



**UNIVERSITE ABDELMALEK ESSAADI  
ECOLE NATIONALES DES SCIENCES APPLIQUEES  
TANGER**

**Centre d'Etudes Doctorales: Sciences et Techniques de l'Ingénieur  
Formation Doctorale : Sciences et Techniques de l'Ingénieur**

**THESE DE DOCTORAT**

**Présentée**

**Pour l'obtention du**

**DOCTORAT EN SCIENCES ET TECHNIQUES DE L'INGENIEUR**

**Par :**

**YAZIDI ALAOUI Youssef**

**Discipline : Image et systèmes**

**Spécialité : Imagerie numérique**

**Titre de la Thèse : Nouvelles méthodes basées sur les ondelettes  
pour l'analyse des images numériques : applications à la  
détection des panneaux routiers de signalisation et à la  
mise en correspondance stéréoscopique.**

**Soutenue le 7 novembre 2015 devant le jury suivant :**

**Pr. DIOURI Jaouad, *Faculté des Sciences de Tétouan***

**Pr. FENNAN Abdelhadi, *Faculté des Sciences Techniques Tanger***

**Pr. EL HADDAD Mohammed, *ENSA Tanger***

**Pr. LAHJOMRI Fouad, *ENSA Tanger***

**Pr. LAAZIZ Yassin, *ENSA Tanger***

**Pr. FILALI MEKNASSI Otman, *ENSA Tanger***

**Pr. EZZINE Abdelhak, *ENSA Tanger***

***Président***

***Rapporteur***

***Rapporteur***

***Rapporteur***

***Examineur***

***Co-Directeur de thèse***

***Co-Directeur de thèse***

**Structure de recherche accréditée d'accueil :**

**Laboratoire des Technologies de l'Information et Communications (LabTIC) de l'ENSA  
Tanger**

# *Avant-propos*

*Dans le but d'approfondir mes connaissances scientifiques et enrichir mon parcours professionnel, je me suis inscrit à l'université Abdelmalek Essadi de Tanger et particulièrement au laboratoire des technologies d'information et de la communication de l'ENSAT pour poursuivre mes études doctorales.*

*Etant titulaire d'un diplôme d'ingénieur d'état en génie informatique et de collaborateur dans un organisme de formation professionnelle (OFPPT), j'ai pris conscience de l'importance de développer mes connaissances scientifiques et académiques. Mon travail m'a permis aussi de saisir l'importance de la recherche scientifique dans le développement scientifique, économique et sociale.*

*Cette thèse a été entreprise au sein du laboratoire des technologies de l'information et de la communication LABTIC de l'ENSAT de Tanger. Le laboratoire LABTIC est l'une des structures de recherche qui fait partie du centre des études doctorales CED de l'université Abdelmalek Essaadi. Ce laboratoire traite plusieurs axes de recherche et on trouve, entre autres, l'axe de système de vision pour l'automobile et qui traite la détection de panneaux routiers et des obstacles par des méthodes de traitement des images numériques basée sur la couleur.*

*Le problème central de tout système de vision artificielle consiste à trouver des méthodes et des algorithmes capables de simuler le comportement de la vision humaine : extraire le sens de l'information capturée et apercevoir le monde. Les systèmes de vision actuels utilisent conjointement deux techniques fondamentales : des techniques numériques pour la classification et la quantification des données, et des techniques mathématiques et de l'intelligence artificielle pour l'analyse sémantique.*

*Tout au long de notre travail, notre problème central a été l'analyse et la segmentation des images numériques couleurs pour extraire les régions d'intérêt dans l'image. En effet, et avant de chercher à segmenter une image, on se fixe souvent comme objectif d'identifier et de localiser les régions pertinentes dans l'image. C'est pour cette raison, et relativement à notre*

*champ d'application, qu'une bonne partie de notre travail a d'abord consisté à trouver les outils et les méthodes convenables pour extraire et caractériser l'information présente dans l'image.*

*Le problème de quantification d'une image est un problème de transformation puis interprétation du résultat de cette transformation. Une grande partie de ce mémoire se propose d'explorer, sur ce principe, de nouvelles méthodes pour extraire l'information la plus pertinente dans une image. Notre but est de chercher les bons outils et les bonnes transformations des images pour développer une approche claire orientée vers la caractérisation des objets d'intérêt dans l'image.*

*La conduite automobile assistée étant le champ d'application visé. A ce jour, et chez tous les acteurs dans le domaine qu'ils soient constructeurs ou chercheurs, la conduite assistée s'appuie sur le montage d'un système embarqué qui est soit actif (radar, lidar,...), soit passif (camera). Et dans les deux cas, il est appelé à offrir des fonctionnalités d'assistance de conduite pour la sécurité routière et amener à un confort au conducteur.*

*Nos travaux sont plutôt basés sur un système passif : Le système embarqué pour l'automobile, et après acquisition (matériel : caméra), analyse et interprète le contenu des scènes routières à l'aide de sa couche logicielle pour assurer des fonctionnalités de détection des objets et des obstacles. On trouve deux axes de préoccupations pour les concepteurs de tels systèmes : l'axe de la vision 3D et l'interprétation de l'entourage (détection des obstacles) et l'axe de reconnaissance des signaux routiers.*

*Nous avons traité les deux axes, en partant de ce qui est déjà réalisé dans le domaine et en étudiant les outils et les techniques disponibles. Les travaux relatifs au premier axe de préoccupation ont fait de notre publication dans le journal IJIP, « Adaptive multiscale Stereo images Matching based on wavelet transform modulus maxima » [97]. Et les travaux relatifs au deuxième axe ont fait de notre publication dans le journal IJETED, « Images segmentation method based on a multi scale analysis of 2D histograms » [98]. Nous reportons les deux versions publiées pour les deux axes dans le présent document.*

# Remerciements

*Le travail présenté dans cette thèse a été effectué au sein du Laboratoire des Technologies de l'Information et de la Communication (LabTIC) de l'Ecole Nationale des Sciences Appliquées de Tanger (ENSAT) de l'Université Abdelmalek Assaadi.*

