant loss of color information.

On the other hand, the retrieval performance comparison of the QDHMs' descriptor and its competitors using the COIL-100 and the ALOI-VIEW is illustrated in the recall-precision curves that are plotted in Figure 2.31 (c) and Figure 2.31 (d), respectively.

The obtained results show that the QDHMs have the best discriminative power compared to the other orthogonal quaternion-based moments as well as Clifford-based descriptors using both the COIL-100 database and the ALOI-VIEW database. As a conclusion, our descriptor is effective in CBIR regardless of the viewing direction of the object in the image.

On the other hand, we present a comparison between the QDHMs-Healpix and some state-of-the-art quaternion moments including the basic QDHMs. Fig. 2.32 shows that the retrieval quality obtained by the QDHMs-Healpix, in terms of recall-precision curves, remains comparable to that of the basic version QDHMs. Besides, QDHMs-Healpix is roughly 12 times faster, as has been reported in Table 2.7. We recall that the motivation behind the Healpix-based QDHMs was to ensure an efficient computation while at the same time preserve the retrieval quality with a good tradeoff quality/time. We also remark from Table. 2.8 that QDHMs-Healpix outperforms some well-known quaternary moments for color image retrieval based on the recall-precision curves using the Chi-squared distance.

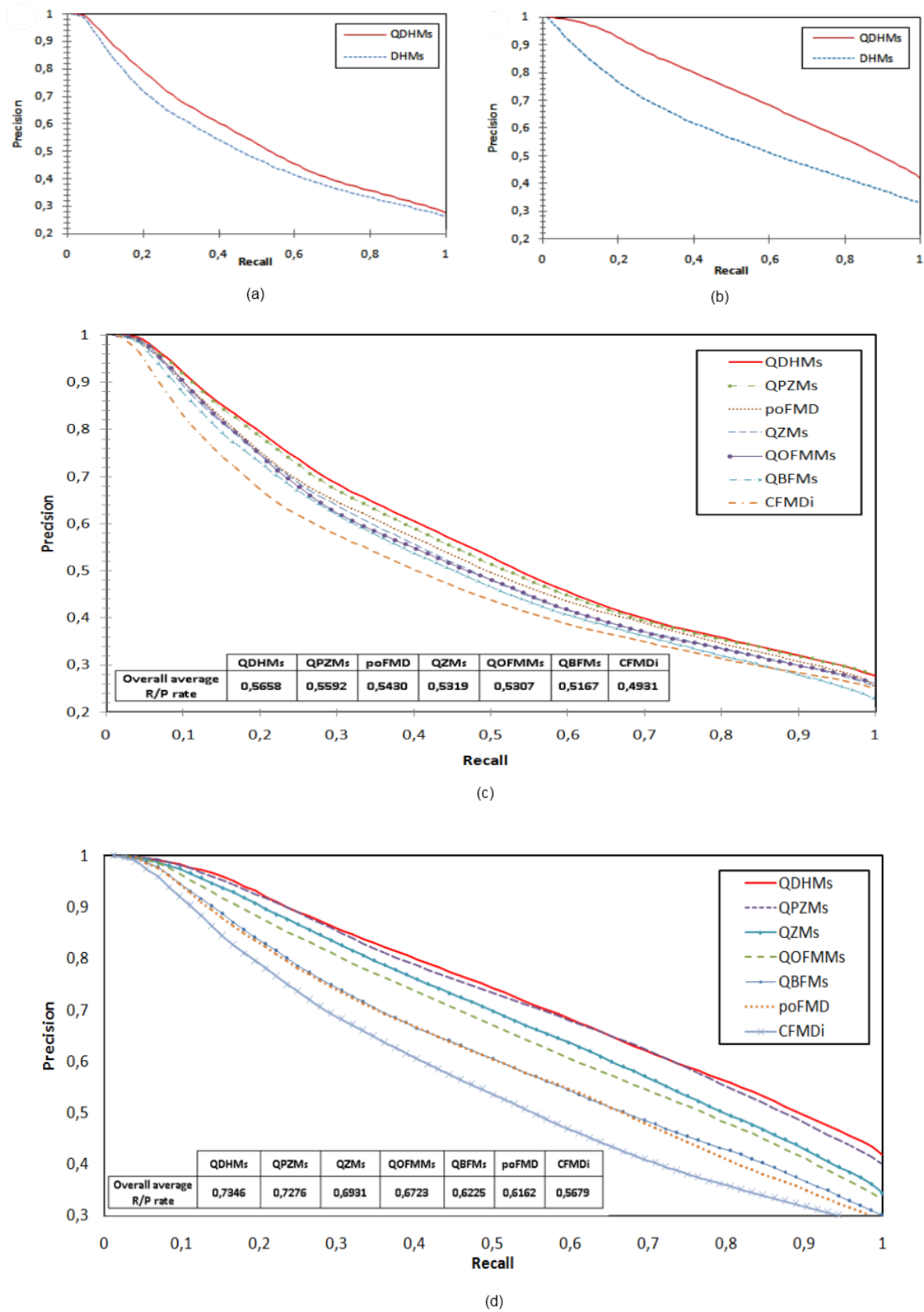


Figure 2.31 Comparative analysis of the discriminative power (recall-precision rates) of (a) the DHMs and QDHMs using the COIL-100, (b) the DHMs and QDHMs using the ALOI-VIEW, (c) the quaternion and Clifford based moments using the COIL-100 and (d) the quaternion and Clifford based moments using the ALOI-VIEW.

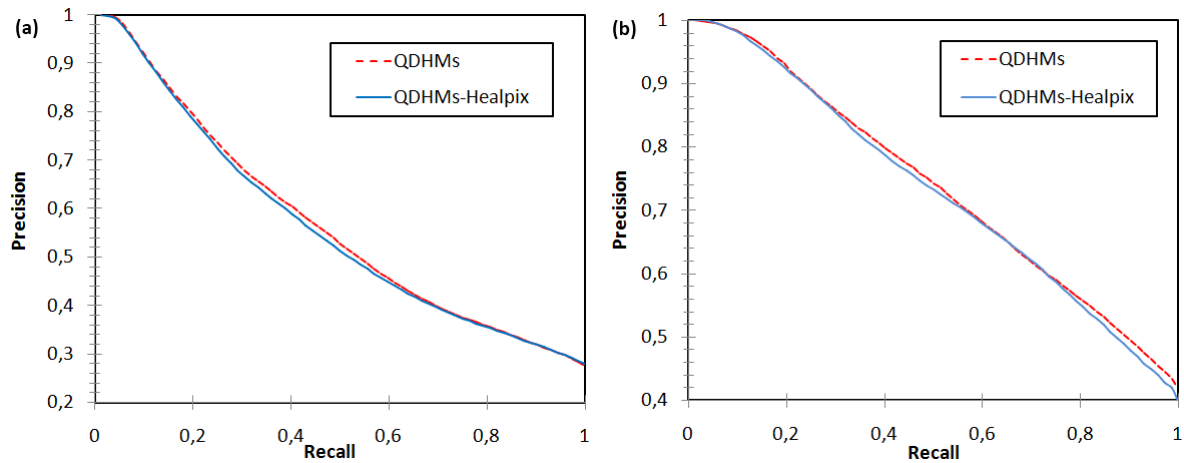


Figure 2.32 Recall-Precision curves based on Chi-squared distance for the QDHMs and the QDHMs-Healpix using (a) COIL-100 and (b) ALOI-VIEW

Table 2.8 Average precision based on the Chi-squared distance comparing the QDHMs-Healpix and some well-known quaternion moments using both the COIL-100 and ALOI-VIEW databases

	QDHMs-Healpix	QDHMs	QPZMs	QZMs	QBFMs
COIL-100	0.5698	0.5658	0.5592	0.5391	0.5167
ALOI-VIEW	0.7385	0.7346	0.7276	0.6931	0.6225

2.5 Conclusion

In this chapter, we presented a set of quaternary moments as an adaptation of the spherical harmonic transform (SHT) for color image description by using the algebra of quaternions.

On the one hand, the proposed moments use the associated Legendre polynomials which are orthogonal on the unit disc. On the other hand, using the quaternions for representing the color information is beneficial in the way that the three color components are integrated in a single feature vector while preserving the correlation between them.

Three versions were proposed: (1) The quaternion disc-harmonic moments (QDHMs), which are derived from the SHT using the same mapping sphere-disc as by Ennahahi et al., 2010, (2) the modified version of the QDHMs based on a more adequate mapping sphere-disc, and (3) the healpix-based version for further enhancing the computational time.

The computation of the three QDHMs versions has been established via their conventional versions DHMs of each color channel. Versions (1) and (2) calculation rely on doing a bunch of discrete Legendre transforms, followed by a bunch of "regular" fast Fourier transforms (FFT). This results in a fast computation of the moments. On the other hand, the healpix-based version allows a 12-time faster computation.

The translation invariance was achieved by locating the origin of the coordinate system at the color image centroid. The rotation invariance was easily obtained by taking the modulus of the moments. And where image scaling is concerned, the normalization process is needed.

We conducted several experiments in order to evaluate the performances of the proposed moments in terms of color image retrieval based on the k-nearest neighbor. We chose some databases that are challenging in the way they present view-point changes, color illumination and illumination direction changes, etc.

We first studied the impact of the maximum order $nMax$, the similarity measure and the color space on the retrieval performance of our QDHMs. Then, we compared the discriminative powers of the proposed descriptors with some well-known color image descriptors. It has been shown that the QDHMs have better discriminative power compared to the DHMs and to most of the existing quaternary and Clifford descriptors.

Another experiment has shown that the QDHMs-Healpix has comparable retrieval results as the basic version QDHMs. The former is advantageous as it is much faster.

The next chapter will be devoted to our parameter-free quaternary orthogonal moments for color image retrieval and classification.

Chapter 3

Proposed Parameter-Free Reformulation of the Quaternary Moments

3.1 Introduction

All the state-of-the-art quaternary moments depend in their computation on a unit pure quaternion parameter $\mu = ai + bj + ck$ such that $a, b, c \in \mathbb{R}$ and $\|\mu\| = 1$. Accordingly, an infinite number of μ values can be used to construct the quaternary moments, which leads to various results. Chen et al., 2014a discussed the choice of the parameter μ using 100 different values including 94 random values as well as 6 values that are common in the literature, i.e.,

- i (Qi et al., 2008),
- j (Wei and Weijiang, 2010),
- $(i + k)/\sqrt{2}$ (Guo and Zhu, 2011),
- $(-2j + 8k)/\sqrt{68}$ (Bas et al., 2003),
- $(i + j + k)/\sqrt{3}$ (Subakan and Vemuri, 2011; Bihan and Sangwine, 2003a; Bihan and Sangwine, 2003b; Ell and Sangwine, 2007; Alexiadis and Sergiadis, 2009; Chen et al., 2012; Shao et al., 2014),
- $(i - j - k)/\sqrt{3}$ (Moxey et al., 2003).

It has been shown that the optimal choice of μ is different for each type of quaternary moments; however, $\mu = (i - j - k)/\sqrt{3}$ presents the overall best performance.

Note that color image reconstruction is independent of the choice of the unit pure quaternion parameter μ as the terms including μ in the inverse transform are eliminated by the corresponding conjugate terms in the forward transform. This has been empirically validated by Li, 2013 who repeated the image reconstruction experiment 100 times using randomly generated transform axis μ resulting in almost identical root mean square errors (RMSE). Note also that $\mu = (i + j + k)/\sqrt{3}$, the gray line in the RGB space or the luminance axis, is chosen in the majority of the quaternary transforms due to its explicit physical meaning and symmetry.

Generally speaking, it is difficult to experimentally set the unit pure quaternion μ to its optimal value, i.e., consuming in terms of computation and time. Furthermore, the estimated parameter value may be dependent on the application and the type of quaternary moments.

Enlightened by Gough's work (Gough, 1984) on quaternions and spherical harmonics, we propose a simple formulation of the quaternary orthogonal moments, which is free from any parameters, thus characterized by its simplicity and generality. We first provide the formula for the computation of our parameter-free quaternary orthogonal moments (PFQOMs) as well as their color image reconstruction capability. Then, we derive their invariants that are also robust to photometric transformations. The performances of the proposed moments are evaluated against the μ -based quaternary moments, the chromaticity moments (CM) (Paschos et al., 2003), the

Legendre chromaticity distribution moments (LCDM) (Yap and Paramesran, 2006) as well as the Clifford-based parallel-orthogonal Fourier-Mellin Descriptor (poFMD) (Mennesson et al., 2014).

3.2 Parameter-Free Quaternary Moments

3.2.1 Quaternion-Valued Spherical Harmonics

Hamilton, 1844 introduced and developed the quaternion in the middle of the 19th century. Although his invention was praised at that time, it was little practiced mainly because of the emergence of vector algebra. In the 20th century, quaternions have been reawakened owe to the many attempts to investigate their application in the theory of relativity and relativistic quantum mechanics. For instance, Finkelstein et al., 1962 were the first to try to develop a kind of quantum mechanics using quaternary coefficients instead of complex.

Later on, Gough, 1984 attempted to discuss the significant implications and applications of the quaternions in three-dimensional quantum mechanics. For this purpose, he considered the spherical harmonics (SHs), detailed in subsection 2.2.1. In his work, Gough, 1984 first presented some quaternion operators related to quantum mechanics including but not limited to the quaternion equivalent of the vector ∇ operator, namely ∇_q , the laplacian operator $\nabla_q \nabla_q$ and angular momentum operators L_x, L_y, L_z , which are respectively given by:

$$\nabla_q = i \frac{\partial}{\partial x} + j \frac{\partial}{\partial y} + k \frac{\partial}{\partial z}, \quad (3.1)$$

$$\nabla_q \nabla_q = -\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}, \quad (3.2)$$

$$L_x = i \left(z \frac{\partial}{\partial y} - y \frac{\partial}{\partial z} \right), \quad L_y = j \left(x \frac{\partial}{\partial z} - z \frac{\partial}{\partial x} \right), \quad L_z = k \left(y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y} \right). \quad (3.3)$$

The author showed that $\nabla_q^2, L = L_x + L_y + L_z$ and $L_z^2 - L_z$ are operators which all commute with each other. This leads him to use the spherical harmonics (SHs), which are eigenfunctions of the laplacian ∇_q^2 , as an example to compound a complete orthonormal set of functions. These latter, namely $U_{n,m}$ and $V_{n,m}$, are expressible as linear combinations, with quaternion coefficients, of the spherical harmonics as follows:

$$U_{n,m} = -(-1)^n \left(\frac{n-m}{2n+1} \right)^{\frac{1}{2}} B_{n,m+1}^{q\text{-QSH}} + \left(\frac{n+m+1}{2n+1} \right)^{\frac{1}{2}} B_{n,-m}^{q\text{-QSH}}, \quad (3.4)$$

$$V_{n,m} = (-1)^n \left(\frac{n+m+1}{2n+1} \right)^{\frac{1}{2}} B_{n,m+1}^{q\text{-QSH}} + \left(\frac{n-m}{2n+1} \right)^{\frac{1}{2}} B_{n,-m}^{q\text{-QSH}}, \quad (3.5)$$

where $B_{n,-m}^{q\text{-QSH}}$ is the quaternion spherical harmonics (QSHs), defined, as the name indicates on the sphere, as follows:

$$B_{n,m}^{q\text{-QSH}} = \sqrt{\frac{(2n+1)(n-m)!}{4\pi(n+m)!}} P_{n,m}(\cos(\theta)) e^{km\varphi} k^{n-m} i^m, \quad (3.6)$$

where θ denotes the co-latitude, φ the longitude and $P_{n,m}$ is the real-valued associated Legendre polynomial. Note that $e^{km\varphi}$, rather than $e^{im\varphi}$, must be used in this expression.

Further theorems on the functions $U_{n,m}$ and $V_{n,m}$ have been established to highlight the power of quaternion algebra over the vector algebra.

Eqs. (3.7) show the first $B_{n,m}^{q\text{-QSH}}$, which can be expressed in spherical coordinates or be alternately written as functions of the Cartesian coordinates x , y and z .

$$\begin{aligned}
 n = 0 & \left\{ B_{0,0}^{q\text{-QSH}}(\theta, \varphi) = \frac{1}{2} \frac{1}{\sqrt{\pi}} \right. \\
 n = 1 & \left\{ \begin{aligned} B_{1,-1}^{q\text{-QSH}}(\theta, \varphi) &= \frac{1}{2} \sqrt{\frac{3}{2\pi}} \sin \theta e^{-k\varphi} i = \frac{1}{2r} \sqrt{\frac{3}{2\pi}} (xi - yj) \\ B_{1,0}^{q\text{-QSH}}(\theta, \varphi) &= \frac{1}{2} \sqrt{\frac{3}{\pi}} \cos \theta k = \frac{1}{2r} \sqrt{\frac{3}{\pi}} zk \\ B_{1,1}^{q\text{-QSH}}(\theta, \varphi) &= -\frac{1}{2} \sqrt{\frac{3}{2\pi}} \sin \theta e^{k\varphi} i = \frac{1}{2r} \sqrt{\frac{3}{2\pi}} (-xi - yj) \end{aligned} \right. \\
 n = 2 & \left\{ \begin{aligned} B_{2,-2}^{q\text{-QSH}}(\theta, \varphi) &= -\frac{1}{4} \sqrt{\frac{15}{2\pi}} \sin^2 \theta e^{-2k\varphi} = \frac{1}{4r^2} \sqrt{\frac{15}{2\pi}} (-x^2 + y^2 + 2xyk) \\ B_{2,-1}^{q\text{-QSH}}(\theta, \varphi) &= \frac{1}{2} \sqrt{\frac{15}{2\pi}} \sin \theta \cos \theta e^{-k\varphi} j = \frac{1}{2r^2} \sqrt{\frac{15}{2\pi}} (xzi + xzj) \\ B_{2,0}^{q\text{-QSH}}(\theta, \varphi) &= -\frac{1}{4} \sqrt{\frac{5}{\pi}} (3 \cos^2 \theta - 1) = \frac{1}{4r^2} \sqrt{\frac{5}{\pi}} (x^2 + y^2 - 2z^2) \\ B_{2,1}^{q\text{-QSH}}(\theta, \varphi) &= -\frac{1}{2} \sqrt{\frac{15}{2\pi}} \sin \theta \cos \theta e^{k\varphi} j = \frac{1}{2r^2} \sqrt{\frac{15}{2\pi}} (xzi - xzj) \\ B_{2,2}^{q\text{-QSH}}(\theta, \varphi) &= -\frac{1}{4} \sqrt{\frac{15}{2\pi}} \sin^2 \theta e^{2k\varphi} = \frac{1}{4r^2} \sqrt{\frac{15}{2\pi}} (-x^2 + y^2 - 2xyk) \end{aligned} \right. \quad (3.7)
 \end{aligned}$$

3.2.2 Quaternary Orthogonal Moments μ -based Computation

In this subsection, we provide the general formula for the computation of the state-of-the-art quaternary moments. Note that in this thesis we focus on the quaternary moments with orthogonal basis function (QOMs) for color image representation. We recall that the advantages of these moments is that they preserve the maximum information of shape and color. In fact, the orthogonal basis functions are employed to deal with shape information while the color information is integrated by using quaternions. Thus, the quaternary moments of order n and repetition m of a color image $f(r, \varphi)$, defined in polar coordinates, are formulated as follows:

$$QOM_{n,m}(f) = \int_0^{2\pi} \int_0^1 f(r, \varphi) \otimes (B_{n,m}^\mu)^* r dr d\varphi, \quad (3.8)$$

where $\otimes$ is the quaternion cross product and $*$ is the quaternion conjugate, respectively given in Eqs. (2.19;2.16), $[0, 1] \times [0, 2\pi]$ refers to the unit disc and $B_{n,m}^\mu$ is the basis function expressed as:

$$B_{n,m}^\mu = R_{n,m}(r) e^{\mu m \varphi}, \quad (3.9)$$

where R are the real-valued radial polynomials corresponding to the moments being calculated

and $e^{\mu m \varphi}$ is the quaternary kernel where μ is a unit pure quaternion parameter, such as $\mu = ai + bj + ck$ and $||\mu|| = 1$.

Table 3.1 presents some radial orthogonal polynomials R needed for the computation of their corresponding moments as well as the conditions over n and m .

Table 3.1 Real-valued radial polynomials R needed for the computation of some orthogonal moments.

Disc-harmonic moments
$R_{n,m}(r) = \sqrt{\frac{(2n+1)(n-m)!}{4\pi(n+m)!}} P_{n,m} \left(\cos \left(\frac{\pi}{2} r \right) \right)$ where $P_{n,m}(x)$ is the real-valued associated Legendre polynomial. The order n and repetition m verify the condition: $m \leq n$
Zernike moments
$R_{n,m}(r) = \frac{n+1}{\pi} \sum_{k=0}^{(n- m)/2} \frac{(-1)^k (n-k)!}{k! \left(\frac{n+ m }{2} - k \right)! \left(\frac{n- m }{2} - k \right)!} r^{n-2k}$ The order n and repetition m verify the conditions: $m \leq n$ and $n - m$ being even
Bessel-Fourier moments
$R_{n,m}(r) = \frac{1}{2\pi a_n} J_\nu(\lambda_n r)$ where $J_\nu(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k! \Gamma(\nu+k+1)} \left(\frac{x}{2} \right)^{\nu+2k}$ is the 1st kind Bessel polynomial of order ν, $\Gamma(x)$ is the gamma function, λ_n is the n^{th} zero of the Bessel polynomial and $a_n = [J_{\nu+1}(\lambda_n)]^2 / 2$ is the normalization constant.

3.2.3 Parameter-Free Quaternion Orthogonal Moments

The state-of-the-art quaternary orthogonal moments are computed using basis functions that depend on the choice of a unit pure quaternion parameter μ as shown in Eq. (3.9). Setting this parameter to its optimal value is very consuming in terms of computation and time. Besides the choice of μ is purely empirical and does not lie on any mathematical basis.

On the other hand, Gough, 1984 managed to mathematically construct a set of spherical harmonics with quaternary coefficients (QSH) for application in three-dimensional quantum mechanics. The formulation of these QSH is simple as it consists of the introduction of post-multiplicative quaternion phase factors into the complex-valued spherical harmonics. Inspired by Gough's work, we propose a set of parameter-free quaternary moments for describing color images. In fact, the μ -dependent angular part $e^{\mu m \varphi}$ in the definition of the conventional quaternary moments [Eq. (3.9)] is replaced by the parameter-free angular part $e^{km\varphi} k^{n-m} i^m$. This latter is then multiplied by the corresponding moments' real part radial orthogonal polynomials. This results in a simple reformulation of the quaternary moments' basis function, which is free from any parameter as:

$$B_{n,m}^q = R_{n,m}(r) e^{km\varphi} k^{n-m} i^m. \quad (3.10)$$

Obviously, we could also have defined $B_{n,m}^q$ to be $R_{n,m}(r) i^m k^{n-m} e^{km\varphi}$, as the multiplication of quaternions is not commutative. We note that, in this work, we adopted the expression in Eq. (3.10). This latter is free from any parameter, unlike the basis function used by the state-of-the-art quaternary moments given in Eq. (3.9).

We show in Figure 3.1 visual representations of the four parts of the quaternion-valued harmonic function (By replacing R in Eq. (3.10) with the dis-harmonic radial polynomials given in Table 3.1) of order up to $n = 8$ and $0 \leq m \leq n$.

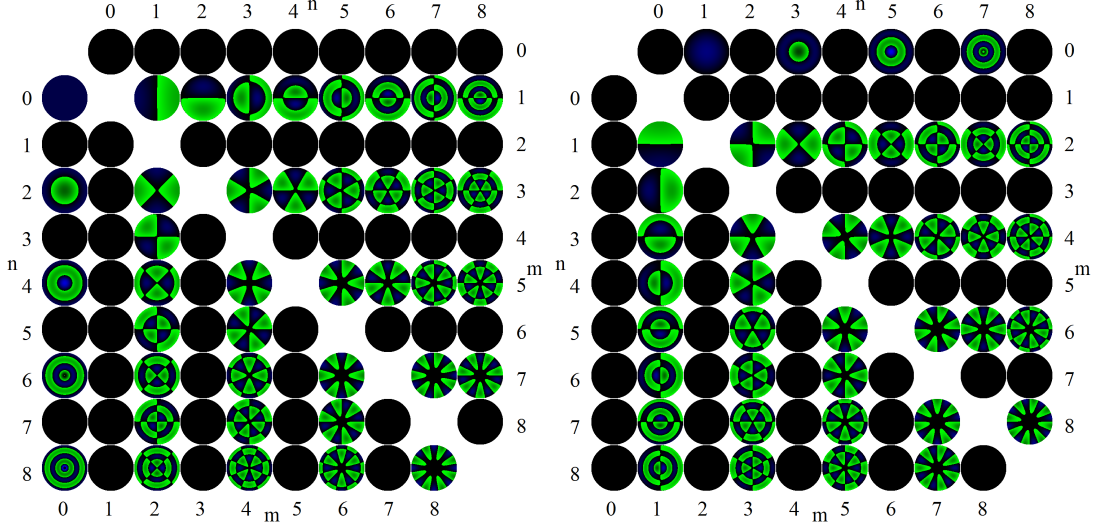


Figure 3.1 Parameter-free harmonic function up to order $n = 8$. At left, the top triangle represents the first imaginary parts whilst the bottom triangle represents the real parts. At right, the top triangle represents the third imaginary parts whilst the bottom triangle represents the second imaginary parts. The repetition m varies between 0 and n . Green pixels represent regions where the function is negative, blue ones represent where it is positive and the black ones are where the function is zero.

Table 3.2 depicts the functions $B_{n,m}^q$ for an order n up to 3, where $B_{n,m}^c$ is the complex-valued basis function expressed with $e^{km\varphi}$ instead of $e^{im\varphi}$.

Table 3.2 Parameter-free quaternary orthogonal basis function $B_{n,m}^q$ up to order $n = 3$ in case $|m| \leq n$.

$B_{0,0}^c$					
$B_{1,-1}^c i$		$B_{1,0}^c k$	$B_{1,1}^c i$		
$-B_{2,-2}^c$	$B_{2,-1}^c j$	$-B_{2,0}^c$	$B_{2,1}^c j$	$-B_{2,2}^c$	
$-B_{3,-3}^c i$	$-B_{3,-2}^c k$	$-B_{3,-1}^c i$	$-B_{3,0}^c k$	$-B_{3,1}^c i$	$-B_{3,2}^c k$
$-B_{3,3}^c i$					

Consequently, the parameter-free quaternary orthogonal moments (PFQOMs) of order n and repetition m of a color image $f(r, \varphi)$ [Eq. (1.26)], defined in polar coordinates, are formulated as follows:

$$PFQM_{n,m}(f) = \int_0^{2\pi} \int_0^1 f(r, \varphi) \otimes (B_{n,m}^q)^* r dr d\varphi, \quad (3.11)$$

Based on the Zero-Order Approximation, the PFQOMs for an image of size $N \times N$ can be numerically computed as follows:

$$PFQM_{n,m}(f) = \frac{\sqrt{2\pi}}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(r, \varphi) \otimes (B_{n,m}^q)^*. \quad (3.12)$$

The proposed PFQOMs inherit the property of completeness and orthogonality of their basis function on the unit disc according to the following formula:

$$\int_0^{2\pi} \int_0^1 [B_{n,m}^q]^* \otimes B_{n',m'}^q r dr d\varphi = \delta_{nn'} \delta_{mm'}, \quad (3.13)$$

where δ is the Kronecker symbol.

This property makes the proposed moments free of shape information redundancy. Thus the RGB image reconstruction from them is possible and it is obtained by:

$$f^r(r, \varphi) = \sum_{n=0}^{n=nMax} \sum_{m=-n}^{m=n} PFQOM_{n,m} \otimes B_{n,m}^q. \quad (3.14)$$

where $f^r(r, \varphi)$ is the reconstructed image up to the maximum order $nMax$.

In this work, we particularly address the case of the disc-harmonic moments, Zernike moments and the Bessel Fourier moments.

3.3 Experimental Results

3.3.1 Color Image Sets

In addition to Columbia Object Image Library (COIL-100) (Nene et al., 1996) and Amsterdam Library of Object Images (ALOI) Viewing Direction, Color Illumination and Illumination Direction change (Geusebroek et al., 2005), whose structures are explained in subsection. 2.4.1, two color face images are employed in the following experiments:

- The Faces96 collection of the University of Essex (Spacek, 2007), which is categorized into 152 individuals, each with 20 images introducing significant head variations as shown in Figure 3.2.

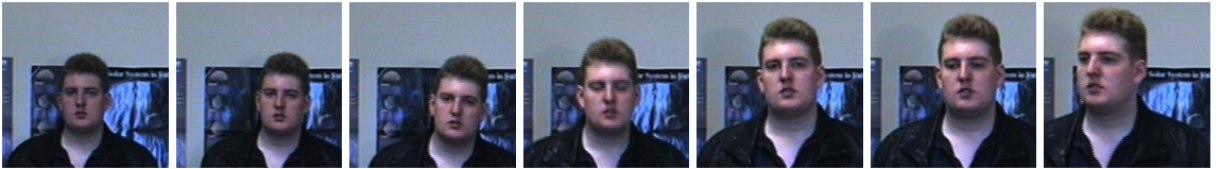


Figure 3.2 Some head variations of one individual from Faces96 collection.

- The Faces95 collection of the University of Essex (Spacek, 2007) (Figure 3.3), which consists of 72 individuals each with 20 images of size 180×200 pixels (portrait format) where the background is a red curtain. This collection is characterized by a large head scale variation.



Figure 3.3 Six images for one individual from Faces95 collection, showing some major variation of head scale.

3.3.2 Color Image Reconstruction

We provide some color image reconstruction examples of the proposed PFQOMs. Particularly, we consider the PFQDHMs, which are based on the associated Legendre polynomials. We first compute the PFQDHMs up to $nMax = 80$ for four images, i.e., the standard Barbara and Lena, an object image from Columbia Object Image Library (COIL-100) (Nene et al., 1996) and a face image from Max Planck Institute faces database (MPIK) (Troje and Bülthoff, 1996). Then we depict in Figure 3.5 the reconstructed color images and their NIRE [Eq. (2.35)] values up to different values of $nMax$.

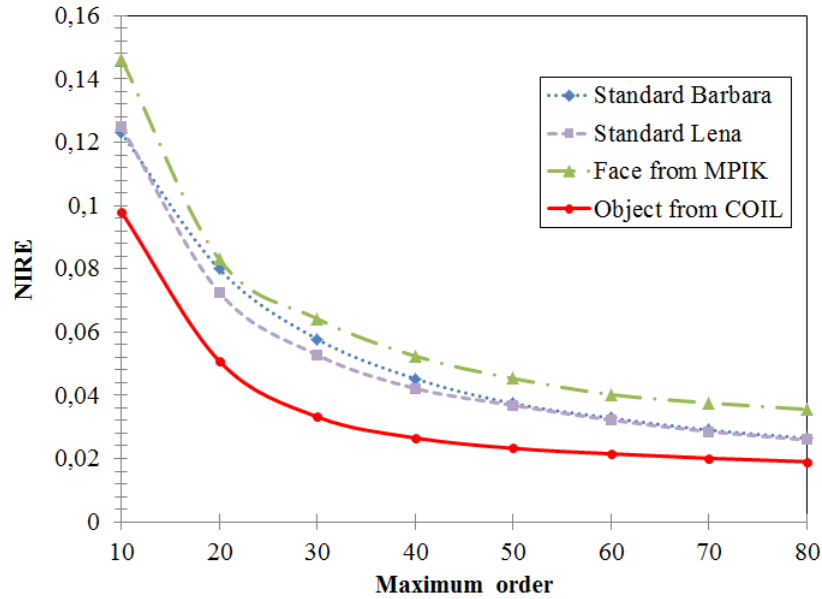


Figure 3.4 Reconstruction errors (NIRE) of the PFQDHMs up to orders from 10 to 80 with a step of 10 for the standard Lena and Barbara, a face image from MPIK and an object image from COIL-100, all of size 128×128 .

It can be seen that the quality of the reconstructed color images from the PFQDHMs depends on the size of the feature vector, i.e., the more quaternary coefficients the feature vector contains the more the reconstructed image is faithful to the original image. Besides, even at low orders, the reconstructed images are close to the original image and one can easily detect the shape of the object contained in the image. Also from Figure 3.4, which represents the reconstruction errors up to different orders, we remark that the lowest NIRE values are those of the object image. Generally speaking, the PFQOMs possess a high capacity for representing the color images content, especially object images.

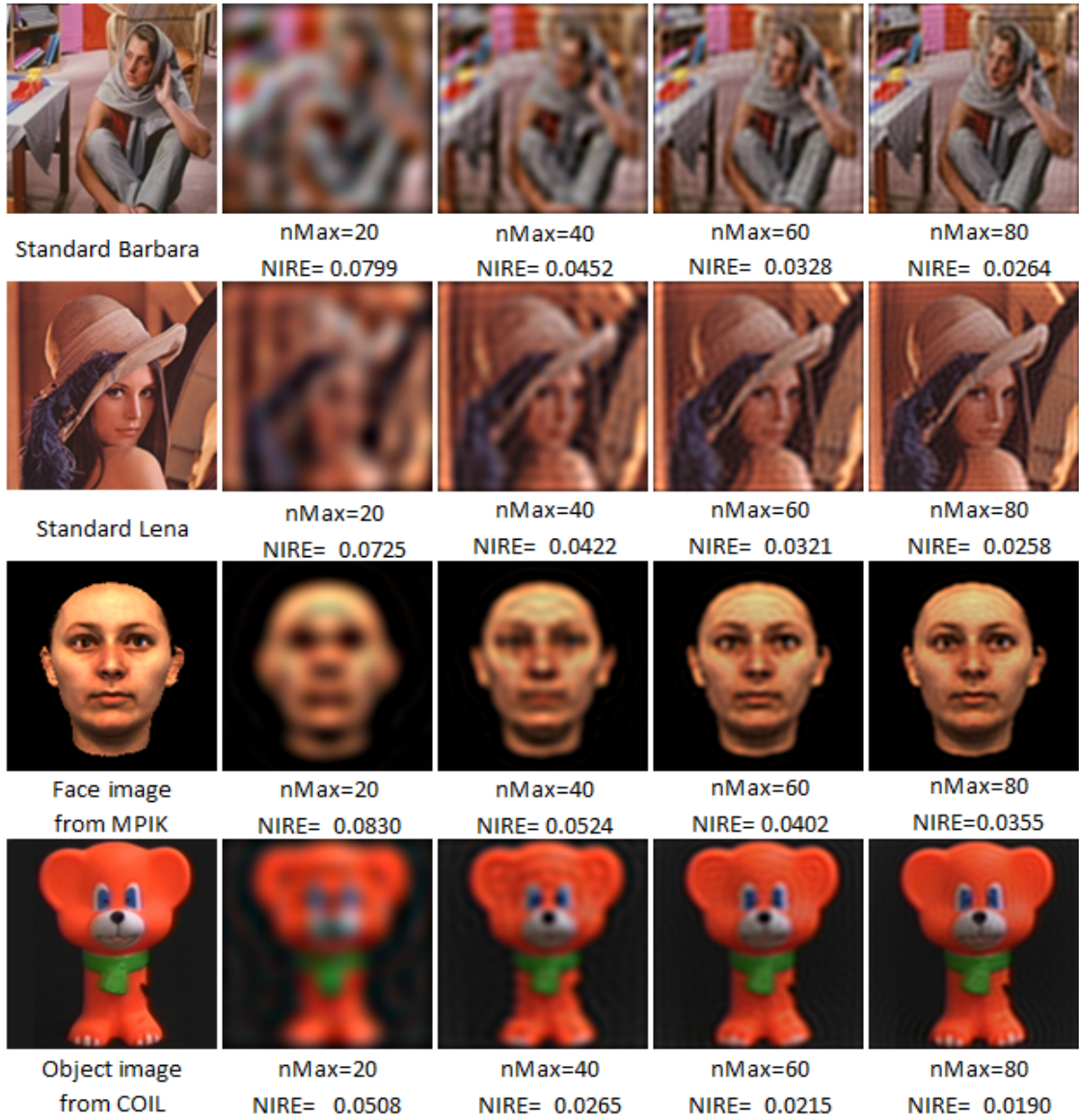


Figure 3.5 Example of reconstructed images of size 128×128 from the PFQDHMs. The orders are 20,40,60 and 80.

3.3.3 Robustness and Invariance

The first 6 modulus-based PFQZMs invariants (up to $nMax = 2$) are reported in Table 3.3 for rooster image, selected from Caltech 101 dataset (Fei-Fei et al., 2004), under various photometric and geometric transformations, i.e., Salt and pepper noise (3% and 5%), blur (15 and 30), rotation (90°) and scaling (50% and 150%). The corresponding mean squared errors (MSE) for $nMax$ varying between 5 and 45 with an increment of 5 are shown in Figure 3.6.

In fact, the MSE measures the difference between the modulus-based PFQOMIs of the original image f and those of its transformed versions f^t as follows:

$$MSE = \frac{1}{S} \sum_{n=0}^{n=nMax} \sum_{m=-n}^{m=n} (||PFQOM_{n,m}(f)|| - ||PFQOM_{n,m}(f^t)||)^2. \quad (3.15)$$

where S is the size of the feature vector. Note that in case of the PFQZMs, the condition $n - |m|$ being even, is to be considered.

Table 3.3 Modulus-based PFQZMs invariants up to $nMax = 2$ for rooster image (Fei-Fei et al., 2004) of size 300×300 under various photometric and geometric transformations. The last two columns represent the corresponding MSE for $nMax = 2$ and $nMax = 8$, respectively.

	Image	$M_{0,0}$	$M_{1,-1}$	$M_{1,1}$	$M_{2,-2}$	$M_{2,0}$	$M_{2,2}$	MSE (nMax=2)	MSE (nMax=8)
Original		221.59	9.71	10.40	13.92	168.94	11.86	—	—
SP=3%		218.93	9.36	10.01	13.45	167.81	11.49	1.49	1.58
SP=5%		217.33	9.18	9.84	13.11	167.20	11.20	3.80	3.66
Blur=15%		221.56	9.74	10.40	14.10	169.01	12.04	0.01	0.03
Blur=30%		221.57	9.75	10.42	14.10	169.30	12.04	0.03	0.15
Rotation=90°		221.71	9.76	10.39	13.86	169.34	11.87	0.03	0.03
Scale=50%		221.99	10.17	10.78	14.08	170.20	12.14	0.36	0.50
Scale=150%		221.66	9.65	10.30	13.94	169.03	11.90	0.004	0.02

One can see from these results that the PFQZMI coefficients of the original image and its blur, rotation and scale transformed versions are very similar regardless of the maximum order ($MSE \leq 0.64$). On the other hand, the SP noisy versions are a bit far from the original image compared to the other transformed images, but the MSE is still sufficient ($MSE \leq 4.62$) to tell that the PFQZMIs are robust against SP noise. In general, our low-order modulus-based PFQOMIs are invariant to geometric transformations and are robust against various attacks.

Figure 3.6 also shows that generally the MSE decreases with the increase of the maximum order. However, when it comes to highly blurred images, one can remark the opposite, i.e., as the maximum order increases, the MSE slightly increases. This behavior is obvious since low-order moments have better tolerance to blur distortions of the image.