*J'adresse mes sincères remerciements à mon directeur de thèse Monsieur Otman FILALI MEKNASSI, Professeur à l'ENSAT, pour m'avoir donné l'opportunité de faire partie de l'équipe LabTIC.*

*J'exprime ma profonde gratitude pour mon co-encadrant Monsieur Abdelhak EZZINE, professeur à l'ENSAT, pour l'intérêt particulier qu'il a apporté à mes travaux de recherche et pour la confiance qu'il m'a accordé. Je le remercie pour ses encouragements durant cette période, ses conseils et sa constante disponibilité. Je lui resterai toujours reconnaissant de m'avoir transmis ses connaissances, inculqué ses méthodes de travail et sa rigueur scientifique. Grace à ses qualités humaines et professionnelles, j'ai pu mener à bien ce travail et ce fût un honneur pour moi de travailler avec lui.*

*Mes vifs remerciements s'adressent également à tous les membres de jury pour avoir accepté d'examiner ce travail et qui se sont déplacés pour me faire l'honneur de siéger dans ce jury.*

*Ma profonde gratitude et ma reconnaissance vont également à toutes les personnes qui ont contribué de près ou de loin au bon déroulement de ce travail.*

*Enfin, mes remerciements vont aussi à ma famille. J'adresse un remerciement particulier à ma femme et je lui exprime ma profonde gratitude pour son soutien constant, ses sacrifices et ses encouragements.*

# Résumé

Ce mémoire présente nos travaux de recherche relatifs à l'utilisation de l'imagerie numérique, pour implémenter certaines fonctionnalités de la vision artificielle d'un système embarqué dans l'automobile. Nous nous intéressons principalement aux fonctionnalités de détection des panneaux routiers de signalisation et à la vision 3D.

Dans ce but, nous avons mis au point une approche originale de détection basée sur la couleur des objets présents dans la scène. Nous supposons que les panneaux routiers de signalisation et les objets dans les images capturées constituent des régions homogènes dont les pixels ont des couleurs proches. Nous avons travaillé aussi sous l'hypothèse que les couleurs des panneaux routiers de signalisation sont bien distinctes.

Dans notre approche, la détection et l'identification de la couleur sont basées sur l'histogramme de l'image qui est une structure probabiliste représentant la distribution des couleurs dans l'image. De ce fait, l'idée avancée est que les régions homogènes de l'image créent des domaines de valeurs dans l'histogramme, domaines que nous identifierons par une analyse multi-échelle en ondelettes.

Nous avons opté pour l'utilisation des ondelettes car elles constituent un outil mathématique qui s'est montré très puissant et très adapté au traitement des images numériques. En effet, les ondelettes permettent de mettre l'accent sur différentes structures de l'image par une analyse multi-échelle. Cette analyse donne des approximations tout en préservant des informations sur les détails qui sont utiles pour caractériser les pixels isolés ou les pixels de transitions.

Nous avons utilisé les ondelettes pour extraire et caractériser les singularités significatives rencontrées dans les images. Le comportement des singularités est étudié à partir des lignes maximales multi-échelles.

Nous avons exploité les résultats précédents pour l'adapter à la vision 3D en concevant une méthode de mise en correspondance dans le cadre de cette vision. Ainsi, nous avons pu évaluer la dimension profondeur dans l'image à partir de calculs de cartes de disparité. Ces cartes sont générées à partir des primitives de deux images stéréoscopiques après avoir réalisé un appariement multi-échelle des points singuliers des deux images.

# Table des matières

<b>INTRODUCTION GENERALE.....</b>	<b>12</b>
<b>I. CONTEXTE .....</b>	<b>13</b>
<b>II. ORGANISATION DU MÉMOIRE.....</b>	<b>16</b>
<b>CHAPITRE 1 : CADRE DE TRAVAIL .....</b>	<b>18</b>
<b>I. INTRODUCTION .....</b>	<b>19</b>
<b>I. PROBLÈME DE DÉTECTION DES PANNEAUX ROUTIERS .....</b>	<b>22</b>
1. Préambule.....	22
2. Axe abordé.....	23
3. Nos contributions.....	25
<b>II. MISE EN CORRESPONDANCE STÉRÉOSCOPIQUE.....</b>	<b>28</b>
1. Avant-propos .....	28
2. Problématique de la mise en correspondance.....	28
3. Nos contributions.....	29
<b>III. CONCLUSION .....</b>	<b>30</b>
<b>CHAPITRE 2 : SYNTHÈSE DES OUTILS UTILISÉS.....</b>	<b>32</b>
<b>I. INTRODUCTION .....</b>	<b>33</b>
<b>II. IMAGES NUMÉRIQUES COULEURS.....</b>	<b>36</b>
<b>1. Les systèmes colorimétriques.....</b>	<b>37</b>
1.1 La vision humaine .....	37
1.2 La colorimétrie.....	37
1.3 Les espaces de représentation de la couleur.....	39
<b>2. Les histogrammes couleurs .....</b>	<b>44</b>
2.1 Calcul des histogrammes .....	44

2.2 Projection de l'histogramme .....	47
2.3 Les histogrammes bidimensionnels.....	49
<b>III. LA VISION 3D.....</b>	<b>54</b>
1. Introduction .....	54
2. Principe de la stéréovision .....	55
3. Aspects géométriques de la stéréovision .....	56
3.1 Calibrage stéréoscopique .....	57
3.2 Rectification.....	57
4. Mise en correspondance .....	59
5. Triangulation.....	60
<b>IV. CONCLUSION.....</b>	<b>60</b>
<b>CHAPITRE 3 : APPROCHE MATHEMATIQUE.....</b>	<b>63</b>
<b>I. ONDELETTES COMME OUTIL D'ANALYSE.....</b>	<b>64</b>
1. Introduction .....	64
2. Transformée de Fourier.....	64
3. Transformée en ondelettes .....	66
3.1 Présentation .....	66
3.2 Exemples d'ondelettes .....	67
3.3 Transformée en ondelettes bidimensionnelle.....	69
3.4 Transformée en ondelette dyadique .....	69
4. Analyse multi résolution et ondelettes .....	70
4.1 Exemple : ondelettes de Haar.....	72
4.2 L'analyse multi résolution dans le cas 2D .....	74
<b>II. TRANSFORMÉ EN ONDELETTES ET CONTOURS MULTI-ÉCHELLES .....</b>	<b>76</b>
1. Modules maxima de la transformée en ondelettes.....	76
2. Modules maxima et détecteur de Canny .....	78
3. Maximas des ondelettes et contours .....	80
<b>III. CONCLUSION .....</b>	<b>83</b>
<b>CHAPITRE 4: ANALYSE MILTIECHELLE DES HISTOGRAMMES 2D.....</b>	<b>84</b>

<b>I. INTRODUCTION</b> .....	<b>85</b>
<b>II. ANALYSE DES HISTOGRAMMES ET DÉTECTION DES PANNEAUX</b> .....	<b>86</b>
1. Analyse des histogrammes .....	86
2. Approche de détection .....	91
<b>III. RÉSULTATS DE L'APPROCHE</b> .....	<b>94</b>
1. Evaluation qualitative .....	95
2. Evaluation quantitative .....	97
3. Résultats dans l'espace $L^a * b^*$ .....	98
<b>IV. CONCLUSION</b> .....	<b>101</b>
<b>CHAPITRE 5 : MISE EN CORRESPONDANCE MULTIECHELLE</b> .....	<b>105</b>
<b>I. INTRODUCTION</b> .....	<b>106</b>
1. Avant-propos .....	106
2. Problématique de la mise en correspondance .....	106
<b>II. MISE EN CORRESPONDANCE MULTIECHELLE</b> .....	<b>108</b>
1. Introduction .....	108
2. Approche de mise en correspondance .....	110
2.1 Construction des chaînes maximales .....	110
2.2 Mise en correspondance multi échelle .....	111
3. Résultats de l'approche .....	114
<b>III. CONCLUSION</b> .....	<b>117</b>
<b>CONCLUSIONS ET PERSPECTIVES</b> .....	<b>118</b>
<b>BIBLIOGRAPHIE</b> .....	<b>122</b>



# Liste des figures

Figure 1: Signes de danger, d'interdiction et d'obligation	15
Figure 2: Image gauche : une scène 2D ; image droite : sa carte de profondeur	16
Figure 3: Processus de traitement d'une image numérique	22
Figure 4: Images couleur avec ces trois plans dans l'espace RGB	37
Figure 5: variation du tri-stimuli en fonction des longueurs d'ondes correspondantes aux primaires RGB	40
Figure 6: L'espace de couleur RGB	40
Figure 7: Les composants S, T, L	41
Figure 8 : L'espace couleur HSV	42
Figure 9 : Espace couleur La*b*	43
Figure 10: Image avec quatre régions de des différentes couleurs et ces histogrammes 3D dans les espaces RGB, HSV et La*b*	45
Figure 11: Image réelle avec ces histogrammes 3D dans les espace RGB, HSV et La*b*	46
Figure 12 : une image de régions et son histogramme couleur RGB	47
Figure 13: Histogrammes monodimensionnels dans l'espace RGB	48
Figure 14: Approche de segmentation marginale d'images couleur	49
Figure 15: Image synthèse est ces histogrammes 2D dans l'espace RGB	51
Figure 16: Zones panneaux dans l'histogramme 2D	52
Figure 17 : Configuration générale du système stéréoscopique	57
Figure 18 : Mise en œuvre des épipôles	58
Figure 19 : Mise en œuvre de la rectification géométrique	58
Figure 20 : Représentation de la transformée en ondelettes ( $\psi_{x1,a1}$ et $\psi_{x2,a2}$ )	67
Figure 21 : Dérivée seconde d'une gaussienne, chapeau mexicain	68
Figure 22: ondelette de Haar	69
Figure 23 : Analyse à l'aide d'un banc de filtres	72
Figure 24 : Transformée en ondelettes orthogonale 2D	75
Figure 25 : Cône d'influence d'un point x	77
Figure 26 : Image 1, ses cartes de modules aux niveaux 1,2 ses cartes de maximas aux niveaux 1,2 et 3	78

Figure 27: Images contours dans les échelles : 1,2 et 3	82
Figure 28 : Vue 3D des histogrammes 2D	87
Figure 29 : Les histogrammes 2D dans les espaces RGB et La*b*	89
Figure 30 : Approximation de l'histogramme RB dans différents niveaux d'échelles	90
Figure 31 : Classification des pixels dans les niveaux de résolution 6,7 et 8	93
Figure 32 : Processus de segmentation à base d'analyse des histogrammes 2D	94
Figure 33 : L'image 1, son histogramme RB, une vue 3D de l'histogramme	96
Figure 34 : Des segmentations de l'image 1, pour les niveaux d'analyse 8,9 et 10	96
Figure 35 : Segmentations de l'image 2 dans les niveaux d'analyse 6,7 et 10	97
Figure 36 : Segmentation de l'image synthèse par analyse de l'histogramme a*b* dans les niveaux 5,6 et 7	99
Figure 37: Segmentation de l'image 1 par analyse de l'histogramme a*b* dans les niveaux 2,3 et 4	99
Figure 38 : Segmentations de l'image 1 par analyse de l'histogramme La* dans les niveaux 2,3 et 4	100
Figure 39 : Evolution des maximas de l'image tsukuba de niveau 1 à 4(avec seuillage)	111
Figure 40 : Maximas appariés dans les images droite et gauche de l'image tsukuba	113
Figure 41 : Image cones (gauche, droite), les chaines maximas de niveau 1 à 4, les cartes des maximas appariées, les points d'intérêt correspondants	115
Figure 42 : Image teddy (gauche, droite), les chaines maximas de niveau 1 à 4, les cartes des maximas appariées, les points d'intérêt correspondants	115
Figure 43 : Image tsukuba(droite, gauche), les chaines maximas de niveau 1 à 4, les cartes des maximas appariées, les points d'intérêt correspondants	116
Figure 44 : Cartes de disparités des images tsukuba,teddy et cones sans post traitement	116

# Liste des tableaux

Tableau 1 : Le nombre de modes et de modules maximaux de l'histogramme de l'image $I$ de niveau 0 au niveau 9	.....91
Tableau 2 : Valeurs des critères d'évaluation: ligne 1 (approche de segmentation proposée) ; ligne 2 (approche de segmentation marginale)	.....98
Tableau 3 : Valeurs des critères d'évaluation dans l'espace RGB	.....100
Tableau 4 : Valeurs des critères d'évaluation dans l'espace La*b	.....101

# **INTRODUCTION GENERALE**

## I. Contexte

Cette thèse s'inscrit dans le cadre d'un projet de recherche entrepris par le Laboratoire des Technologies de l'Information et de la Communication (LabTIC) de l'ENSAT autour du thème de la sécurité routière et de la conduite intelligente. Les perspectives de ce projet sont, entre autres, de développer un Framework proposant une large panoplie de fonctions liées au traitement et à l'analyse des images numériques utilisées pour l'interprétation des objets d'une scène routière et pour la détection des obstacles. Le but est la réalisation d'un système automobile embarqué.