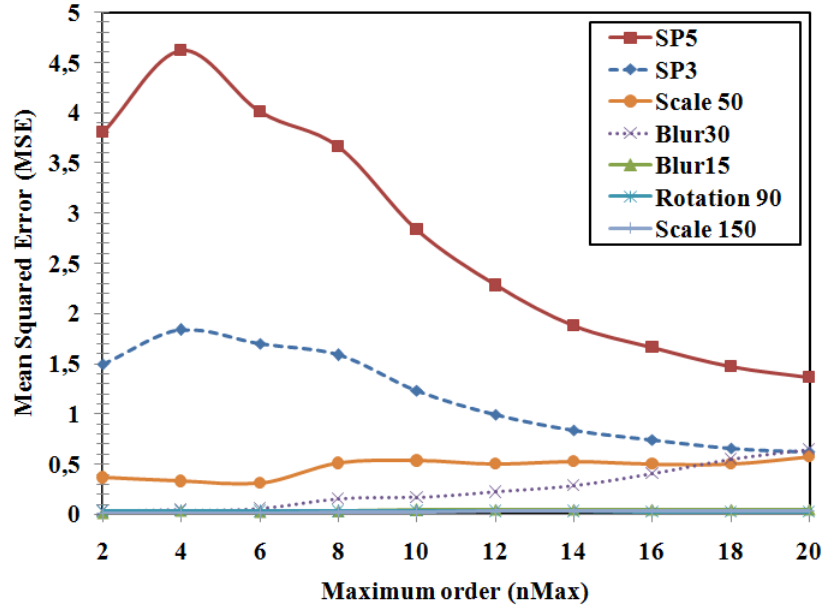


Figure 3.6 Mean squared error (MSE) of the different transformed images up to different orders.

3.3.4 Color Image Retrieval

In this subsection, we evaluate the retrieval performance, based on k-nearest-neighbor (k-NN), of our PFQOMs, i.e., PFQDHMs, PFQBFMs, PFQZMs against:

- Their μ -based versions,
- The chromaticity moments (CMs) (Paschos et al., 2003),
- The Legendre chromaticity distribution moments (LCDMs) (Yap and Paramesran, 2006),
- The Clifford-based parallel-orthogonal Fourier moments (poFMD) (Mennesson et al., 2014).

Recall that the CMs (Paschos et al., 2003) are the regular moments of both the trace and the distribution of the xy chromaticity space of a color image; therefore, require the construction of the chromaticity histogram by quantizing the chromaticity space into histogram bins. On the other hand, the LCDMs (Yap and Paramesran, 2006) are computed directly from the opponent chromaticity space without the need to construct the chromaticity histogram. Both CMs and LCDMs were applied for content-based image retrieval based on the L_1 metric for similarity measurement.

In this work, we compute the CMs for two quantization levels $Q = 10, 20$ and the poFMD for the red bivector B_{red} . The quaternary moments we address are based on the Zernike radial polynomials, the 1st kind Bessel moments and the associated Legendre polynomials (See Table. 3.1). For the μ -based quaternary moments, we set μ to $(i - j - k)/\sqrt{3}$ as established by Chen et al., 2014a.

First, we extract the feature vectors up to $nMax = 8$ ($nMax = 8$ is chosen based on the results obtained in our first contribution detailed in Chapter 2), from the color images in the database. Then, we compute the similarity between the query color image descriptor (each image in the database) and the database color images' descriptors. Finally, we compute the precision at each value of k ($1 \leq k \leq 72$ for the COIL-100 and $1 \leq k \leq 20$ for the Faces96 collection). We report the precision using the first nearest neighbor (1-NN) in Table 3.4, then the overall average

precision in Figs. 3.7 and 3.8.

Table 3.4 Precision in % using the first nearest neighbor (1-NN) for the different color image moments when processing the COIL-100 and the Faces96 collections.

Bold characters refer to the results obtained by our moments.

	PFQDHMs	QDHMs	PFQZMs	QZMs	PFQBFMs	QBFMs	poFMD	LCDM	CM ₁₀	CM ₅₀
COIL-100	100	99.91	99.91	99.89	99.46	99.38	99.56	96.51	92.96	86.38
Faces96	99.12	99.09	98.94	98.84	94.62	94.43	97.70	86.89	71.22	65.21

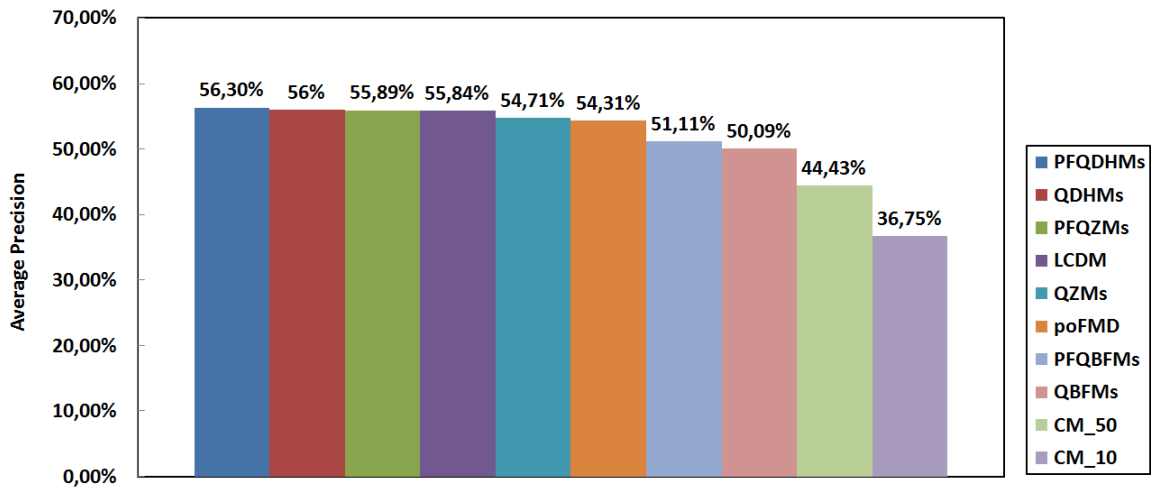


Figure 3.7 Average precision using different color image moments when processing the COIL-100 database.

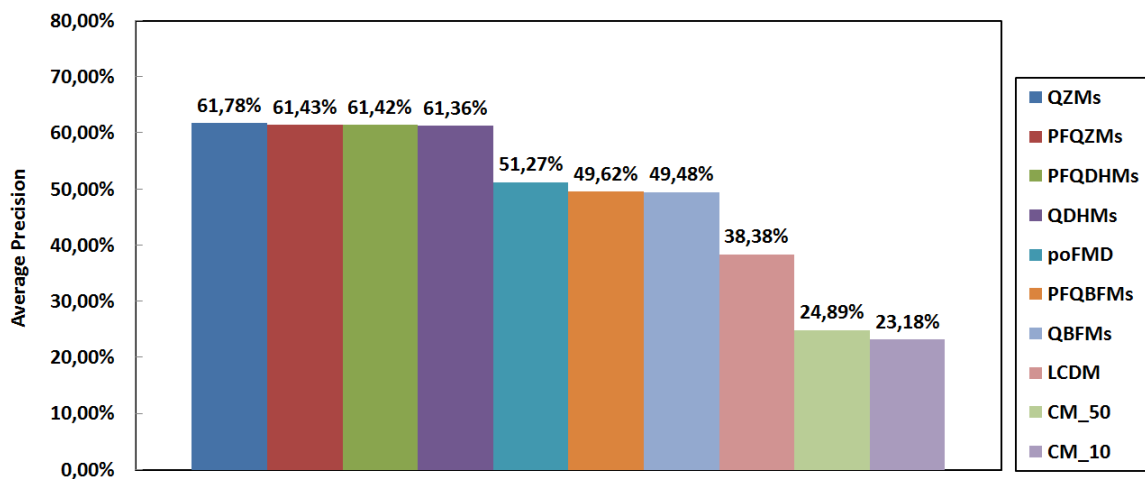


Figure 3.8 Average precision using different color image moments when processing the Faces96 collection.

Note that for similarity measurement, the Chi-square distance is used with our parameter-free quaternary moments, their μ -based versions and the poFMD while the L_1 distance is employed

with the CMs and the LCDMs as established in their original work (Paschos et al., 2003; Yap and Paramesran, 2006).

According to Figure 3.7 and 3.8, we can see that our PFQOMs have discriminative powers that are acceptable and comparable to their μ -based versions. This is because that the quaternary orthogonal moments are based on orthogonal polynomials and take advantage of the quaternary cross product; therefore, can efficiently capture the color and shape information in an image. It is also seen that both the quaternary moments based on the associated Legendre polynomials, i.e., PFQDHMs and QDHMs, and the Zernike radial polynomials yield the highest retrieval rates compared to the moments based on the 1st kind of Bessel Fourier moments especially when dealing with images of faces captured on a complex background. The poFMD and LCDM also allow satisfactory retrieval results comparable to those of the orthogonal moments. Furthermore, the results of CMs are the worst and those with $Q = 50$ are slightly superior to those with $Q = 10$.

In this experiment, we also measure the performance of the proposed methods based on the area under the receiver operating characteristic (ROC) curve (AUC). In fact, the ROC is a plot of the true acceptance rate (TAR=1-FRR, the false rejection rate) against the false acceptance rate (FAR). These two latter metrics are computed at each threshold (20 different threshold values are considered in this experiment) as follows:

$$\text{FRR} = \frac{\text{No. of false rejected genuine}}{\text{No. of genuine accesses}}, \quad (3.16)$$

$$\text{FAR} = \frac{\text{No. of false accepted impostors}}{\text{No. of impostor accesses}}. \quad (3.17)$$

Figure 3.9 and 3.10 show the ROC curves generated for the different color image moments using the k-NN, whilst Table 3.5 tabulates the corresponding AUC values.

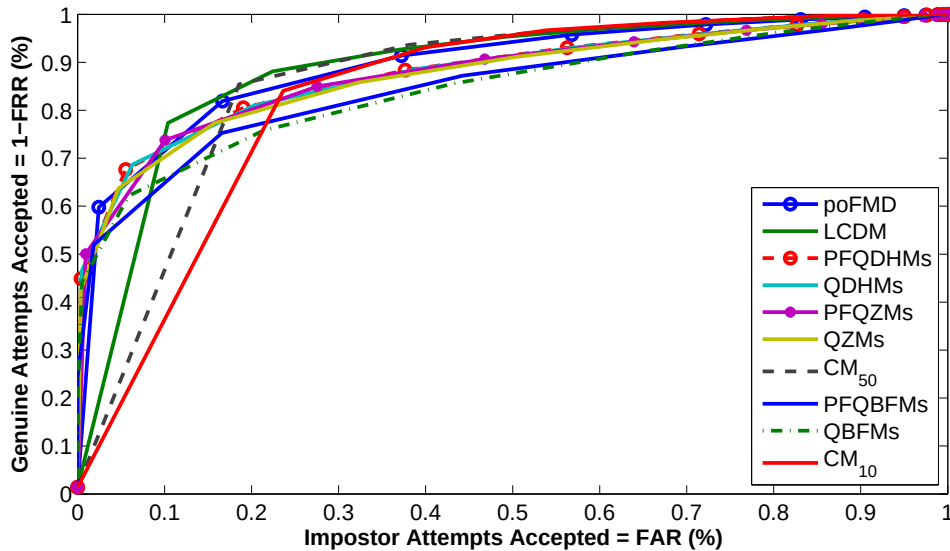


Figure 3.9 ROC curves of different color image moments when processing COIL-100 database.

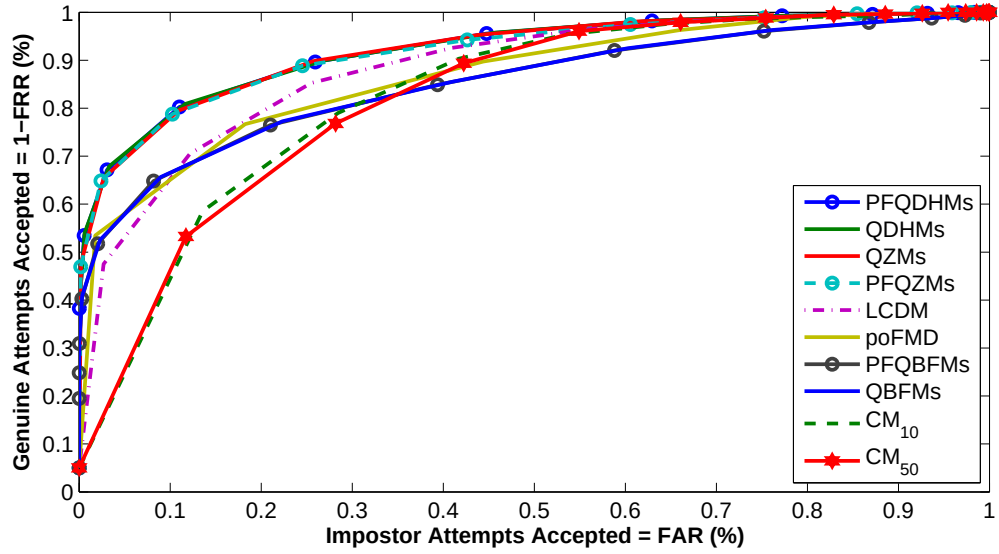


Figure 3.10 ROC curves of different color image moments when processing Faces96 collection.

Table 3.5 Area under curve (AUC) in % for the different color image moments when processing the COIL-100 and the Faces96 collections. Bold characters refer to the results obtained by our moments.

	PFQDHMs	QDHMs	PFQZMs	QZMs	PFQBFMs	QBFMs	poFMD	LCDM	CM ₁₀	CM ₅₀
COIL-100	87.95	87.81	87.63	87.33	84.77	84.46	89.41	88.55	83.33	86.10
Faces96	92.14	92.13	91.72	91.99	85.27	85.25	86.69	88.10	82.47	82.04

The AUC values obtained by the proposed quaternary moments confirm that they are effective in CBIR regardless of the viewing direction of the object in the image or of the scale head variations.

3.3.5 Color Image Classification

In this experiment, we evaluate the color image classification performance of our modulus-based PFQOMIs fed into the Extreme Learning Machine classifier (ELM) (Ding et al., 2015). First, we divide the database into a training set (70%) and a testing set (30%), for each of which, we extract the color image moments up to $nMax = 5$ and normalize them using min-max. Then, for training, we feed the color image features of the training set into an ELM classifier to learn its optimal parameters. As for testing, we use the learned weights to compute the output and provide the final decision based on the maximum of these outputs.

Note that the number of ELM's input layer nodes is the number of the moments' coefficients, e.g., 36 for PFQDHMs, PFQBFMs, QDHMs and QBFMs and 21 for PFQZMs and QZMs, whilst the number of output layer nodes is the number of classes (100 classes for both COIL-100 and Faces96). Also, the classification accuracies, presented hereafter, are average values in 10 trials using the same combinations of training/test samples. In fact, the ELM gives slightly different classification accuracies every time it is trained even though the training and test data is the same. This is due mainly to the random weights assigned every time it is trained.

The classification accuracy is measured as follows:

$$\text{Classification accuracy} = \frac{\text{No. of correctly classified images}}{\text{total number of images in testing set}}. \quad (3.18)$$

The classification results using an ELM with different numbers of hidden layer nodes are illustrated in Figure 3.11. Also in Table 3.6, we recapitulate the highest classification results.

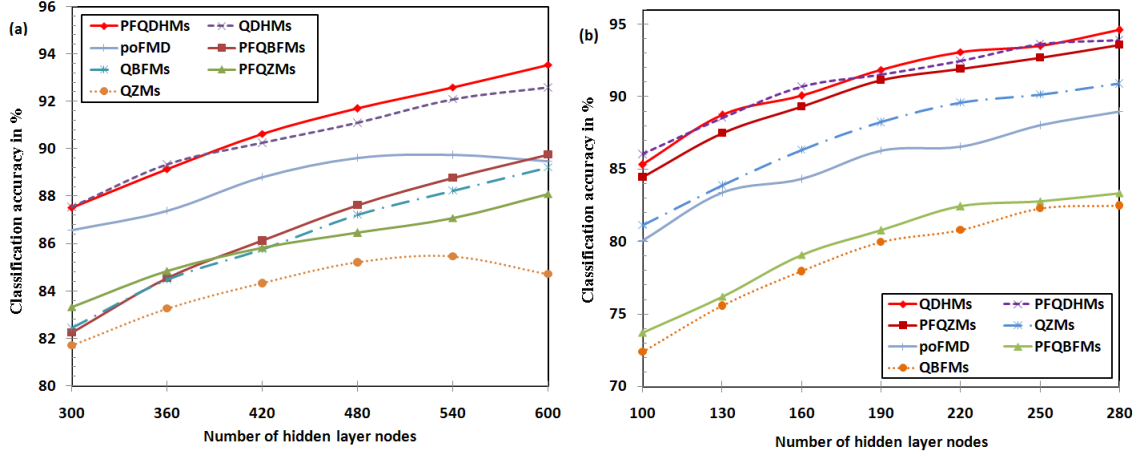


Figure 3.11 Classification accuracies in % of different color image moments using an ELM with different numbers of hidden layer nodes when processing (a) COIL-100 database and (b) Faces96 collection.

Table 3.6 Classification accuracies in % for the different color image moments using the ELM with 600 hidden layer nodes when processing the COIL-100 and with 280 hidden layer nodes when processing Faces96 collection. Bold characters refer to the results obtained by our moments.

	PFQDHMs	QDHMs	PFQBFMs	QBFMs	PFQZMs	QZMs	poFMD	LCDM	CM ₅₀	CM ₁₀
COIL-100	93.55	92.57	89.76	89.21	88.09	84.72	89.74	22.52	17.40	17.77
Faces96	93.92	94.61	83.34	82.5	93.6	90.6	89.01	74.74	64.34	55.58

One can see from the results that (1) Generally, our PFQOMs yield higher classification accuracies compared to their μ -based version dealing with both the COIL-100 and the Faces96 collections regardless of the ELM's hidden layer nodes number. (2) The PFQDHMs and the QDHMs are the most performing in recognizing objects and are comparable to the PFQZMs when dealing with faces captured on complex background. (3) The classification accuracies generally increase with the increase of the ELM's hidden layer nodes number. (4) The orthogonality property of the quaternary orthogonal moments make them capable of capturing both the shape and the color information; thus, they all allow significant classification accuracies (≥ 84.72 for COIL-100 and ≥ 82.5 for Faces96).

In another test, we fix the number of hidden layer nodes of the ELM to 600 and 280 for COIL-100 and Faces96, respectively, then we vary the maximum order between 4 and 8 with a step of 1. The classification results are shown in Figure 3.12.

The results show that there is a significant difference between the classification accuracies of the quaternary moments at $nMax = 4$ and at $nMax = 5$. On the other hand, when the

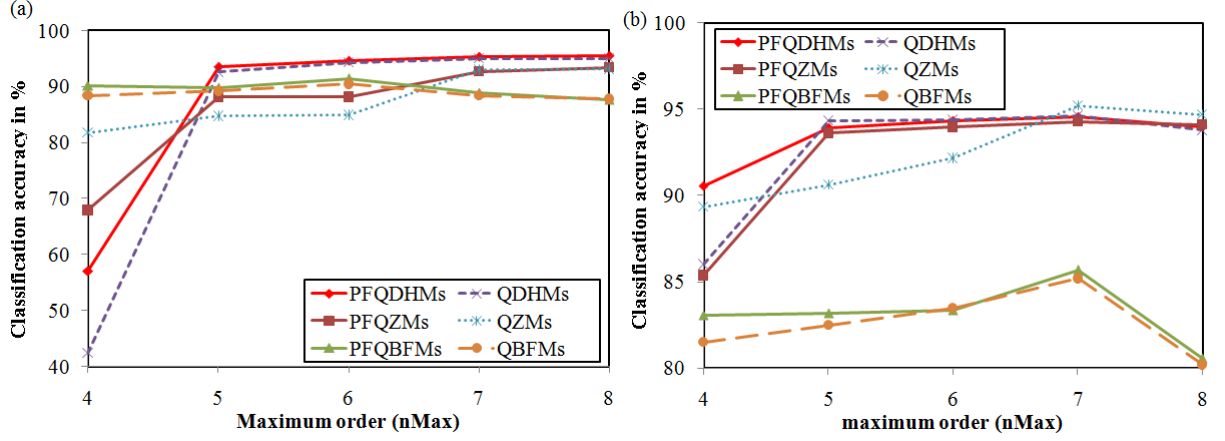


Figure 3.12 Classification accuracies in % of different color image moments up to different maximum orders $nMax$ using the ELM when processing (a) COIL-100 database and (b) Faces96 collection.

maximum order exceeds 5 ($nMax \geq 5$), the classification accuracies increase slightly and then begin to decrease. This is due to over-fitting, which is usually caused by a bad sizing of the structure used to classify, i.e., wrong choice of the number of hidden layer nodes and input nodes. Therefore, in these experiments, $nMax = 5$ with 600 and 280 as numbers of hidden layer nodes for COIL-100 and Faces96, respectively, seems to be an optimal combination to define the size of the ELM structure.

3.3.6 Computational Time

In this experiment, we investigate the time consumed in the computation of the different color image moments for an image of size 128×128 . At each maximum order $nMax$ varying between 5 and 30 with a step of 5, we report, in Table 3.7 and Figure 3.13, the times in seconds elapsed in the computation of different color image moments and the corresponding feature vector size, respectively. The experiment is carried out on a personal computer i3-2310M 2.10 GHz CPU, 4 Go RAM.

Table 3.7 Elapsed CPU times in seconds for the computation of different color image moments of an image of size 128×128 up to different maximum orders $nMax$.

$nMax$	PFQDHMs	QDHMs	PFQBFMs	QBFMs	PFQZMs	QZMs	LCDM	CM ₅₀	CM ₁₀	poFMD
5	1.03	1.37	1.63	1.97	0.78	0.98	1.63	1	0.94	0.13
10	3.96	4.68	5.53	6.63	2.78	3.44	4.74	3.17	3.13	0.13
15	8.26	10.31	11.62	14.03	9.01	10.87	9.96	6.65	6.38	0.14
20	15.08	18.26	19.98	24.06	18.14	20.38	18.27	11.3	10.96	0.19
25	24.95	29.00	30.64	36.88	47.69	51.22	28.78	16.82	16.67	0.25
30	33.53	42.21	43.64	52.49	140.09	144.68	44.32	23.99	23.4	0.33

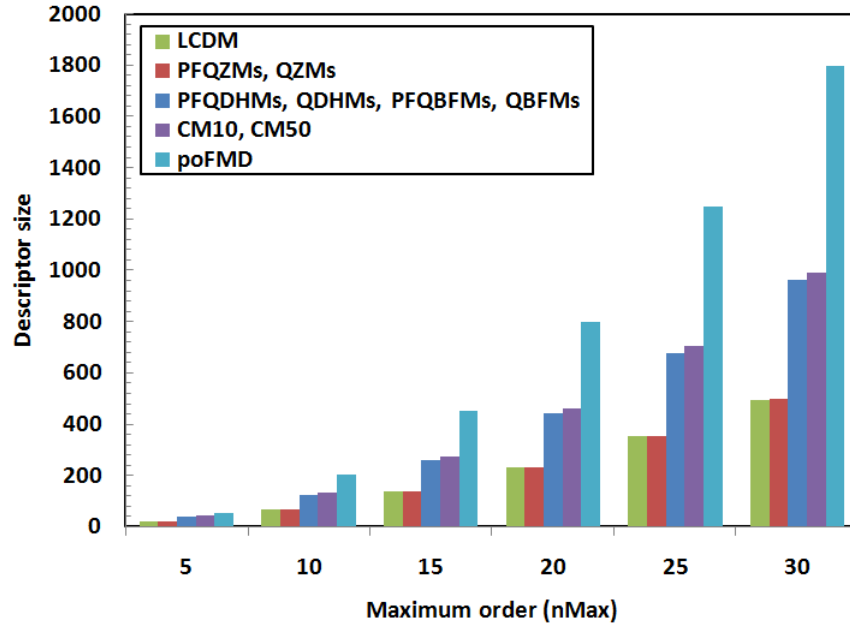


Figure 3.13 Feature vector size for each of the color image moments up to different maximum orders $nMax$. The size of the PFQDHMs, PFQBFMs, QDHMs and QBFMs is $(nMax + 1)^2$ moments, of the PFQZMs and QZMs is $(nMax + 1)(nMax + 2)/2$ moments, of LCDM is $nMax(nMax + 3)/2$ moments, for CM is $(nMax + 1)(nMax + 2)$ moments and for poFMD is $2nMax^2$ moments.

From the results, one can remark that:

(1) the computation of the proposed PFQOMs is generally slightly faster than their μ -based version especially for high-order moments. And since the basis functions $B_{n,m}^\mu$ [Eq. (3.9)] and $B_{n,m}^q$ [Eq. (3.10)] are independent from the image content, they can be pre-computed and stored; therefore, not be considered. On the other hand, since $f(r, \varphi)$ is a pure quaternion, i.e., the real part is set to zero and the μ -based $B_{n,m}^\mu$, where $\mu = (i - j - k)/\sqrt{3}$ is a full quaternion, i.e., all the four parts are non-zero; their quaternionary inner product $f(r, \varphi) \otimes B_{n,m}^\mu$ necessitates 12 multiplications and 8 additions. For an image with size $N \times N$, the total number of operations is $12N^2$ multiplications and $8N^2$ additions. Regarding our PFQOMs, one can remark from Table 3.2 that the parameter-free $B_{n,m}^q$ is a quaternion where only two parts are non-zero. Thus, the quaternionary inner product $f(r, \varphi) \otimes B_{n,m}^q$ involves only 6 multiplications and 2 additions resulting in a total of $6N^2$ multiplications and $2N^2$ additions for an image with size $N \times N$. From this analysis, our PFQOMs can be 2 times faster than their μ -based version if we do not consider the computation of their basis functions.

It is also worth noting that $\mu = (i - j - k)/\sqrt{3}$ could not be the optimal value for a given application. Finding optimal μ among a set of infinite number can be very time consuming and computationally expensive. For instance, Chen et al., 2014a discussed the choice of this parameter for different applications (color image reconstruction, face recognition and image registration) and using 100 different values of μ , and found out that the performances of these moments seem to be very close. Therefore, a good choice of μ could slightly increase the moments retrieval performance but to the detriment of the computational time; therefore, the retrieval performance/Computational time is a trade-off to be examined when deciding which

method is more suitable for a given retrieval setting.

In sum, our moments have an advantage over the μ -based moments in terms of time requirement while allowing satisfactory color image retrieval performances.

- (2) The computational time of the moments is generally proportional to size of the feature vector.
- (3) The radial polynomials needed for the computation of the PFQZMs and the QZMs are based on factorial terms which gives them a high-computational complexity; Thus, they are slow compared to the others.
- (4) The PFQDHMs and the QDHMs are generally fast as they are based on the associated Legendre polynomials expressed by recurrences rather than factorial terms.
- (5) The computational time for the CM is slightly affected by the quantization level Q .
- (6) The Clifford-based poFMD is the fastest among the moments and its calculation is not proportional to the vector size.

3.4 Conclusion

In this chapter, we have introduced a parameter-free formulation of the quaternary orthogonal moments (PFQOMs), whose modulus-based invariants were investigated for both color image retrieval and classification.

The proposed PFQOMs have a significant color image reconstruction capability, and are robust to various photometric transformations as well as invariant under geometric transformations. Besides, their performances were compared to those of the state-of-the-art μ -based quaternary moments, the chromaticity moments, the Legendre chromaticity distribution moments as well as the Clifford-based parallel-orthogonal Fourier Mellin descriptor. With regard to color image retrieval, the k-nearest-neighbor has been used together with the Chi-squared distance as similarity measure. It has been shown that the discriminative powers of the proposed PFQOMs are satisfactory and comparable to their μ -based versions dealing with both COIL-100 and Faces96 collections.

As for color image classification, we have employed the extreme learning machine classifier (ELM) as it is known for its efficiency. High classification results were generally observed thanks to the orthogonality property of the moments. Particularly the parameter-free quaternary orthogonal moments yield good results in recognizing objects and faces.

Furthermore, the parameter-free formulation of the quaternary moments allows a faster computation compared to their μ -based versions.

Chapter 4

Proposed Color Image Classification Method

4.1 Introduction

Image classification is an active research topic in the field of pattern recognition and computer vision. This task is challenging in the way that the image is subject to various changes including linear transformations and photometric transformations. For the purpose of pattern recognition, numerous classification methods have been proposed. These methods can be divided into two categories: shallow methods and deep learning methods.

The shallow methods consist of (1) a hand-crafted feature extraction, in which the image feature descriptor is extracted allowing to reduce dimensionality as well as guarantee invariance to the different image variations and transformations, then (2) feature classification where the hand-crafted descriptor is fed into a certain classifier that outputs the class to which this pattern or image belongs. This kind of methods are achieving great success in object recognition, face recognition and texture classification.

On the other hand, the deep learning methods rely more on fully automatic learning by adopting a multi-level architecture. In fact, deep learning methods are directly fed with images, rather than feature vectors, which allows both feature extraction and classification.

Recently, many researchers have been investigating the extension of these methods (Both categories) to deal with multi-dimensional signals, such as color images, by using the algebra of quaternions. In this context, we propose a shallow color image classification method, which is based on our parameter-free quaternion disc-harmonic moments (PFQDHMs), presented in chapter 3 and a quaternary neural network that is the quaternary extreme learning machine (QELM) (Minemoto et al., 2016a). The objective is to confirm the relevance of our quaternary descriptor and validate it as part of an image classification system. The basic idea is to keep the quaternary form of the coefficients composing the PFQDHMs descriptor instead of being restricted to the moment's modulus. In addition, it is expected that neurons with quaternary compact representation and operations would be useful for classifying color images based on four-dimensional descriptors.

Our quaternary method is distinguished from the marginal methods, i.e., feeding the four parts of the quaternary coefficients separately into real-valued classifiers. In fact, the marginal methods face many issues such as degrading performance, dimension curse as well as the loss of correlation between the four parts, real and imaginary, of the quaternary descriptor coefficients. This chapter first presents a review of the most well-known real and quaternary machine-learning algorithms. Then, it details our contribution in which we show how these classifiers can be adapted to cope with the quaternary invariants that can be either real valued, i.e., based

on the modulus, or quaternion valued. Afterward, we present our color image classification method, which will be evaluated using different challenging color image collections.

4.2 Review of Machine-Learning Algorithms

Dealing with multi-dimensional signals, such as quaternion-valued descriptors, is deemed to be an important issue in artificial neural networks (ANN). These latter consist of many interconnected neurons that accept only real-valued signals for their input, hidden states, and output. Intuitively, neural networks can cope with multi-dimensional signals by configuring neurons so that each of them covers each part of these signals. However, this configuration does not take into account the natural dependency between the parts of these multi-dimensional signals, which should be processed as a single entity.

For 2-dimensional signals, many applications of complex-valued neural networks have been extensively investigated (Hirose, 2004; Hirose, 2006; Nitta, 2009). These works have motivated the extension of real-valued neural networks to handle signals with more than two dimensions. For instance, one can opt for quaternion neuron representations and operations, which is expected to be useful for processing three and four-dimensional signals. In addition, Clifford numbers can be employed for dealing with more than four-dimensional signals.

This section provides an extensive review regarding real-valued, quaternion-valued and Clifford-valued neural networks.

4.2.1 Real-valued Machine Learning Algorithms

4.2.1.1 Feed-Forward Neural Networks

Artificial neural networks (ANNs) are being widely used in a range of applications with prediction or classification problems in fields such as finance, medicine, engineering, geology and physics. Inspired by the biological nervous system, McCulloch and Pitts, 1943 were the first to introduce the notion of artificial neural networks. In fact, the nervous system is a net of cells, or neurons, which work by sending electrical signals from one to another. Figure 4.1 shows a neuron, which contains a cell body or soma and an axon. The main task of soma is to receive signals from other neurons while the axon communicates a chemical message to another nerve or muscle.

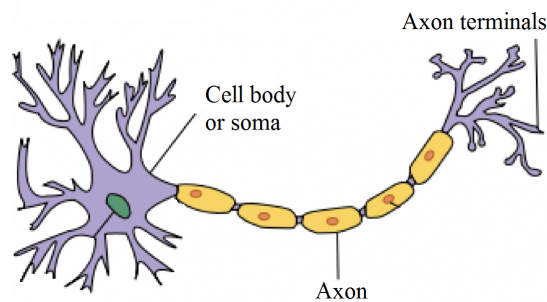


Figure 4.1 Structure of a typical neuron.

According to McCulloch and Pitts, 1943, a single neuron passes the weighted sum of its inputs via an activation function to produce the output as shown in Figure 4.2. The weighted sum is:

$$x_i = \sum_{l=1}^{l=L} w_{li} I_l, \quad (4.1)$$

where I is the input, w is the weight and L is the number of nodes from the layer that feeds the node i .

The output of a node i , in turn, is computed as:

$$y_i = f(x_i), \quad (4.2)$$

where $f(x)$ is the activation function that can take many forms such as linear, threshold, sigmoid or with radial basis as follows (Huang and Babri, 1998):

$$\text{Linear activation } f(x) = \beta x, \quad (4.3)$$

$$\text{Threshold activation } f(x) = \begin{cases} 1 & x \geq \theta \\ 0 & x < \theta \end{cases} \quad (4.4)$$

$$\text{Sigmoid activation } f(x) = \frac{1}{1 + e^{-\gamma x}}, \quad (4.5)$$

$$\text{Radial-basis activation } f(x) = \exp -\frac{(x - \mu)^2}{\sigma^2}, \quad (4.6)$$

where $\beta, \theta, \gamma, \sigma, \mu$ are the activation parameters.

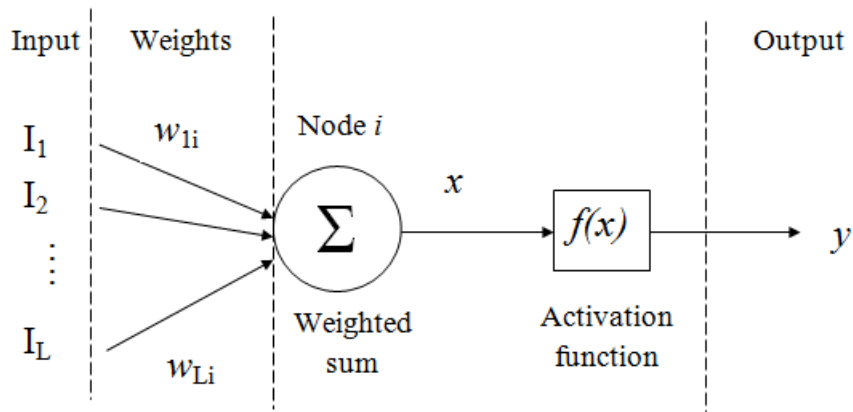
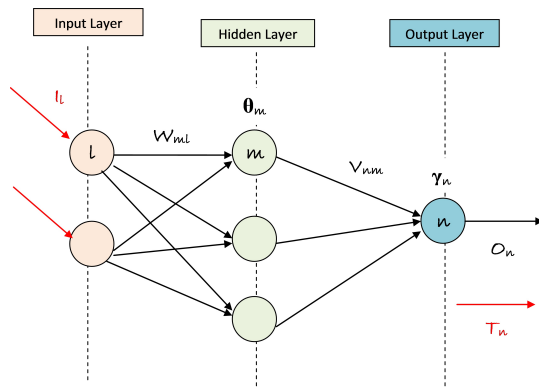


Figure 4.2 Single neuron proposed by McCulloch and Pitts, 1943.

Feed forward neural networks (FFNN) consist of layers classified into input, hidden and output. There is only one input layer and one output layer while the number of hidden layers can be unlimited. Each layer is composed of nodes that activate the nodes in the subsequent layer. Indeed, the activation flows from the input layer to the hidden layers until reaching the output layer. Figure 4.3 shows the structure of a three-layer neural network.

Obviously, each layer has different numbers of neurons. In image classification, the input layer contains a number of neurons that is equal to the number of image pixels or bands. The output layer, in turn, has a number of nodes that is equal to the number of classes. On the other hand,



I_l : The input signal of neuron l .

H_m and θ_m : The output signals and threshold of the hidden neuron m .

O_n, T_n and γ_n : The output signals, target output signals and threshold of the output neuron n .

W_{ml} : The weight between the input neuron l and the hidden neuron m .

V_{nm} : The weight between the hidden neuron m and the output neuron n .

Figure 4.3 Structure of a three-layered neural network.

the number of neurons in the hidden layer is determined through experiments. In this context, many researchers have tried to propose formulas for empirically setting the number of hidden layer nodes. Sheela and Deepa, 2013 established a review of the literature on the methods to fix the number of hidden layer nodes in artificial neural networks. Note that increasing the number of nodes in the hidden layer slows down the training process and can also result in overfitting. This latter occurs when the amount of information contained in the training set is not enough to train all of the neurons in the hidden layer.

Assume L, M, N are the number of nodes in the input layer, hidden layer and output layer respectively, then, one can start the training by fixing the number of hidden layer nodes M using one of these methods:

- $L \leq M \leq N$,
- $M = \frac{2}{3}(L + N)$,
- $M \leq 2L$

As for the initial values of the weights, they should be both unequal and small. From the one hand, the inequality of the initial weights is required to make the network learn adequately. On the other hand, the initial weights should be small so that the weights can be updated proportionally to the derivative of the activation function.

Artificial neural networks are trained by examples before they can be used. In case of supervised learning, the connection weights and the neurons' thresholds are adjusted so as to minimize the error between the output signal and the target signal in the training stage and using the training data set. Then, in the execution or testing stage, the trained networks with the learned parameters will be used to directly compute the output corresponding to the input pattern. Obviously, this latter stage is much faster.

In this work, we investigate two supervised learning algorithms, i.e. the back-propagation algorithm and the extreme learning machine, which aim to minimize the error between the actual network outputs and the desired or target output by changing the values of the weights in the network. Note that thresholds can also be affected to the nodes of the neural network; therefore, learning will also adjust these threshold parameters. For more details about artificial neural networks, we refer to the book by Rojas, 1996.

4.2.1.2 Back-Propagation Neural Networks

The back-propagation (BP) learning algorithm was first introduced by Werbos, 1974 in his PhD thesis, then by Rumelhart et al., 1988. After that BP became extensively used in a wide range of applications.

The back-propagation neural network (BPNN) is a learning multilayer neural network where each connection has a weight assigned to it. All the weights and thresholds are randomly initialized. The inputs are first propagated forward until reaching the output layer, then the network parameters are adjusted so as to minimize the mean squared error between the predicted (output) and the attribute class using gradient descent, for example. This is repeated until the terminating conditions are verified, e.g., a given number of iterations or a given error threshold.

We assume that no thresholds are affected to nodes in the BPNN, then the output O_n of a node n is defined as:

$$O_n = f \left(\sum_{m=1}^M V_{nm} f \left(\sum_{l=1}^L W_{ml} I_l \right) \right), \quad (4.7)$$

where $f(x)$ is an activation function.