La conduite intelligente a fait l'objet de thèmes de recherche très soutenus depuis quelques années. Cela est dû à la croissance considérable des performances de calcul des processeurs des systèmes embarqués dans l'automobile. La détection des obstacles et la conduite automobile assistée au moyen de systèmes embarqués (radar, caméra, sonar,...) sont devenues d'un grand intérêt pour les constructeurs automobiles. Ces systèmes doivent être capables d'aider, voir de palier les défauts du conducteur dans le cas d'un manque de concentration ou dans des conditions de conduite difficile (brouillard, pluie, nuit, etc.).

Dans le cas de la détection des panneaux routiers de signalisation, le système embarqué doit être en mesure d'analyser des scènes routières et de les interpréter [1]. La richesse du contenu de ces scènes routières est très variable selon qu'il s'agisse de scènes dans le milieu urbain ou dans le milieu rural. En effet, une scène du milieu urbain est beaucoup plus riche en information à cause de la diversité des nombreux objets de la scène, d'où la difficile tâche de détection des panneaux routiers de signalisation dans un tel décor.

La difficulté de détection des panneaux routiers réside aussi dans les problèmes liés aux conditions d'acquisition des scènes et aux effets de dégradation subis par les panneaux de signalisation. La reconnaissance des panneaux routiers doit donc tenir compte de plusieurs difficultés et les contourner.

Les systèmes d'aide à la conduite assistée utilisent des méthodes qui analysent les scènes restituées à l'aide d'un capteur, et décident automatiquement de leur contenu. La plupart de ces méthodes privilégient d'analyser la scène sous forme d'une séquence d'images prises dans de petites intervalles [2]. Ces méthodes réalisent un traitement sur les images capturées afin d'identifier les panneaux routiers en passant, dans un premier temps, par une étape de détection et, dans un deuxième temps, par une étape de reconnaissance.

Dans la littérature, plusieurs travaux utilisant la vision artificielle ont été effectués pour la détection et la reconnaissance des panneaux routiers [3, 4, 5,6]. Ces travaux utilisent soit des approches basées sur la couleur des panneaux, soit des approches basées sur la forme des panneaux, ou bien des approches combinant les deux.

Les méthodes basées sur la couleur des panneaux utilisent le paramètre couleur pour la détection [5, 6,7] : les panneaux en rouge indiquent l'interdiction ou le danger, les panneaux en bleu indiquent l'obligation, les panneaux en jaune indiquent l'interdiction temporaire,...Ces méthodes ont l'avantage d'être plus rapides que les autres méthodes mais souffrent d'une sensibilité aux variations des conditions d'éclairage.

Les méthodes basées sur la forme exploitent le fait que les panneaux ont des formes normalisées (triangulaire, circulaire, rectangulaire, ...) [8]. Ces méthodes se basent essentiellement sur des techniques de reconnaissance de formes généralement très robustes. Ces méthodes présentent certains inconvénients : la sensibilité aux déformations dues surtout aux conditions d'acquisition et l'éventuelle confusion avec d'autres objets de la scène ayant la même forme. La *figure 1* montre des exemples de panneaux de signalisation routiers de différentes couleurs et de différentes formes.

Certaines méthodes combinent les deux approches précédentes afin d'améliorer les résultats de détection des panneaux, alors que d'autres se basent sur des méthodes dites d'apprentissage généralement très coûteuses en termes de traitement et de temps de calcul [9].

Dans le présent travail, nous présentons une nouvelle méthode de détection des panneaux routiers avec comme paramètre de détection, la couleur des panneaux. Nous avons choisi le formalisme mathématique des ondelettes pour faire une analyse multi-échelle des scènes enregistrées et aboutir à la détection des panneaux routiers. Ces ondelettes ont été exploitées selon un deuxième axe pour résoudre le problème de la mise en correspondance des scènes 2D dans le cadre de la vision 3D.

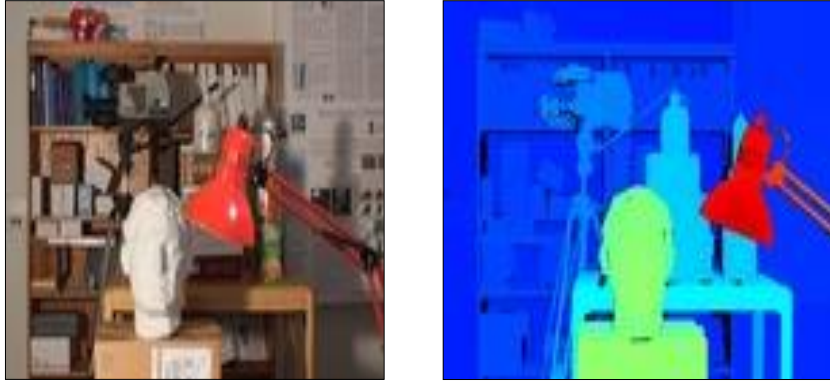
Différents modèles de la vision artificielle ou de la vision par ordinateur capables de reproduire des fonctions de la vision humaine ont été développés. Ces systèmes sont souvent autonomes et capables d'effectuer des tâches telles que la reconnaissance des formes ou la navigation, d'où la réalisation de robots capables de se déplacer de manière autonome dans le monde réel et, la conception d'un véhicule intelligent capable de discriminer son entourage.

Ces robots, et ces véhicules disposent d'un système de vision dont les performances sont dues à son système d'acquisition, son processus de traitement d'image et ses fonctions de reconnaissance. La fonction principale de telles machines peut se résumer dans la fonction de description de la scène dans son milieu naturel, et qui est généralement en 3D. Ceci passe par la détermination de la position de points dans l'espace à partir d'une ou plusieurs images fournies par le système d'acquisition. Cette action constitue un des objectifs de la vision artificielle, et c'est dans ce but que la stéréoscopie intervient comme une technique utilisée pour reconstruire une scène 3D à partir des images prises sous des angles de vue différents.

La stéréovision ou la vision 3D est l'ensemble des techniques et procédures permettant d'apercevoir le relief 3D à partir des images 2D (mise en correspondance stéréoscopique cf. *figure 2*). Différentes méthodes existent pour la réalisation de la mise en correspondance, mais elles sont souvent conçues sur mesure pour répondre à une situation particulière [10, 11, 12]. L'approche que nous avons développée se veut d'être globale. Elle est basée sur l'utilisation des modules maximaux des coefficients des ondelettes comme points d'intérêts, et permet d'aboutir à une mise en correspondance entre deux images stéréoscopiques.



Figure 1 : Signes de danger, d'interdiction et d'obligation



Scène 2D

Reconstruction 3D de la scène

Figure 2 : Image gauche : une scène 2D ; Image droite : sa carte de profondeur

## II. Organisation du mémoire

Le présent rapport est organisé en quatre chapitres rapportant les deux axes de recherche abordés. Le premier axe présente nos travaux relatifs à la détection des panneaux de signalisation routiers, le deuxième axe présente nos contributions dans la vision 3D.

Le premier chapitre de ce mémoire explicite notre cadre de travail. Nous présentons les deux axes de recherche abordés, et nous donnons des idées générales sur les contributions faites dans les deux sujets.

Dans le deuxième chapitre, nous présentons une synthèse des différentes approches de traitement et de segmentation des images numériques (état de l'art). Nous abordons aussi la théorie de la couleur et les systèmes de représentation de la couleur. En effet, l'étude de la couleur est étroitement liée au calcul des histogrammes couleurs. On parlera également dans ce chapitre des techniques de calcul des histogrammes.

La théorie des ondelettes est rappelée dans le deuxième chapitre en se rapprochant de la relation entre cette théorie et le traitement et l'analyse des images numériques. D'ailleurs, Les images numériques sont des signaux bidimensionnels pouvant subir des analyses par la théorie des ondelettes.

Le chapitre donne aussi un aperçu de la stéréovision et rappelle les méthodes existantes pour la mise en correspondance entre deux images 2D pour une vision 3D. On présente



également dans ce chapitre, la notion de modules maximaux des ondelettes que nous utiliserons par la suite, et le lien existant entre les modules maximaux avec les contours et les singularités des images.

Nous présentons dans le troisième chapitre la méthode de segmentation originale que nous avons mise au point. Nous étudions principalement le résultat de l'analyse multi-résolution faites avec les ondelettes sur les histogrammes 2D des images couleurs. Les résultats de simulation obtenus sont exposés et discutés. Les améliorations à envisager pour notre méthode et les perspectives à entreprendre seront discutées.

Ce rapport présente nos contributions dans le domaine de la stéréovision. Notre approche innovante de mise en correspondance multi-échelle pour une vision 3D, est exposée dans le chapitre quatre. Les résultats de notre méthode sur des exemples de simulation sont présentés et discutés.

Enfin, nous concluons ce rapport de thèse par une discussion des différents résultats obtenus durant nos travaux en s'attachant aux améliorations possibles à envisager et aux perspectives à entreprendre dans le futur.

# **CHAPITRE 1 :**

# **CADRE DE TRAVAIL**

## I. Introduction

Dans ce travail de recherche, nous nous sommes intéressés au traitement des images numériques afin d'interpréter le contenu d'une scène routière capturée à l'aide d'un système de vision, en vue d'une aide à la décision dans le cadre d'une conduite intelligente.

Les outils de traitement et d'analyse des images numériques sont très nombreux et tellement variés qu'on ne peut pas facilement les classer [13,14]. Une classification possible peut être faite en se basant sur le formalisme mathématique utilisé. Les différentes approches mathématiques ont conduit à une variété de méthodes et d'outils d'analyse et de traitement des images numériques, qu'on utilise souvent d'une façon complémentaire pour tenter d'avoir les résultats escomptés.

Toutefois, les méthodes mathématiques de traitement des images numériques sont des solutions qui sont loin d'être des solutions générales qu'on peut appliquer dans n'importe quelle situation. Chaque méthode est développée pour un cas d'étude particulier. De ce fait, le choix de la méthode mathématique à retenir dépend de plusieurs facteurs ou paramètres telles que le champ d'application, le contenu des images traitées, la performance des algorithmes de calcul,...

L'objectif de notre travail s'intègre dans la conception d'un système de vision capable d'extraire des informations et des caractéristiques sémantiques à partir des images d'entrée représentant des scènes routières. La segmentation des images numériques est notre outil de travail primordial et doit se faire d'une façon automatique en donnant des résultats performants en temps réel.

Généralement, on peut classer les approches de segmentation des images en deux grandes catégories. La première catégorie utilise un formalisme mathématique probabiliste basé sur les distributions des couleurs dans un espace couleur choisi pour l'analyse. La deuxième catégorie de segmentation analyse les pixels directement dans le plan image et peuvent utiliser des connaissances en amont sur les structures de l'image. Toutes les approches utilisées cherchent à mettre en évidence des régions cohérentes de l'image aussi bien spatialement que par leur contenu [13][14][15].

Dans le cadre de notre travail de recherche, la segmentation des images a été faite selon une approche probabiliste. Les approches probabilistes reposent sur une classification et un

regroupement des pixels (individus) de l'image qui ont des propriétés identiques. L'information couleur de chaque pixel est la métrique la plus évidente et la plus utilisée pour ces approches probabilistes. Des classes de pixels de l'image peuvent être définies et construites à partir de certains paramètres prédéfinis (centres de classes par exemple). Les classes des pixels peuvent aussi être construites à partir d'une étude de l'histogramme couleur de l'image. En effet, les histogrammes des images couleurs constituent un outil mathématique puissant permettant d'évaluer les classes des pixels, et extraire les caractéristiques de ces classes.