Once the output layer is reached, we need to define the total error. This latter is the sum of the errors, i.e., difference between the output signal O_n and the target signal T_n , of each output node. Intuitively, we need to take the square of these differences before summing them up so that the negative and positive errors do not cancel each other out. Finally, we scale this quantity by $1/2$ to obtain the final total error E as follows:

$$E := \frac{1}{2} \sum_{n=1}^N (T_n - O_n)^2, \quad (4.8)$$

where N is the number of output layer nodes.

Figure 4.4 illustrates data classification using a back-propagation neural network (BPNN).

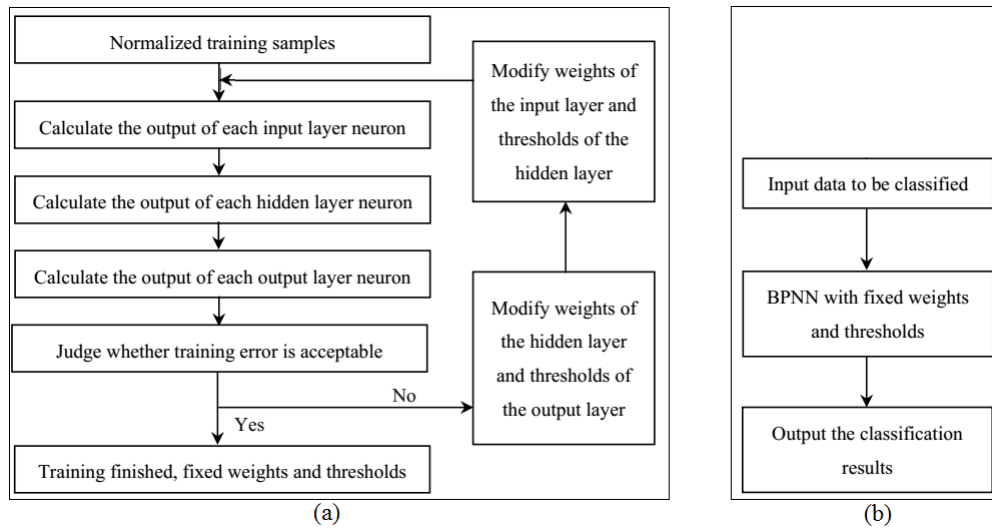


Figure 4.4 Classification using the back-propagation neural networks (BPNN): (a) training process to find the optimal parameters, i.e., weights and thresholds of BPNN; (b) the classification process using the trained BPNN.

Thanks to its high classification performance, the back-propagation algorithm is considered to be one of the most well-known artificial neural networks learning algorithms. However, it still encounters several limitations, recapitulated in the following:

- The BPNN requires a long time to learn and converge.
- As the BPNN is based on a supervised learning algorithm, numerous input-output training examples are required. In addition, the selected training data affects significantly the accuracy of the output. Consequently, a well-selected training data is required for high-quality classification results.
- The stochastic gradient descent, employed by the BPNN, exists on a surface which is not flat (cf. Figure 4.5). Therefore, the BPNN may achieve any local minimum, i.e., an error that is lower than the surrounding but is not the smallest possible one, on the error surface. In addition, the network is likely to get stuck in this local minimum, thus, it is not guaranteed that the network will converge to an acceptable output.

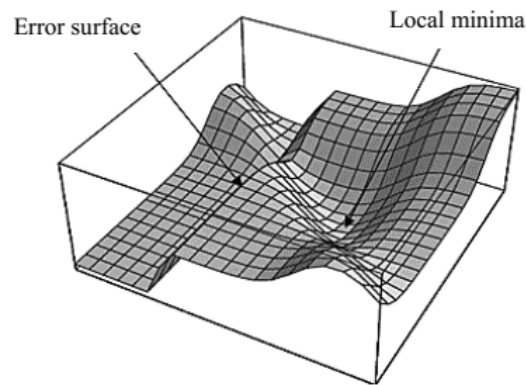


Figure 4.5 Example of local minima of an error function.

Many improvements have been made in the literature to overcome the two main drawbacks of the BPNN, i.e., learning time and local minimum convergence. The first modification consists of introducing the notion of momentum into the weights correction task. The purpose of the momentum is to orient the gradient of the error function to the center of the error surface. This term will help the network weights to pass the local minimum as well as speed up the learning process as well.

Another BPNN improvement consists of using the Levenburg-Marquardt training algorithm (LM algorithm) (Hagan and Menhaj, 1994), which is more stable. In fact, like the aim of gradient descent, Levenburg-Marquardt provides a numerical solution to the problem of minimizing a nonlinear function over its space parameters.

Other improvements in the back-propagation algorithm's efficiency and optimality have been proposed. We cite, for example, the second order optimization methods (Hagan and Menhaj, 1994; Wilamowski and Yu, 2010), subset selection methods (Chen et al., 1991; Li et al., 2005) as well as global optimization methods (Yao, 1993; Branke, 1995).

In the same line, Huang et al., 2006b proposed the extreme learning machine (ELM) for single-hidden layer feed-forward neural networks (SLFNs) which tends to provide good generalization performance at extremely fast learning speed. The ELM will be detailed in the next subsection.

4.2.1.3 Extreme Learning Machine

Traditionally, all the parameters (weights and thresholds) of the feed-forward neural networks need to be tuned for training. Besides, gradient descent-based learning methods are generally very slow due to the many iterative learning steps required to obtain better learning performance and may easily converge to local minimum. All these drawbacks of gradient descent-based learning methods have motivated Huang et al., 2004; Huang et al., 2006b to propose the extreme learning machine (ELM). This latter is a single-hidden layer feed-forward neural network (SLFN) whose learning speed can be thousands of times faster than traditional feed-forward network learning algorithms like the back-propagation algorithm while obtaining better generalization performance.

In fact, the extreme learning machine (ELM) has the same architecture as a three-layered back-propagation neural network except that the weights between the input layer and the hidden layer as well as the thresholds of the hidden neurons are fixed randomly and never changed, whereas the remaining weights are calculated by solving a least squares optimization problem. The extreme learning machine needs only a single feed-forward pass and no gradient operators; therefore, known for its efficiency.

Huang et al., 2004; Huang et al., 2006b drew up the conclusion that their ELM has several interesting advantages compared to the traditional gradient-based learning algorithms, e.g., back-propagation algorithm, for feed-forward neural networks. These advantages can be recapitulated in the following:

- The extreme learning machine learning speed is extremely fast.
- The extreme learning machine has better generalization performance compared to the popular gradient-based learning methods.
- The extreme learning machine tends to reach the solutions straightforward and prevents the issues of the gradient-based learning algorithms including local minimum, inadequate learning rate and overfitting.
- The extreme learning machine learning algorithm could be used to train single-hidden layer feed-forward neural networks with many non-differentiable activation functions (Huang et al., 2006a), unlike the conventional gradient-based learning algorithms, which require differentiable activation functions.

In the work by Chazal et al., 2015, a classification performance comparison between the back-propagation neural networks and the extreme learning machine was conducted using the MNIST (Modified National Institute of Standards and Technology) database processing raw pixel values. Giving the same topology to both back-propagation neural networks and extreme learning machine, the authors found out that a larger number of hidden nodes was required for the ELM in order to give a comparable classification performance as the back-propagation neural networks.

Many variations of the extreme learning machine were introduced, such as the regularized extreme learning machine (RELM) (Deng et al., 2009b) which combines structural risk minimization theory and weighted least squares method and extreme learning machine with kernel (KELM) (Huang and Siew, 2004; Frénay and Verleysen, 2011; Huang et al., 2012; Iosifidis et al., 2015) which introduced kernel function into ELM.

For more information about the extreme learning machine learning algorithm, we refer to the works by Huang, 2015; Huang et al., 2015.

4.2.1.4 Least Squares Support Vector Machines

The least squares support vector machine (LS-SVM), which is a kernel-based learning method, is a least squares version of the well-known support vector machine (SVM) (Cortes and Vapnik, 1995). The LS-SVM was first introduced by Suykens and Vandewalle, 1999 and follows from solving a set of linear equations, instead of quadratic programming (QP) for SVM.

Suykens and Vandewalle, 1999 showed that a least squares SVM with radial basis function (RBF) kernel allows excellent generalization performance for a complicated two-spiral classification problem. This is illustrated in Figure 4.6. In addition, LS-SVM is characterized by a low computational cost and by being free of local minima as it is the solution to a convex optimization problem.

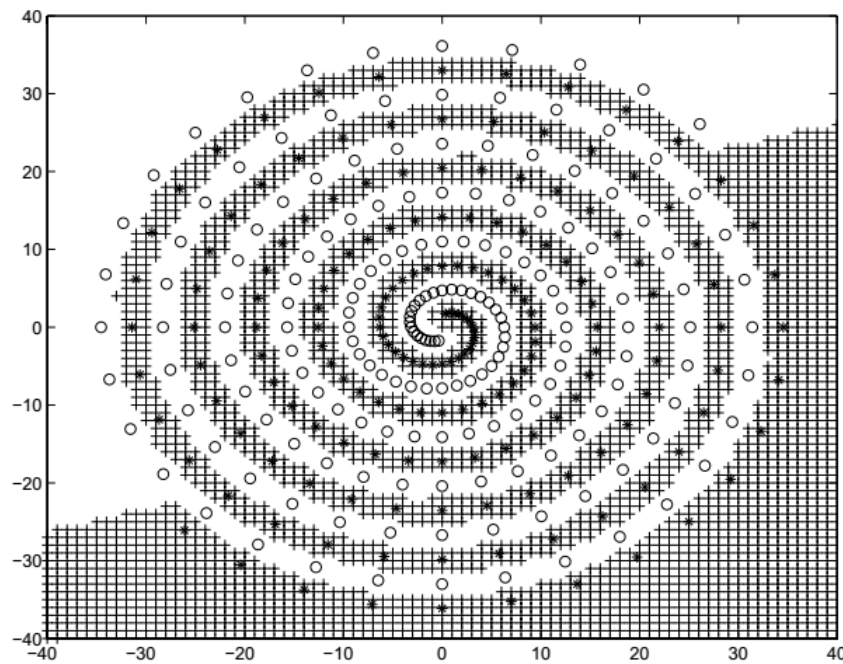


Figure 4.6 The excellent generalization performance of the least squares support vector machine (LS-SVM) with RBF kernel on a two-spiral benchmark problem (Suykens and Vandewalle, 1999).

Gestel et al., 2004 employed twenty UCI (University of California, Irvine) binary and multi-class benchmark datasets (Lichman, 2013) to evaluate the LS-SVM classifier performance using different kernels, e.g., linear, polynomial and radial basis function (RBF) in comparison with different classification methods, i.e., the SVM classifier with linear and RBF kernel, decision tree C4.5 (Quinlan, 1993), Holte's one-rule classifier (Holte, 1993), linear discriminant analysis (LDA) (Bishop, 1995), quadratic discriminant analysis (QDA) (Duda and Hart, 1973), logistic regression (Ripley and Hjort, 1995), instance based learners (Aha et al., 1991) and Naive Bayes (John and Langley, 1995). It has been shown that LS-SVM with RBF kernel and cross-validation tuning of the hyperparameters perform better than the other classifiers. In addition, the LS-SVM sparse approximation procedure can be successfully applied.

4.2.1.5 Deep Learning Algorithms

Deep learning classification methods are multi-layer networks, which have been introduced for both feature extraction and classification. In fact, the key idea behind such methods is to

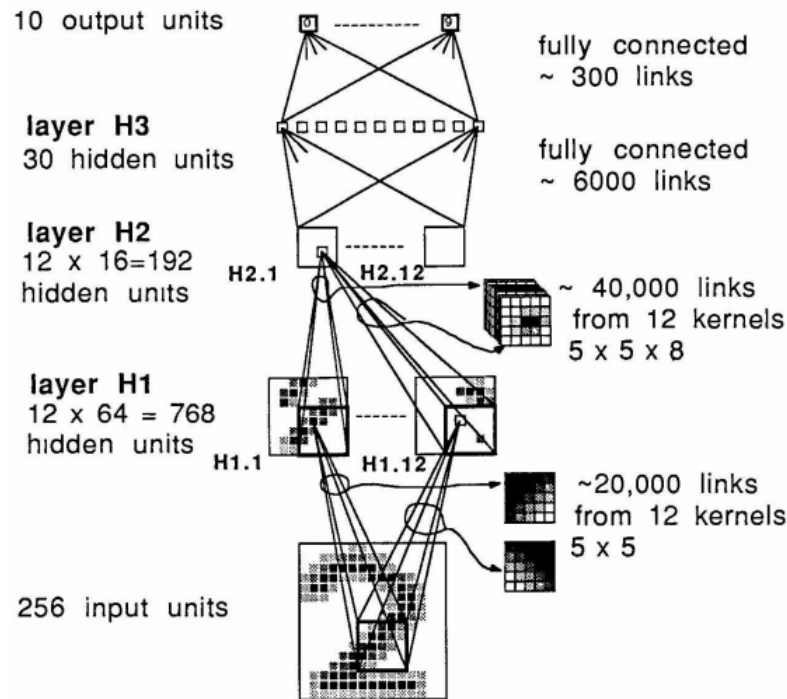


Figure 4.7 Network architecture proposed by LeCun et al., 1989 for recognizing handwritten zip codes.

build better pattern recognition systems that rely more on automatic learning unlike the shallow networks which require hand-crafted feature extraction.

The most well-known deep learning methods are the convolutional neural networks (CNN), which have known a great success since their introduction by LeCun et al., 1989, thanks to three main factors:

1. The availability of low-cost machines with fast arithmetic units such as graphic processing units (GPU) (Steinkraus et al., 2005), which yield a faster training and testing.
2. The availability of large databases, such as databases for handwritten recognition.
3. The availability of high-performing machine learning methods when fed with large datasets.

Unlike shallow networks, convolutional neural networks are directly fed with images, rather than feature vectors. Indeed, (1) the input image is passed through convolution layers to extract highly-dimensional features, then (2) the pooling layers reduce feature dimensionality. Finally (3) this feature representation is fed into a fully connected neural network for classification.

Many gradient-based convolutional neural networks variations have been proposed in the literature for many computer vision tasks. We cite the pioneer work by LeCun et al., 1989 who managed to incorporate the 2D shape invariance by using local connection patterns and imposing constraints on the weights. Figure 4.7 illustrates the three-hidden-layer network proposed by LeCun et al., 1989 for handwritten zip code recognition. Later on, Lecun et al., 1998, Krizhevsky et al., 2012 and Kavukcuoglu et al., 2010 introduced variations of CNNs.

Many other deep learning architectures have been proposed such as deep belief networks (DBN) (Hinton et al., 2006), deep Boltzmann machines (DBM) (Salakhutdinov and Hinton, 2009), deep autoencoder (Hinton and Salakhutdinov, 2006), sparse coding (Olshausen and Field, 1996), local coordinate coding (LCC) (Lin et al., 2010; Yu et al., 2009), principal

component analysis networks (PCANet) (Chan et al., 2015), recurrent convolutional neural networks (RCNN) (Liang and Hu, 2015) and convolutional extreme learning machines with kernel (CKELM) (Ding et al., 2017).

4.2.2 Quaternion-Valued Machine Learning Algorithms

Many quaternary and Clifford versions of well-known classification algorithms (either shallow or deep learning) have been proposed in the literature allowing to deal with multi-dimensional signals. We cite for example the quaternionic Hopfield neural networks and its variations (Isokawa et al., 2006; Minemoto et al., 2016b; Valle and Castro, 2017; Kobayashi, 2017a; Kobayashi, 2017b), the twisted quaternary neural networks (Kobayashi and Nakajima, 2012), the quaternion and Clifford support vector machine classifier (López-González et al., 2014; Bayro-Corrochano and Arana-Daniel, 2010), quaternion softmax classifier (Zeng et al., 2014), Clifford neural networks (Buchholz et al., 2007; Buchholz and Sommer, 2008), quaternion-based deep belief networks (QDBN) (Papa et al., 2017) and quaternion principal component analysis network (QPCANet) (Zeng et al., 2016). Note that all the parameters, input and output of these quaternary classification algorithms are quaternion-valued and all the mathematical operations are governed by the quaternion algebra.

In this work, we describe two state-of-the-art quaternary classifiers that are the quaternion back-propagation neural networks (QBPNN) (Nitta, 1995; Arena et al., 1997) and the quaternion extreme learning machine (QELM) (Minemoto et al., 2016a). Both of them are multilayer neural networks, they have the same structure and perform almost the same feed forward algorithm but use different learning algorithms.

4.2.2.1 Quaternion Back-Propagation Neural Networks

The quaternion back-propagation neural network (QBPNN) is an extended version of the back-propagation neural network (BPNN) where all inputs, outputs, weights and thresholds are quaternary-valued. The pioneer works of the quaternary version of the back-propagation neural network were introduced by Nitta, 1995; Arena et al., 1997. Using a simple example on quaternary patterns, the authors concluded that the QBPNN converges faster than its conventional version BPNN. Furthermore, the quaternion back-propagation neural network was recently applied in color face recognition using the quaternion Zernike moments Shu et al., 2014 and other quaternion-type moments (Chen et al., 2014a) as input.

The quaternion back-propagation neural network consists of:

(1) feed forward propagation: The output O_n of a node n is defined as:

$$O_n = f \left(\sum_{m=1}^M V_{nm} \otimes f \left(\sum_{l=1}^L W_{ml} \otimes I_l + \theta_m \right) + \gamma_n \right), \quad (4.9)$$

where $f(x)$ is an activation function, the sigmoid function in this work, defined as:

$$f(x) = \frac{1}{1 + \exp(-x)} \quad (4.10)$$

As x is a quaternion, one can compute $f(x)$ simply by splitting the activation function as follows:

$$f(x) = f(x^a) + f(x^b)i + f(x^c)j + f(x^d)k, \quad (4.11)$$

(2) Back-propagation: This step allows to learn the new quaternion back-propagation neural network parameters (thresholds and weights) that help minimizing the square error between the obtained output O_n and the wanted output T_n . We note that a pattern can have only one positive target $1 + i + j + k$ corresponding to the class to which it belongs, and the others are all 0.

For each pattern, one can compute the square error E between the output signals of the N output neurons and the corresponding targets as follows:

$$E = \frac{1}{2} \sum_{n=1}^N |T_n - O_n|^2 \quad (4.12)$$

For a sufficiently small constant learning rate α , and using a steepest descent method, one can compute the aforementioned parameters variations as follows:

$$\Delta\gamma_n = \alpha(T_n - O_n) \bullet (1 - O_n) \bullet O_n, \quad (4.13)$$

$$\Delta V_{nm} = H_m^* \otimes \Delta\gamma_n, \quad (4.14)$$

$$\Delta\theta_m = (1 - H_m) \bullet H_m \bullet \left[\sum_n (T_n - O_n) \otimes V_{n,m}^* \right], \quad (4.15)$$

$$\Delta W_{ml} = I_l^* \otimes \Delta\theta_m. \quad (4.16)$$

Then, the old parameters *old* are substituted with new ones *new* according to the following:

$$new = old + \Delta old. \quad (4.17)$$

As far as testing is concerned, the predicted class of a pattern is computed by retaining the minimum quaternion vector distance between the output signal $O_n, 1 \leq n \leq N$ and all the existing classes.

Although the quaternion back-propagation neural network may lead to high classification accuracies, it suffers, just like its conventional version BPNN, from its dependance on several parameters, i.e., the learning rate, the number of iterations and the number of hidden layer nodes, as well as from being time consuming.

4.2.2.2 Quaternion Extreme Learning Machine

Unlike the quaternion back-propagation neural network (QBPNN), the quaternion extreme learning machine (QELM) (Minemoto et al., 2016a), an extension of the three layered network with random weights (ELM), depends only on the number of hidden layer nodes and needs only one feed-forward shot, thus known for its rapidity.

Minemoto et al., 2016a tested the performances of their proposed quaternion extreme learning machine for color image classification and reconstruction in the CIFAR-10 dataset (“Canadian Institute for Advanced Research”). They found out that the QELM performs better than the ELM in both applications. As for the color image classification task, the input for the network was the set of pixel values in the image without passing through the feature extraction stage.

This network consists as well of a feed forward propagation where the input signal O_n of a node n is given by:

$$O_n = \sum_{m=1}^M f \left(\sum_{l=1}^L I_l \otimes W_{ml} + \theta_m \right) \otimes V_{nm}, \quad (4.18)$$

where $f(x)$ is the sigmoid function, defined in Eq. (4.10). W_{ml} and θ_m are set at random values and never changed, whereas V_{nm} are learned by solving the following least squares norm minimization problem:

$$\min_B \|HB - T\|^2 \quad (4.19)$$

where $\|\cdot\|$ is the Frobenius norm, $H \in \mathbb{H}^{L \times M}$ is a matrix of elements $f(I_l \otimes W_{ml} + \theta_m)$, $B \in \mathbb{H}^{M \times N}$ is the transpose of matrix of elements V_{nm} and $T \in \mathbb{H}^{L \times N}$ is transpose of matrix of elements T_n . A pattern can have only one positive target $1 + 0i + 0j + 0k$ corresponding to the class to which it belongs, and the others are all 0.

The solution for Eq. (4.19) is:

$$\tilde{B} = H^+ T \quad (4.20)$$

where H^+ denotes the Moore-Penrose pseudo inverse matrix of the quaternion matrix H . This solution can be rewritten as:

$$\widetilde{adj(B)} = \begin{cases} H_c^h (H_c H_c^h)^{-1} T_c, & (L \leq M) \\ (H_c^h H_c)^{-1} H_c^h T_c, & (L > M) \end{cases} \quad (4.21)$$

where Q_c is the complex adjoint matrix of a quaternion matrix Q . In fact, $Q \in \mathbb{H}^{L \times M}$ can be represented, using Cayley-Dickson notation, as:

$$Q = (Q^a + Q^b i) + (Q^c + Q^d i)j = Q_A + H_Q j, \quad (4.22)$$

where $Q_A, Q_B \in \mathbb{C}^{L \times M}$ are complex matrices. This notation is needed for the computation of the complex adjoint matrix $Q_c \in \mathbb{C}^{2L \times 2M}$ as follows:

$$Q_c = \begin{pmatrix} Q_A & Q_B \\ -Q_B^h & Q_A^h \end{pmatrix} \quad (4.23)$$

where $(\cdot)^h$ is the Hermitian transpose.

For testing, the maximum of real parts of $O_n, 1 \leq n \leq N$ is retained to decide the class to which the pattern belongs.

4.3 Image Classification Method based on the Quaternary Approach

From the review on the existing classification methods, presented in the previous section, one can recapitulate the differences between deep learning methods and shallow methods in the following:

- Deep-learning methods necessitate a huge amount of data to unsupervisedly extract the image features.
- Deep learning methods are constrained by the availability of computing resources since processing higher resolution images requires larger and more layers.
- Shallow methods are based on extracting vectors via mathematical transformations. This means that the rotation, scale and translation invariance in the image feature extraction stage can be mathematically obtained by the image descriptors.

In this work, we propose a shallow color image classification method, which takes advantage of our hand-crafted color image quaternary descriptors. But first, we show how we can adapt shallow learning-machine algorithms to take as input quaternary features. Recall that the quaternary color image features allow maintaining the correlation between the three color channels thanks to the cross quaternary product.

For color image classification, when extracting the color image features, one can use the quaternary moments modulus-based invariants (real-valued invariants) or quaternion-valued invariants which amounts to using either real or quaternary machine-learning algorithms. Therefore, three situations, which are illustrated in Figure 4.8, arise:

- Situation (1) consists of feeding the modulus-based quaternary moment invariants (real-valued) into a real-valued learning algorithm. The drawback of this situation is that obtaining rotation invariance by taking the modulus of the quaternion moments leads to the loss of the phase information and only provides one real-valued invariant.
- Situation (2), also known as the marginal approach, consists of feeding the four parts of the quaternion-valued moment invariants separately into a real-valued machine-learning algorithm. The drawback of this situation is that the classifiers' input is four times larger, which slows down the training time of the classifier. In addition, the correlation between the four parts of the quaternary feature vector is lost. Figure 4.9 illustrates the situation (2).
- Situation (3) consists of feeding the quaternion-valued moment invariants as a whole into a quaternary classifier. This situation is advantageous in the way that phase information of the quaternary feature vector as well as the correlation between their four parts are preserved. Recall that the operations for updating the connections' weights and neurones' states in the quaternary neural networks are governed by the quaternion algebra. Situation (3) is illustrated in Figure 4.10.

In this work, we adopt the situation (3) for color image classification. The basic idea is to keep the quaternary form of the coefficients composing the proposed parameter-free quaternion disc-harmonic moment (PFQDHMs) descriptor instead of being restricted to the moment's modulus. This is established through the following formula:

$$PFQDHMI_{n,m}(f) = PFQDHM_{n,m}(f) \otimes (PFQDHM_{l,m}(f))^*, \quad (4.24)$$

$$|m| \leq n, |m| \leq l,$$

which has been mathematically proven by Chen et al., 2014a. In this work, $l = |m|$ is chosen. In fact, obtaining rotation invariance by taking the modulus of the quaternary moments leads to the loss of the phase information and only provides one real-valued invariant.

To maintain the correlation between the four parts, real and imaginary, of the quaternary descriptor coefficients, we take advantage of a quaternary classifier that is the quaternary extreme learning machine (QELM).

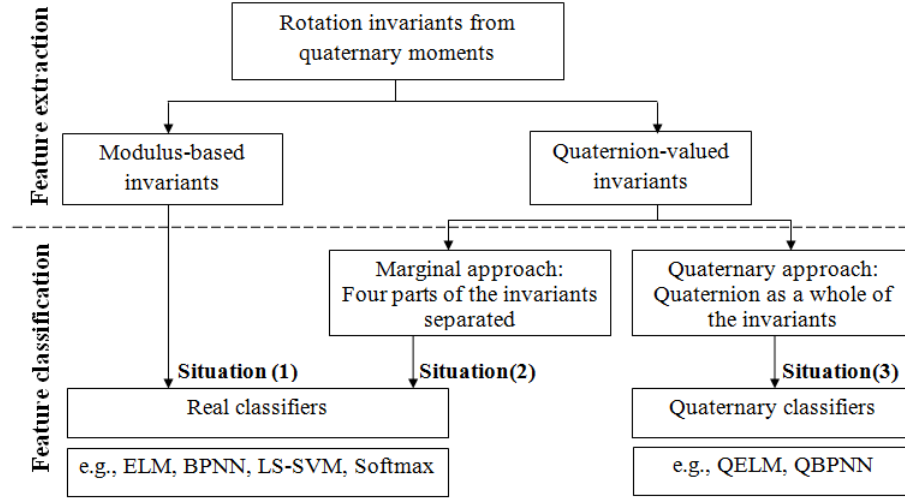


Figure 4.8 The different rotation invariants from quaternary moments and the different classification approaches to deal with them.

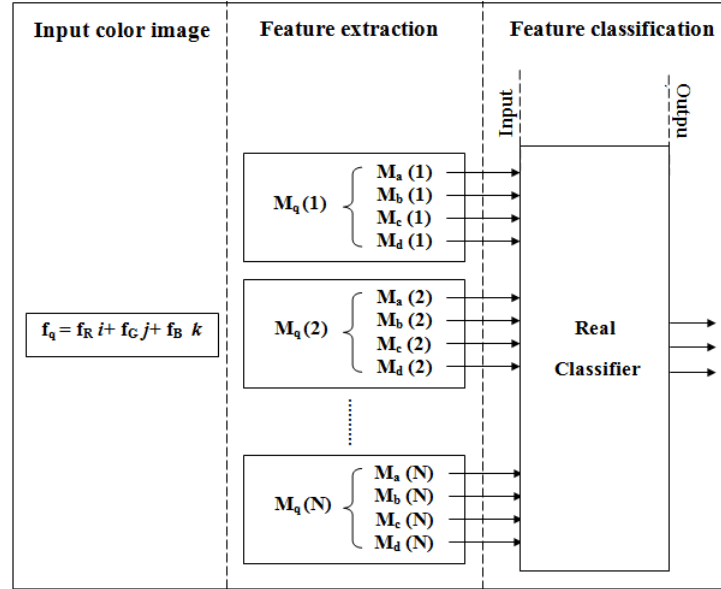


Figure 4.9 Marginal approach for color image classification, which consists of feeding a real-valued classifier with the four parts of the quaternion-valued moment invariants separately (the classifiers' input is four times larger).

As shown in Figure 4.11, our color image classification method is divided into two stages: Training and testing.

For the training stage, we proceed as follows:

1. We represent each color image f from the training set by means of a pure quaternion number as $f = f_R i + f_G j + f_B k$, where f_R , f_G , f_B are the red, green and blue components of the image.
2. We compute the quaternion-valued parameter-free disc-harmonic moment invariants (PFQDHMs), from Eq. (4.24), for each of the quaternary color images obtained in step (a). Recall that the PFQDHMs are detailed in our second contribution (Chapter 3).

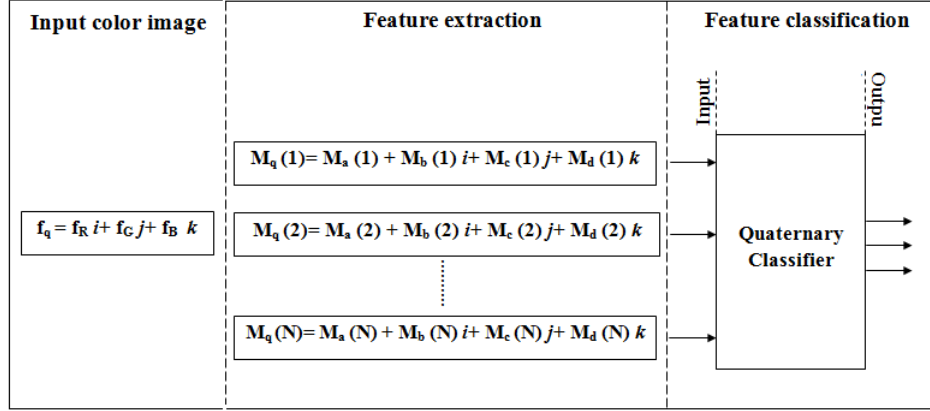


Figure 4.10 Quaternary approach for color image classification, which consists of using a quaternary classifier fed with the quaternion-valued moment invariants as a whole.

3. We feed the quaternion-valued feature vectors into the quaternion extreme learning machine (QELM), detailed in subsection 4.2.2.2, in order to learn its parameters.

The testing stage consists in the following steps:

1. We represent each color image f from the testing set by means of a pure quaternion number as $f = f_R i + f_G j + f_B k$, where f_R, f_G, f_B are the red, green and blue components of the image.
2. We compute the quaternion-valued parameter-free disc-harmonic moment invariants (PFQDHMs), from Eq. (4.24), for each of the quaternary color images obtained in step (a) of this stage.
3. We feed quaternion-valued feature vectors issued from the previous step into the quaternion extreme learning machine (QELM) for classification using the parameters learned in the training stage.

Our method aims to (1) investigate the quaternion-valued invariants of the proposed moments instead of the modulus-based invariants in the application of color image classification and (2) take advantage of the quaternion extreme learning machine (QELM) (Minemoto et al., 2016a), which is known for its high classification accuracies as well as its fast training.

4.4 Experimental Results

In the following experiments, we evaluate the color image classification performance of the proposed method compared with methods derived from both the situation (1) and the situation (3) of Figure 4.8.

For feature extraction, we narrow the experimental evaluation to our PFQDHMs, the quaternion disc-harmonic moments (QDHMs) and the well-known quaternion Zernike moments (QZMs) (Chen et al., 2010; Chen et al., 2012). In fact, it has been previously shown, in chapter 2, that our QDHMs perform better than its traditional version disc-harmonic moments (DHMs) for gray-scale images, some competing quaternary moment descriptors, i.e., the quaternion Bessel-Fourier moments (QBFMs) (Shao et al., 2014), the QZMs, the quaternion Pseudo-Zernike moments (QPZMs) (Bei-Jing et al., 2012) and the quaternion orthogonal Fourier-Mellin moments (QOFMMs), as well as Clifford-based moments (Mennesson et al., 2014) in terms of color image reconstruction, robustness, invariance as well as image retrieval.

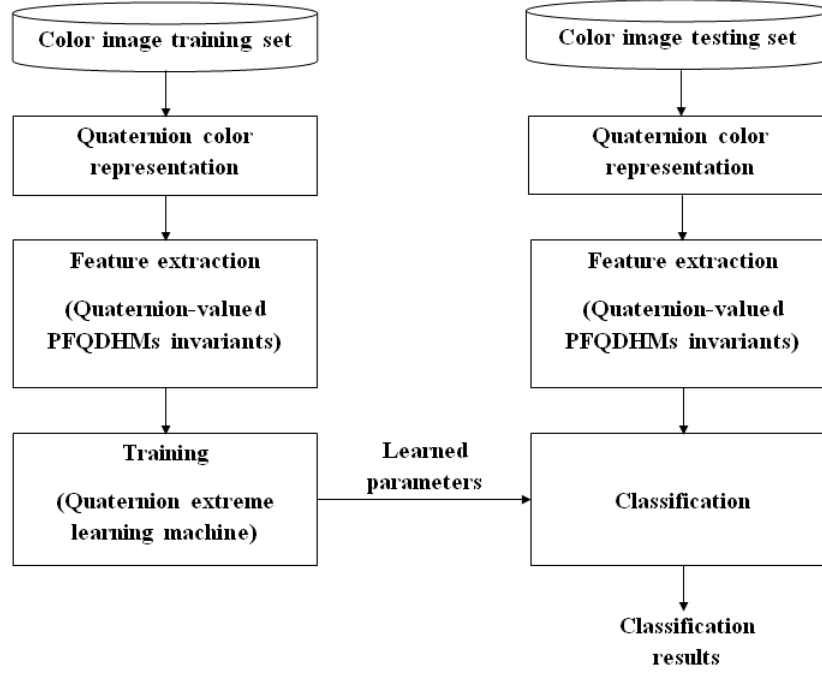


Figure 4.11 Flowchart of the proposed method for color image classification.

As for feature classification, the two rotation-invariants of these quaternary moments are fed into many classifiers, either quaternion-valued, e.g., QBPNN and QELM, or real-valued ones, e.g., ELM, BPNN, Softmax, LS-SVM. The literature review of these classifiers, previously given in subsections 4.2.1 and 4.2.2, proves that they are still very used in many fields thanks to their high classification accuracies.

4.4.1 Classification using Faces95 Collection

This experiment aims at comparing the classification accuracies of our method against feeding the PFQDHMs, QDHMs and QZMs into several classifiers, such as, QBPNN, QELM, BPNN, ELM, LS-SVM and Softmax, as in the situations (1) and (3) of Figure 4.8. The 72 individuals from Faces95 collection are used in this experiment. Recall that Faces95 contains a total of $72 \times 20 = 1440$ images as described in subsection 3.3.1. It is worth noting that:

- The images are re-sized to 128×128 before computing their quaternary descriptors.
- The maximum order $nMax = 5$ was retained for the proposed PFQDHMs, the QDHMs and QZMs, which corresponds to 36, 36 and 21 coefficients, respectively.
- For the QBPNN and the BPNN, we set the learning rate to $\alpha = 0.8$ and the number of iterations to 300. The number of hidden layer nodes is set to $M = 2(L + N)/3$ as reported in Table 4.1.
- The number of hidden layer nodes M for the QELM and the ELM is set empirically for each of the moments PFQDHMs, QDHMs and QZMs. M was varied between 30 and 540 with a step of 30. The values of M that give the best classification accuracies, which are reported in Figure 4.16 (a) and Table 4.1, are retained for the rest of this experiment. It is worth noting that the classification accuracies of the QELM and ELM, presented hereafter, are average values in 10 trials.
- The Softmax classifier is given 100,000 learning epochs as termination condition.
- The Gaussian RBF kernel is set for the LS-SVM. This latter is tuned to find out the best

tuning parameters: the regularization parameter and the squared bandwidth.

Table 4.1 Number of hidden layer nodes M for each method processing faces95 database with no corrupted images. For QBPNN and BPNN, M is set to $2(L + N)/3$, whilst for QELM and ELM, M is set empirically according to Figure 4.16 (a).

	QBPNN	BPNN	QELM	ELM
PFQDHMs	72	72	330	270
QDHMs	72	72	330	180
QZMs	62	62	330	210

Figure 4.12 illustrates the color image classification results facing the proposed method against feeding the quaternion-valued invariants from the quaternary moments (QDHMs-qua and QZMs-qua) into the quaternary classifiers (QELM and QBPNN), processing Faces95 collection.

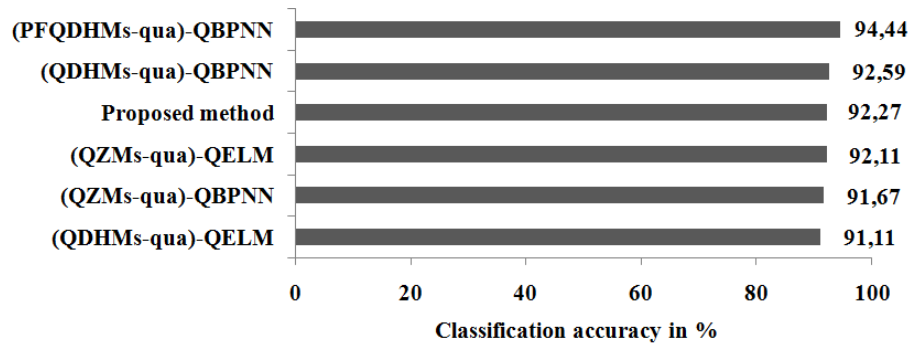


Figure 4.12 Classification accuracy in % of the proposed method and other quaternary methods using the quaternion-valued invariants of the quaternary orthogonal moments with quaternary classifiers, processing the noise-free Faces95 collection.

One can see from these results that: (1) The quaternion-valued PFQDHMs, QDHMs and QZMs invariants, using either the QELM or the QBPNN, yield close classification accuracies ($\geq 90\%$). (2) The QELM needs a number of hidden layer nodes greater than the QBPNN in order to achieve their best performances as can be seen from Table 4.1. The QELM algorithm is, however, advantageous in the way its training stage is much faster than the QBPNN as will be demonstrated in subsection 4.4.5.

On the other hand, Figure 4.13 shows the classification results using the modulus of the quaternary moments (PFQDHMs-mod, QDHMs-mod and QZMs-mod) fed into the real-valued classifiers, processing the Faces95 collection.

We can see that: (1) The modulus of the PFQDHMs, QDHMs and QZMs yield comparable classification accuracies depending on the classifier. (2) Generally, employing the full form of the quaternary moments (QMs) invariants (QMs-qua) leads to better classification results ($91.11\% \leq QMs - qua \leq 94.44\%$) compared to the use of the modulus (QMs-mod) ($78.7\% \leq QMs - mod \leq 88.19\%$). (3) The BPNN yields almost similar results as the ELM. This confirms

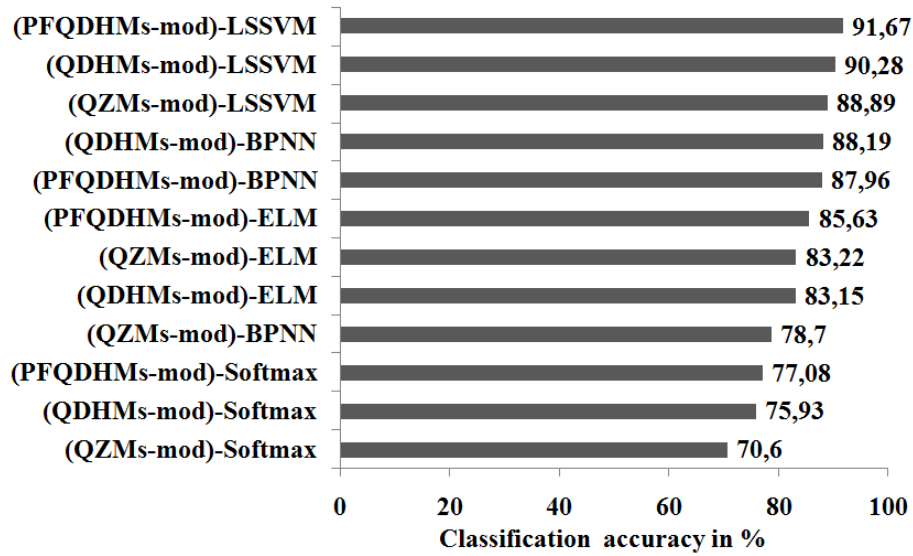


Figure 4.13 Classification accuracy in % for different methods using the modulus of the quaternary orthogonal moments with real-valued classifiers, processing the noise-free Faces95 collection.

the results drawn up by Chazal et al., 2015, processing the CIFAR database, who found out that the ELM requires a larger number of hidden layer nodes in order to show similar results as the BPNN. (4) The LS-SVM outperforms the other real-valued classifiers since it is tuned to find out the best parameters for each type of moments. (5) It seems that the Softmax classifier needs a greater number of iterations ($\geq 100,000$) in order to give results close to those obtained by the other real-valued classifiers.

4.4.2 Classification using Corrupted Faces95 Collection

This experiment aims at evaluating the performance of the proposed method in terms of robustness to both blur and noise. For this purpose, we consider the Gaussian blur (GB) as well as salt and pepper noise (SP). The first 10 images of each individual from faces95 collection were added Gaussian blur of kernel from 1 up to 19 with step 2, and the remaining 10 images were added intensities of salt and pepper noise from 1% up to 10%. Figure 4.14 shows some images of the faces95 collection after noise adding.

Note that for each individual, we keep the 20 original images and we add the 20 corrupted images, which results in a total of 40 images per individual. Therefore, the corrupted Faces95 contains a total of $72 \times 40 = 2880$ images. We consider randomly 27 images from each individual for training and the rest 13 for testing, which corresponds to $72 \times 27 = 1944$ images in the training set and $72 \times 13 = 936$ images in the testing set.

We report in Table 4.2 the optimal number of hidden layer nodes for each method.

The results concerning the corrupted Face95 collection are illustrated in Figure 4.15 and Figure 4.17. The classification accuracies in this experiment, dealing with corrupted images, are higher than those obtained when dealing with noise-free images (Figure 4.12 and Figure 4.13). This can be explained by the fact that: (1) 40 images per class are used instead of just 20 images, which gives the classifiers the opportunity to learn better. (2) Low-order orthogonal moments are proved to be highly insensitive to noise and blur as shown in subsection 3.3.2.

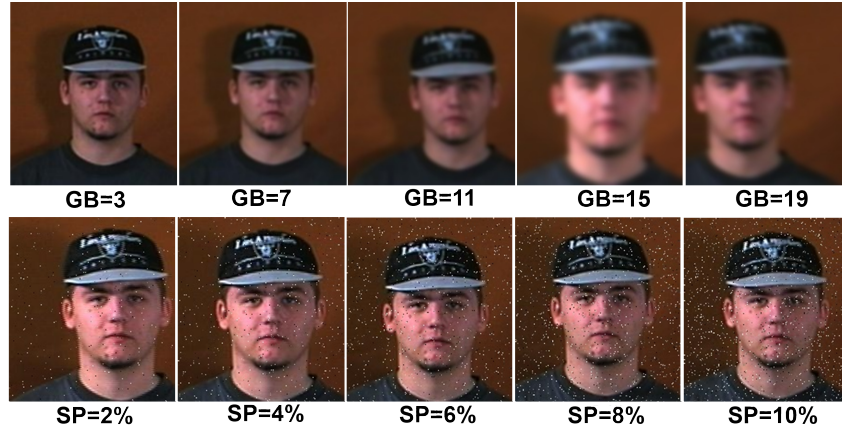


Figure 4.14 Ten images for one individual from faces95 after adding Gaussian blur and salt and pepper noise.

Table 4.2 Number of hidden layer nodes M for each method processing faces95 database with corrupted images. For QBPNN and BPNN, M is set to $2(L+N)/3$, whilst for QELM and ELM, M is concluded from Figure 4.16 (b).

	QBPNN	BPNN	QELM	ELM
PFQDHMs	72	72	480	360
QDHMs	72	72	480	180
QZMs	62	62	360	330

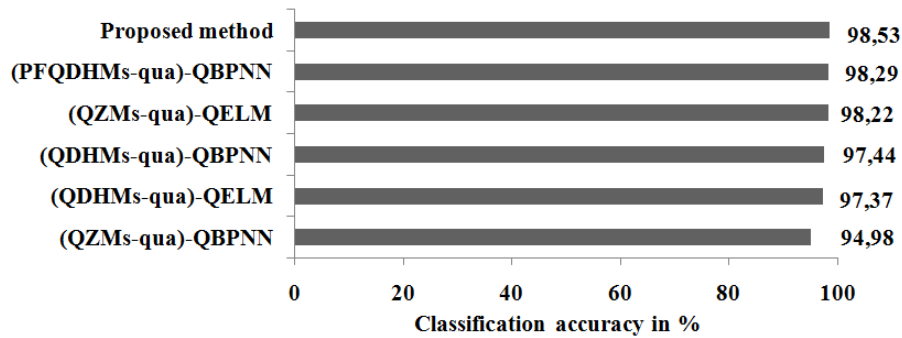


Figure 4.15 Classification accuracy in % of the proposed method and other quaternary methods using the quaternion-valued invariants of the quaternary orthogonal moments with quaternary classifiers, processing corrupted Faces95 collection.

4.4.3 Classification using ALOI-COL Collection

This experiment aims at comparing the classification performances of the quaternary moments fed into the QELM, ELM, QBPNN, BPNN classifiers processing the ALOI-COL collection, whose structure is explained above in subsection 2.4.1. Recall that ALOI-COL consists of classes varied in configurations by changing the illumination color temperature. The results are depicted in Table 4.3. Note that the QELM and ELM were tuned to find their optimal number of hidden layer nodes, and that their results are averaged from 10 repeats.

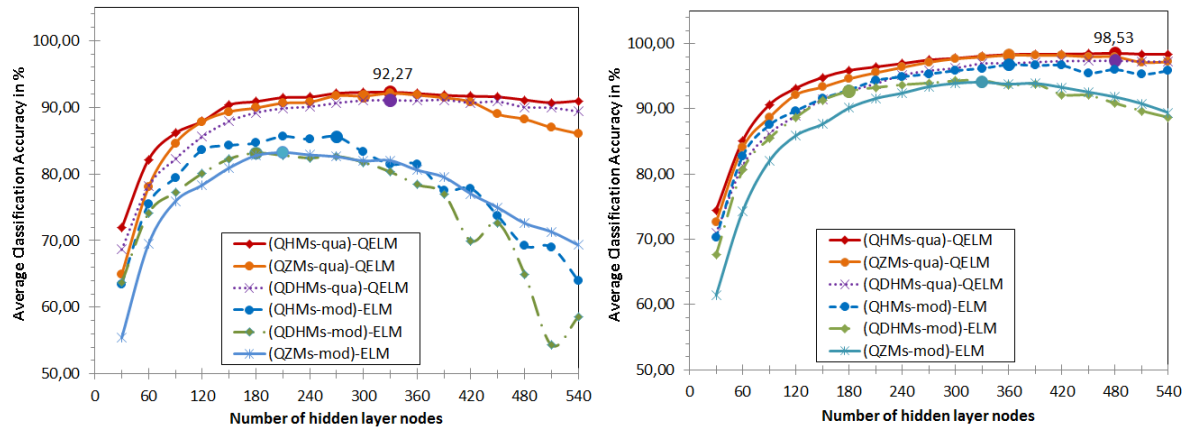


Figure 4.16 Classification accuracies of the QELM and ELM using different numbers of hidden layer nodes. (a) processing the Faces95 collection without corrupted images. (b) Processing Faces95 with corrupted images.

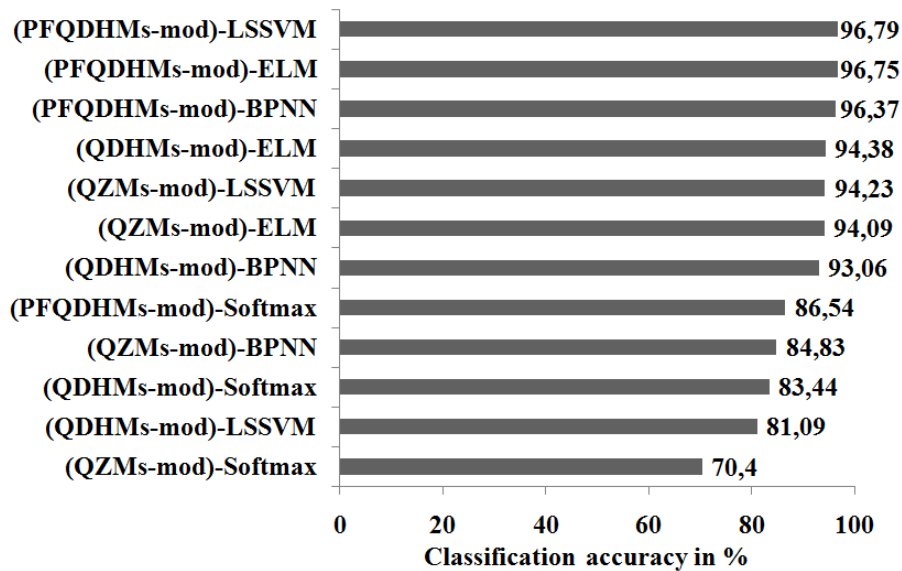


Figure 4.17 Classification accuracy in % for different methods using the modulus of the quaternary orthogonal moments with real-valued classifiers, processing corrupted Faces95 collection.

Table 4.3 Comparison between the methods based on QELM, ELM, QBPNN and BPNN, including our method, in terms of classification accuracy (in %) processing ALOI-COL.

	QBPNN	BPNN	QELM	ELM
PFQDHMs	96.25	88.5	97.7	81.85
QDHMs	95	86.75	96.27	81.87
QZMs	89.75	85	95.7	80.05

Unlike the comparable results remarked for the Face95 collection, processing ALOI-COL shows a significant difference between the quaternion-valued moment invariants and

the modulus-based moment invariants in terms of classification accuracy. For example, (PFQDHMs-qua)-QELM and (PFQDHMs-mod)-ELM show a difference of 15.85%. In fact, achieving rotation invariance with the modulus results in a loss of the phase information, which is unsuitable when dealing with images under illumination color change. Besides, our method outperforms the others with a classification accuracy of 97.7%; therefore, the PFQDHMs are the less influenced by the illumination color change.

4.4.4 Classification using ALOI-ILL Collection

In this experiment we compare the classification performances of the quaternary moments fed into the QELM, ELM, QBPNN, BPNN classifiers processing the ALOI-ILL collection. This latter, as explained in subsection 2.4.1, is partitioned into classes with configurations taken under different illumination angles. The results are depicted in Table 4.4.

Table 4.4 Comparison between the methods based on QELM, ELM, QBPNN and BPNN, including our method, in terms of classification accuracy (in %) processing ALOI-ILL.

	QBPNN	BPNN	QELM	ELM
PFQDHMs	81.5	47.25	88.9	62
QDHMs	75	44.37	83.2	60.31
QZMs	54	46.37	82.35	64.41

The results, processing ALOI-ILL, also show a significant difference between the quaternion-valued moment invariants and the modulus-based moment invariants in terms of classification accuracy. Besides, our method outperforms the others with a classification accuracy of 88.9% with ALOI-ILL, which is much smaller than the accuracy obtained with ALOI-COL (97.7%). This large difference can be explained by the fact that, in some cases, parts of the object become invisible for the illumination direction change (nonlinear transformation), unlike the illumination color change, which does not affect the shape of the object (linear transformation). It is also seen that the classification accuracies of the quaternion-valued PFQDHMs and QDHMs invariants go far beyond the QZMs when utilizing the QBPNN as classifier.

4.4.5 Training Time

In this experiment, we compare the training time by the proposed method against the other methods based on QELM, ELM, QBPNN and BPNN processing the corrupted Faces95 database. The experiments were carried out on a personal computer i3-2310M 2.10 GHz, 4 Go RAM. Note that we used the same numbers of hidden layer nodes as given in Table 4.2.

From the results shown in Table 4.5, one can see that training the QELM and ELM is many hundred times faster than the QBPNN and BPNN. This is explained by the fact that they need only one feed-forward pass to learn, unlike the QBPNN and BPNN, which learn through iterations (300 iterations in this experiment). Furthermore, the slight difference between the results of the QELM and QBPNN, remarked in the previous experiments, is compensated by the speed of QELM training.

Table 4.5 Comparison between the methods based on QELM, ELM, QBPNN and BPNN, including our method, in terms of training time.

	QBPNN	BPNN	QELM	ELM
PFQDHMs	5169.81 s	1784.98 s	17.70 s	2.59 s
QDHMs	5190.92 s	1787.20 s	18.54 s	1.88 s
QZMs	2587.06 s	863.89 s	15.37 s	3.30 s

4.5 Conclusion

In this chapter, we showed how we adapted some machine-learning algorithms to deal with quaternary color features for image classification. In fact, the need for handling multi-dimensional signals, such as color images, was the motivation behind introducing the algebra of quaternions and generally the algebra of Clifford into these conventional machine-learning algorithms.

We designed a color image classification method that consists of feature extraction using our previously proposed parameter-free quaternion disc-harmonic moments (PFQDHMs) along with the quaternion extreme learning machine. In fact, the PFQDHMs rotation invariance was, this time, obtained with a quaternary product (cross product) that allows us to preserve the quaternary form of the invariants (quaternion-valued). On the other hand, we take advantage of the quaternion extreme learning machine, which is known for its high classification accuracies as well as its fast training.

In the experiments section, we compared our color image classification method with methods using different real and quaternary classifiers (ELM, BPNN, LS-SVM, Softmax, QELM and QBPNN) fed with the modulus-based and quaternion-valued invariants of different quaternary moments (QDHMs and QZMs). The results demonstrated that the proposed PFQDHMs give classification accuracies very close to the QDHMs, and both PFQDHMs and QDHMs go far beyond the QZMs especially when dealing with color illumination change. Furthermore, the quaternion-valued moment invariants yield better classification accuracies than the modulus-based invariants. It was also drawn up that although the QELM needs a greater number of hidden layer nodes than the QBPNN to achieve similar performances, one can opt for the QELM owe to its rapidity.

Conclusion and Perspectives

In this PhD thesis work, we focused on investigating and improving color shape description for image retrieval and classification purposes. This improvement concerns several aspects that are image reconstruction capability, robustness and invariance, computational time as well as discriminative power.

Based on the literature review established in Chapter 1, we concluded that among the region-based shape descriptors, the orthogonal moments have some very important properties yielding to preserving the maximum amount of information on the shape. Therefore, they have been widely employed in many applications including content-based image retrieval (CBIR). We also concluded that employing the quaternion algebra for embedding color information into these conventional descriptors is beneficial in the way the correlation between the three channels is maintained. This technique proved its performances in many applications dealing with color images such as, watermarking, registration, enhancement, medical image fusion, retrieval and recognition, etc.

In Chapter 2 we have described our first contribution in which we propose an adaptation of the spherical harmonic transform (SHT) for color image description by using both the algebra of quaternions and an appropriate sphere-to-disc transformation. We first provided a detailed overview about the SHT including their properties and applications dealing with both 3D and 2D objects. In fact, many works have been investigating the employment of the 3D SHT for 2D image retrieval. Among them, there is the disc-harmonic moments (DHMs) that have been previously introduced by the same research team. The DHMs are based on orthonormal and complete basis functions and were established by performing a sphere-to-disc transformation. The initial idea behind our first contribution was to generalize the DHMs to color images using quaternions. In addition, we further improved the computational time of our descriptor, namely quaternion dischismic moments (QDHMs), by employing the Healpix spherical sampling technique. At the end, we have conducted exhaustive empirical tests to evaluate our descriptor performances in terms of image reconstruction capability, robustness to photometric transformations, invariance under geometric transformations, computational time as well as discriminative power. The proposed moments have shown significant improvement in image reconstruction compared to the existing orthogonal moments extended by quaternions (the QZMs, QPZMs, QBFMs, and QOFMMs). Our moments are also the most robust to SP noise and Gaussian blur compared to quaternion-based moments and show, generally, satisfactory results concerning robustness to illumination color change, illumination direction change, and invariance under rotation. As for CBIR, we have studied the influence of the maximum order, color space and similarity measure. Our quaternary moments have better discriminative power compared to their conventional version (the DHMs) since they take into account the color information. In addition, they slightly outperform most of the existing quaternary and Clifford descriptors. This slight difference in retrieval results can be compensated by the fast computation of our moments, especially when performing the Healpix technique which allows a 12-time faster computation. Thus, our moments could be useful as a feature vector in color image retrieval systems.

Chapter 3 was devoted to our second contribution, which consists of introducing a parameter-free reformulation of the quaternary moments. We were motivated by the fact that all the existing quaternary moments depend in their computation on a unit pure quaternion parameter μ whose optimal value can be fixed only with the help of experiments. Furthermore, the estimated value of this parameter is strongly dependent on the application domain. Our parameter-free quaternary orthogonal moments (PFQOMs) were defined by substituting the μ -dependent angular part by a parameter-free angular part. The experimental results showed that the proposed PFQOMs have a significant color image reconstruction capability. Besides, their modulus-based invariants are robust to various photometric transformations as well as invariant under geometric transformations. The discriminative performances of our descriptors were compared to those of the state-of-the-art μ -based quaternary moments, the chromaticity moments, the Legendre chromaticity distribution moments as well as the Clifford-based parallel-orthogonal Fourier Mellin. Both the recall-precision and ROC curves have shown that the discriminative powers of the proposed PFQOMs are satisfactory and comparable to their μ -based versions dealing with both COIL-100 and Faces96 collections. As for color image classification, we have used the Extreme Learning Machine classifier (ELM) as it is known for its efficiency. The obtained results showed the performance of the proposed descriptors.

Chapter 4, in turn, detailed our third contribution dedicated to our color image classification method based on the parameter-free QDHMs (PFQDHMs). The objective was to confirm the relevance of our quaternary descriptor and validate it as part of an image classification system. We first presented a literature review regarding the real and quaternary classification algorithms. In this contribution, we derived our PFQDHMs invariants so as to keep the quaternary form of its coefficients instead of the real form obtained by the modulus. This also preserves the phase of the quaternary moments. In order to maintain the correlation between the four parts, real and imaginary, of our quaternary descriptor coefficients, we took advantage of a quaternary classifier that is the quaternary extreme learning machine (QELM). We have compared the performances of our method to several techniques combining real or quaternary classifiers (ELM, BPNN, LS-SVM, Softmax, QBPNN and QELM) with the quaternary orthogonal moments (QDHMs and QZMs) in either their modulus-based invariants (real valued) or quaternion-valued invariants. Then obtained results showed that: (1) Our method yields satisfactory classification accuracies. (2) The quaternion-valued moment invariants yield better classification accuracies than the modulus-based invariants. (3) Although the QELM needs a greater number of hidden layer nodes than the QBPNN to achieve similar performances, one can opt for the QELM owe to its rapidity.

Generally speaking, we have demonstrated that our color region-based shape descriptor (QDHMs) has several properties including image reconstruction capability, invariance to photometric and geometric transformations, an easy and fast computation as well as a satisfactory discriminative power. We have also highlighted the advantages of computing the quaternary moments using our parameter-free reformulation. In addition, our color image classification method validates and confirms the relevance of our quaternary descriptor. The promising results drawn up in this thesis work arise several perspectives to be explored:

- We intend to make a comparative analysis between the global and local color image quaternary descriptors. In fact, local descriptors are more robust than the global ones when it comes to images containing more complex changes, e.g., large pose or illumination variations. Currently, many researchers proposed local descriptors for color images based on the quaternion representation. We cite, for example, the quaternionic local ranking binary pattern (QLRBP) (Lan et al., 2016), the quaternionic weber local descriptor (QWLD)

- (Lan et al., 2017) and the quaternion-Michelson descriptor (QMD) (Lan and Zhou, 2016).
- The performances of the proposed set of moments are to be evaluated in applications dealing with color images, other than retrieval, such as image watermarking, template matching and image enhancement. We can also go behind color image description and employ our quaternary moments in applications where multi-dimensional (3D or 4D) representations are required. For instance, it would be interesting to investigate the representation of RGB-D images by means of quaternions. In fact, RGB-D images/video have been widely used to enhance computer vision algorithms (Han et al., 2013), especially with the development of low-cost RGB-D sensors (such as Microsoft Kinect and Xtion Pro-Live), thanks to which high-quality RGB-D data can be easily acquired.
 - A possible extension of our work is to describe three-dimensional (3D) color objects. In recent years, the analysis of three-dimensional images has gained more and more importance, especially with the availability of acquisition techniques, in the medical and biological sciences, that produce an enormous amount of 3D data. In this context, the spherical harmonics (SHs) have been extensively used for three-dimensional image analysis (Kazhdan et al., 2003; Skibbe and Reisert, 2017; Zhou and Wang, 2017; Bales et al., 2017; Radvilaitė et al., 2017).
 - We are motivated to investigate color face recognition using quaternary deep learning architectures on a GPU-embedded machine. The success of deep learning methods in computer vision depends heavily on the availability of big data as well as machines embedded with fast arithmetic units. Since large scale training face datasets (Labeled Faces in the Wild (LFW) of the University of Massachusetts (Learned-Miller, 2014) and Youtube faces database (YTF) (Wolf et al., 2011)) are available, face recognition is considered one of the areas that have known a significant progress in recent years thanks to using the convolutional neural network (CNN). Besides, the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) (Deng et al., 2009a) has played an important role in providing data for general image classification.
 - It would be useful to extend some deep learning methods, convolutional networks for example, using the quaternion or Clifford algebras. In our knowledge, very few deep learning methods have been proposed in the literature for dealing with multi-dimensional signals based on the algebra of quaternions, e.g., the quaternion principal component analysis network (QPCANet) (Zeng et al., 2016) and the quaternion-based deep belief networks (QDBN) (Papa et al., 2017).
 - We intend to generalize the proposed set of moments using the Clifford algebra. In fact, the Clifford algebra, in which the quaternion algebra is contained, can be employed to handle problems that require multi-dimensional ($> 4D$) representations.
 - For real-time color image retrieval, it would be useful to extend some well-known clustering algorithms by means of the quaternion algebra. Until now, very few studies have been addressing quaternion clustering methods. For instance, Kumar and Dwivedi, 2016 proposed a quaternion version of the k -means clustering algorithm.

Publications of the Author

International Journals (#3)

1. Nisrine Dad, Noureddine En-nahnahi, Said El Alaoui Ouatik and Mohammed Oumsis. "Adaptation of spherical harmonic transform for color shape reconstruction and retrieval using quaternion algebra". In: Journal of Electronic Imaging 25.5. DOI: 10.1117/1.JEI.25.5.053026. 2016.
2. Nisrine Dad, Noureddine En-nahnahi and Said El Alaoui Ouatik. "Parameter-free quaternary orthogonal moments for color image retrieval and recognition". In: Journal of Electronic Imaging 27. DOI: 10.1117/1.JEI.27.1.011007. 2018.
3. Nisrine Dad, Noureddine En-nahnahi and Said El Alaoui Ouatik. "2D color image recognition by quaternion-valued spherical harmonics and a quaternary version of extreme learning machine". **Submitted to a journal on 31 August 2017.**

International Conferences (#3)

1. Nisrine Dad, Noureddine En-nahnahi, Said El Alaoui Ouatik and Mohammed Oumsis. "Quaternion des moments disque-harmoniques pour la description des formes 2D colorées". In: Coresa. Le Creusot. France. 2013.
2. Nisrine Dad, Noureddine En-nahnahi, Said El Alaoui Ouatik and Mohammed Oumsis. "2D color shapes description by quaternion Disc-Harmonic moments". In: IEEE/ACS 11th International Conference on Computer Systems and Applications (AICCSA). Doha. Qatar, pp. 266–272. DOI: 10.1109/AICCSA.2014.7073208. 2014.
3. Nisrine Dad, Noureddine En-nahnahi, Said El Alaoui Ouatik and Mohammed Oumsis. "Quaternion DISC-harmonic moments based on heulpix sampling for fast 2D colour shape retrieval". In: International Symposium on Signal, Image, Video and Communications (ISIVC). Tunis. Tunisia, pp. 131–136. DOI: 10.1109/ISIVC.2016.7893975. 2016.

National Conferences (#4)

1. Nisrine Dad, Noureddine En-nahnahi, Said El Alaoui Ouatik and Mohammed Oumsis. "Quaternion Disc Harmonic Moments for Describing 2D Color Shapes". In: 2ème Edition des Journées Doctorales en Systèmes d'Information, Réseaux et Télécommunications (JDSIRT). Rabat. Morocco. 2013.
2. Nisrine Dad, Noureddine En-nahnahi, Said El Alaoui Ouatik and Mohammed Oumsis. "Quaternion des moments disc-harmonic pour la description des formes 2D colorées: Pertinence et tolérance aux bruits". In: 5th Workshop on Information Technologies and Communication (WOTIC). Fez. Morocco. 2013.

3. Nisrine Dad, Noureddine En-nahnahi and Said El Alaoui Ouatik. “Description des images couleur par des moments orthogonaux quaternioniques”. In: Journées scientifiques nationales en Signal, Image, Multimédia et Applications (SIGMA). Fez. Morocco. 2016.
4. Nisrine Dad, Noureddine En-nahnahi and Said El Alaoui Ouatik. “Moments orthogonaux quaternaires sans paramètres pour la recherche et la reconnaissance des images couleur”. **Accepted on 16 February 2018 in:** Journées scientifiques nationales en Signal, Image, Multimédia et Applications (SIGMA). Rabat. Morocco. 2018.

Bibliography

- Abbasi, S., F. Mokhtarian, and J. Kittler (2000). "Enhancing CSS-based shape retrieval for objects with shallow concavities". In: *Image and Vision Computing* 18.3, pp. 199–211. ISSN: 0262-8856. DOI: [http://doi.org/10.1016/S0262-8856\(99\)00019-0](http://doi.org/10.1016/S0262-8856(99)00019-0). URL: <http://www.sciencedirect.com/science/article/pii/S0262885699000190>.
- Abbasi, Sadegh, Farzin Mokhtarian, and Josef Kittler (1999). "Curvature scale space image in shape similarity retrieval". In: *Multimedia Systems* 7.6, pp. 467–476. ISSN: 1432-1882. DOI: 10.1007/s005300050147. URL: <http://dx.doi.org/10.1007/s005300050147>.
- Abdallah, Samer M., Eduardo M. Nebot, and David C. Rye (1997). "Object recognition and orientation via Zernike moments". In: *Computer Vision — ACCV'98: Third Asian Conference on Computer Vision Hong Kong, China, January 8–10, 1998 Proceedings, Volume I*. Ed. by Roland Chin and Ting-Chuen Pong. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 386–393. ISBN: 978-3-540-69669-8. DOI: 10.1007/3-540-63930-6_145. URL: http://dx.doi.org/10.1007/3-540-63930-6_145.
- Abdeslam Hmimid Mhamed Sayyouri, Hassan Qjidaa (2014). "Image classification using a new set of separable two-dimensional discrete orthogonal invariant moments". In: *Journal of Electronic Imaging* 23, pp. 23–23–16. DOI: 10.1117/1.JEI.23.1.013026. URL: <http://dx.doi.org/10.1117/1.JEI.23.1.013026>.
- Abramowitz, Milton (1974). *Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables*, ISBN: 0486612724.
- Aha, David W., Dennis Kibler, and Marc K. Albert (1991). "Instance-based learning algorithms". In: *Machine Learning* 6.1, pp. 37–66. ISSN: 1573-0565. DOI: 10.1007/BF00153759. URL: <https://doi.org/10.1007/BF00153759>.
- Albuz, Elif, Erturk Kocalar, and Ashfaq A. Khokhar (2001). "Scalable Color Image Indexing and Retrieval Using Vector Wavelets". In: *IEEE Trans. on Knowl. and Data Eng.* 13.5, pp. 851–861. ISSN: 1041-4347. DOI: 10.1109/69.956109. URL: <http://dx.doi.org/10.1109/69.956109>.
- Alexiadis, D. S. and G. D. Sergiadis (2009). "Estimation of Motions in Color Image Sequences Using Hypercomplex Fourier Transforms". In: *IEEE Transactions on Image Processing* 18.1, pp. 168–187. ISSN: 1057-7149. DOI: 10.1109/TIP.2008.2007603.
- Alhichri, Haikel Salem and Mohamed Kamel (2003). "Virtual Circles: A New Set of Features for Fast Image Registration". In: *Pattern Recogn. Lett.* 24.9-10, pp. 1181–1190. ISSN: 0167-8655. DOI: 10.1016/S0167-8655(02)00300-8. URL: [http://dx.doi.org/10.1016/S0167-8655\(02\)00300-8](http://dx.doi.org/10.1016/S0167-8655(02)00300-8).
- Althof, R. J., M. G. J. Wind, and J. T. Dobbins (1997). "A rapid and automatic image registration algorithm with subpixel accuracy". In: *IEEE Transactions on Medical Imaging* 16.3, pp. 308–316. ISSN: 0278-0062. DOI: 10.1109/42.585765.
- Ananth Raj, P. (2017). "Quaternion Circularly Semi-orthogonal Moments for Invariant Image Recognition". In: *Proceedings of International Conference on Computer Vision and Image*

- Processing: CVIP 2016, Volume 2*. Ed. by Balasubramanian Raman, Sanjeev Kumar, Partha Pratim Roy, and Debashis Sen. Singapore: Springer Singapore, pp. 509–522. ISBN: 978-981-10-2107-7. DOI: 10.1007/978-981-10-2107-7_46. URL: http://dx.doi.org/10.1007/978-981-10-2107-7_46.
- Arena, P., L. Fortuna, G. Muscato, and M.G. Xibilia (1997). “Multilayer Perceptrons to Approximate Quaternion Valued Functions”. In: *Neural Networks* 10.2, pp. 335–342. ISSN: 0893-6080. DOI: [http://dx.doi.org/10.1016/S0893-6080\(96\)00048-2](http://dx.doi.org/10.1016/S0893-6080(96)00048-2). URL: <http://www.sciencedirect.com/science/article/pii/S0893608096000482>.
- Bales, Ben, Tresa Pollock, and Linda Petzold (2017). “Segmentation-free image processing and analysis of precipitate shapes in 2D and 3D”. In: *Modelling and Simulation in Materials Science and Engineering* 25.4, p. 045009. URL: <http://stacks.iop.org/0965-0393/25/i=4/a=045009>.
- Ballard, Dana Harry and Christopher M. Brown (1982). *Computer Vision*. 1st. Prentice Hall Professional Technical Reference. ISBN: 0131653164.
- Bas, P., N. Le Bihan, and J. M. Chassery (2003). “Color image watermarking using quaternion Fourier transform”. In: *Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). 2003 IEEE International Conference on*. Vol. 3, III–521–4 vol.3. DOI: 10.1109/ICASSP.2003.1199526.
- Batard, Thomas, Michel Berthier, and Christophe Saint-Jean (2010). “Clifford–Fourier Transform for Color Image Processing”. In: *Geometric Algebra Computing: in Engineering and Computer Science*. Ed. by Eduardo Bayro-Corrochano and Gerik Scheuermann. London: Springer London, pp. 135–162. ISBN: 978-1-84996-108-0. DOI: 10.1007/978-1-84996-108-0_8. URL: http://dx.doi.org/10.1007/978-1-84996-108-0_8.
- Bayro-Corrochano, E. J. and N. Arana-Daniel (2010). “Clifford Support Vector Machines for Classification, Regression, and Recurrence”. In: *IEEE Transactions on Neural Networks* 21.11, pp. 1731–1746. ISSN: 1045-9227. DOI: 10.1109/TNN.2010.2060352.
- Bei-Jing, CHEN, SUN Xing-Ming, WANG Ding-Cheng, and ZHAO Xiao-Ping (2012). “Color Face Recognition Using Quaternion Representation of Color Image”. In: *Acta Automatica Sinica* 38.11, 1815, p. 1815.
- Beijing, Chen, Huazhong Shu, Hui Zhang, Gouenou Coatrieux, Limin Luo, and Jean-Louis Coatrieux (2011). “Combined invariants to similarity transformation and to blur using orthogonal Zernike moments.” In: *IEEE Transactions on Image Processing* 20.2, pp. 345–60. DOI: 10.1109/TIP.2010.2062195. URL: <http://www.hal.inserm.fr/inserm-00522932>.
- Belkasim, S.O., M. Shridhar, and M. Ahmadi (1991). “Pattern recognition with moment invariants: A comparative study and new results”. In: *Pattern Recognition* 24.12, pp. 1117–1138. ISSN: 0031-3203. DOI: [http://dx.doi.org/10.1016/0031-3203\(91\)90140-Z](http://dx.doi.org/10.1016/0031-3203(91)90140-Z). URL: <http://www.sciencedirect.com/science/article/pii/003132039190140Z>.
- Bihan, N. Le and S. J. Sangwine (2003a). “Color image decomposition using quaternion singular value decomposition”. In: *2003 International Conference on Visual Information Engineering VIE 2003*, pp. 113–116. DOI: 10.1049/cp:20030500.
- Bihan, N. Le and S. J. Sangwine (2003b). “Quaternion principal component analysis of color images”. In: *Proceedings 2003 International Conference on Image Processing (Cat. No.03CH37429)*. Vol. 1, I–809–12 vol.1. DOI: 10.1109/ICIP.2003.1247085.

- Billmeyer, Fred W. (1983). "Color Science: Concepts and Methods, Quantitative Data and Formulae, 2nd ed., by Gunter Wyszecki and W. S. Stiles, John Wiley and Sons, New York, 1982, 950 pp." In: *Color Research and Application* 8.4, pp. 262–263. ISSN: 1520-6378. DOI: 10.1002/col.5080080421. URL: <http://dx.doi.org/10.1002/col.5080080421>.
- Bishop, Christopher M. (1995). *Neural Networks for Pattern Recognition*. New York, NY, USA: Oxford University Press, Inc. ISBN: 0198538642.
- Blahut, Richard E. (1985). "Fast Algorithms for Signal Processing and Error Control". In: *The Impact of Processing Techniques on Communications*. Ed. by J. K. Skwirzynski. Dordrecht: Springer Netherlands, pp. 123–144. ISBN: 978-94-009-5113-6. DOI: 10.1007/978-94-009-5113-6_8. URL: http://dx.doi.org/10.1007/978-94-009-5113-6_8.
- Branke, Jürgen (1995). "Evolutionary Algorithms for Neural Network Design and Training". In: *IN PROCEEDINGS OF THE FIRST NORDIC WORKSHOP ON GENETIC ALGORITHMS AND ITS APPLICATIONS*, pp. 145–163.
- Buchholz, Sven and Gerald Sommer (2008). "On Clifford neurons and Clifford multi-layer perceptrons". In: *Neural Networks* 21.7, pp. 925–935. ISSN: 0893-6080. DOI: <https://doi.org/10.1016/j.neunet.2008.03.004>. URL: <http://www.sciencedirect.com/science/article/pii/S0893608008000592>.
- Buchholz, Sven, Kanta Tachibana, and Eckhard M. S. Hitzler (2007). "Optimal Learning Rates for Clifford Neurons". In: *Artificial Neural Networks – ICANN 2007: 17th International Conference, Porto, Portugal, September 9-13, 2007, Proceedings, Part I*. Ed. by Joaquim Marques de Sá, Luís A. Alexandre, Włodzisław Duch, and Danilo Mandic. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 864–873. ISBN: 978-3-540-74690-4. DOI: 10.1007/978-3-540-74690-4_88. URL: https://doi.org/10.1007/978-3-540-74690-4_88.
- Cha, Sung-Hyuk (2007). "Comprehensive Survey on Distance/Similarity Measures between Probability Density Functions". In: *International Journal of Mathematical Models and Methods in Applied Sciences* 1.4, pp. 300–307.
- Chan, T. H., K. Jia, S. Gao, J. Lu, Z. Zeng, and Y. Ma (2015). "PCANet: A Simple Deep Learning Baseline for Image Classification?" In: *IEEE Transactions on Image Processing* 24.12, pp. 5017–5032. ISSN: 1057-7149. DOI: 10.1109/TIP.2015.2475625.
- Chang, Shih-Hsu, Fang-Hsuan Cheng, Wen-Hsing Hsu, and Guo-Zua Wu (1997). "Fast algorithm for point pattern matching: Invariant to translations, rotations and scale changes". In: *Pattern Recognition* 30.2, pp. 311–320. ISSN: 0031-3203. DOI: [http://dx.doi.org/10.1016/S0031-3203\(96\)00076-3](http://dx.doi.org/10.1016/S0031-3203(96)00076-3). URL: <http://www.sciencedirect.com/science/article/pii/S0031320396000763>.
- Chazal, P. de, J. Tapon, and A. van Schaik (2015). "A comparison of extreme learning machines and back-propagation trained feed-forward networks processing the mnist database". In: *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2165–2168. DOI: 10.1109/ICASSP.2015.7178354.
- Chen, Beijing, Huazhong Shu, Hui Zhang, Gang Chen, and Limin Luo (2010). "Color Image Analysis by Quaternion Zernike Moments". In: *Pattern Recognition (ICPR), 2010 20th International Conference on*, pp. 625–628. DOI: 10.1109/ICPR.2010.158.