D'un point de vue statistique, les pixels portant des couleurs très proches seront regroupés dans une même classe (une région de l'image). Ces pixels définissent donc une zone de valeurs très élevées dans l'histogramme avec la présence d'un sommet (extremum). Les méthodes de classification à partir de l'histogramme, traitent cette zone en la traduisant en une seule région homogène et non pas en plusieurs régions homogènes connexes ou non. Cette dernière remarque met en évidence une des difficultés rencontrées en utilisant cette méthode. En effet, la mise en évidence ou le discernement d'objets ayant des couleurs proches mais pas identiques dans une image, devient très difficile par l'utilisation de l'histogramme. Ceci a été vérifié par les expériences faites sur le traitement des images acquises par un système de vision embarqué dans l'automobile [16].

Les méthodes de détection des panneaux routiers développées jusqu'à nos jours et qui sont basées sur la couleur, tentent de réduire la sensibilité aux variations de l'intensité de la scène. C. Maaoui et al. [17] ont proposé une technique de segmentation couleur qui a été utilisée afin de détecter un panneau de signalisation routier dans une image. L'algorithme *mean-shift* développé par les auteurs, est basé sur le calcul des valeurs moyennes des pixels, permet de distinguer les différents objets présents dans une image couleur. U. Zakir et al. [3] ont présenté une méthode de segmentation originale réalisée dans différents espaces couleurs (RVB, HSV, YCbCr, La\*b\*,...) en faisant un seuillage couleur de l'image. Les valeurs des différents seuils ont ainsi été déterminées en prenant en compte des conditions de luminosité de la scène analysée, la variation de luminosité pouvant être due aux conditions météorologiques ou à la variation de l'ensoleillement durant la journée. De façon similaire, M. Bénallal et al. [4] ont proposé de fixer à l'avance les seuils couleurs pour les différents pixels pour pouvoir déterminer les différentes classes des pixels d'une image. C.F. Paulo et als. [5] utilisent une technique de binarisation basée sur un algorithme dit algorithme d'Otsu, et arrivent avec cette technique à fixer des seuils couleurs dans l'espace couleur HSV. Dans [6], une technique

basée sur la détermination de la vraie couleur indépendamment de l'effet de la luminance au niveau de chaque pixel (technique de la constance des couleurs) a été utilisée, afin de diminuer l'effet de la luminosité pendant la segmentation de l'image. En effet, la constance des couleurs permet de représenter l'image dans un espace couleur où la reconnaissance de la vraie couleur d'un pixel n'est pas affectée par la luminosité et on obtient ainsi une certaine stabilité chromatique et donc une meilleure reconnaissance des objets panneaux de signalisation routiers.

L'utilisation de méthodes basées sur l'exploitation de l'histogramme dans le cas de détection des panneaux routiers présente plusieurs difficultés, car un panneau routier (qui constitue une petite région) située à côté d'une grande région homogène (véhicule, arbre, route, immeuble ou panneau publicitaire) serait rarement restitué [18,19]. Ce type de difficultés est intensifié par la confusion d'une part des panneaux routiers avec d'autres objets de la route et, d'autre part, de la variation des paramètres chrominance et luminance des panneaux qui est due à leur exposition aux facteurs naturels.

La stéréoscopie est un processus de reconstruction de la profondeur d'une scène à partir de deux images de la même scène. En effet, la profondeur d'une scène crée une disparité géométrique entre deux images acquises selon deux angles différentes.

La difficulté majeure de la stéréoscopie consiste à déterminer le décalage entre les images prises, et cela conduit systématiquement à une opération de mise en correspondance, difficile à mettre en place, entre les primitives des images (pixels, régions, contours, etc.). Pour le faire de manière simple et similaire au système visuel humain, le système stéréoscopique binoculaire n'utilise que deux images prises par deux capteurs légèrement décalés pour créer l'image gauche et l'image droite qui correspondent respectivement à l'image vue par l'œil gauche et l'image vue par l'œil droit.

Dans la littérature, il existe des méthodes qui traitent l'ensemble de l'image et construisent les cartes de disparité en réalisant un appariement pixel par pixel, et des méthodes qui cherchent à apparier d'abord les points caractéristiques (points d'intérêts) et utiliser le résultat pour reconstruire des cartes de correspondances [10, 11, 12,71, 96].

# I. Problème de détection des panneaux routiers

## 1. Préambule

La segmentation des images numériques est une opération qui vise à trouver des régions dans l'image correspondant à des objets d'intérêt. Cette segmentation est utilisée par le système visuel humain qui est capable d'identifier et distinguer les objets d'une scène.

Différentes applications utilisant la segmentation existent. La segmentation est utilisée, entre autres, pour des applications de reconnaissance de texte et de formes, dans des applications de recherche et d'indexation des images numériques, dans le domaine de la médecine, la télésurveillance, etc. [20]

Nos travaux de recherches concernent la détection des panneaux routiers par un système de vision embarqué et relèvent ainsi du domaine de la conduite intelligente. On est confronté à analyser une séquence d'images capturées à l'aide d'un capteur (caméra numérique) et à décider après cette analyse de l'existence ou non d'un panneau routier : c'est l'étape de détection. Vient ensuite une étape de reconnaissance et d'identification des panneaux routiers réalisée souvent à partir d'une base de données qui utilise des techniques d'indexation ou de mise en correspondance [1,4]. La segmentation est une des étapes du processus utilisé dans une chaîne de traitement d'une image numérique en vue d'interpréter son contenu (cf. *figure 3*).