- Chen, Beijing, Huazhong Shu, Gouenou Coatrieux, Gang Chen, Xingming Sun, and JeanLouis Coatrieux (2014a). “Color Image Analysis by Quaternion-Type Moments”. English. In: *Journal of Mathematical Imaging and Vision*, pp. 1–21. ISSN: 0924-9907. DOI: 10 . 1007 / s10851-014-0511-6.
- Chen, Beijing, Gouenou Coatrieux, Gang Chen, Xingming Sun, Jean Louis Coatrieux, and Huazhong Shu (2014b). “Full 4-D quaternion discrete Fourier transform based watermarking for color images”. In: *Digital Signal Processing* 28, pp. 106 –119. ISSN: 1051-2004. DOI: <http://doi.org/10.1016/j.dsp.2014.02.010>. URL: <http://www.sciencedirect.com/science/article/pii/S1051200414000475>.
- Chen, B.J., H.Z. Shu, H. Zhang, G. Chen, C. Toumoulin, J.L. Dillenseger, and L.M. Luo (2012). “Quaternion Zernike moments and their invariants for color image analysis and object recognition”. In: *Signal Processing* 92.2, pp. 308 –318. ISSN: 0165-1684.
- Chen, Qin-Sheng, M. Defrise, and F. Deconinck (1994). “Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16.12, pp. 1156–1168. ISSN: 0162-8828. DOI: 10.1109/34.387491.
- Chen, S., C. F. N. Cowan, and P. M. Grant (1991). “Orthogonal least squares learning algorithm for radial basis function networks”. In: *IEEE Transactions on Neural Networks* 2.2, pp. 302–309. ISSN: 1045-9227. DOI: 10.1109/72.80341.
- Chong, Chee-Way, P. Raveendran, and R. Mukundan (2003). “A comparative analysis of algorithms for fast computation of Zernike moments”. In: *Pattern Recognition* 36.3, pp. 731 –742. ISSN: 0031-3203. DOI: [http://doi.org/10.1016/S0031-3203\(02\)00091-2](http://doi.org/10.1016/S0031-3203(02)00091-2). URL: <http://www.sciencedirect.com/science/article/pii/S0031320302000912>.
- Cortes, Corinna and Vladimir Vapnik (1995). “Support-Vector Networks”. In: *Machine Learning* 20.3, pp. 273–297. ISSN: 1573-0565. DOI: 10 . 1023 / A : 1022627411411. URL: <https://doi.org/10.1023/A:1022627411411>.
- Deng, An-Wen, Chia-Hung Wei, and Chih-Ying Gwo (2016). “Stable, Fast Computation of High-order Zernike Moments Using a Recursive Method”. In: *Pattern Recogn.* 56.C, pp. 16–25. ISSN: 0031-3203. DOI: 10.1016/j.patcog.2016.02.014. URL: <http://dx.doi.org/10.1016/j.patcog.2016.02.014>.
- Deng, J., W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei (2009a). “ImageNet: A Large-Scale Hierarchical Image Database”. In: *CVPR09*.
- Deng, W., Q. Zheng, and L. Chen (2009b). “Regularized Extreme Learning Machine”. In: *2009 IEEE Symposium on Computational Intelligence and Data Mining*, pp. 389–395. DOI: 10 . 1109/CIDM.2009.4938676.
- Denis, Patrice, Philippe Carre, and Christine Fernandez-Maloigne (2007). “Spatial and spectral quaternionic approaches for colour images”. In: *Computer Vision and Image Understanding* 107.1–2. Special issue on color image processing, pp. 74 –87. ISSN: 1077-3142. DOI: <http://dx.doi.org/10.1016/j.cviu.2006.11.019>. URL: <http://www.sciencedirect.com/science/article/pii/S1077314206002177>.
- Dilts, Gary A (1985). “Computation of spherical harmonic expansion coefficients via FFT’s”. In: *Journal of Computational Physics* 57.3, pp. 439 –453. ISSN: 0021-9991.
- Ding, Shifei, Han Zhao, Yanan Zhang, Xinzheng Xu, and Ru Nie (2015). “Extreme learning machine: algorithm, theory and applications”. In: *Artificial Intelligence Review* 44.1, pp. 103–115. ISSN: 1573-7462. DOI: 10 . 1007 / s10462-013-9405-z. URL: <http://dx.doi.org/10.1007/s10462-013-9405-z>.

- Ding, Shifei, Lili Guo, and Yanlu Hou (2017). “Extreme learning machine with kernel model based on deep learning”. In: *Neural Computing and Applications* 28.8, pp. 1975–1984. ISSN: 1433-3058. DOI: 10.1007/s00521-015-2170-y. URL: <https://doi.org/10.1007/s00521-015-2170-y>.
- Driscoll, J. R. and D. M. Healy (1994). “Computing Fourier Transforms and Convolutions on the 2-Sphere”. In: *Adv. Appl. Math.* 15.2, pp. 202–250. ISSN: 0196-8858. DOI: 10.1006/aama.1994.1008. URL: <http://dx.doi.org/10.1006/aama.1994.1008>.
- Duda, R. O. and P. E. Hart (1973). *Pattern classification and scene analysis*.
- Dunkl, Charles F. and Yuan Xu (2014). *Orthogonal Polynomials of Several Variables*. 2nd ed. Encyclopedia of Mathematics and its Applications. Cambridge University Press. DOI: 10.1017/CBO9781107786134.
- El Fadili, Hakim, Khalid Zenkouar, and Hassan Qjidaa (2003). “Lapped Block Image Analysis via the Method of Legendre Moments”. In: *EURASIP Journal on Advances in Signal Processing* 2003.9, p. 624754. ISSN: 1687-6180. DOI: 10.1155/S1110865703305062. URL: <https://doi.org/10.1155/S1110865703305062>.
- Ell, T. A. (1993). “Quaternion-Fourier transforms for analysis of two-dimensional linear time-invariant partial differential systems”. In: *Proceedings of 32nd IEEE Conference on Decision and Control*, 1830–1841 vol.2. DOI: 10.1109/CDC.1993.325510.
- Ell, T. A. and S. J. Sangwine (2007). “Hypercomplex Fourier Transforms of Color Images”. In: *IEEE Transactions on Image Processing* 16.1, pp. 22–35. ISSN: 1057-7149. DOI: 10.1109/TIP.2006.884955.
- Ell, Todd, Nicolas Le Bihan, and Stephen Sangwine (2014). *Quaternion Fourier Transforms for Signal and Image Processing*. Wiley-ISTE, p. 160. URL: <https://hal.archives-ouvertes.fr/hal-00987367>.
- Ennahnahi, N., M. Oumsis, A. Bouhouch, and M. Meknassi (2010). “Fast shape description based on a set of moments defined on the unit disc and inspired by three-dimensional spherical harmonics”. In: *Image Processing, IET* 4.2, pp. 120–131. ISSN: 1751-9659. DOI: 10.1049/iet-ipr.2009.0318.
- Fei-Fei, L., R. Fergus, and Pietro Perona (2004). “Learning Generative Visual Models From Few Training Examples: An Incremental Bayesian Approach Tested on 101 Object Categories”. In:
- Feng, Wei, Bo Hu, and Cheng Yang (2008). “A subpixel color image registration algorithm using quaternion phase-only correlation”. In: *2008 International Conference on Audio, Language and Image Processing*, pp. 1045–1049. DOI: 10.1109/ICALIP.2008.4590016.
- Fernández, Lidia, Teresa E. Pérez, Miguel A. Piñar, and Yuan Xu (2010). “Krall-type orthogonal polynomials in several variables”. In: *Journal of Computational and Applied Mathematics* 233.6. Special Functions, Information Theory, and Mathematical Physics. Special issue dedicated to Professor Jesus Sanchez Dehesa on the occasion of his 60th birthday, pp. 1519 – 1524. ISSN: 0377-0427. DOI: <https://doi.org/10.1016/j.cam.2009.02.067>. URL: <http://www.sciencedirect.com/science/article/pii/S0377042709001563>.
- Fernández, Lidia, Teresa E. Pérez, and Miguel A. Piñar (2011). “Orthogonal polynomials in two variables as solutions of higher order partial differential equations”. In: *Journal of Approximation Theory* 163.1. Proceedings of the International Workshop on Orthogonal Polynomials and Approximation Theory dedicated to the 60th Anniversary of Guillermo Lopez Lagomasino, pp. 84 –97. ISSN: 0021-9045. DOI: <https://doi.org/10.1016/j.jat>.

- 2009.08.007. URL: <http://www.sciencedirect.com/science/article/pii/S0021904509001658>.
- Fernández, Lidia, Teresa E. Pérez, and Miguel A. Piñar (2012). "On Koornwinder classical orthogonal polynomials in two variables". In: *Journal of Computational and Applied Mathematics* 236.15. Proceedings of the Fifteenth International Congress on Computational and Applied Mathematics (ICCAM-2010), Leuven, Belgium, 5-9 July, 2010, pp. 3817–3826. ISSN: 0377-0427. DOI: <https://doi.org/10.1016/j.cam.2011.08.017>. URL: <http://www.sciencedirect.com/science/article/pii/S0377042711004638>.
- Finkelstein, David, Josef M. Jauch, Samuel Schiminovich, and David Speiser (1962). "Foundations of Quaternion Quantum Mechanics". In: *Journal of Mathematical Physics* 3.2, pp. 207–220. DOI: 10.1063/1.1703794. eprint: <https://doi.org/10.1063/1.1703794>. URL: <https://doi.org/10.1063/1.1703794>.
- Flickner, M. et al. (1995). "Query by image and video content: the QBIC system". In: *Computer* 28.9, pp. 23–32. ISSN: 0018-9162. DOI: 10.1109/2.410146.
- Flusser, Jan and Tomás Suk (1993). "Pattern recognition by affine moment invariants". In: *Pattern Recognition* 26.1, pp. 167–174. ISSN: 0031-3203. DOI: [http://dx.doi.org/10.1016/0031-3203\(93\)90098-H](http://dx.doi.org/10.1016/0031-3203(93)90098-H). URL: <http://www.sciencedirect.com/science/article/pii/003132039390098H>.
- Flusser, Jan, Tomáš Suk, and Barbara Zitová (2009). "Orthogonal Moments". In: *Moments and Moment Invariants in Pattern Recognition*. John Wiley and Sons, Ltd, pp. 165–211. ISBN: 9780470684757. DOI: 10.1002/9780470684757.ch6.
- Frénay, Benoît and Michel Verleysen (2011). "Parameter-insensitive kernel in extreme learning for non-linear support vector regression". In: *Neurocomputing* 74.16. Advances in Extreme Learning Machine: Theory and Applications Biological Inspired Systems. Computational and Ambient Intelligence, pp. 2526–2531. ISSN: 0925-2312. DOI: <https://doi.org/10.1016/j.neucom.2010.11.037>. URL: <http://www.sciencedirect.com/science/article/pii/S0925231211002591>.
- Funkhouser, Thomas, Patrick Min, Michael Kazhdan, Joyce Chen, Alex Halderman, David Dobkin, and David Jacobs (2003). "A Search Engine for 3D Models". In: *ACM Trans. Graph.* 22.1, pp. 83–105. ISSN: 0730-0301. DOI: 10.1145/588272.588279. URL: <http://doi.acm.org/10.1145/588272.588279>.
- Gao, Guangyong and Guoping Jiang (2015). "Bessel-Fourier moment-based robust image zero-watermarking". In: *Multimedia Tools and Applications* 74.3, pp. 841–858. ISSN: 1573-7721. DOI: 10.1007/s11042-013-1701-8. URL: <http://dx.doi.org/10.1007/s11042-013-1701-8>.
- Gestel, Tony van, Johan A.K. Suykens, Bart Baesens, Stijn Viaene, Jan Vanthienen, Guido Dedene, Bart de Moor, and Joos Vandewalle (2004). "Benchmarking Least Squares Support Vector Machine Classifiers". In: *Machine Learning* 54.1, pp. 5–32. ISSN: 1573-0565. DOI: 10.1023/B:MACH.0000008082.80494.e0. URL: <http://dx.doi.org/10.1023/B:MACH.0000008082.80494.e0>.
- Geusebroek, Jan-Mark, Gertjan J Burghouts, and Arnold WM Smeulders (2005). "The Amsterdam library of object images". In: *International Journal of Computer Vision* 61.1, pp. 103–112.
- Goshtasby, A., G. C. Stockman, and C. V. Page (1986). "A Region-Based Approach to Digital Image Registration with Subpixel Accuracy". In: *IEEE Transactions on Geoscience and*

- Remote Sensing* GE-24.3, pp. 390–399. ISSN: 0196-2892. DOI: 10.1109/TGRS.1986.289597.
- Gough, W (1984). “Quaternions and spherical harmonics”. In: *European Journal of Physics* 5.3, p. 163. URL: <http://stacks.iop.org/0143-0807/5/i=3/a=008>.
- Green, Robin (2003). “Spherical harmonic lighting: The gritty details”. In: *Archives of the Game Developers Conference*. Vol. 56, p. 4.
- Grigoryan, Artyom M., John Jenkinson, and Sos S. Agaian (2015). “Quaternion Fourier transform based alpha-rooting method for color image measurement and enhancement”. In: *Signal Processing* 109, pp. 269–289. ISSN: 0165-1684. DOI: <http://doi.org/10.1016/j.sigpro.2014.11.019>. URL: <http://www.sciencedirect.com/science/article/pii/S0165168414005581>.
- Górski, K. M., E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann (2005). “HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere”. In: *The Astrophysical Journal* 622.2, p. 759. URL: <http://stacks.iop.org/0004-637X/622/i=2/a=759>.
- Guo, Li-Qiang and Ming Zhu (2011). “Quaternion Fourier–Mellin moments for color images”. In: *Pattern Recognition* 44.2, pp. 187–195. ISSN: 0031-3203. DOI: <http://doi.org/10.1016/j.patcog.2010.08.017>. URL: <http://www.sciencedirect.com/science/article/pii/S0031320310004036>.
- Hagan, M. T. and M. B. Menhaj (1994). “Training feedforward networks with the Marquardt algorithm”. In: *IEEE Transactions on Neural Networks* 5.6, pp. 989–993. ISSN: 1045-9227. DOI: 10.1109/72.329697.
- Hamilton, William Rowan (1844). “II. On quaternions; or on a new system of imaginaries in algebra”. In: *Philosophical Magazine Series 3* 25.163, pp. 10–13. DOI: 10.1080/14786444408644923.
- Han, Ju and Kai-Kuang Ma (2002). “Fuzzy Color Histogram and Its Use in Color Image Retrieval”. In: *Trans. Img. Proc.* 11.8, pp. 944–952. ISSN: 1057-7149. DOI: 10.1109/TIP.2002.801585. URL: <http://dx.doi.org/10.1109/TIP.2002.801585>.
- Han, Jungong, Ling Shao, Dong Xu, and Jamie Shotton (2013). “Enhanced Computer Vision with Microsoft Kinect Sensor: A Review”. In: *IEEE Trans. Cybernetics* 43. URL: <https://www.microsoft.com/en-us/research/publication/enhanced-computer-vision-with-microsoft-kinect-sensor-a-review/>.
- Healy, D.M., D.N. Rockmore, P.J. Kostelec, and S. Moore (2003). “FFTs for the 2-Sphere: Improvements and Variations”. In: *Journal of Fourier Analysis and Applications* 9.4, pp. 341–385. ISSN: 1531-5851. DOI: 10.1007/s00041-003-0018-9. URL: <https://doi.org/10.1007/s00041-003-0018-9>.
- Hestenes, David and Garret Sobczyk (1987). *Clifford Algebra to Geometric Calculus: A Unified Language for Mathematics and Physics (Fundamental Theories of Physics)*. Kluwer Academic Publishers.
- Hinton, Geoffrey and Ruslan Salakhutdinov (2006). “Reducing the Dimensionality of Data with Neural Networks”. In: *Science* 313.5786, pp. 504–507.
- Hinton, Geoffrey E., Simon Osindero, and Yee-Whye Teh (2006). “A Fast Learning Algorithm for Deep Belief Nets”. In: *Neural Comput.* 18.7, pp. 1527–1554. ISSN: 0899-7667. DOI: 10.1162/neco.2006.18.7.1527. URL: <http://dx.doi.org/10.1162/neco.2006.18.7.1527>.
- Hirose, A. (2006). *Complex-Valued Neural Networks*.

- Hirose, Akira (2004). *Complex-Valued Neural Networks: Theories and Applications (Series on Innovative Intelligence, 5)*. World Scientific Press. ISBN: 9812384642.
- Hmimid, Abdeslam, Mhamed Sayyouri, and Hassan Qjidaa (2015). “Fast Computation of Separable Two-dimensional Discrete Invariant Moments for Image Classification”. In: *Pattern Recogn.* 48.2, pp. 509–521. ISSN: 0031-3203. DOI: 10.1016/j.patcog.2014.08.020. URL: <http://dx.doi.org/10.1016/j.patcog.2014.08.020>.
- Holte, Robert C. (1993). “Very Simple Classification Rules Perform Well on Most Commonly Used Datasets”. In: *Machine Learning* 11.1, pp. 63–90. ISSN: 1573-0565. DOI: 10.1023/A:1022631118932. URL: <https://doi.org/10.1023/A:1022631118932>.
- Hosny, Khalid M. and Mohamed M. Darwish (2016). “A kernel-based method for fast and accurate computation of PHT in polar coordinates”. In: *Journal of Real-Time Image Processing*, pp. 1–13. ISSN: 1861-8219. DOI: 10.1007/s11554-016-0622-y. URL: <http://dx.doi.org/10.1007/s11554-016-0622-y>.
- Hosny, Khalid-M. and Mohamed M. Darwish (2017). “Accurate computation of quaternion polar complex exponential transform for color images in different coordinate systems”. In: *Journal of Electronic Imaging* 26.2, p. 023021. DOI: 10.1117/1.JEI.26.2.023021. URL: <http://dx.doi.org/10.1117/1.JEI.26.2.023021>.
- Hosny, Khalid Mohamed, Mohamed A. Shouman, and Hayam M. Abdel Salam (2011). “Fast computation of orthogonal Fourier–Mellin moments in polar coordinates”. In: *Journal of Real-Time Image Processing* 6.2, pp. 73–80. ISSN: 1861-8219. DOI: 10.1007/s11554-009-0135-z. URL: <http://dx.doi.org/10.1007/s11554-009-0135-z>.
- Hu, Hai tao, Ya dong Zhang, Chao Shao, and Quan Ju (2014). “Orthogonal moments based on exponent functions: Exponent-Fourier moments”. In: *Pattern Recognition* 47.8, pp. 2596–2606. ISSN: 0031-3203. DOI: <https://doi.org/10.1016/j.patcog.2014.02.014>. URL: <http://www.sciencedirect.com/science/article/pii/S0031320314000661>.
- Hu, Hai tao, Quan Ju, and Chao Shao (2016). “Errata and comments on “Errata and comments on Orthogonal moments based on exponent functions: Exponent-Fourier moments””. In: *Pattern Recognition* 52, pp. 471–476. ISSN: 0031-3203. DOI: <https://doi.org/10.1016/j.patcog.2015.10.005>. URL: <http://www.sciencedirect.com/science/article/pii/S0031320315003726>.
- Hu, Ming-Kuei (1962). “Visual pattern recognition by moment invariants, computer methods in image analysis”. In: *IRE Transactions on Information Theory* 8.
- Huang, G. B., H. Zhou, X. Ding, and R. Zhang (2012). “Extreme Learning Machine for Regression and Multiclass Classification”. In: *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 42.2, pp. 513–529. ISSN: 1083-4419. DOI: 10.1109/TSMCB.2011.2168604.
- Huang, Gao, Guang-Bin Huang, Shiji Song, and Keyou You (2015). “Trends in extreme learning machines: A review”. In: *Neural Networks* 61. Supplement C, pp. 32–48. ISSN: 0893-6080. DOI: <https://doi.org/10.1016/j.neunet.2014.10.001>. URL: <http://www.sciencedirect.com/science/article/pii/S0893608014002214>.
- Huang, Guang-Bin (2015). “What are Extreme Learning Machines? Filling the Gap Between Frank Rosenblatt’s Dream and John von Neumann’s Puzzle”. In: *Cognitive Computation* 7.3, pp. 263–278. ISSN: 1866-9964. DOI: 10.1007/s12559-015-9333-0. URL: <https://doi.org/10.1007/s12559-015-9333-0>.
- Huang, Guang-Bin and H. A. Babri (1998). “Upper bounds on the number of hidden neurons in feedforward networks with arbitrary bounded nonlinear activation functions”. In: *IEEE*

- Transactions on Neural Networks* 9.1, pp. 224–229. ISSN: 1045-9227. DOI: 10.1109/72.655045.
- Huang, Guang-Bin and Chee-Kheong Siew (2004). “Extreme learning machine: RBF network case”. In: *ICARCV 2004 8th Control, Automation, Robotics and Vision Conference, 2004*. Vol. 2, 1029–1036 Vol. 2. DOI: 10.1109/ICARCV.2004.1468985.
- Huang, Guang-Bin, Qin-Yu Zhu, and Chee-Kheong Siew (2004). “Extreme learning machine: a new learning scheme of feedforward neural networks”. In: *2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541)*. Vol. 2, 985–990 vol.2. DOI: 10.1109/IJCNN.2004.1380068.
- Huang, Guang-Bin, Qin-Yu Zhu, K. Z. Mao, Chee-Kheong Siew, P. Saratchandran, and N. Sundararajan (2006a). “Can threshold networks be trained directly?” In: *IEEE Transactions on Circuits and Systems II: Express Briefs* 53.3, pp. 187–191. ISSN: 1549-7747. DOI: 10.1109/TCSII.2005.857540.
- Huang, Guang-Bin, Qin-Yu Zhu, and Chee-Kheong Siew (2006b). “Extreme learning machine: Theory and applications”. In: *Neurocomputing* 70.1. Neural Networks, pp. 489–501. ISSN: 0925-2312. DOI: <https://doi.org/10.1016/j.neucom.2005.12.126>. URL: <http://www.sciencedirect.com/science/article/pii/S0925231206000385>.
- Huang, Jing, S. Ravi Kumar, Mandar Mitra, Wei-Jing Zhu, and Ramin Zabih (1997). “Image Indexing Using Color Correlograms”. In: *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97)*. CVPR '97. Washington, DC, USA: IEEE Computer Society, pp. 762–. ISBN: 0-8186-7822-4. URL: <http://dl.acm.org/citation.cfm?id=794189.794514>.
- Iosifidis, Alexandros, Anastasios Tefas, and Ioannis Pitas (2015). “On the kernel Extreme Learning Machine classifier”. In: *Pattern Recognition Letters* 54, pp. 11–17. ISSN: 0167-8655. DOI: <http://dx.doi.org/10.1016/j.patrec.2014.12.003>. URL: <http://www.sciencedirect.com/science/article/pii/S0167865514003705>.
- Isokawa, T., H. Nishimura, N. Kamiura, and N. Matsui (2006). “Fundamental Properties of Quaternionic Hopfield Neural Network”. In: *The 2006 IEEE International Joint Conference on Neural Network Proceedings*, pp. 218–223. DOI: 10.1109/IJCNN.2006.246683.
- Jayadeva, R. Khemchandani, and S. Chandra (2007). “Twin Support Vector Machines for Pattern Classification”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29.5, pp. 905–910. ISSN: 0162-8828. DOI: 10.1109/TPAMI.2007.1068.
- Jing, Jiang Yong (2011). “Exponent Moments and Its Application in Pattern Recognition”. dissertation. Beijing University of Posts and Telecommunications.
- John, George H. and Pat Langley (1995). “Estimating Continuous Distributions in Bayesian Classifiers”. In: *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*. UAI'95. Montré#233;l, Qu#233;, Canada: Morgan Kaufmann Publishers Inc., pp. 338–345. ISBN: 1-55860-385-9. URL: <http://dl.acm.org/citation.cfm?id=2074158.2074196>.
- Kavukcuoglu, Koray, Pierre Sermanet, Y-Lan Boureau, Karol Gregor, Michaël Mathieu, and Yann LeCun (2010). “Learning Convolutional Feature Hierarchies for Visual Recognition”. In: *Proceedings of the 23rd International Conference on Neural Information Processing Systems - Volume 1*. NIPS'10. Vancouver, British Columbia, Canada: Curran Associates Inc., pp. 1090–1098. URL: <http://dl.acm.org/citation.cfm?id=2997189.2997311>.

- Kazhdan, Michael, Thomas Funkhouser, and Szymon Rusinkiewicz (2003). "Rotation Invariant Spherical Harmonic Representation of 3D Shape Descriptors". In: *Proceedings of the 2003 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing*. SGP '03. Aachen, Germany: Eurographics Association, pp. 156–164. ISBN: 1-58113-687-0. URL: <http://dl.acm.org/citation.cfm?id=882370.882392>.
- Khotanzad, A. and Yaw Hua Hong (1990). "Invariant image recognition by Zernike moments". In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 12.5, pp. 489–497. ISSN: 0162-8828. DOI: 10.1109/34.55109.
- Kim, Hyung Shin and Heung-Kyu Lee (2003). "Invariant image watermark using Zernike moments". In: *IEEE Transactions on Circuits and Systems for Video Technology* 13.8, pp. 766–775. ISSN: 1051-8215. DOI: 10.1109/TCSVT.2003.815955.
- Kobayashi, Masaki (2017a). "Quaternionic Hopfield neural networks with twin-multistate activation function". In: *Neurocomputing* 267.Supplement C, pp. 304–310. ISSN: 0925-2312. DOI: <https://doi.org/10.1016/j.neucom.2017.06.013>. URL: <http://www.sciencedirect.com/science/article/pii/S0925231217310688>.
- Kobayashi, Masaki (2017b). "Symmetric quaternionic Hopfield neural networks". In: *Neurocomputing* 240.Supplement C, pp. 110–114. ISSN: 0925-2312. DOI: <https://doi.org/10.1016/j.neucom.2017.02.044>. URL: <http://www.sciencedirect.com/science/article/pii/S0925231217303351>.
- Kobayashi, Masaki and Akihiro Nakajima (2012). "Twisted quaternary neural networks". In: *IEEE Transactions on Electrical and Electronic Engineering* 7.4, pp. 397–401. ISSN: 1931-4981. DOI: 10.1002/tee.21746. URL: <http://dx.doi.org/10.1002/tee.21746>.
- Koornwinder, Tom (1975). "Two-Variable Analogues of the Classical Orthogonal Polynomials". In: *Theory and Application of Special Functions*. Ed. by Richard A. Askey. Academic Press, pp. 435–495. ISBN: 978-0-12-064850-4. DOI: <https://doi.org/10.1016/B978-0-12-064850-4.50015-X>. URL: <https://www.sciencedirect.com/science/article/pii/B978012064850450015X>.
- Kostelec, Peter J. and Daniel N. Rockmore (2004). *S2Kit: A Lite Version of SpharmonicKit*.
- Krizhevsky, Alex, Vinod Nair, and Geoffrey Hinton. "Canadian Institute for Advanced Research". In: URL: <http://www.cs.toronto.edu/~kriz/cifar.html>.
- Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton (2012). "ImageNet Classification with Deep Convolutional Neural Networks". In: *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume I*. NIPS'12. Lake Tahoe, Nevada: Curran Associates Inc., pp. 1097–1105. URL: <http://dl.acm.org/citation.cfm?id=2999134.2999257>.
- Kumar, R. and R. Dwivedi (2016). "Quaternion Domain k -Means Clustering for Improved Real Time Classification of E-Nose Data". In: *IEEE Sensors Journal* 16.1, pp. 177–184. ISSN: 1530-437X. DOI: 10.1109/JSEN.2015.2475640.
- Lan, R. and Y. Zhou (2016). "Quaternion-Michelson Descriptor for Color Image Classification". In: *IEEE Transactions on Image Processing* 25.11, pp. 5281–5292. ISSN: 1057-7149. DOI: 10.1109/TIP.2016.2605922.
- Lan, R., Y. Zhou, and Y. Y. Tang (2016). "Quaternionic Local Ranking Binary Pattern: A Local Descriptor of Color Images". In: *IEEE Transactions on Image Processing* 25.2, pp. 566–579. ISSN: 1057-7149. DOI: 10.1109/TIP.2015.2507404.

- Lan, R., Y. Zhou, and Y. Y. Tang (2017). “Quaternionic Weber Local Descriptor of Color Images”. In: *IEEE Transactions on Circuits and Systems for Video Technology* 27.2, pp. 261–274. ISSN: 1051-8215. DOI: 10.1109/TCSVT.2015.2492839.
- Learned-Miller, Gary B. Huang Erik (2014). *Labeled Faces in the Wild: Updates and New Reporting Procedures*. Tech. rep. UM-CS-2014-003. University of Massachusetts, Amherst.
- LeCun, Y., B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel (1989). “Backpropagation Applied to Handwritten Zip Code Recognition”. In: *Neural Computation* 1.4, pp. 541–551. ISSN: 0899-7667. DOI: 10.1162/neco.1989.1.4.541.
- Lecun, Y., L. Bottou, Y. Bengio, and P. Haffner (1998). “Gradient-based learning applied to document recognition”. In: *Proceedings of the IEEE* 86.11, pp. 2278–2324. ISSN: 0018-9219. DOI: 10.1109/5.726791.
- Lee, J.-M. and W.-Y. Kim (2012). “A New Shape Description Method Using Angular Radial Transform”. In: *IEICE Transactions on Information and Systems* 95, pp. 1628–1635. DOI: 10.1587/transinf.E95.D.1628.
- Lewanowicz, Stanisław and Paweł Woźny (2010). “Two-variable orthogonal polynomials of big q -Jacobi type”. In: *Journal of Computational and Applied Mathematics* 233.6. Special Functions, Information Theory, and Mathematical Physics. Special issue dedicated to Professor Jesus Sanchez Dehesa on the occasion of his 60th birthday, pp. 1554–1561. ISSN: 0377-0427. DOI: <https://doi.org/10.1016/j.cam.2009.02.070>. URL: <http://www.sciencedirect.com/science/article/pii/S0377042709001617>.
- Li, Kang, Jian-Xun Peng, and G. W. Irwin (2005). “A Fast Nonlinear Model Identification Method”. In: *IEEE Transactions on Automatic Control* 50.8, pp. 1211–1216. ISSN: 0018-9286. DOI: 10.1109/TAC.2005.852557.
- Li, Xuelong (2003). “Image retrieval based on perceptive weighted color blocks”. In: *Pattern Recognition Letters* 24.12, pp. 1935–1941. ISSN: 0167-8655. DOI: [https://doi.org/10.1016/S0167-8655\(03\)00032-1](https://doi.org/10.1016/S0167-8655(03)00032-1). URL: <http://www.sciencedirect.com/science/article/pii/S0167865503000321>.
- Li, Y. N. (2013). “Quaternion Polar Harmonic Transforms for Color Images”. In: *IEEE Signal Processing Letters* 20.8, pp. 803–806. ISSN: 1070-9908. DOI: 10.1109/LSP.2013.2267775.
- Liang, Ming and Xiaolin Hu (2015). “Recurrent convolutional neural network for object recognition”. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3367–3375. DOI: 10.1109/CVPR.2015.7298958.
- Lichman, M. (2013). *UCI Machine Learning Repository*. URL: <http://archive.ics.uci.edu/ml>.
- Lin, Yuanqing, Tong Zhang, Shenghuo Zhu, and Kai Yu (2010). “Deep Coding Network”. In: *Advances in Neural Information Processing Systems* 23. Ed. by J. D. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel, and A. Culotta. Curran Associates, Inc., pp. 1405–1413. URL: <http://papers.nips.cc/paper/3929-deep-coding-network.pdf>.
- López-González, G., Nancy Arana-Daniel, and Eduardo Bayro-Corrochano (2014). “Quaternion Support Vector Classifier”. In: *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications: 19th Iberoamerican Congress, CIARP 2014, Puerto Vallarta, Mexico, November 2-5, 2014. Proceedings*. Cham: Springer International Publishing, pp. 722–729. ISBN: 978-3-319-12568-8.

- Lucchese, L., G. Doretto, and G. M. Cortelazzo (2002). "A frequency domain technique for range data registration". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24.11, pp. 1468–1484. ISSN: 0162-8828. DOI: 10.1109/TPAMI.2002.1046160.
- Ma, Zi ping, Bao sheng Kang, and B. Xiao (2011). "A study of Bessel Fourier moments invariant for image retrieval and classification". In: *2011 International Conference on Image Analysis and Signal Processing*, pp. 316–320. DOI: 10.1109/IASP.2011.6109054.
- McCulloch, Warren S. and Walter Pitts (1943). "A logical calculus of the ideas immanent in nervous activity". In: *The bulletin of mathematical biophysics* 5.4, pp. 115–133. ISSN: 1522-9602. DOI: 10.1007/BF02478259. URL: <https://doi.org/10.1007/BF02478259>.
- Mehtre, Babu M., Mohan S. Kankanhalli, and Wing Foon Lee (1997). "Shape measures for content based image retrieval: A comparison". In: *Information Processing & Management* 33.3, pp. 319–337. ISSN: 0306-4573. DOI: [http://dx.doi.org/10.1016/S0306-4573\(96\)00069-6](http://dx.doi.org/10.1016/S0306-4573(96)00069-6). URL: <http://www.sciencedirect.com/science/article/pii/S0306457396000696>.
- Mennesson, J., C. Saint-Jean, and L. Mascarilla (2010). "New geometric fourier descriptors for color image recognition". In: *2010 IEEE International Conference on Image Processing*, pp. 2685–2688. DOI: 10.1109/ICIP.2010.5651939.
- Mennesson, J., C. Saint-Jean, and L. Mascarilla (2014). "Color Fourier–Mellin descriptors for image recognition". In: *Pattern Recognition Letters* 40, pp. 27–35. ISSN: 0167-8655. DOI: <http://dx.doi.org/10.1016/j.patrec.2013.12.014>. URL: <http://www.sciencedirect.com/science/article/pii/S0167865513005126>.
- Mesbah, Abderrahim, Aissam Berrahou, Mostafa El Mallahi, and Hassan Qjidaa (2016). "Fast and efficient computation of three-dimensional Hahn moments". In: *Journal of Electronic Imaging* 25, p. 12. DOI: 10.1117/1.JEI.25.6.061621.
- Mindru, Florica, Tinne Tuytelaars, Luc Van Gool, and Theo Moons (2004). "Moment invariants for recognition under changing viewpoint and illumination". In: *Computer Vision and Image Understanding* 94.1–3. Special Issue: Colour for Image Indexing and Retrieval, pp. 3–27. ISSN: 1077-3142. DOI: <http://doi.org/10.1016/j.cviu.2003.10.011>. URL: <http://www.sciencedirect.com/science/article/pii/S1077314203001930>.
- Minemoto, Toshifumi, Tejiro Isokawa, Haruhiko Nishimura, and Nobuyuki Matsui (2016a). "Feed forward neural network with random quaternionic neurons". In: *Signal Processing*, pp. –. ISSN: 0165-1684. DOI: <http://dx.doi.org/10.1016/j.sigpro.2016.11.008>. URL: <http://www.sciencedirect.com/science/article/pii/S016516841630322X>.
- Minemoto, Toshifumi, Tejiro Isokawa, Haruhiko Nishimura, and Nobuyuki Matsui (2016b). "Quaternionic multistate Hopfield neural network with extended projection rule". In: *Artificial Life and Robotics* 21.1, pp. 106–111. ISSN: 1614-7456. DOI: 10.1007/s10015-015-0247-4. URL: <https://doi.org/10.1007/s10015-015-0247-4>.
- Mokhtarian, F. and A. K. Mackworth (1992). "A theory of multiscale, curvature-based shape representation for planar curves". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14.8, pp. 789–805. ISSN: 0162-8828. DOI: 10.1109/34.149591.
- Mokhtarian, Farzin, Sadegh Abbasi, and Josef Kittler (1996). "Efficient and Robust Retrieval by Shape Content through Curvature Scale Space". In: pp. 35–42.
- Morales, María Álvarez de, Lidia Fernández, Teresa E. Pérez, and Miguel A. Piñar (2009). "Bivariate orthogonal polynomials in the Lyskova class". In: *Journal of Computational and*

- Applied Mathematics* 233.3. 9th OPSFA Conference, pp. 597–601. ISSN: 0377-0427. DOI: <https://doi.org/10.1016/j.cam.2009.02.027>. URL: <http://www.sciencedirect.com/science/article/pii/S0377042709001010>.
- Moxey, C. E., S. J. Sangwine, and T. A. Ell (2003). “Hypercomplex correlation techniques for vector images”. In: *IEEE Transactions on Signal Processing* 51.7, pp. 1941–1953. ISSN: 1053-587X. DOI: 10.1109/TSP.2003.812734.
- Nagarajan, B. and V. Devendran (2012). “Vehicle Classification under Cluttered Background and Mild Occlusion Using Zernike Features”. In: *Procedia Engineering* 30, pp. 201–209. ISSN: 1877-7058. DOI: <http://dx.doi.org/10.1016/j.proeng.2012.01.852>. URL: <http://www.sciencedirect.com/science/article/pii/S1877705812008624>.
- Nawaz, Qamar, Bin Xiao, Isma Hamid, and Du Jiao (2016). “Multi-modal Color Medical Image Fusion Using Quaternion Discrete Fourier Transform”. In: *Sensing and Imaging* 17.1, p. 7. ISSN: 1557-2072. DOI: 10.1007/s11220-016-0131-x. URL: <http://dx.doi.org/10.1007/s11220-016-0131-x>.
- Nene, S., S. Nayar, and H. Murase (1996). *Columbia Object Image Library (COIL-100)*. Tech. rep. Columbia University.
- Nitta, T. (1995). “A quaternary version of the back-propagation algorithm”. In: *IEEE International Conference on Neural Networks Proceedings*. Vol. 5, 2753–2756 vol.5. DOI: 10.1109/ICNN.1995.488166.
- Nitta, Tohru (2009). *Complex-valued Neural Networks: Utilizing High-dimensional Parameters*. Hershey, PA: Information Science Reference - Imprint of: IGI Publishing. ISBN: 1605662143, 9781605662145.
- Ogle, V. E. and M. Stonebraker (1995). “Chabot: retrieval from a relational database of images”. In: *Computer* 28.9, pp. 40–48. ISSN: 0018-9162. DOI: 10.1109/2.410150.
- Olshausen, B. and D. Field (1996). “Emergence of Simple-Cell Receptive Field Properties by Learning a Sparse Code for Natural Images”. In: *Nature* 381, pp. 607–609.
- Ouyang, Junlin, Gouenou Coatrieux, Beijing Chen, and Huazhong Shu (2015). “Color image watermarking based on quaternion Fourier transform and improved uniform log-polar mapping”. In: *Computers & Electrical Engineering* 46, pp. 419–432. ISSN: 0045-7906. DOI: <http://doi.org/10.1016/j.compeleceng.2015.03.004>. URL: <http://www.sciencedirect.com/science/article/pii/S0045790615000725>.
- Pan, Juntao, Shumin Fei, Wenlin Zou, and Yudong Ni (2010). “Color image registration via quaternion Fourier transform and parametric template method”. In: *2010 8th World Congress on Intelligent Control and Automation*, pp. 1006–1011. DOI: 10.1109/WCICA.2010.5554658.
- Papa, João Paulo, Gustavo H. Rosa, Danillo R. Pereira, and Xin-She Yang (2017). “Quaternion-based Deep Belief Networks fine-tuning”. In: *Applied Soft Computing* 60. Supplement C, pp. 328–335. ISSN: 1568-4946. DOI: <https://doi.org/10.1016/j.asoc.2017.06.046>. URL: <http://www.sciencedirect.com/science/article/pii/S1568494617303897>.
- Park, M., J. S. Jin, and L. S. Wilson (2002). “Fast content-based image retrieval using quasi-Gabor filter and reduction of image feature dimension”. In: *Proceedings Fifth IEEE South-west Symposium on Image Analysis and Interpretation*, pp. 178–182. DOI: 10.1109/IAI.2002.999914.
- Paschos, George, Ivan Radev, and Nagarajan Prabakar (2003). “Image Content-Based Retrieval Using Chromaticity Moments”. In: *IEEE Trans. on Knowl. and Data Eng.* 15.5, pp. 1069–

1072. ISSN: 1041-4347. DOI: 10.1109/TKDE.2003.1232264. URL: <http://dx.doi.org/10.1109/TKDE.2003.1232264>.
- Pass, G. and R. Zabih (1996). "Histogram refinement for content-based image retrieval". In: *Applications of Computer Vision, 1996. WACV '96., Proceedings 3rd IEEE Workshop on*, pp. 96–102. DOI: 10.1109/ACV.1996.572008.
- Pass, Greg, Ramin Zabih, and Justin Miller (1996). "Comparing Images Using Color Coherence Vectors". In: *Proceedings of the Fourth ACM International Conference on Multimedia. MULTIMEDIA '96*. Boston, Massachusetts, USA: ACM, pp. 65–73. ISBN: 0-89791-871-1. DOI: 10.1145/244130.244148. URL: <http://doi.acm.org/10.1145/244130.244148>.
- Pratt, W. K. (1974). "Correlation Techniques of Image Registration". In: *IEEE Transactions on Aerospace and Electronic Systems* AES-10.3, pp. 353–358. ISSN: 0018-9251. DOI: 10.1109/TAES.1974.307828.
- Press, William H., Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery (2007). *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. 3rd ed. New York, NY, USA: Cambridge University Press. ISBN: 0521880688, 9780521880688.
- Prokop, Richard J and Anthony P Reeves (1992). "A survey of moment-based techniques for unoccluded object representation and recognition". In: *CVGIP: Graphical Models and Image Processing* 54.5, pp. 438–460. ISSN: 1049-9652. DOI: [http://dx.doi.org/10.1016/1049-9652\(92\)90027-U](http://dx.doi.org/10.1016/1049-9652(92)90027-U). URL: <http://www.sciencedirect.com/science/article/pii/104996529290027U>.
- Pu, Jiantao and Karthik Ramani (2006). "On visual similarity based 2D drawing retrieval". In: *Computer-Aided Design* 38.3, pp. 249–259. ISSN: 0010-4485. DOI: <http://doi.org/10.1016/j.cad.2005.10.009>. URL: <http://www.sciencedirect.com/science/article/pii/S0010448505001855>.
- Qi, Q., Y. Hu, and C. Zhang (2008). "A New Color Pattern Recognition Algorithm Based On Quaternion Correlation". In: *2008 IEEE International Conference on Networking, Sensing and Control*, pp. 462–466. DOI: 10.1109/ICNSC.2008.4525261.
- Qjidaa, H. and L. Radouane (1999). "Robust line fitting in a noisy image by the method of moments". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21.11, pp. 1216–1223. ISSN: 0162-8828. DOI: 10.1109/34.809115.
- Quinlan, J. Ross (1993). *C4.5: Programs for Machine Learning*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. ISBN: 1-55860-238-0.
- Radvilaitė, Urtė, Rimantas Kačianauskas, Dainius Rusakevičius, and Arūnas Jaras (2017). "Modelling Soil Particles by Low-resolution Spherical Harmonics". In: *Procedia Engineering* 172. Modern Building Materials, Structures and Techniques, pp. 913–921. ISSN: 1877-7058. DOI: <https://doi.org/10.1016/j.proeng.2017.02.100>. URL: <http://www.sciencedirect.com/science/article/pii/S1877705817306069>.
- Reddy, B. S. and B. N. Chatterji (1996). "An FFT-based technique for translation, rotation, and scale-invariant image registration". In: *IEEE Transactions on Image Processing* 5.8, pp. 1266–1271. ISSN: 1057-7149. DOI: 10.1109/83.506761.
- Ren, Haiping, Ziliang Ping, Wurigen Bo, Wenkai Wu, and Yunlong Sheng (2003). "Multidistortion-invariant image recognition with radial harmonic Fourier moments". In: *J. Opt. Soc. Am. A* 20.4, pp. 631–637. DOI: 10.1364/JOSAA.20.000631. URL: <http://josaa.osa.org/abstract.cfm?URI=josaa-20-4-631>.
- Ripley, Brian D. and N. L. Hjort (1995). *Pattern Recognition and Neural Networks*. 1st. New York, NY, USA: Cambridge University Press. ISBN: 0521460867.

- Rojas, Raúl (1996). *Neural Networks: A Systematic Introduction*. New York, NY, USA: Springer-Verlag New York, Inc. ISBN: 3-540-60505-3.
- Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams (1988). “Neurocomputing: Foundations of Research”. In: ed. by James A. Anderson and Edward Rosenfeld. Cambridge, MA, USA: MIT Press. Chap. Learning Representations by Back-propagating Errors, pp. 696–699. ISBN: 0-262-01097-6. URL: <http://dl.acm.org/citation.cfm?id=65669.104451>.
- Sajanhar, A., Guojun Lu, and Dengsheng Zhang (2004). “Discriminating shape descriptors based on connectivity [shape representation for content-based image retrieval]”. In: *2004 IEEE International Conference on Multimedia and Expo (ICME) (IEEE Cat. No.04TH8763)*. Vol. 3, 2075–2078 Vol.3. DOI: 10.1109/ICME.2004.1394674.
- Sajjanhart, A., Guojun Lu, and Dengsheng Zhang (2005). “Spherical Harmonics Descriptor for 2D-Image Retrieval”. In: *2005 IEEE International Conference on Multimedia and Expo*, pp. 105–108. DOI: 10.1109/ICME.2005.1521371.
- Salakhutdinov, Ruslan and Geoffrey Hinton (2009). “Deep Boltzmann Machines”. In: *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics*. Ed. by David van Dyk and Max Welling. Vol. 5. Proceedings of Machine Learning Research. Hilton Clearwater Beach Resort, Clearwater Beach, Florida USA: PMLR, pp. 448–455. URL: <http://proceedings.mlr.press/v5/salakhutdinov09a.html>.
- Sangwine, S. J. (1996). “Fourier transforms of colour images using quaternion or hypercomplex numbers”. In: *Electronics Letters* 32.21, pp. 1979–1980. ISSN: 0013-5194. DOI: 10.1049/el:19961331.
- Sayyouri, Mhamed, Abdeslam Hmimid, and Hassan Qjidaa (2016). “Image analysis using separable discrete moments of Charlier-Hahn”. In: *Multimedia Tools and Applications* 75.1, pp. 547–571. ISSN: 1573-7721. DOI: 10.1007/s11042-014-2307-5. URL: <https://doi.org/10.1007/s11042-014-2307-5>.
- Shao, Zhuhong, Huazhong Shu, Jiasong Wu, Beijing Chen, and Jean Louis Coatrieux (2014). “Quaternion Bessel–Fourier moments and their invariant descriptors for object reconstruction and recognition”. In: *Pattern Recognition* 47.2, pp. 603–611. ISSN: 0031-3203.
- Sheela, K. Gnana and S. N. Deepa (2013). “Review on Methods to Fix Number of Hidden Neurons in Neural Networks”. In: *Mathematical Problems in Engineering* 2013, 11 pages. DOI: doi:10.1155/2013/425740.
- Sheng, Yunlong and Henri H. Arsenault (1986). “Experiments on pattern recognition using invariant Fourier–Mellin descriptors”. In: *J. Opt. Soc. Am. A* 3.6, pp. 771–776. DOI: 10.1364/JOSAA.3.000771. URL: <http://josaa.osa.org/abstract.cfm?URI=josaa-3-6-771>.
- Sheng, Yunlong and Lixin Shen (1994). “Orthogonal Fourier–Mellin moments for invariant pattern recognition”. In: *J. Opt. Soc. Am. A* 11.6, pp. 1748–1757. DOI: 10.1364/JOSAA.11.001748. URL: <http://josaa.osa.org/abstract.cfm?URI=josaa-11-6-1748>.
- Shu, Hua Zhong, Bei Jing Chen, Gang Chen, and Jun Ge (2014). “Color Face Recognition Based on Quaternion Zernike Moment Invariants and Quaternion BP Neural Network”. In: *Advanced Research in Material Science and Mechanical Engineering*. Vol. 446. Applied Mechanics and Materials. Trans Tech Publications, pp. 1034–1039. DOI: 10.4028/www.scientific.net/AMM.446-447.1034.

- Singh, C. and R. Upneja (2012a). "Accuracy and numerical stability of high-order polar harmonic transforms". In: *IET Image Processing* 6.6, pp. 617–626. ISSN: 1751-9659. DOI: 10.1049/iet-ipr.2011.0510.
- Singh, Chandan and Rahul Upneja (2012b). "Accurate Computation of Orthogonal Fourier-Mellin Moments". In: *Journal of Mathematical Imaging and Vision* 44.3, pp. 411–431. ISSN: 1573-7683. DOI: 10.1007/s10851-012-0335-1. URL: <http://dx.doi.org/10.1007/s10851-012-0335-1>.
- Skibbe, Henrik and Marco Reisert (2017). "Spherical Tensor Algebra: A Toolkit for 3D Image Processing". In: *Journal of Mathematical Imaging and Vision* 58.3, pp. 349–381. ISSN: 1573-7683. DOI: 10.1007/s10851-017-0715-7. URL: <https://doi.org/10.1007/s10851-017-0715-7>.
- Sloan, Peter-Pike, Jan Kautz, and John Snyder (2002). "Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments". In: *ACM Transactions on Graphics (TOG)*. Vol. 21. 3. ACM, pp. 527–536.
- Spacek, L. (2007). *Collection of Facial Images: Grimace and Faces95*. <http://cswww.essex.ac.uk/mv/allfaces>.
- Steinkraus, D., I. Buck, and P. Y. Simard (2005). "Using GPUs for machine learning algorithms". In: *Eighth International Conference on Document Analysis and Recognition (ICDAR'05)*, 1115–1120 Vol. 2. DOI: 10.1109/ICDAR.2005.251.
- Subakan, Özlem N. and Baba C. Vemuri (2011). "A Quaternion Framework for Color Image Smoothing and Segmentation". In: *International Journal of Computer Vision* 91.3, pp. 233–250. ISSN: 1573-1405. DOI: 10.1007/s11263-010-0388-9. URL: <https://doi.org/10.1007/s11263-010-0388-9>.
- Suk, Tomáš and Jan Flusser (2009). "Affine Moment Invariants of Color Images". In: *Computer Analysis of Images and Patterns: 13th International Conference, CAIP 2009, Münster, Germany, September 2-4, 2009. Proceedings*. Ed. by Xiaoyi Jiang and Nicolai Petkov. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 334–341. ISBN: 978-3-642-03767-2. DOI: 10.1007/978-3-642-03767-2_41. URL: http://dx.doi.org/10.1007/978-3-642-03767-2_41.
- Suykens, J.A.K. and J. Vandewalle (1999). "Least Squares Support Vector Machine Classifiers". In: *Neural Processing Letters* 9.3, pp. 293–300. ISSN: 1573-773X. DOI: 10.1023/A:1018628609742. URL: <http://dx.doi.org/10.1023/A:1018628609742>.
- Swain, Michael J. and Dana H. Ballard (1991). "Color indexing". In: *International Journal of Computer Vision* 7.1, pp. 11–32. ISSN: 1573-1405. DOI: 10.1007/BF00130487. URL: <https://doi.org/10.1007/BF00130487>.
- Taubin, Gabriel and David B. Cooper (1991). *Recognition and positioning of rigid objects using algebraic moment invariants*. DOI: 10.1117/12.48423. URL: <http://dx.doi.org/10.1117/12.48423>.
- Teague, Michael Reed (1980). "Image analysis via the general theory of moments*". In: *J. Opt. Soc. Am.* 70.8, pp. 920–930. DOI: 10.1364/JOSA.70.000920. URL: <http://www.osapublishing.org/abstract.cfm?URI=josa-70-8-920>.
- Teh, C. H. and R. T. Chin (1988). "On image analysis by the methods of moments". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 10.4, pp. 496–513. ISSN: 0162-8828. DOI: 10.1109/34.3913.
- Toxqui-Quitl, C., L. Gutiérrez-Lazcano, A. Padilla-Vivanco, and C. Camacho-Bello (2011). *Gray-level image reconstruction using Bessel-Fourier moments*. DOI: 10.1117/12.903391. URL: <http://dx.doi.org/10.1117/12.903391>.

- Troje, Nikolaus F. and Heinrich H. Bülthoff (1996). "Face recognition under varying poses: The role of texture and shape". In: *Vision Research* 36.12, pp. 1761–1771. ISSN: 0042-6989. DOI: [http://dx.doi.org/10.1016/0042-6989\(95\)00230-8](http://dx.doi.org/10.1016/0042-6989(95)00230-8). URL: <http://www.sciencedirect.com/science/article/pii/0042698995002308>.
- Tsougenis, E. D., G. A. Papakostas, and D. E. Koulouriotis (2013). "Introducing the separable moments for image watermarking in a Totally moment-oriented framework". In: *2013 18th International Conference on Digital Signal Processing (DSP)*, pp. 1–6. DOI: 10.1109/ICDSP.2013.6622813.
- Tsougenis, E. D., G. A. Papakostas, and D. E. Koulouriotis (2015). "Image watermarking via separable moments". In: *Multimedia Tools and Applications* 74.11, pp. 3985–4012. ISSN: 1573-7721. DOI: 10.1007/s11042-013-1808-y. URL: <https://doi.org/10.1007/s11042-013-1808-y>.
- Ueda, Kazuhiro and Shin-Ichi Takahashi (1996). "Realization of two transfer functions with complex coefficients by a transfer function with hypercomplex coefficients". In: *Electronics and Communications in Japan (Part III: Fundamental Electronic Science)* 79.7, pp. 65–80. ISSN: 1520-6440. DOI: 10.1002/ecjc.4430790707. URL: <http://dx.doi.org/10.1002/ecjc.4430790707>.
- Valle, M. E. and F. Z. de Castro (2017). "On the Dynamics of Hopfield Neural Networks on Unit Quaternions". In: *IEEE Transactions on Neural Networks and Learning Systems* PP.99, pp. 1–8. ISSN: 2162-237X. DOI: 10.1109/TNNLS.2017.2691462.
- Varian, Ioana, Patrick Lambert, and Salim Bettahar (2013). "Un indicateur de fausses couleurs pour évaluer le traitement des images couleurs". In: *24ème colloque GRETSI*. Brest, France, non définies. URL: <https://hal.archives-ouvertes.fr/hal-00875045>.
- Wang, Chun peng, Xing yuan Wang, Zhi qiu Xia, Chuan Zhang, and Xing jun Chen (2016a). "Geometrically resilient color image zero-watermarking algorithm based on quaternion Exponent moments". In: *Journal of Visual Communication and Image Representation* 41, pp. 247–259. ISSN: 1047-3203. DOI: <https://doi.org/10.1016/j.jvcir.2016.10.004>. URL: <http://www.sciencedirect.com/science/article/pii/S1047320316302103>.
- Wang, Wen-Hao and Yung-Chang Chen (1997). "Image registration by control points pairing using the invariant properties of line segments". In: *Pattern Recognition Letters* 18.3, pp. 269–281. ISSN: 0167-8655. DOI: [http://doi.org/10.1016/S0167-8655\(97\)00010-X](http://doi.org/10.1016/S0167-8655(97)00010-X). URL: <http://www.sciencedirect.com/science/article/pii/S016786559700010X>.
- Wang, Xiang yang, Chun peng Wang, Hong ying Yang, and Pan pan Niu (2013). "A robust blind color image watermarking in quaternion Fourier transform domain". In: *Journal of Systems and Software* 86.2, pp. 255–277. ISSN: 0164-1212. DOI: <http://doi.org/10.1016/j.jss.2012.08.015>. URL: <http://www.sciencedirect.com/science/article/pii/S0164121212002312>.
- Wang, Xiang yang, Pan pan Niu, Hong ying Yang, Chun peng Wang, and Ai long Wang (2014a). "A new robust color image watermarking using local quaternion exponent moments". In: *Information Sciences* 277, pp. 731–754. ISSN: 0020-0255. DOI: <https://doi.org/10.1016/j.ins.2014.02.158>. URL: <http://www.sciencedirect.com/science/article/pii/S002002551400293X>.
- Wang, Xiang yang, Yu nan Liu, Meng meng Han, and Hong ying Yang (2016b). "Local quaternion {PHT} based robust color image watermarking algorithm". In: *Journal of Visual Communication and Image Representation* 38, pp. 678–694. ISSN: 1047-3203.

- DOI: <https://doi.org/10.1016/j.jvcir.2016.04.011>. URL: <http://www.sciencedirect.com/science/article/pii/S1047320316300463>.
- Wang, Xiang-Yang, Bei-Bei Zhang, and Hong-Ying Yang (2014b). "Content-based image retrieval by integrating color and texture features". In: *Multimedia Tools and Applications* 68.3, pp. 545–569. ISSN: 1573-7721. DOI: 10.1007/s11042-012-1055-7. URL: <http://dx.doi.org/10.1007/s11042-012-1055-7>.
- Wang, Xiang-Yang, Zhi-Fang Wu, Liang Chen, Hong-Liang Zheng, and Hong-Ying Yang (2016c). "Pixel classification based color image segmentation using quaternion exponent moments". In: *Neural Networks* 74, pp. 1–13. ISSN: 0893-6080. DOI: <https://doi.org/10.1016/j.neunet.2015.10.012>. URL: <http://www.sciencedirect.com/science/article/pii/S0893608015002129>.
- Wang, Xiang-yang, Yu-nan Liu, Huan Xu, Pei Wang, and Hong-ying Yang (2016d). "Robust copy-move forgery detection using quaternion exponent moments". In: *Pattern Analysis and Applications*, pp. 1–17. ISSN: 1433-755X. DOI: 10.1007/s10044-016-0588-1. URL: <http://dx.doi.org/10.1007/s10044-016-0588-1>.
- Wang, Xuan, Tengfei Yang, and Fangxia Guo (2016e). "Image analysis by circularly semi-orthogonal moments". In: *Pattern Recognition* 49, pp. 226–236. ISSN: 0031-3203. DOI: <https://doi.org/10.1016/j.patcog.2015.08.005>. URL: <http://www.sciencedirect.com/science/article/pii/S0031320315002939>.
- Wei, Z. and W. Weijiang (2010). "Color Watermarking Algorithm Based on Modified QFFT and DCT". In: *2010 International Conference on Optoelectronics and Image Processing*. Vol. 1, pp. 328–331. DOI: 10.1109/ICOIP.2010.315.
- Werbos, P. (1974). "Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences". PhD thesis. Cambridge, MA: Harvard University.
- Wilamowski, B. M. and H. Yu (2010). "Neural Network Learning Without Backpropagation". In: *IEEE Transactions on Neural Networks* 21.11, pp. 1793–1803. ISSN: 1045-9227. DOI: 10.1109/TNN.2010.2073482.
- Wolf, L., T. Hassner, and I. Maoz (2011). "Face recognition in unconstrained videos with matched background similarity". In: *CVPR 2011*, pp. 529–534. DOI: 10.1109/CVPR.2011.5995566.
- Xiang-yang, Wang, Li Wei-yi, Yang Hong-ying, Niu Pan-pan, and Li Yong-wei (2015a). "Invariant quaternion radial harmonic Fourier moments for color image retrieval". In: *Optics and Laser Technology* 66.0, pp. 78–88. ISSN: 0030-3992.
- Xiang-yang, Wang, Li Wei-yi, Yang Hong-ying, Wang Pei, and Li Yong-wei (2015b). "Quaternion polar complex exponential transform for invariant color image description". In: *Applied Mathematics and Computation* 256, pp. 951–967. ISSN: 0096-3003. DOI: <https://doi.org/10.1016/j.amc.2015.01.075>. URL: <http://www.sciencedirect.com/science/article/pii/S0096300315001071>.
- Xiao, Bin, Jian-Feng Ma, and Xuan Wang (2010). "Image analysis by Bessel-Fourier moments". In: *Pattern Recognition* 43.8, pp. 2620–2629. ISSN: 0031-3203.
- Xiao, Bin, Wei sheng Li, and Guo yin Wang (2015). "Errata and comments on "Orthogonal moments based on exponent functions: Exponent-Fourier moments"". In: *Pattern Recognition* 48.4, pp. 1571–1573. ISSN: 0031-3203. DOI: <https://doi.org/10.1016/j.patcog.2014.10.030>. URL: <http://www.sciencedirect.com/science/article/pii/S0031320314004476>.
- Yang, Hong ying, Yan Zhang, Pei Wang, Xiang yang Wang, and Chun peng Wang (2014). "A geometric correction based robust color image watermarking scheme using quaternion

- Exponent moments". In: *Optik - International Journal for Light and Electron Optics* 125.16, pp. 4456–4469. ISSN: 0030-4026. DOI: <https://doi.org/10.1016/j.ijleo.2014.02.028>. URL: <http://www.sciencedirect.com/science/article/pii/S0030402614003908>.
- Yang, Hong-Ying, Xiang-Yang Wang, Pan-Pan Niu, and Ai-Long Wang (2015). "Robust Color Image Watermarking Using Geometric Invariant Quaternion Polar Harmonic Transform". In: *ACM Trans. Multimedia Comput. Commun. Appl.* 11.3, 40:1–40:26. ISSN: 1551-6857. DOI: 10.1145/2700299. URL: <http://doi.acm.org/10.1145/2700299>.
- Yao, Xin (1993). "A review of evolutionary artificial neural networks". In: *International Journal of Intelligent Systems* 8.4, pp. 539–567. ISSN: 1098-111X. DOI: 10.1002/int.4550080406. URL: <http://dx.doi.org/10.1002/int.4550080406>.
- Yap, P. T. and R. Paramesran (2006). "Content-based image retrieval using Legendre chromaticity distribution moments". In: *IEE Proceedings - Vision, Image and Signal Processing* 153.1, pp. 17–24. ISSN: 1350-245X. DOI: 10.1049/ip-vis:20045064.
- Yap, P. T., X. Jiang, and A. Chichung Kot (2010). "Two-Dimensional Polar Harmonic Transforms for Invariant Image Representation". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32.7, pp. 1259–1270. ISSN: 0162-8828. DOI: 10.1109/TPAMI.2009.119.
- Young, Thomas (1802). "The Bakerian Lecture: On the Theory of Light and Colours". In: *Philosophical Transactions of the Royal Society of London* 92, pp. 12–48. ISSN: 02610523. URL: <http://www.jstor.org/stable/107113>.
- Yu, Kai, Tong Zhang, and Yihong Gong (2009). "Nonlinear Learning using Local Coordinate Coding". In: *Advances in Neural Information Processing Systems* 22. Ed. by Y. Bengio, D. Schuurmans, J. D. Lafferty, C. K. I. Williams, and A. Culotta. Curran Associates, Inc., pp. 2223–2231. URL: <http://papers.nips.cc/paper/3875-nonlinear-learning-using-local-coordinate-coding.pdf>.
- Zeng, R., J. Wu, Z. Shao, L. Senhadji, and H. Shu (2014). "Quaternion softmax classifier". In: *Electronics Letters* 50.25, pp. 1929–1931. ISSN: 0013-5194. DOI: 10.1049/el.2014.2526.
- Zeng, Rui, Jiasong Wu, Zhuhong Shao, Yang Chen, Beijing Chen, Lotfi Senhadji, and Huazhong Shu (2016). "Color image classification via quaternion principal component analysis network". In: *Neurocomputing* 216.Supplement C, pp. 416–428. ISSN: 0925-2312. DOI: <https://doi.org/10.1016/j.neucom.2016.08.006>. URL: <http://www.sciencedirect.com/science/article/pii/S092523121630827X>.
- Zhang, Dengsheng and Guojun Lu (2001). "A Comparison of Shape Retrieval Using Fourier Descriptors and Short-Time Fourier Descriptors". In: *Advances in Multimedia Information Processing — PCM 2001: Second IEEE Pacific Rim Conference on Multimedia Beijing, China, October 24–26, 2001 Proceedings*. Ed. by Heung-Yeung Shum, Mark Liao, and Shih-Fu Chang. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 855–860. ISBN: 978-3-540-45453-3. DOI: 10.1007/3-540-45453-5_111. URL: http://dx.doi.org/10.1007/3-540-45453-5_111.
- Zhang, Dengsheng and Guojun Lu (2002). "A comparative Study of Fourier Descriptors for Shape Representation and Retrieval". In: *Proc. of 5th Asian Conference on Computer Vision (ACCV)*. Springer, pp. 646–651.
- Zhang, Dengsheng and Guojun Lu (2004). "Review of shape representation and description techniques". In: *Pattern Recognition* 37.1, pp. 1–19. ISSN: 0031-3203.

- Zhang, H., H.Z. Shu, P. Haigron, B.S. Li, and L.M. Luo (2010). “Construction of a complete set of orthogonal Fourier–Mellin moment invariants for pattern recognition applications”. In: *Image and Vision Computing* 28.1, pp. 38–44. ISSN: 0262-8856. DOI: <http://doi.org/10.1016/j.imavis.2009.04.004>. URL: <http://www.sciencedirect.com/science/article/pii/S0262885609000730>.
- Zhou, B. and J. Wang (2017). “Generation of a realistic 3D sand assembly using X-ray micro-computed tomography and spherical harmonic-based principal component analysis”. In: *International Journal for Numerical and Analytical Methods in Geomechanics* 41.1. NAG-15-0295.R1, pp. 93–109. ISSN: 1096-9853. DOI: 10.1002/nag.2548. URL: <http://dx.doi.org/10.1002/nag.2548>.
- Zhu, Hongqing (2012). “Image representation using separable two-dimensional continuous and discrete orthogonal moments”. In: *Pattern Recognition* 45.4, pp. 1540–1558. ISSN: 0031-3203. DOI: <https://doi.org/10.1016/j.patcog.2011.10.002>. URL: <http://www.sciencedirect.com/science/article/pii/S0031320311004055>.

Résumé détaillé en Français

Contributions à la description d'images couleur pour la recherche d'images par le contenu

Introduction

Ces dernières années, nous assistons à une quantité énorme d'images numériques qui sont utilisées dans un large éventail de domaines d'application tels que la médecine, la vision robotique, l'astronomie, l'automatisation industrielle, l'environnement, l'artisanat, le trafic militaire et routier, etc. Cette quantité excessive d'images, exploitée dans une multitude de domaines, requiert des systèmes de recherche efficaces. Ceux-ci permettent de retrouver des images pertinentes à partir d'une grande base en réponse à une requête utilisateur exprimée sous forme textuelle (méta-données) ou d'une image exemple.

Pour cela, des systèmes de recherche d'images par le contenu (*Content-based image retrieval* (CBIR)) ont été proposés et largement utilisés dans de nombreux domaines, notamment la prévention des crimes, la conception architecturale, la conception de mode, le diagnostic médical, les systèmes de télédétection, le patrimoine culturel et la recherche sur le Web.

Généralement, un système de recherche d'images par le contenu consiste en deux étapes: (1) Une étape d'indexation, hors ligne, où les contenus visuels des images sont extraits sous forme de vecteurs caractéristiques pour construire la base des descripteurs. (2) Une étape de recherche, en ligne, au cours de laquelle un utilisateur fournit au système une image exemple. Le système transforme cette image requête en un vecteur descripteur. Ensuite, il calcule les mesures de similarité entre ce descripteur vectoriel et ceux des images stockées dans la base de données. Par conséquent, le système renvoie les images les mieux classées en fonction de cette mesure de similarité. Plusieurs facteurs influent la performance d'un système CBIR, nous citons par exemple, le descripteur employé, la mesure de similarité, la structure d'indexation, la technique de recherche, etc.

Cependant, l'extraction des caractéristiques d'images est considérée l'un des problèmes les plus importants dans la CBIR. Trois attributs de bas niveau sont généralement utilisés pour l'extraction des caractéristiques d'images: la texture, la forme et la couleur.

- **Texture:** Une représentation de texture est utilisée pour définir les motifs visuels dans les images en fonction du voisinage. Cette représentation est utile pour décrire le contenu de différentes images du monde réel telles que les empreintes digitales, les briques, les arbres et les tissus. Parmi les descripteurs de texture destinés à la CBIR, nous citons les travaux de Wang et al., 2014b et Park et al., 2002.

- **Forme:** Une représentation de forme pourrait être utilisée pour identifier l'objet contenu dans une image. Dans ce cas, la robustesse aux transformations photométriques (flou, bruit, couleur d'éclairage, direction d'illumination, etc.) et l'invariance aux transformations géométriques (translation, changement d'échelle, rotation, etc.) sont nécessaires pour une meilleure reconnaissance de l'objet. Selon la revue des techniques de description de formes établie par Zhang and Lu, 2004, il existe deux classes de méthodes pour la description des formes:

(1) les méthodes basées-contour, qui utilisent uniquement l'information sur le contour pour décrire la forme,

(2) les méthodes basées sur la région, qui exploitent la totalité de la forme, c'est-à-dire tous les pixels contenus dans la forme.

Chacune de ces deux classes est constituée de deux approches: l'approche structurale, où la forme est représentée par des segments, et l'approche globale, où la forme est représentée dans son ensemble.

Il est généralement admis que les descripteurs de formes globaux basés-région sont les plus utiles car ils sont moins affectés par le bruit et les faibles variations contrairement aux méthodes basées-contour. Les méthodes basées-région incluent les moments orthogonaux, qui ont des propriétés très importantes permettant de capturer le maximum d'information sur la forme.

- **Couleur:** La couleur, qui peut être perçue sous forme d'une combinaison de trois composants, est un attribut très important pour la description d'une image. Plusieurs travaux ont étudié la description à base de la couleur, nous citons à titre d'exemples les travaux de Swain and Ballard, 1991; Pass et al., 1996; Albuz et al., 2001; Paschos et al., 2003.

De nombreux travaux ont porté sur la construction de nouveaux descripteurs, l'amélioration de la précision et du temps de calcul, les propriétés d'invariance, etc. Dans ce travail, nous nous intéressons spécifiquement aux descripteurs d'images basés sur l'information forme et couleur. Cette dernière peut être représentée suivant différentes approches et formalismes mathématiques, par exemple des approches marginales et vectorielles.

En fait, de nombreux travaux de traitement d'images couleur impliquent l'adaptation des algorithmes de traitement d'images en niveaux de gris pour les images couleur. Cette approche, souvent appelée marginale, consiste à appliquer un traitement séparé sur chacun des trois composants chromatiques puis à combiner les résultats pour produire la sortie finale. Par conséquent, l'information couleur n'est pas considérée globalement, mais plutôt répartie en trois composants couleur indépendamment. Cette stratégie pose des problèmes tels que la perte de corrélation entre les canaux couleur et dans la plupart des applications, elle conduit à l'apparition de fausses couleurs.

Une alternative à l'approche marginale consiste à considérer la couleur comme une information vectorielle. Dans ce cas, l'information couleur est vue comme un vecteur tridimensionnel se référant à une coordonnée dans un espace colorimétrique choisi. Bien que la représentation des couleurs au moyen de vecteurs permette de traiter globalement l'information couleur, certains problèmes persistent. Par exemple, il est difficile de trier des vecteurs couleur dans l'espace tridimensionnel dans des applications telles que le filtrage médian. De plus, il est nécessaire de redéfinir des traitements vectoriels complexes pour les adapter aux vecteurs couleur. D'une manière générale, bien que la méthode vectorielle permette de traiter globalement l'information couleur, elle est désavantageuse dans la mesure où elle induit des traitements plus complexes par rapport à l'approche marginale.

Dans ce contexte, nous nous intéressons à l'opportunité d'encoder l'information couleur en utilisant des représentations algébriques. Ces représentations utilisent l'information couleur vectorielle comme un nombre de l'algèbre utilisée. Ceci nous permettra de tirer parti des opérations standards fournies uniquement par l'algèbre, au lieu des opérations vectorielles, qui sont très complexes à mettre en œuvre. Par exemple, l'algèbre des nombres complexes est largement utilisée dans une variété d'applications de traitement de signal et d'images. Précisément, les moments disque-harmoniques (DHMs), qui nécessitent l'utilisation des nombres complexes, est un descripteur de forme basé-région qui a prouvé ses performances pour la recherche d'images par le contenu. En effet, les moments disque-harmoniques ont été proposés par notre équipe de recherche (Ennahahi et al., 2010) en s'inspirant de la transformée en harmonique sphérique (SHT). Cette dernière se caractérise par la complétude et l'orthogonalité sur la sphère unitaire. L'objectif initial de notre travail est de proposer une extension de ces moments disque-harmoniques conventionnels pour traiter les images couleur. Dans cette optique, nous nous intéressons à exploiter le formalisme des quaternions associé aux images couleur (Sangwine, 1996). En fait, les quaternions, qui sont une généralisation des nombres complexes, peuvent capturer l'information couleur. De plus, les images couleur peuvent être manipulées au moyen d'opérations simples définies dans l'algèbre des quaternions. Comme un nombre quaternion se compose d'une partie réelle et de trois parties imaginaires, un pixel couleur peut être exprimé par un quaternion pur où la partie réelle est mise à zéro et les trois parties imaginaires abritent les composants couleur.

Dans ce travail de thèse, nous proposons différents descripteurs quaternaires pour la recherche d'images par le contenu. La performance de l'ensemble des descripteurs proposés est évaluée en termes de capacité de reconstruction d'images couleur, de robustesse aux changements photométriques, d'invariance aux transformations géométriques, de pouvoir discriminatoire et de temps de calcul.

Ce rapport de thèse se compose d'une introduction générale, ainsi que de quatre chapitres, et d'une conclusion:

L'**introduction générale** souligne l'importance du sujet traité et le situe dans le contexte de la CBIR puis elle décrit les motivations et les objectifs des travaux de recherche élaborés.

Le **chapitre 1** de notre rapport de thèse est consacré à une revue concernant les descripteurs d'images couleur existants et leurs applications notamment pour la recherche d'images par le contenu. Ce chapitre présente également les avantages et les limitations de chaque famille de descripteurs.

Dans le **chapitre 2**, nous présentons notre première contribution qui consiste à proposer un descripteur quaternaire, appelé (*quaternion disc-harmonic moments*), inspiré des harmoniques sphériques (SH) pour la recherche d'images couleur par le contenu. Ce chapitre présente d'abord des préliminaires sur les SH et leurs diverses applications. Ensuite, il détaille les différentes étapes de calcul du descripteur quaternaire invariant proposé. La représentation de l'image à trois composants couleur par l'algèbre des quaternions s'avère intéressante puisqu'elle garantit à la fois la conservation de l'information couleur et la corrélation entre les trois canaux chromatiques composant l'image. De plus, cette contribution améliore le temps de calcul du descripteur proposé en introduisant la technique Healpix, du laboratoire JPL de la NASA. En fin, ce chapitre dresse les résultats obtenus des différentes évaluations du descripteur proposé. Ces évaluations comprennent la capacité de reconstruction d'images couleur, l'invariance aux changements géométriques, la robustesse aux transformations photométriques, le temps de calcul ainsi que le pouvoir discriminatoire. Pour la recherche des k-plus-proches voisins, nous étudions l'effet de la mesure de similarité, de l'espace couleur ainsi que celui de

l'ordre maximal.

Le **chapitre 3** décrit notre deuxième contribution qui consiste à introduire une reformulation sans paramètre des moments orthogonaux quaternaires, appelée *parameter-free quaternary orthogonal moments* (PFQOM), dont les invariants à base du module sont évalués dans le contexte de la recherche et la classification d'images couleur. Les résultats expérimentaux obtenus montrent l'efficacité des descripteurs proposés.

Le **chapitre 4** décrit notre troisième contribution qui consiste à introduire une méthode de classification d'images couleur basée sur la combinaison des moments quaternaires disques-harmoniques et d'un classificateur de type quaternaire, qui est la machine d'apprentissage extrême quaternaire (QELM). Cette méthode est évaluée en utilisant différentes collections d'images couleur. Les résultats obtenus montrent que la méthode proposée assure une bonne précision en un temps raisonnable.

La **conclusion** récapitule les contributions réalisées durant ce travail de thèse et discute les différentes ouvertures et **perspectives** possibles qui peuvent faire objet de futurs travaux de recherche connexes.

1 Etat de l'Art sur la Description d'Images Couleur

Les descripteurs d'images couleur peuvent être classifiés selon trois catégories: (1) Les méthodes basées sur l'histogramme, (2) les méthodes marginales et (3) les méthodes algébriques.

1.1 Méthodes Basées sur l'Histogramme Couleur

Pour la description d'images, les méthodes basées sur l'histogramme couleur peuvent être classifiées en deux catégories: (1) Les méthodes qui manipulent l'image directement à partir de l'espace colorimétrique en utilisant des conversions adéquates entre ces espaces. (2) Les méthodes qui agissent sur les espaces de chromaticité afin d'éliminer l'impact du composant représentant la luminance.

Dans la catégorie des méthodes basées sur l'espace colorimétrique, l'histogramme couleur est le plus populaire. Celui-ci est utilisé dans les systèmes de recherche d'images par le contenu tels que QBIC (Flickner et al., 1995) et Chabot (Ogle and Stonebraker, 1995). L'histogramme est obtenu à partir de l'espace colorimétrique en le divisant en régions uniformes selon un niveau de quantification Q et les comptages de pixels sont affectés aux classes d'histogrammes associées. Plusieurs espaces colorimétriques, tels RGB, HSV et XYZ, peuvent être utilisés. Par exemple, Swain and Ballard, 1991 ont proposé d'exploiter l'histogramme couleur extrait à partir des trois composants de l'espace colorimétrique *opponent* pour la recherche d'images. Celle-ci est basée sur l'intersection d'histogrammes comme mesure de similarité. Bien que les histogrammes couleur réussissent à extraire l'aspect couleur de l'image et soient robustes aux petites transformations, ils ignorent l'information spatiale.

Afin de surmonter cette limitation, Pass et al., 1996 ont introduit des vecteurs de cohérence des couleurs (CCV), ou *color coherence vectors*, comme étant une méthode de raffinement de l'histogramme couleur en incorporant l'information spatiale. Une extension améliorée de ces vecteurs de cohérence, où le centre de l'image est utilisé comme caractéristique supplémentaire, a été introduite par les mêmes auteurs (Pass and Zabih, 1996). Ensuite, Han and Ma, 2002 ont introduit la notion d'histogrammes couleur flous (FCH), ou *fuzzy color histograms*, pour atténuer le problème de la sensibilité des histogrammes aux changements d'illumination et aux erreurs de quantification.

D'autres travaux basés sur l'espace colorimétrique ont été proposés, incluant les corrélogrammes couleur (Huang et al., 1997), les vecteurs de cohérence des couleurs (Pass et al., 1996), les histogrammes couleur flous (Han and Ma, 2002), etc.

D'autre part, la chromaticité représente la qualité d'une couleur sans comprendre sa luminance. Certains espaces colorimétriques, tels que xyY , $YCbCr$, YIQ , YUV et YCC , séparent les trois composants couleur en un composant de luminance et une paire de composants chromatiques. Les descripteurs d'images basés sur l'espace de chromaticité incluent les moments de chromaticité (Paschos et al., 2003) et les moments de distribution de chromaticité de Legendre (Yap and Paramesran, 2006).

1.2 Méthodes Marginales

Pour décrire une image couleur, une solution intuitive consiste à convertir cette image en niveaux de gris, puis à calculer son descripteur conventionnel. Cependant, cette approche peut entraîner une perte significative d'information couleur. Une alternative est de calculer séparément le descripteur conventionnel (technique de niveaux de gris) pour chaque canal couleur de l'image et de combiner les résultats individuels. Cependant, cette direction, appelée méthode marginale, ne réussit pas à conserver la corrélation inhérente entre les canaux couleur.

Notons que les moments sont des grandeurs scalaires utilisées pour décrire une image et pour capturer ses caractéristiques discriminantes. Parmi les descripteurs de formes basés-région (conventionnel de niveaux de gris), les moments géométriques et les moments orthogonaux sont les plus connus.