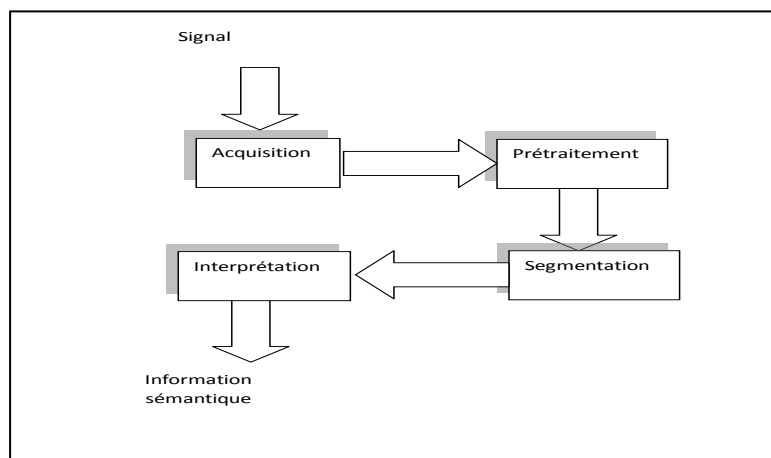


Figure 3 : Processus de traitement d'une image numérique

La difficulté de la tâche de segmentation vient du fait que cette étape ne tient pas compte des informations concernant les objets à identifier. En effet, les différents algorithmes de segmentation n'utilisent pas ces informations dans leurs calculs. De plus, privilégier un algorithme ou un autre peut être lié au champ d'application d'intérêt ou lié à la nature des images traitées. La majorité des algorithmes de segmentation sont basées sur deux principes : la recherche de régions homogènes dans l'image selon un critère d'homogénéité (couleur, texture,...) ou la recherche de transitions dans l'image (variation brusque d'un paramètre qui caractérise les pixels) ce qui permet de définir les frontières des objets.

## 2. Axe abordé

Les travaux scientifiques ayant trait à la reconnaissance des panneaux routiers peuvent être classés en trois catégories selon la méthode de détection utilisée. La première catégorie se base sur la couleur dû au fait que les panneaux ont des couleurs standards prédéfinies (rouge, bleu, ou jaune). Les méthodes de détection employées dans cette première catégorie ont l'avantage d'être basées sur des algorithmes rapides qui donnent des résultats satisfaisants pour des scènes comportant peu d'objets comme les scènes capturées dans le milieu rural. Toutefois, ces méthodes souffrent de certains inconvénients : le degré de la luminosité de l'image affecte le résultat de la détection comme c'est le cas des images sombres prises dans des conditions d'éclairage difficiles [4]. La difficulté de l'opération de détection qui se base sur la couleur est de pouvoir, en détectant deux objets ayant des couleurs proches, les distinguer.

La deuxième catégorie des méthodes de détection des panneaux routiers, regroupent les travaux qui utilisent une approche basée sur la forme des objets. En effet, les formes des panneaux routiers sont normalisées par la législation internationale en vigueur, et ils ont soit la forme d'un cercle, d'un triangle, d'un hexagone, etc. Les différents contours de l'image obtenus après segmentation seront ainsi analysés par une approche structurelle ou globale afin d'extraire les formes présentes dans l'image. Ces formes extraites seront comparées avec les formes connues des panneaux routiers. La difficulté de l'opération de détection par cette méthode, résulte de l'effet de perspective présent dans l'image et qui peut déformer les objets d'où le conflit possible lors de l'identification avec d'autres objets ayant la même forme. [2,8]

Plus récemment, ont vu le jour des approches hybrides exploitant conjointement les informations données par la couleur et par la forme. Ces approches constituent la troisième

catégorie de travaux menés dans le domaine de la détection des panneaux routiers. La couleur est exploitée pour réduire les zones de recherches de la présence de panneaux routiers dans les images analysées et ensuite extraire l'information sur la forme des panneaux. [21]

Nos travaux de recherche dans le domaine de la détection des panneaux routiers s'inscrivent dans le cadre des travaux regroupés dans la première catégorie décrite précédemment. Nos travaux ont abouti à la mise au point d'une méthode de segmentation basée sur le paramètre couleur de l'image étudiée. Le principe de cette approche est de diviser l'image en régions connexes homogènes dont le critère d'homogénéité est la couleur.

Les méthodes de détection par approche couleur sont basées sur le calcul des attributs de l'image (valeurs des pixels) de façon à pouvoir regrouper ces informations en classes et pouvoir ensuite, définir des ensembles connexes de pixels. Un objet quelconque de la scène se retrouve donc dans un ou plusieurs de ces ensembles connexes. Cette approche a l'avantage de permettre l'identification de la classe des objets formant l'image à partir de sa classe de couleur. Néanmoins, cette approche présente l'inconvénient d'être sensible aux variations de la luminosité des scènes analysées. En effet, un objet de l'image peut se trouver, après classification, dans une classe ou une autre selon l'éclairage de la scène, et, d'autre part, deux objets ayant des valeurs de chrominance proches, ne seraient pas discernés.

Compte tenu du champ d'application de nos travaux, notre méthode de segmentation des images doit tenir compte d'un certain nombre de contraintes:

- Les images à analyser peuvent avoir été enregistrées dans des conditions climatiques difficiles ou à n'importe quel moment de la journée, donc variation de l'éclairage de la scène.
- Notre méthode devant faire partie d'une application intégrée dans un système embarqué travaillant en temps réel, la solution recherchée doit être peu gourmande en temps de calcul.
- Bien que les couleurs de panneaux de signalisation routiers soient bien définies, il n'empêche que d'autres objets ayant des couleurs proches peuvent être présents dans les images étudiées et on doit pouvoir les discriminer.
- Les panneaux routiers à détecter peuvent correspondre à plusieurs petites régions dans l'image à analyser. De ce fait, notre méthode de segmentation doit pouvoir séparer les petites régions et les grandes régions de l'image.



### 3. Nos contributions

Pour ce sujet, nous présentons des solutions pour surmonter les contraintes énumérées précédemment. Notre approche est basée sur une technique de segmentation simple permettant une classification des pixels de l'image à partir d'une analyse des histogrammes couleurs 2D.

L'histogramme couleur d'une image représente une estimation de la loi de densité de probabilité d'apparition des couleurs dans l'image [22]. Les méthodes de segmentation basées sur les histogrammes couleurs sont des approches de classification probabilistes qui reposent sur une analyse globale des pixels à partir d'une distribution des couleurs de l'image dans l'histogramme. L'étude de l'histogramme couleur met en évidence les différentes couleurs des pixels de l'image. L'histogramme couleur montre l'existence de la quantité de pixels d'une couleur donnée dans l'image. Les pixels les plus nombreux correspondant à une couleur donnée forment des modes (maximums) dans l'histogramme. Ces modes sont séparés par des zones où le nombre de pixels est moins important (vallées). La construction des classes de pixels revient donc à détecter les modes de l'histogramme.