Les moments géométriques ont été introduits par Hu, 1962 pour la reconnaissance des formes bidimensionnelles, puis largement appliqués grâce à leurs propriétés intéressantes (Hu, 1962; Teh and Chin, 1988; Taubin and Cooper, 1991; Belkasim et al., 1991; Prokop and Reeves, 1992; Flusser and Suk, 1993; Mehtre et al., 1997). Par exemple, les moments d'ordres zéro représentent l'énergie moyenne de l'image, les moments du premier ordre localisent le centre de l'image et les moments du second ordre caractérisent la taille et l'orientation de l'image.

Mindru et al., 2004 ont adopté l'approche marginale basée sur les moments géométriques afin de proposer un ensemble de descripteurs d'images couleur invariants aux transformations géométriques et aux changements photométriques (GPDI). De plus, Suk and Flusser, 2009 ont introduit la notion du centroïde commun pour dériver des moments centraux invariants à la translation.

Teague, 1980 a montré que le calcul des moments géométriques d'ordre élevé est très complexe car il repose sur des puissances standards qui croissent rapidement avec l'ordre. Il a également été démontré que la reconstruction de l'image à partir de ces moments d'ordres élevés ne reste pas fidèle à l'image originale à cause de la redondance. D'autre part, Zhang and Lu, 2004 a prouvé que les invariants des moments géométriques d'ordre réduit sont appropriés pour décrire des formes simples sans bruit (MPEG-7) et ne sont donc pas suffisants pour décrire avec précision les objets réels.

Pour pallier ces inconvénients, les moments orthogonaux ont été proposés. En effet, Teague, 1980 a proposé de remplacer le noyau des moments géométriques, c'est-à-dire les puissances standards, par des fonctions de base complètes et orthogonales. C'est le cas des moments de Zernike (Teague, 1980), des moments orthogonaux de Fourier-Mellin (Sheng and Shen, 1994), des moments de Bessel-Fourier (Xiao et al., 2010), des moments disque-harmoniques (Ennahahi et al., 2010) et des moments séparables (Zhu, 2012; Tsougenis et al., 2013; Tsougenis et al., 2015; Abdeslam Hmimid, 2014; Hmimid et al., 2015; Sayyouri et al., 2016).

Les moments orthogonaux sont utilisés dans plusieurs applications de vision par ordinateur et

de traitement d'images, y compris le tatouage numérique (Kim and Lee, 2003; Gao and Jiang, 2015), la reconnaissance et la classification d'objets (Khotanzad and Hong, 1990; Abdallah et al., 1997; Zhang et al., 2010; Ma et al., 2011; Nagarajan and Devendran, 2012), la reconstruction d'images (Toxqui-Quitl et al., 2011) ainsi que la correspondance des modèles (Beijing et al., 2011).

1.3 Méthodes Algébriques

Une autre alternative pour représenter l'image couleur exploite les algèbres des quaternions (Hamilton, 1844) et Clifford (Hestenes and Sobczyk, 1987). Hamilton a introduit l'algèbre des quaternions, qui est une généralisation des nombres complexes. En fait, un quaternion consiste en une partie réelle et trois parties imaginaires. Notons que l'algèbre de Clifford est plus générale et englobe celle des quaternions.

Dans la littérature, les moments conventionnels traitant les images binaires ou en niveaux de gris ont été généralisés pour la description des images couleur. Cette technique permet de coder les trois canaux de couleur dans un seul vecteur caractéristique sans passer par une transformation en niveaux de gris, ce qui permet également de prendre en considération la corrélation entre ces trois composants.

Initialement, Sangwine, 1996 a introduit l'analyse d'images couleur par l'algèbre quaternaire. Il a proposé la version quaternaire de la transformée de Fourier discrète/rapide (QDFT / QFFT), qui rend possible la manipulation d'images couleur dans le domaine fréquentiel, sans séparation des composants couleur. L'auteur a proposé d'exprimer un pixel couleur par un quaternion pur où la partie réelle est mise à zéro et les trois parties imaginaires représentent chacune des trois canaux couleur. Par exemple, un pixel couleur RVB (rouge, vert, bleu) $f(x, y)$ peut être représenté comme suit:

$$f(x, y) = f_R(x, y)i + f_V(x, y)j + f_B(x, y)k, \quad (\text{R. 1})$$

où $f_R(x, y)$, $f_V(x, y)$ et $f_B(x, y)$ représentent respectivement le canal rouge, vert et bleu de l'image couleur à la position (x, y) .

La version quaternaire de la transformée de Fourier discrète/rapide (QDFT/QFFT) est une extension par l'algèbre des quaternions de la DFT/FFT classique. Elle a été employée dans plusieurs applications, à savoir le recalage d'images (Feng et al., 2008; Pan et al., 2010), le tatouage numérique (Bas et al., 2003; Wang et al., 2013; Chen et al., 2014b; Ouyang et al., 2015), la correction d'images (Grigoryan et al., 2015) et la fusion d'images médicales couleur (Nawaz et al., 2016).

Des descripteurs quaternaires, autres que les transformées de Fourier quaternaires (QFT), ont été proposés. Nous citons, par exemple, les moments de Fourier-Mellin quaternaires (QFMM) (Guo and Zhu, 2011) qui ont été utilisés dans plusieurs applications sur les images couleur. Cependant, les QFT et QFMM ne sont pas basés sur des fonctions orthogonales et souffrent d'une forte présence de redondance d'information. Pour pallier cet inconvénient, plusieurs moments orthogonaux quaternaires ont été suggérés. Nous citons à titre d'exemples, le quaternion des moments de Zernike (QZM) (Chen et al., 2010; Chen et al., 2012), des moments de pseudo-Zernike (QPZM) (Bei-Jing et al., 2012), des moments de Bessel-Fourier (QBFM) (Shao et al., 2014), des moments harmoniques radiaux de Fourier (QRHFM) (Xiang-yang et al., 2015a) et le quaternion de la transformée polaire harmonique (QPHT) (Li, 2013; Xiang-yang et al., 2015b). Quant aux moments orthogonaux quaternaires de Zernike (QZM), de pseudo-Zernike (QPZM)

et de Fourier-Mellin (QOFMM), ils souffrent de l'instabilité numérique ainsi que d'une complexité algorithmique élevée lorsque l'ordre des moments devient grand. Ces limitations peuvent s'expliquer par le fait que le calcul de ces moments se base sur des termes factoriels (Xiangyang et al., 2015a).

Puisque l'algèbre des quaternions est contenue dans l'algèbre de Clifford, cette dernière constitue une autre direction pour représenter les images couleur. Par exemple, Batard et al., 2010 ont introduit la transformée de Clifford-Fourier (CFT) pour la description des images couleur et ont prouvé que la transformée de Fourier quaternaire (QFT) de Sangwine, 1996 n'est qu'un cas particulier de la CFT.

2 Moments Quaternaires Proposés pour la Description d'Images Couleur

Dans notre première contribution, nous proposons un ensemble de moments quaternaires inspirés de la transformée en harmoniques sphériques (SHT) et des moments disque-harmoniques (DHM) (Ennahnahi et al., 2010) pour la description d'images couleur.

Dans le chapitre 2 de notre rapport de thèse, nous avons d'abord mis en lumière les nombreux avantages et propriétés de la transformée en harmoniques sphériques (SHT), par exemple, l'invariance à la rotation, l'orthogonalité ainsi que la rapidité de calcul. La SHT a été exploitée dans un large éventail d'applications relatives à différentes disciplines, y compris la physique, la chimie et la vision par ordinateur. En fait, la SHT a été une source d'inspiration pour de nombreux chercheurs (Sloan et al., 2002; Green, 2003; Kazhdan et al., 2003; Funkhouser et al., 2003; Pu and Ramani, 2006; Sajjanhar et al., 2004; Sajjanhart et al., 2005). En particulier, Ennahnahi et al., 2010 ont introduit les moments disque-harmoniques (DHM) comme une adaptation de la transformée en harmoniques sphériques pour la description des formes 2D en adoptant une transformation sphère-disque appropriée.

2.1 Préliminaires

2.1.1 Opérations sur les Quaternions

Un quaternion (Hamilton, 1844) est une généralisation d'un nombre complexe. Il se compose d'une partie réelle et de trois parties imaginaires comme $q = q^a + q^b i + q^c j + q^d k$. En analyse d'image couleur, une image RVB $f(x, y)$ est représentée par un quaternion pur comme dans Eq. (R. 1):

Les parties imaginaires i, j et k vérifient les combinaisons: $i^2 = j^2 = k^2 = -1, ij = k, jk = i, ki = j, ji = -k, kj = -i, ik = -j$.

Le conjugué et le module d'un quaternion q sont définis, respectivement, par :

$$q^* = q^a - q^b i - q^c j - q^d k, \quad (\text{R. 2})$$

$$||q|| = \sqrt{(q^a)^2 + (q^b)^2 + (q^c)^2 + (q^d)^2}. \quad (\text{R. 3})$$

Le produit scalaire et le produit croisé entre deux quaternions q_1 et q_2 sont définis, respectivement, par:

$$q_1 \bullet q_2 = q_1^a q_2^a + q_1^b q_2^b + q_1^c q_2^c + q_1^d q_2^d, \quad (\text{R. 4})$$

$$\begin{aligned} q_1 \otimes q_2 = & (q_1^a q_2^a - q_1^b q_2^b - q_1^c q_2^c - q_1^d q_2^d) + (q_1^a q_2^b + q_1^b q_2^a + q_1^c q_2^d - q_1^d q_2^c) i \\ & + (q_1^a q_2^c - q_1^b q_2^d + q_1^c q_2^a + q_1^d q_2^b) j + (q_1^a q_2^d + q_1^b q_2^c - q_1^c q_2^b + q_1^d q_2^a) k. \end{aligned} \quad (\text{R. 5})$$

Notons que le produit croisé des quaternions $\otimes$ n'est pas commutatif.

2.1.2 Moments Disque-Harmoniques

Les moments disque-harmoniques (DHM) (Ennahnahi et al., 2010) sont une adaptation de la transformée en harmoniques sphériques pour la description des formes 2D en utilisant le passage sphère-disque suivant:

$$\sin(\theta) = \sqrt{1 - r^2} \quad , \quad \cos(\theta) = r \quad , \quad \varphi_{sphere} = \varphi_{disc}, \quad (R. 6)$$

où $(\theta, \varphi_{sphere})$ représente les coordonnées sphériques et (r, φ_{disque}) les coordonnées polaires, tel que $0 \leq \varphi \leq 2\pi$, $0 \leq \theta \leq \pi/2$ et $0 \leq r \leq 1$. Sur la sphère, θ varie entre 0 et π , néanmoins, comme r doit être positif, la mise en correspondance avec le disque unitaire a été limitée seulement à l'hémisphère nord ($0 \leq \theta \leq \pi/2$) comme indiqué sur Figure 1.

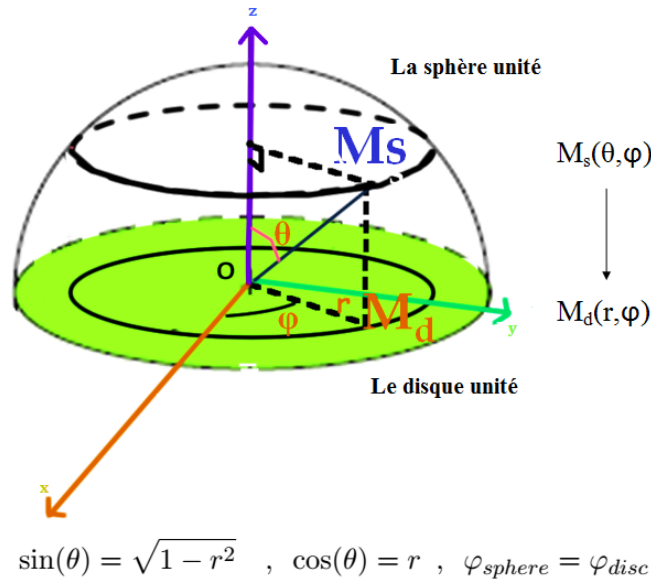


Figure 1 Passage sphère-disque adopté par Ennahnahi et al., 2010.

Les moments disque-harmoniques complexes d'ordre n et de répétition m sont exprimés pour une image en niveaux de gris $f_C(r, \varphi)$ comme suit:

$$CDHM_{n,m}(f_C) = \int_0^{2\pi} \int_0^1 f_C(r, \varphi) N_{n,m} P_{n,m}(r) [e^{jm\varphi}]^* r dr d\varphi, \quad |m| \leq n, \quad (R. 7)$$

où X^* est le conjugué du nombre complexe X , $P_{n,m}$ est le polynôme de Legendre associé et $N_{n,m}$ est le coefficient de normalisation donné par:

$$N_{n,m} = \sqrt{\frac{(2n+1)(n-m)!}{4\pi(n+m)!}}. \quad (R. 8)$$

Selon Ennahnahi et al., 2010, les modules de ces moments complexes ont permis d'obtenir un descripteur invariant, robuste et facile à calculer.

Malgré tous les avantages des moments DHM, ces derniers ne peuvent manipuler que les images binaires ou en niveaux de gris vu qu'ils sont à valeur complexe. Afin d'incorporer l'information couleur, nous étendons les moments disque-harmoniques en utilisant l'algèbre des quaternions.

2.2 Moments Disque-Harmoniques Quaternaires Proposés

Dans cette contribution, nous proposons (cf. 2.2.1) une version de base qui consiste à étendre les moments disque-harmoniques complexes pour les images couleur en exploitant l'algèbre des quaternions. Ensuite (cf. 2.2.2), nous apportons une rectification relative au calcul des moments quaternaires proposés en adoptant une nouvelle transformation sphère-disque. En fin (cf. 2.2.3), nous présentons notre amélioration concernant le temps de calcul.

2.2.1 Version Quaternaire de Base

En exploitant l'algèbre des quaternions, nous formulons une généralisation des moments disque-harmoniques complexes pour tenir compte des images couleur. En effet, le noyau $e^{jm\varphi}$ dans Eq. (R. 7) est remplacé par $e^{\mu m\varphi}$, comme suit:

$$QDHM_{n,m}(f) = \int_0^{2\pi} \int_0^1 f(r, \varphi) \otimes N_{n,m} P_{n,m}(r) [e^{\mu m\varphi}]^* r dr d\varphi, \quad \text{with } |m| \leq n, \quad (\text{R. 9})$$

où $f(r, \varphi)$ est le pixel couleur représenté par Eq. (R. 1),

Q^* est le conjugué du quaternion Q et $\otimes$ est le produit quaternaire croisé, donnés respectivement par Eq. (R. 2) et Eq. (R. 5),

μ est un quaternion unitaire pur, dépourvu de sa partie réelle, que nous fixons à $\mu = (i + j + k)/\sqrt{3}$.

Le produit quaternaire croisé permet de combiner les valeurs R, V et B d'un pixel de sorte à préserver leurs corrélations. Cependant, le calcul direct de l'Eq. (R. 9) engendre un nombre élevé d'opérations de multiplication. Pour remédier à ce problème, nous avons élaboré une relation mathématique entre les moments quaternaires QDHM et les moments complexes DHM relatifs à chacune des trois canaux RVB, comme suit:

$$QDHM_{n,m}(f) = A_{n,m} + iB_{n,m} + jC_{n,m} + kD_{n,m}, \quad (\text{R. 10})$$

where

$$\begin{aligned} A_{n,m} &= -\frac{1}{\sqrt{3}} [\Im(CDH M_{n,m}(f_R)) + \Im(CDH M_{n,m}(f_V)) + \Im(CDH M_{n,m}(f_B))], \\ B_{n,m} &= \Re(CDH M_{n,m}(f_R)) + \frac{1}{\sqrt{3}} [\Im(CDH M_{n,m}(f_V)) - \Im(CDH M_{n,m}(f_B))], \\ C_{n,m} &= \Re(CDH M_{n,m}(f_V)) + \frac{1}{\sqrt{3}} [\Im(CDH M_{n,m}(f_B)) - \Im(CDH M_{n,m}(f_R))], \\ D_{n,m} &= \Re(CDH M_{n,m}(f_B)) + \frac{1}{\sqrt{3}} [\Im(CDH M_{n,m}(f_R)) - \Im(CDH M_{n,m}(f_V))]. \end{aligned} \quad (\text{R. 11})$$

où $CDH M_{n,m}(f_R)$, $CDH M_{n,m}(f_V)$ et $CDH M_{n,m}(f_B)$ représentent respectivement les DHM complexes [Eq. (R. 7)] pour le composant rouge, vert et bleu. $\Re(X)$ et $\Im(X)$ sont respectivement la partie réelle et la partie imaginaire du nombre complexe X .

Les équations (R. 11) sont simples en termes de calcul et ne comportent que peu d'opérations de multiplication par rapport à Eq. (R. 9).

2.2.2 Moments Disque-Harmoniques Quaternaires Rectifiés

Bien que les formules de passage sphère-disque dans Eq. (R. 6) permettent des performances satisfaisantes des QDHM en termes de précision de recherche d'images, elles souffrent d'un

problème d'irrégularité du pas d'échantillonnage suivant le rayon r (qui varie de manière sinusoïdale). Pour cela, nous proposons la modification suivante au niveau de ces formules:

$$\theta = \frac{\pi}{2}r, \quad \varphi_{sphere} = \varphi_{disc}, \quad (\text{R. 12})$$

où $0 \leq \theta \leq \pi/2$ et $0 \leq r \leq 1$.

Par conséquent, les équations (R. 7) et (R. 9) permettant le calcul des DHM et QDHM sont rectifiées comme suit :

$$CDHM_{n,m}(f_C) = \int_0^{2\pi} \int_0^1 f_C(r, \varphi) N_{n,m} P_{n,m}(\cos(\frac{\pi}{2}r)) [e^{jm\varphi}]^* r dr d\varphi, \quad |m| \leq n, \quad (\text{R. 13})$$

$$QDHM_{n,m}(f) = \int_0^{2\pi} \int_0^1 f(r, \varphi) \otimes N_{n,m} P_{n,m}(\cos(\frac{\pi}{2}r)) [e^{jm\varphi}]^* r dr d\varphi, \quad \text{with } |m| \leq n. \quad (\text{R. 14})$$

La figure 2 montre la modification apportée sur la régularité de la distribution des points d'échantillonnage. En effet, les points d'échantillonnage obtenus par Eq. (R. 12) sont distribués de manière uniforme (Figure 2-a) contrairement à ceux représentés dans Figure 2-b, qui sont générés par Eq. (R. 6).

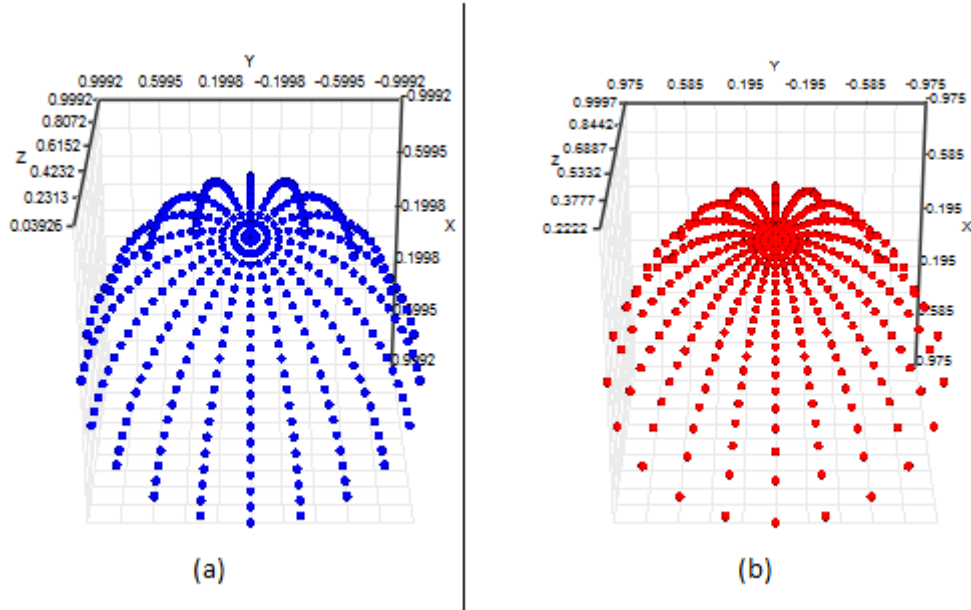


Figure 2 Distribution des pixels sur l'hémisphère obtenue par (a) le nouveau passage sphère-disque, qui considère des pas uniformes pour φ et $r = \frac{2 \times \theta}{\pi}$, et (b) l'ancien passage avec des pas uniformes pour φ et $r = \cos(\theta)$.

La figure 3 montre l'amélioration de la qualité de reconstruction de l'image par rapport à la version de base proposée des moments QDHM.

Rappelons que le calcul des moments quaternaires QDHM rectifiés est obtenu à base des équations (R. 11), (R. 13) et (R. 14), en d'autres termes à base des moments DHM complexes.

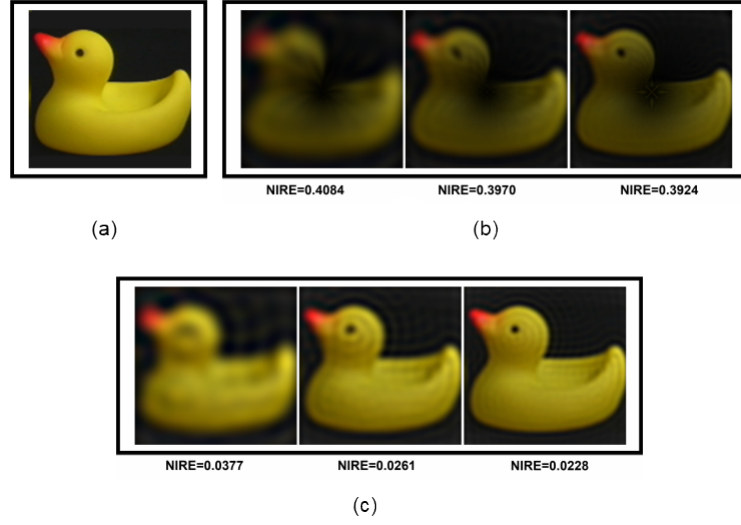


Figure 3 Reconstruction d'images couleur à partir des QDHM de (a) l'image originale pour les ordres 20, 40 et 60 en utilisant (b) l'ancien passage sphère-disque et (c) le nouveau passage.

L'approximation numérique de l'intégration pour le calcul des DHM, donnée par Eq. (R. 13), est effectuée sur une grille d'échantillonnage régulière ($N \times N$) tel que $r_i = (2i + 1)/2N$ et $\varphi_j = 2\pi j/N$ où $i = 0, \dots, N - 1$ et $j = 0, \dots, N - 1$. Ainsi, nous obtenons une formulation de calcul rapide des moments DHMs complexes qui se présente comme suit:

$$\begin{aligned}
 CDHM_{n,m}(f_C) &= \frac{\sqrt{2\pi}N_{n,m}}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} w_i P_{n,m}(\cos(\frac{\pi}{2}r_i)) f_C(r_i, \varphi_j) e^{-jm\varphi_j}, \\
 &= \frac{\sqrt{2\pi}N_{n,m}}{N} \left(\sum_{i=0}^{N-1} w_i P_{n,m}(\cos(\frac{\pi}{2}r_i)) \left(\sum_{j=0}^{N-1} f_C(r_i, \varphi_j) e^{-jm\varphi_j} \right) \right). \quad (R. 15)
 \end{aligned}$$

Les poids de quadrature w_i sont ceux de Chebyshev-Gauss (Abramowitz, 1974).

L'équation (R. 15) se base sur la séparation des variables et le passage par la transformée de Fourier rapide. Ainsi, l'évaluation des polynômes de Legendre associés $P_{n,m}$ n'est nécessaire que pour N valeurs du rayon r .

Il a été démontré par Driscoll and Healy, 1994 que le calcul des harmoniques sphériques à base de Eq. (R. 15) a une complexité algorithmique de l'ordre de $N(\log N)^2$, contrairement à l'algorithme naïf qui est de l'ordre quadratique (N^2), où N est le nombre de points échantillonnés.

2.2.3 Version Rapide Basée-Healpix

Pour améliorer davantage le temps de calcul des moments QDHM proposés, nous adaptons la technique Healpix élaborée par le laboratoire JPL (*Jet Propulsion Laboratory*) de la NASA.

La technique Healpix (*Hierarchical Equal Area Iso-Latitude Pixelation*) (Górski et al., 2005) a été proposée pour l'analyse rapide et précise des fonctions sphériques en les transformant à base des harmoniques sphériques. En effet, Healpix repose sur une technique d'échantillonnage particulière permettant d'obtenir des partitions d'égale aire et des poids de quadrature égaux

pour l'intégration numérique. Figure 4 illustre cette technique d'échantillonnage qui consiste à effectuer une première division de la sphère en 12 larges pixels. Ceux-ci, sont ensuite subdivisés selon la résolution N_{side} en question. Ainsi, une carte de $N_{pix} = 12 \times N_{side}^2$ pixels de résolution N_{side} est générée.

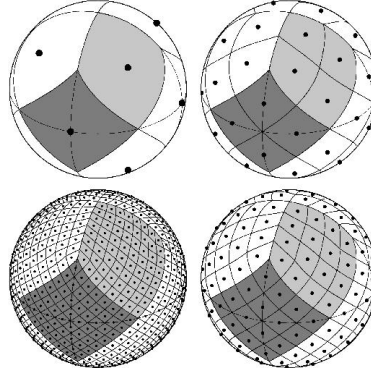


Figure 4 Échantillonnage Healpix de la sphère pour $N_{side} = 1, 2, 4, 8$ (Górski et al., 2005).

Une carte Healpix est construite pour aboutir à des partitions à égale aire où les pixels sont alignés suivant des anneaux iso-latitude. Ainsi, le problème de la forte densité des points au voisinage des pôles est résolu. La répartition sur des anneaux iso-latitude garantit le calcul rapide des harmoniques sphériques à travers la transformée de Fourier rapide.

Nous calculons notre descripteur QDHM au moyen de la technique d'échantillonnage Healpix. Cela nous permet de réduire considérablement la complexité des calculs qui passe à $N_{pix}^{3/2}$.

2.3 Résultats Expérimentaux

Nous évaluons les performances des moments quaternaires proposés en termes de capacité de reconstruction d'images, d'invariance et robustesse, de temps de calcul ainsi que de pouvoir discriminatoire. Notons que les invariants de nos moments à la translation sont mathématiquement établis en localisant l'origine du système de coordonnées au centre commun des trois canaux (Suk and Flusser, 2009), tandis que l'invariance à la rotation est obtenue en retenant le module des moments.

Notre descripteur QDHM est comparé à d'autres descripteurs quaternaires: QZM (Moments de Zernike), QPZM (Moments de pseudo-Zernike), QOFMM (Moments Fourier-Mellin Orthogonaux) et QBFM (Moments Bessel-Fourier), ainsi qu'à des descripteurs basés sur Clifford: poFMD et CFMDi. Les bases d'images standard COIL-100, ALOI-VIEW, ALOI-COL et ALOI-ILL sont utilisées dans les expérimentations.

2.3.1 Reconstruction d'Images Couleur

La figure 5 présente l'erreur moyenne de reconstruction d'un ensemble de 10 images COIL-100 en fonction de l'ordre des différents moments orthogonaux quaternaires. Nous remarquons que lorsque l'ordre de reconstruction augmente, l'erreur de reconstruction (NIRE) relative aux moments QDHM proposés diminue. Cette diminution est plus rapide que celles des autres moments. Ceci s'explique par le fait que les QDHM, même dans le cas des faibles ordres, permettent de capturer une grande partie de l'information contenue dans l'image.

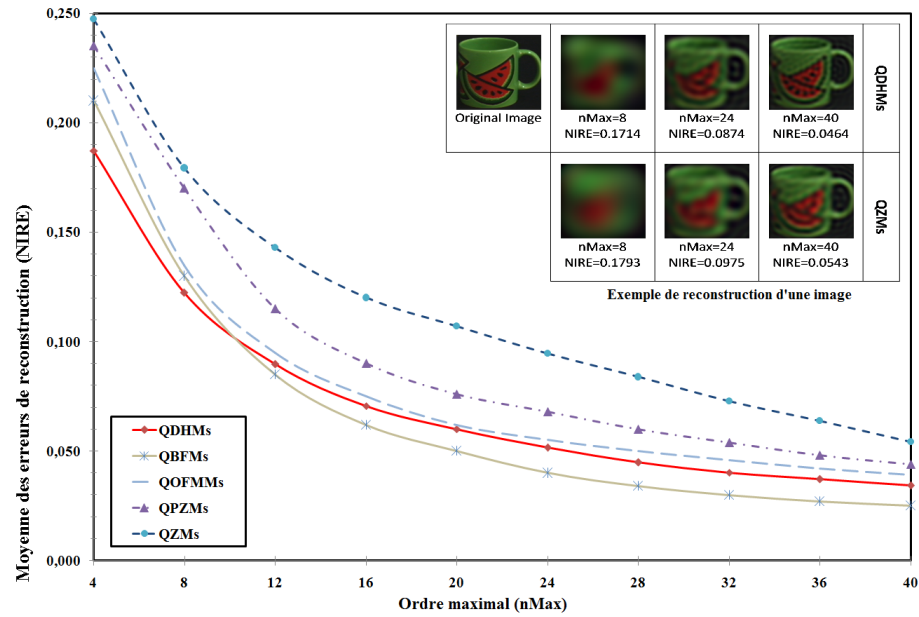


Figure 5 Erreur moyenne de reconstruction (NIRE) d'un ensemble de 10 images COIL-100 en fonction de l'ordre des différents moments quaternaires.

2.3.2 Robustesse et Invariance

Nous évaluons notre descripteur en termes de la robustesse au bruit sel et poivre, au flou gaussien, au changement de la couleur d'illumination et de la direction d'éclairage ainsi qu'à son invariance à la rotation. La figure 6 illustre les résultats obtenus concernant la robustesse au bruit sel et poivre ainsi qu'au flou gaussien.

Nous remarquons que notre descripteur QDHM est le moins sensible aux bruits et flous comparé aux autres descripteurs quaternaires.

Les détails sur les expérimentations menées ainsi que les résultats obtenus sont présentés à la section 2.4.3 du chapitre 2.

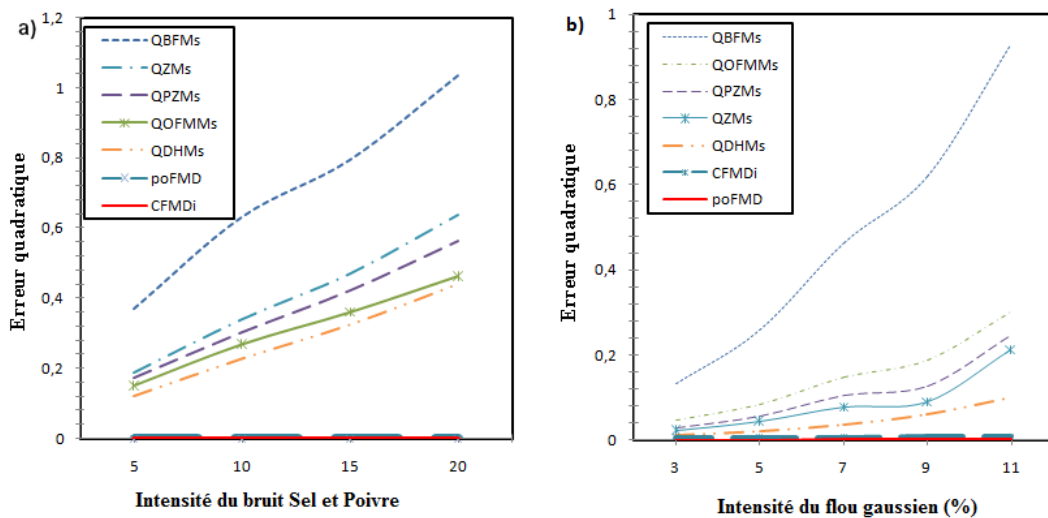


Figure 6 Erreur quadratique des descripteurs d'images affectées par (a) le bruit sel et poivre et (b) par du flou gaussien

2.3.3 Temps de Calcul

La figure 7 montre le temps en secondes effectué pour le calcul des différents descripteurs en fonction du nombre des moments. Nous pouvons remarquer que le temps de calcul des descripteurs Clifford poFMD et du CFMDi n'est pas proportionnel à l'ordre, et il est bien inférieur aux temps de calcul des moments quaternaires. D'un autre côté, le temps nécessaire pour le calcul des QDHM est inférieur aux QOFMM, QZM, QPZM et est comparable aux QBFM.

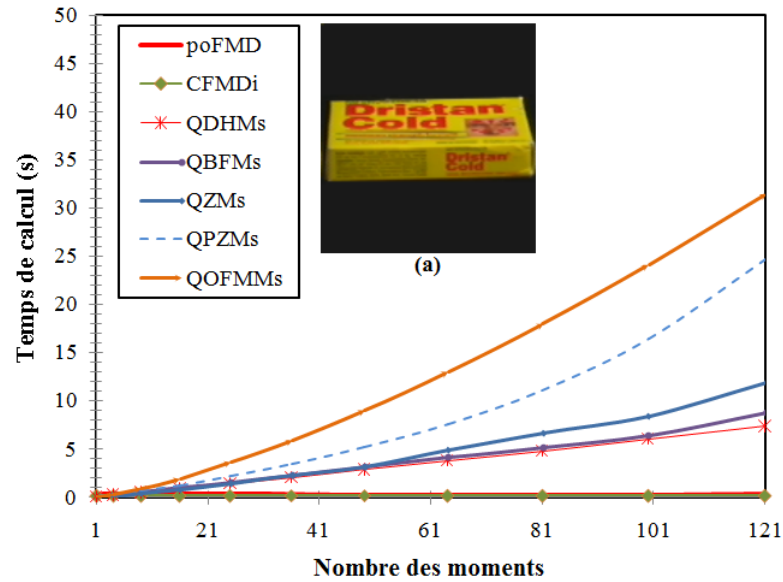


Figure 7 Temps moyen en secondes en fonction du nombre des moments.

De plus, le tableau 1 montre clairement que la version de notre descripteur QDHM basée sur Healpix est environ 12 fois plus rapide que la méthode de base. Ce résultat concerne $nMax = 8$, $N_{side} = 32$ pour QDHM-Healpix et des images de taille 128×128 pour les QDHM de base.

Tableau 1 Temps moyen en seconds pour le calcul de QDHMs-Healpix et différents moments quaternaires (La taille des images est 128×128)

QDHMs-Healpix	QDHMs	QBFMs	QZMs	QPZMs
0.0937s	1.1278s	1.3125s	1.6852s	2.7856s

2.3.4 Recherche d'Images

Pour tester notre descripteur QDHM en termes de la qualité de recherche, nous avons implémenté un système de recherche d'images basé sur les k-plus-proches voisins. Nous fixons empiriquement la valeur de l'ordre maximal $nMax$ à 8, la mesure de similarité à Chi-Carré et l'espace colorimétrique à TSV. Ces valeurs assurent une meilleure précision. La figure 8 montre le rapport rappel-précision des différents descripteurs. Il est démontré que le descripteur QDHM permet un meilleur pouvoir discriminatoire par rapport au descripteur DHM complexe. Ceci est expliqué par sa capacité de la prise en compte de l'information couleur. De plus, notre descripteur QDHM est plus performant que les autres descripteurs quaternaires et Clifford.

Les détails sur les expérimentations menées ainsi que les résultats obtenus sont présentés à la section 2.4.5 du chapitre 2.

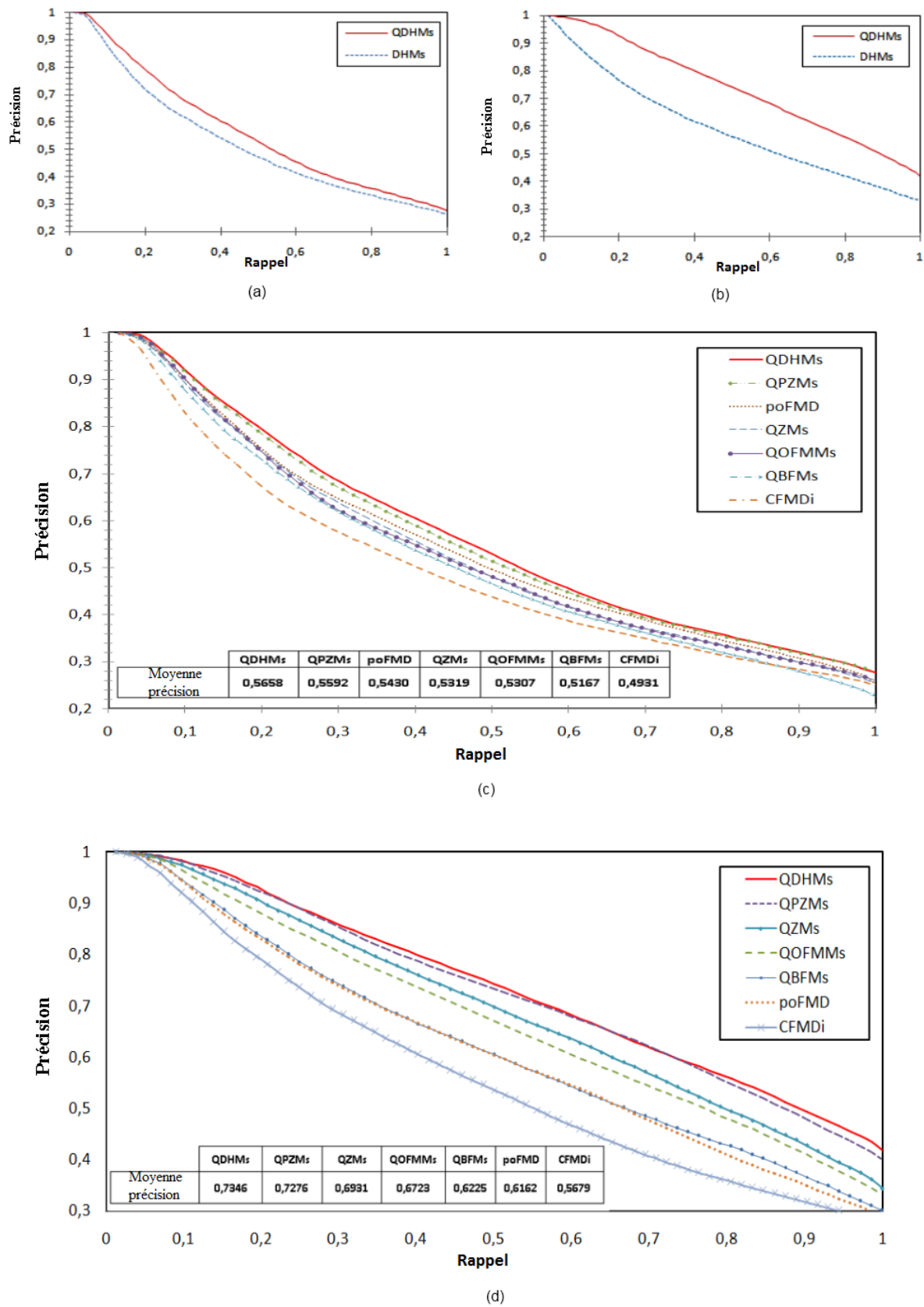


Figure 8 Courbes rappel-précision des descripteurs (a) DHM et QDHM pour la base COIL-100, (b) DHM et QDHM pour la base ALOI-VIEW, (c) moments quaternaires et Clifford pour la base COIL-100 and (d) moments quaternaires et Clifford pour la base ALOI-VIEW.

3 Reformulation Sans Paramètre des Moments Quaternaires

Généralement, le calcul des moments quaternaires existants dépend d'un paramètre quaternion unitaire pur μ dont la valeur optimale ne peut être fixée qu'empiriquement. De plus, la valeur estimée de ce paramètre est fortement liée au domaine d'application. Inspirés par le travail de Gough, 1984 sur les quaternions et les harmoniques sphériques, nous proposons une formulation simple des moments orthogonaux quaternaires, libre de tout paramètre, donc caractérisée par sa simplicité et sa généralité.

3.1 Préliminaires

3.1.1 Harmoniques Sphériques et Quaternions

Gough, 1984 a abordé les applications des quaternions dans la mécanique quantique tridimensionnelle. A cet effet, il a considéré les harmoniques sphériques, et en a établi certaines combinaisons linéaires avec des coefficients quaternaires. Il a également prouvé que ces harmoniques sphériques quaternaires ont des propriétés très intéressantes telles que la complétude et l'orthogonalité.

Les harmoniques sphériques quaternaires (QSHs), comme leur nom l'indique, fonctionnent sur la sphère et sont définies comme suit:

$$B_{n,m}^{q\text{-QSH}} = \sqrt{\frac{(2n+1)(n-m)!}{4\pi(n+m)!}} P_{n,m}(\cos(\theta)) e^{km\varphi} k^{n-m} i^m, \quad (\text{R. 16})$$

où θ indique la co-latitude, φ la longitude et $P_{n,m}$ est le polynôme de Legendre associé.

Il est à noter que $e^{km\varphi}$, plutôt que $e^{im\varphi}$, doit être utilisé dans cette expression.

3.1.2 Moments Orthogonaux Quaternaires

Dans cette deuxième contribution, nous nous intéressons aux moments quaternaires à fonctions de base orthogonales pour la description des images couleur. Les avantages de ces moments résident dans le fait qu'ils permettent de préserver le maximum d'information sur la forme et la couleur.

Les moments quaternaires d'ordre n et de répétition m d'une image couleur $f(r, \varphi)$, définie en coordonnées polaires, sont formulés comme suit:

$$QOM_{n,m}(f) = \int_0^{2\pi} \int_0^1 f(r, \varphi) \otimes (B_{n,m}^\mu)^* r dr d\varphi, \quad (\text{R. 17})$$

où $\otimes$ est le produit croisé du quaternion et $*$ est le conjugué du quaternion, donnés respectivement par Eq. (R. 5) et Eq. (R. 2). $[0, 1] \times [0, 2\pi]$ fait référence au disque unité et $B_{n,m}^\mu$ est la fonction de base exprimée par:

$$B_{n,m}^\mu = R_{n,m}(r) e^{\mu m \varphi}, \quad (\text{R. 18})$$

où R sont les polynômes radiaux réels correspondant aux moments calculés et $e^{\mu m \varphi}$ est le noyau quaternionaire où μ est un paramètre quaternion unitaire pur, tel que $\mu = ai + bj + ck$ et $\|\mu\| = 1$.

3.2 Moments Orthogonaux Quaternaires Sans Paramètre Proposés

Les moments quaternaires existants sont calculés en utilisant des fonctions de base qui dépendent du choix d'un paramètre quaternion unitaire pur μ comme indiqué par Eq. (R. 18). Rappelons que le choix de μ est purement empirique et ne repose sur aucune base mathématique.

D'autre part, Gough, 1984 a réussi à construire mathématiquement un ensemble d'harmoniques sphériques avec des coefficients quaternaires (QSH) pour une application en mécanique quantique tridimensionnelle. La formulation de ces QSH est simple car elle consiste en l'introduction de facteurs quaternaires dans les harmoniques sphériques complexes.

Inspirés par le travail de Gough, nous proposons un ensemble de moments quaternaires sans paramètres pour les images couleur. En effet, la partie angulaire qui dépend de μ , $e^{\mu m \varphi}$, dans la définition des moments quaternaires classiques [Eq. (R. 18)] est remplacée par une partie angulaire sans paramètres $e^{km\varphi} k^{nm} i^m$. Cette partie est ensuite multipliée par les polynômes radiaux orthogonaux à valeur réelle d'autres familles de moments définies sur le disque 2D.

Il en résulte une formulation simple de la fonction de base des moments quaternaires, libre de tout paramètre comme suit:

$$B_{n,m}^q = R_{n,m}(r) e^{km\varphi} k^{n-m} i^m. \quad (\text{R. 19})$$

Évidemment, nous aurions aussi pu définir $B_{n,m}^q$ comme étant $R_{n,m}(r) i^m k^{nm} e^{km\varphi}$ vu que la multiplication des quaternions n'est pas commutative. Nous notons que, dans ce travail, nous avons adopté l'expression dans Eq. (R. 19). Cette expression est libre de tout paramètre, contrairement à la fonction de base utilisée par les moments quaternaires de l'état de l'art donnée par Eq. (R. 18).

Les moments quaternaires sans paramètre proposés d'ordre n et de répétition m , de l'image couleur $f(r, \varphi)$ [Eq. (R. 1)], définie en coordonnées polaires, peuvent alors être calculés comme suit:

$$PFQOM_{n,m}(f) = \frac{\sqrt{2\pi}}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(r, \varphi) \otimes (B_{n,m}^q)^*. \quad (\text{R. 20})$$

L'invariance à la rotation est obtenue en retenant le module, comme suit:

$$PFQOMI = ||PFQOM||. \quad (\text{R. 21})$$

3.3 Résultats expérimentaux

Dans cette deuxième contribution, nous montrons d'abord les capacités de reconstruction d'images couleur des moments proposés. Ensuite, nous évaluons leurs robustesse et invariance. De plus, nous montrons que la formulation sans paramètre des moments quaternaires permet un calcul plus rapide par rapport à leurs versions basées sur μ . Les détails des expérimentations menées sont décrits dans la section 3.3 du chapitre 3.

Dans un deuxième temps, nous évaluons les performances des moments proposés dans deux applications: la recherche d'images couleur par le contenu et la classification d'images couleur. Deux collections d'images couleur sont utilisées pour les expérimentations: (1) La base d'objets Columbia (COIL-100) (Nene et al., 1996) et (2) la collection de visages Faces96 (Spacek, 2007).

Nous comparons les performances des moments sans paramètre proposés avec leurs versions basées sur μ , les moments de chromaticité (CM) (Paschos et al., 2003), les moments de distribution de chromaticité de Legendre (LCDM) (Yap and Paramesran, 2006) et le descripteur basé Clifford poFMD (Mennesson et al., 2014). Notons que les CM sont calculés pour deux niveaux de quantification $Q = 10, 20$ et le poFMD est calculé pour le bi-vecteur rouge B_{rouge} .

Les moments quaternaires que nous abordons sont basés sur les polynômes radiaux de Zernike, les polynômes de Bessel du 1^{er} type et les polynômes associés de Legendre. Pour les moments quaternaires basés sur μ , nous fixons μ à $(i - j - k)/\sqrt{3}$ comme établi par Chen et al., 2014a.

3.3.1 Recherche d'Images

Nous rapportons dans le tableau 2 les résultats de la recherche d'images par le contenu en utilisant le premier plus proche voisin (1-NN).

Tableau 2 Précisions en % obtenues par les différents moments d'images couleur en utilisant le premier plus proche voisin (1-NN) dans les deux collections COIL-100 et Faces96.

	PFQDHM	QDHM	PFQZM	QZM	PFQBFM	QBFM	poFMD	LCDM	CM ₁₀	CM ₅₀
COIL-100	100	99.91	99.91	99.89	99.46	99.38	99.56	96.51	92.96	86.38
Faces96	99.12	99.09	98.94	98.84	94.62	94.43	97.70	86.89	71.22	65.21

D'après ce tableau, nous remarquons que les moments sans paramètres proposés ont des pouvoirs discriminatoires acceptables et comparables à leurs versions basées sur μ . Ceci est expliqué par le fait que les moments orthogonaux quaternaires sont basés sur des polynômes orthogonaux et profitent du produit croisé quaternaire. Par conséquent, ils peuvent capturer efficacement l'information couleur et forme contenue dans une image. Nous remarquons aussi que les moments quaternaires basés sur les polynômes associés de Legendre, PFQDHM et QDHM, ainsi que les polynômes radiaux de Zernike assurent une meilleure précision par rapport aux moments basés sur Bessel Fourier du 1^{er} type, particulièrement lorsqu'il s'agit d'images de visages capturées sur un fond complexe. Le poFMD et le LCDM atteignent également des résultats de recherche satisfaisants. En revanche, les CM sont les moins performants.

Les détails sur les expérimentations menées ainsi que les résultats obtenus sont présentés à la section 3.3.4 du chapitre 3.

3.3.2 Classification d'images

Nous évaluons les performances des PFQOMI proposés dans le contexte de la classification d'images couleur en se basant sur la machine d'apprentissage extrême ou *extreme learning machine* (ELM) (Ding et al., 2015).

Dans un premier temps, la base d'images est partitionnée en deux parties: une base d'apprentissage (70%) et une base de test (30%). Pour chacune des bases, nous calculons les descripteurs d'images couleur jusqu'à l'ordre $nMax = 5$ et les normalisons en utilisant la méthode Min-Max. Ensuite, nous passons les descripteurs de la base d'apprentissage comme entrée au classificateur ELM pour apprendre ses paramètres optimaux du modèle. Les paramètres appris sont utilisés pour la classification des images de la base de test.

Les résultats obtenus sont rapportés dans le tableau 3.

Les détails sur les expérimentations menées ainsi que les résultats obtenus sont présentés à la section 3.3.5 du chapitre 3.

Les résultats obtenus sur les deux collections COIL-100 et Faces96 montrent que: (1) Les moments sans paramètre proposés assurent des précisions de classification plus satisfaisants que leur version basée sur μ . (2) Les PFQDHM et les QDHM sont les plus performants et comparables aux PFQZM lorsqu'ils s'agit d'images de visages capturées sur un fond complexe. (3) Les moments orthogonaux quaternaires possèdent la propriété d'orthogonalité et sont capables de

Tableau 3 Précisions de classification en % obtenues par les différents moments d'images couleur en utilisant une machine d'apprentissage extrême avec une couche cachée à 600 nœuds lors du traitement de COIL-100 et à 280 nœuds lors du traitement de Faces96.

	PFQDHM	QDHM	PFQBFM	QBFM	PFQZM	QZM	poFMD	LCDM	CM ₅₀	CM ₁₀
COIL-100	93.55	92.57	89.76	89.21	88.09	84.72	89.74	22.52	17.40	17.77
Faces96	93.92	94.61	83.34	82.5	93.6	90.6	89.01	74.74	64.34	55.58

capturer à la fois l'information de forme et de couleur. Ainsi, ils permettent tous des précisions significatives (≥ 84.72 pour COIL-100 et ≥ 82.5 pour Faces96).

4 Notre Méthode de Classification d'Images Couleur

Dans cette troisième contribution, nous suggérons une méthode de classification d'images couleur basée sur les moments QDHM sans paramètre. L'objectif est de confirmer la pertinence de notre descripteur quaternaire et de le valider dans le cadre d'un système de classification d'images.

4.1 Revue de la Littérature sur les Algorithmes de Classification

Le chapitre 4 de notre rapport de thèse établit un état-de-l'art concernant les principales méthodes de classification d'images. Ces méthodes peuvent être classifiées en deux catégories: les méthodes superficielles et les méthodes d'apprentissage en profondeur.

Les méthodes superficielles consistent en: (1) l'extraction manuelle des caractéristiques de l'image via un vecteur descripteur permettant de réduire la dimension et d'assurer l'invariance aux différentes variations et transformations d'images, puis (2) la classification de ces descripteurs. Ces derniers sont introduits dans un classificateur qui détermine la classe à laquelle appartient ce modèle ou cette image. Ce type de méthode connaît un grand succès dans les domaines de la reconnaissance d'objets, la reconnaissance faciale et la classification de texture. D'autre part, les méthodes d'apprentissage en profondeur s'appuient davantage sur l'apprentissage automatique en adoptant une architecture à plusieurs niveaux. En fait, ces méthodes traitent directement les images plutôt que les vecteurs caractéristiques, ce qui permet à la fois l'extraction des caractéristiques et la classification.

De nombreux chercheurs ont étudié l'extension de ces méthodes (les deux catégories) pour traiter les signaux multidimensionnels, tels que les images couleur, en utilisant l'algèbre des quaternions.

D'après notre étude bibliographique, nous avons tiré les principales différences entre les méthodes d'apprentissage en profondeur et les méthodes superficielles, résumées comme suit:

- Les méthodes d'apprentissage en profondeur nécessitent des méga-données pour extraire de manière non-supervisée les caractéristiques de l'image.
- Les méthodes d'apprentissage en profondeur sont limitées par la disponibilité des ressources informatiques, car le traitement d'images à plus haute résolution nécessite des couches plus grandes et nombreuses.
- Les méthodes superficielles sont basées sur une extraction de descripteurs à partir des calculs mathématiques. Cela signifie que l'invariance à la rotation, au changement d'échelle et à la translation dans l'étape d'extraction des caractéristiques de l'image est mathématiquement obtenue par les descripteurs d'images.

4.2 Notre Méthode pour la Classification d'Images

Dans cette contribution, nous proposons une méthode superficielle pour la classification d'images couleur. Cette méthode tire partie de notre descripteur quaternaire sans paramètre. D'abord, nous procédons à l'adaptation des algorithmes de classification superficiels pour prendre en entrée des vecteurs caractéristiques quaternaires. Rappelons que les descripteurs d'images couleur quaternaires permettent de maintenir la corrélation entre les trois canaux de couleur grâce au produit quaternaire croisé.

Pour la classification d'images couleur à base de vecteurs descripteurs quaternaires, trois situations, illustrées dans Figure 9, se présentent:

- La situation (1) consiste à introduire les invariants des moments quaternaires basés sur le module (valeur réelle) dans un classificateur superficiel à valeur réelle. En revanche, obtenir l'invariance à la rotation en retenant le module des moments quaternaires entraîne la perte de l'information de phase et fournit seulement un invariant à valeur réelle.
- La situation (2), appelée aussi approche marginale, consiste à introduire séparément les quatre parties des invariants des moments à valeur quaternaire dans un classificateur à valeur réelle. L'inconvénient de cette situation est que la taille de l'entrée du classificateur est quatre fois plus grande, ce qui pénalise le temps de calcul du classificateur. De plus, la corrélation entre les quatre parties du vecteur caractéristique quaternaire n'est pas maintenue.
- La situation (3) consiste à ce que l'entrée du classificateur quaternaire soit constituée des invariants à valeur quaternaire formant ainsi une seule entité. Cette solution permet de conserver les informations de phase du vecteur caractéristique quaternaire ainsi que la corrélation entre les quatre parties.

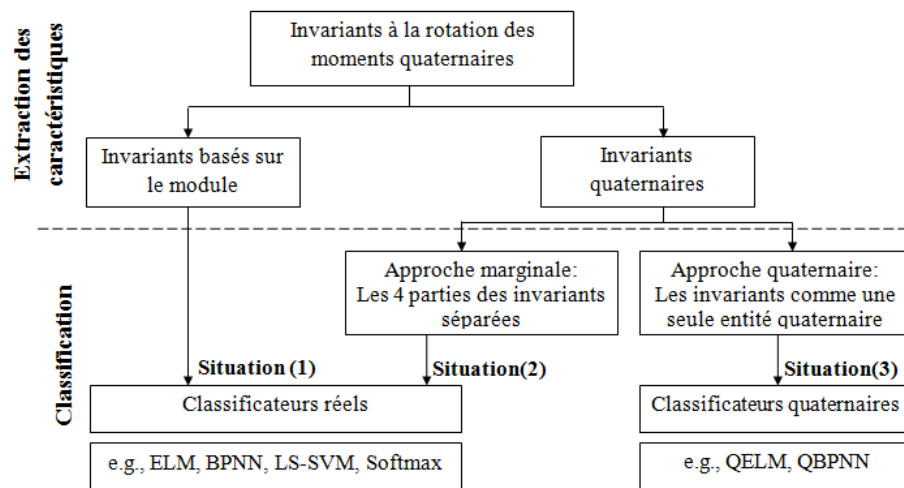


Figure 9 Les différents invariants à la rotation des moments quaternaires et les différentes approches pour les classifier.

Dans ce travail, nous adoptons la situation (3) pour la classification d'images couleur. L'idée de base est de garder la forme quaternaire des coefficients composant le descripteur QDHM au lieu de se restreindre aux modules des moments. Pour maintenir la corrélation entre les quatre parties, réelle et imaginaires, des coefficients du descripteur quaternaire, nous nous sommes basés sur un classificateur de type quaternaire, qui est la machine d'apprentissage extrême quaternaire (QELM). Comme indiqué dans Figure 10, notre méthode est constituée de deux phases: Apprentissage et test.

Dans la phase d'apprentissage, nous procédons comme suit:

1. Chaque image couleur f est représentée au moyen d'un nombre quaternion pur comme $f = f_R i + f_V j + f_B k$, où f_R, f_V, f_B sont les composants rouge, vert et bleu de l'image.
2. Les invariants à valeur quaternaire des PFQDHM (*parameter-free quaternion disharmonic moments*) sont calculés pour chacune des images couleur quaternaires obtenues à l'étape précédente. Rappelons que les PFQDHM sont détaillés dans notre deuxième contribution (Chapitre 3). De plus, les invariants à valeur quaternaire sont obtenus à partir de l'équation suivante:

$$PFQDHMI_{n,m}(f) = PFQDHM_{n,m}(f) \otimes (PFQDHMI_{l,m}(f))^*, \quad (R. 22)$$

$$|m| \leq n, |m| \leq l,$$

Cette équation permet de conserver la forme quaternaire des invariants des PFQDHM. Nous fixons $l = |m|$.

3. Ces vecteurs caractéristiques quaternaires sont introduits dans la machine d'apprentissage extrême quaternaire (QELM) pour en apprendre les paramètres du modèle.

La phase de test est constituée des étapes suivantes:

1. Chaque image couleur f est représentée au moyen d'un nombre quaternion pur comme $f = f_R i + f_V j + f_B k$, où f_R, f_V, f_B sont les composants rouge, vert et bleu de l'image.
2. Les invariants quaternaires des PFQDHM sont calculés pour chacune des images couleur quaternaires obtenues à l'étape précédente selon Eq. (R. 22).
3. Ces vecteurs caractéristiques quaternaires sont introduits dans la machine d'apprentissage extrême quaternaire (QELM) pour la classification en utilisant les paramètres appris au cours de la phase d'apprentissage.

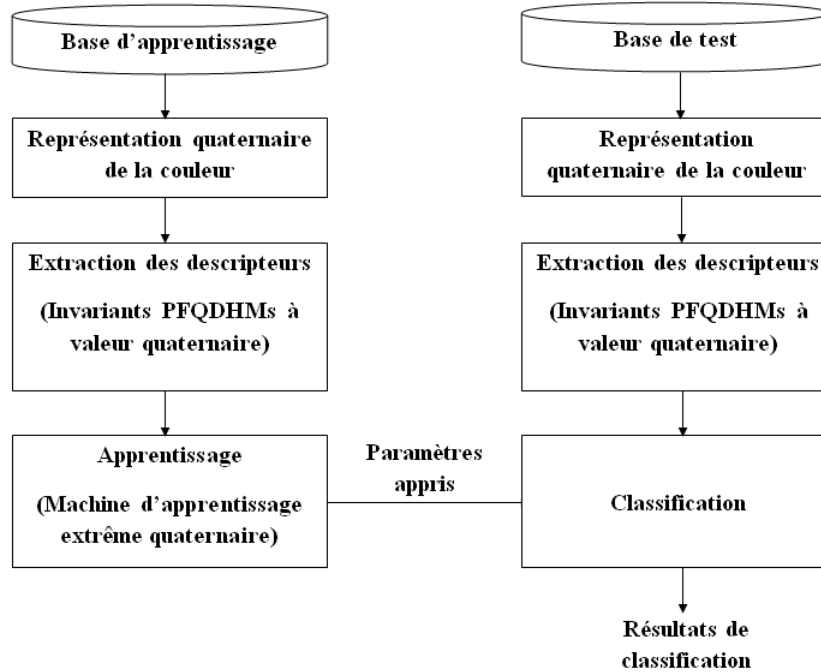


Figure 10 Architecture de la méthode proposée pour la classification d'images couleur.

Notre méthode permet (1) d'étudier les performances des invariants à valeur quaternaire des moments proposés, au lieu des invariants basés sur le module et (2) d'exploiter la machine d'apprentissage extrême quaternaire (QELM) (Minemoto et al., 2016a), distinguée par la qualité de la précision et le temps de calcul réduit.

4.3 Résultats Expérimentaux

Nous évaluons les performances de notre méthode de classification d'images couleur en comparaison avec des méthodes correspondant aux situations (1) et (3) de Figure 9. Nous considérons les descripteurs quaternaires suivant: les PFQDHM (moments disque-harmoniques sans paramètre), les QDHM (moments disque-harmoniques) et les QZM (moments de Zernike). Quant aux classificateurs, nous considérons le QBPNN, QELM, ELM, BPNN, Softmax et LS-SVM.

Il est à noter que:

- L'ordre maximal $nMax = 5$ est retenu pour les PFQDHM, les QDHM et les QZM, ce qui correspond respectivement à 36, 36 et 21 coefficients.
- Pour le QBPNN et le BPNN, nous fixons le taux d'apprentissage à $\alpha = 0.8$ et le nombre d'itérations à 300. Le nombre de nœuds de la couche cachée est fixé à $M = 2(L + N)/3$ où L est le nombre de nœuds dans la couche d'entrée et N est le nombre de nœuds dans la couche de sortie.
- Le nombre de nœuds de la couche cachée M pour le QELM et le ELM est fixé empiriquement pour chacun des moments PFQDHM, QDHM et QZM. Il est à souligner que les précisions de classification des QELM et ELM sont des valeurs moyennes obtenues sur 10 essais.
- Le classificateur Softmax se limite à 100.000 itérations comme condition de terminaison.
- Le noyau RBF gaussien est adopté pour le LS-SVM. Les valeurs optimales du paramètre de régularisation et celui de la bande passante sont automatiquement ajustées.

Nous utilisons plusieurs bases d'images à savoir la base Faces95, Faces95 corrompue, ALOI-ILL et ALOI-COL. Les détails concernant les expérimentations menées ainsi que les résultats obtenus sont présentés à la section 4.4 du chapitre 4. Dans ce résumé, nous rapportons uniquement les résultats concernant la base Faces95 corrompue avec un bruit sel et poivre et un flou gaussien (Figure 11 et Figure 12).

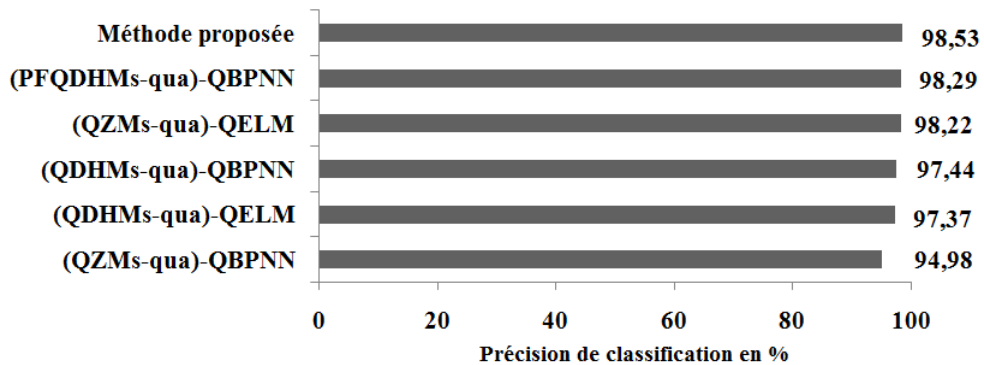


Figure 11 Précision de classification de la méthode proposée et des autres méthodes quaternaires en utilisant la collection Faces95 corrompue.

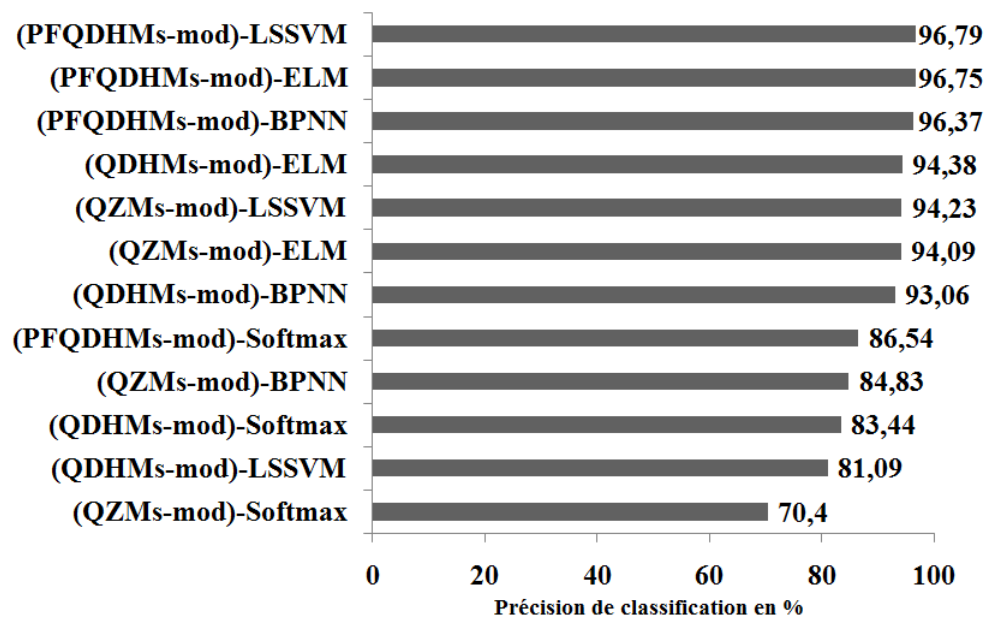


Figure 12 Précision de classification des différentes méthodes basées sur le module des moments orthogonaux quaternaires avec des classificateurs à valeurs réelles, en utilisant la collection Faces95 corrompue.

Les résultats obtenus montrent que les PFQDHM assurent des précisions de classification très proches des QDHM, et que les PFQDHM et les QDHM vont bien au-delà des QZM, en particulier lorsqu'il s'agit de changement d'éclairage de couleur. De plus, les invariants des moments à valeur quaternaire assurent de meilleures précisions de classification que les invariants à base du module.

Notons que, même si le QELM nécessite un nombre important de nœuds dans la couche cachée par rapport au QBPNN pour obtenir des performances similaires, le QELM est plus rapide, comme il est indiqué dans le tableau 4.

Tableau 4 Comparaison entre les méthodes basées sur QELM, ELM, QBPNN et BPNN, y compris notre méthode, en termes de temps d'apprentissage.

	QBPNN	BPNN	QELM	ELM
PFQDHM	5169.81 s	1784.98 s	17.70 s	2.59 s
QDHM	5190.92 s	1787.20 s	18.54 s	1.88 s
QZM	2587.06 s	863.89 s	15.37 s	3.30 s

D'après les résultats présentés dans le tableau 4, nous constatons que le QELM et le ELM sont plusieurs centaines de fois plus rapides que le QBPNN et le BPNN. Ceci s'explique par le fait qu'ils n'ont besoin que d'une seule passe *feed-forward* pour apprendre, contrairement au QBPNN et au BPNN, qui apprennent par itérations (300 itérations dans cette expérience). De plus, la légère différence entre les résultats du QELM et du QBPNN, remarquée dans les expériences précédentes, est compensée par la rapidité de l'apprentissage de QELM.

Conclusion et Perspectives

Dans ce travail de thèse, nous nous sommes penchés sur l'étude et l'amélioration de la description des formes couleur à des fins de recherche et de classification d'images. Cette amélioration concerne plusieurs aspects qui sont la capacité de reconstruction d'images, la robustesse et l'invariance, le temps de calcul ainsi que le pouvoir discriminatoire.

Sur la base de la revue de la littérature établie au chapitre 1, nous avons conclu que parmi les descripteurs de formes basés-région, les moments orthogonaux ont des propriétés très importantes qui permettent de préserver le maximum d'information sur la forme. Par conséquent, ils ont été largement utilisés dans de nombreuses applications, y compris la recherche d'images par le contenu (CBIR). Nous avons également conclu que l'exploitation de l'algèbre des quaternions pour intégrer l'information couleur dans ces descripteurs conventionnels est une solution prometteuse qui permet de conserver la corrélation entre les trois canaux couleur.

Dans le chapitre 2, nous avons décrit notre première contribution qui consiste à proposer une adaptation de la transformée en harmoniques sphériques (SHT) pour la description d'images couleur en utilisant à la fois l'algèbre des quaternions et une transformation sphère-disque appropriée. De plus, nous avons amélioré le temps de calcul de notre descripteur, appelé *quaternion disc-harmonic moments* (QDHM), en se basant sur la technique d'échantillonnage sphérique Healpix. Nous avons mené des tests empiriques exhaustifs pour évaluer les performances des moments proposés en termes de capacité de reconstruction d'images, de robustesse aux transformations photométriques, d'invariance aux transformations géométriques, de temps de calcul et de pouvoir discriminatoire. Les moments proposés ont montré une amélioration significative de la reconstruction d'images par rapport à d'autres moments orthogonaux existants étendus par les quaternions (les QZM, QPZM, QBFM et QOFMM). Les moments proposés sont aussi les plus robustes au bruit Sel et Poivre et au flou gaussien comparés aux autres moments quaternaires. De plus, ils assurent des résultats satisfaisants concernant la robustesse aux changements de la couleur et la direction de l'éclairage ainsi qu'à l'invariance à la rotation. Quant à la recherche d'images par le contenu (CBIR), nous avons étudié l'effet de l'ordre maximal, de l'espace de couleur et de la mesure de similarité. Les moments quaternaires proposés assurent un meilleur pouvoir discriminatoire par rapport à leur version conventionnelle complexe (les DHM) puisqu'ils prennent en compte l'information couleur. De plus, ils dépassent légèrement les autres descripteurs quaternaires et de Clifford. Cette légère différence au niveau des résultats de recherche d'images est compensée par le calcul rapide des moments proposés, notamment en utilisant la technique Healpix qui permet un calcul 12 fois plus rapide. Ainsi, le descripteur proposé pourrait constituer une solution prometteuse pour les systèmes de recherche d'images couleur.

Le chapitre 3 a été consacré à notre deuxième contribution, qui consiste à introduire une reformulation sans paramètre des moments quaternaires. Nous étions motivés par le fait que le calcul de tous les moments quaternaires existants dépend d'un paramètre quaternion unitaire pur μ dont la valeur optimale ne peut être fixée qu'empiriquement. De plus, la valeur estimée de ce paramètre est fortement liée au domaine d'application. Nous avons défini les moments orthogonaux quaternaires sans paramètres (PFQOM) proposés en substituant la partie angulaire dépendant de μ par une partie angulaire sans paramètre. Les résultats expérimentaux ont montré que les PFQOM proposés possèdent une capacité de reconstruction d'images couleur importante, et que leurs invariants à base du module sont robustes à diverses transformations photométriques et géométriques. De plus, leur pouvoir discriminatoire a été comparé à celui de leur version

basée sur μ , des moments de chromaticité, des moments de distribution chromatique de Legendre ainsi que les moments de Fourier Mellin orthogonaux à base de Clifford. Les rapports rappel-précision et les courbes ROC ont montré que les pouvoirs discriminatoires des PFQOM proposés sont satisfaisants et comparables à leurs versions à base de μ . En ce qui concerne la classification d'images couleur, nous nous sommes basés sur la machine d'apprentissage extrême *extreme learning machine* (ELM), distinguée par son efficacité. Les résultats obtenus ont montré la performance des descripteurs proposés.

Le chapitre 4 a détaillé notre troisième contribution qui consiste à suggérer une méthode de classification d'images couleur basée sur les moments QDHM sans paramètre. L'objectif est de confirmer la pertinence de notre descripteur quaternaire et de le valider dans le cadre d'un système de classification d'images. Pour maintenir la corrélation entre les quatre parties, réelle et imaginaires, des coefficients du descripteur quaternaire, nous nous sommes basés sur un classificateur de type quaternaire, qui est la machine d'apprentissage extrême quaternaire (QELM). Nous avons comparé les performances de notre méthode avec plusieurs techniques combinant des classificateurs réels ou quaternaires (ELM, BPNN, LS-SVM, Softmax, QBPNN et QELM) et des moments quaternaires (QDHM et QZM). Les résultats obtenus ont montré que: (1) notre méthode aboutit à des précisions satisfaisantes. (2) Les invariants des moments à valeur quaternaire assurent de meilleures précisions de classification que les invariants à base du module. (3) Même si le QELM nécessite un nombre important de nœuds dans la couche cachée par rapport au QBPNN pour obtenir des performances similaires, le QELM est beaucoup plus rapide.

Les résultats prometteurs de ce travail de thèse soulèvent plusieurs perspectives à explorer:

- Nous envisageons une analyse comparative entre les descripteurs d'images couleur quaternaires globaux et locaux. En effet, les descripteurs locaux sont plus robustes que ceux globaux lorsqu'il s'agit d'images contenant des grandes variations de pose ou d'illumination.
- Il serait intéressant d'explorer et d'évaluer la performance des moments proposés dans d'autres applications, telles que le tatouage numérique, le recalage et la correction d'images.
- Une extension possible de notre travail est de décrire des objets 3D couleur. Les images tridimensionnelles sont de plus en plus exploitées dans les sciences médicales et biologiques, notamment avec la disponibilité de techniques d'acquisition d'objets 3D. Dans ce contexte, les harmoniques sphériques (SH) ont été largement utilisées pour l'analyse des formes tridimensionnelles (Kazhdan et al., 2003; Skibbe and Reiser, 2017; Zhou and Wang, 2017; Bales et al., 2017; Radvilaitė et al., 2017).
- Nous projetons généraliser l'ensemble des moments proposés en utilisant l'algèbre de Clifford. En effet, l'algèbre de Clifford, dans laquelle l'algèbre des quaternions est contenue, peut être utilisée pour traiter des problèmes qui nécessitent des représentations multidimensionnelles ($> 4D$).

Bibliographie

Pour la bibliographie concernant ce résumé, nous référons à la section "Bibliography" du rapport de thèse.

Contributions à la description d'images couleur pour la recherche d'images par le contenu

Résumé: Ce travail de thèse s'inscrit dans le cadre de l'amélioration de la recherche d'images par le contenu en se focalisant sur le processus d'extraction des caractéristiques d'images. Nous nous intéressons particulièrement aux descripteurs d'images basés à la fois sur la forme et la couleur. En effet, en exploitant l'algèbre des quaternions, nous généralisons les moments orthogonaux radiaux permettant de décrire la forme en vue d'intégrer également l'information couleur. Dans un premier temps, nous décrivons l'état de l'art concernant les descripteurs d'images couleur, y compris les moments quaternaires. Ensuite, nous présentons trois contributions concernant la description d'images couleur pour la recherche par le contenu. Dans notre première contribution, nous proposons une adaptation de la transformée en harmoniques sphériques (SHT) pour la description d'images couleur en utilisant à la fois l'algèbre des quaternions et une transformation sphère-disque appropriée. De plus, nous améliorons le temps de calcul de notre descripteur, appelé *Quaternion Disc-Harmonic Moments* (QDHM), en se basant sur la technique d'échantillonnage sphérique *Healpix*. La deuxième contribution consiste à proposer une reformulation sans paramètre des moments orthogonaux quaternaires, appelés *Parameter-free quaternion orthogonal moments* (PFQOM). Ceci est justifié par le fait que le calcul des moments quaternaires existants dépend du choix d'un paramètre quaternion unitaire pur dont la valeur optimale ne peut être fixée qu'empiriquement. De plus, la valeur estimée de ce paramètre est fortement liée au domaine d'application. Dans notre troisième contribution, nous suggérons une méthode de classification d'images couleur basée sur les moments QDHM sans paramètre. L'objectif est de confirmer la pertinence de notre descripteur quaternaire et de le valider dans le cadre d'un système de classification d'images. L'idée de base est de garder la forme quaternaire des coefficients composant le descripteur QDHM au lieu de se restreindre aux modules des moments. Pour maintenir la corrélation entre les quatre parties, réelle et imaginaires, des coefficients du descripteur quaternaire, nous nous sommes basés sur un classificateur de type quaternaire, qui est la machine d'apprentissage extrême quaternaire (QELM). La performance de l'ensemble des descripteurs proposés est évaluée en termes de capacité de reconstruction d'images couleur, de robustesse aux changements photométriques, d'invariance géométrique, de pouvoir discriminatoire et de temps de calcul. Les résultats obtenus montrent l'intérêt de notre propos.

Mots clés: Descripteurs de formes, Moments orthogonaux, Harmoniques sphériques, Image couleur, Algèbre des quaternions, Recherche d'images par le contenu, Classification, Machine d'apprentissage extrême.

Contributions in Color Image Description for Content-Based Image Retrieval

Abstract: This thesis work concerns the enhancement of content-based image retrieval (CBIR) by focusing on the image feature extraction process. We are particularly interested in image description based on both color and shape information. Indeed, by exploiting the algebra of quaternions, we generalize the radial orthogonal moments, meant for describing shapes, in order to also integrate the color information. We first present an extensive state-of-the-art regarding color image descriptors including the quaternion-based moments. Then, we provide three contributions in color image description for CBIR. In our first contribution, we propose an adaptation of the spherical harmonic transform (SHT) for color image description by using both the algebra of quaternions and an appropriate sphere-to-disc transformation. In addition, we improve the computational time of our descriptor, namely the quaternion disc-harmonic moments (QDHMs), by employing the Healpix spherical sampling technique. Our second contribution consists in proposing a parameter-free reformulation of the quaternary orthogonal moments, namely the parameter-free quaternary orthogonal moments (PFQOMs). We were motivated by the fact that the calculation of the existing quaternary moments depend on the choice of a unit pure quaternion parameter whose optimal value can be fixed only with the help of experiments. Furthermore, the estimated value of this parameter is strongly dependent on the application domain. In our third contribution, we introduce a color image classification method based on the parameter-free QDHMs. The objective is to confirm the relevance of our quaternary descriptor and validate it as part of an image classification system. The basic idea is to keep the quaternary form of the coefficients composing the QDHMs descriptor instead of being restricted to the moment's modulus. To maintain the correlation between the four parts, real and imaginary, of the quaternary descriptor coefficients, we take advantage of a quaternary classifier that is the quaternary extreme learning machine (QELM). The performance of all the proposed descriptors is evaluated in terms of color image reconstruction capability, robustness to photometric changes, invariance under geometric transformations, discriminative power as well as computational time. The obtained results illustrate the effectiveness of our methods.

Keywords: Shape descriptors, Orthogonal moments, Spherical harmonics, Color image, Quaternion algebra, Content-based image retrieval, Classification, Extreme learning machine.