L'histogramme est un outil statistique très puissant et utilisé très souvent pour étudier et analyser en particulier les images monochromatiques et les images en différents niveaux de gris. Pour les images monochromatiques, l'histogramme est une fonction unidimensionnelle qui lie chaque valeur de niveau de gris au nombre de pixels portant cette valeur dans l'image. Les algorithmes de calcul des histogrammes de ces images aboutissent à une structure unidimensionnelle facile à exploiter. Par contre, les pixels dans les images couleurs ont des valeurs vectorielles : à chaque pixel de l'image est associé un vecteur ayant comme composantes trois valeurs conformément à l'espace de représentation de la couleur choisi. De ce fait, les histogrammes des images couleurs sont des fonctions qui associent à chaque point de l'espace couleur le nombre de pixels de l'image portant cette valeur couleur. Les algorithmes de calculs des histogrammes couleurs, nécessitent donc des structures de données tridimensionnelles pour tenir compte de la représentation tridimensionnelles de la couleur. Il en résulte que le calcul des histogrammes des images couleurs est beaucoup plus long que pour celui des images monochromatiques, et que ces histogrammes occupent beaucoup plus de place mémoire dans un ordinateur. De plus, des travaux scientifiques ont montré que la détection des modes de l'histogramme des images couleurs nécessite des calculs complexes.

Toutefois, les algorithmes de segmentation des images couleurs qui exploitent l'histogramme couleur tridimensionnel sont rares et ils ont connu peu de réussite [22, 23].

Des auteurs ont contribué à réduire la complexité des opérations de calcul des histogrammes tridimensionnels 3D des images couleurs, en déterminant trois histogrammes unidimensionnels 1D correspondant chacun à l'une des trois composantes couleurs [19, 23]. Cette approche simplifie le calcul et le ramène à une analyse de trois histogrammes monodimensionnels. Toutefois, cette technique présente un inconvénient majeur dans la mesure où elle ignore l'information de corrélation entre les composantes couleurs, et cela se traduit par l'apparition de certains modes dans les histogrammes unidimensionnels.

D'autres auteurs, et toujours dans le but de simplifier les calculs, ont proposé le calcul des histogrammes bidimensionnels 2D. Ces histogrammes sont obtenus en ignorant une composante couleur, et de ce fait, ils permettent d'avoir un renseignement sur la corrélation entre les différentes composantes couleurs de l'espace couleur choisi. Cette technique est une solution intermédiaire entre celles utilisant les histogrammes monodimensionnels et celles utilisant les histogrammes tridimensionnels.[24]

Dans notre travail, nous avons choisi d'exploiter les histogrammes 2D pour concevoir notre approche de détection des panneaux de signalisation routiers. Nous avons exploité les histogrammes 2D des images couleurs en appliquant sur ces histogrammes une analyse mathématique basée sur la théorie des ondelettes. La corrélation partielle entre les composantes couleurs représentée par ces histogrammes 2D suffit pour faire une étude de l'image selon deux composantes couleurs ou pour faire un choix adéquat d'un espace couleur pour l'analyse.

Dans le domaine de l'imagerie numérique, la théorie des ondelettes a prouvé son efficacité et elle a conduit à des approches de traitement des images numériques puissantes, notamment pour la conception des méthodes de filtrage et de compression (norme JPEG2000). Nous avons exploité le fait que l'analyse des images numériques avec les ondelettes est étroitement liée à leurs représentations multi-échelles. Chaque étape d'analyse par les ondelettes appliquée sur une image conduit à une approximation de l'image dans un niveau de résolution plus grossier, et elle génère des coefficients de détails correspondant aux hautes fréquences dans l'image [25,26]. Ces hautes fréquences correspondent à des discontinuités dans les images analysées décrivant des pixels isolés ou des pixels de contours.

L'application de la théorie des ondelettes à l'histogramme aboutit à plusieurs histogrammes correspondant à plusieurs approximations de différents degrés de résolution. Cette représentation multi-échelle de l'histogramme, permet de mieux suivre l'évolution des modes présents dans l'histogramme et d'identifier ceux qui définissent des régions d'intérêt dans les images analysées.

Les domaines modaux sont définis à partir des valeurs maximales des histogrammes et caractérisent une région homogène dans l'image source [22]. Cependant, plusieurs valeurs pics apparaissent généralement dans l'histogramme représentant des pixels isolés dans l'image, ou une faible régularité de la distribution des couleurs dans une même région de l'image. Les approximations multi-échelles des histogrammes 2D avec l'analyse par les ondelettes permettent d'éliminer les pixels isolés et de contrôler les différentes valeurs des maximas de l'histogramme avec une précision maîtrisée.

Les modules maximas sont des grandeurs multi-échelles calculées à partir des coefficients d'ondelettes. Ces grandeurs caractérisent les détails des approximations générés en passant d'un niveau d'analyse à un niveau plus grossier. Pendant la décomposition en ondelettes de l'histogramme 2D, le nombre de pics diminue au fur et à mesure que le niveau de décomposition en ondelettes (échelle) augmente et l'histogramme devient ainsi de plus en plus régulier. Les modules maximas commencent à avoir des valeurs relativement stables à partir d'un certain niveau de décomposition et ne changent pas énormément si on passe à un niveau de résolution supérieur. Les approximations dans les niveaux de résolution générant peu de détails sont les approximations utilisées pour calculer les pics (maximas) significatifs de l'histogramme. Ces pics sont utilisés dans un processus de classification des pixels de l'image pour identifier les différentes régions homogènes de l'image et aboutir à une segmentation.

Cette approche de segmentation a été appliquée sur des images tests. Nous avons obtenu une segmentation appropriée des différentes régions de l'image pour la détection des panneaux de signalisation routiers. Notre choix s'est porté sur l'espace couleur  $La^*b^*$  qui est le plus approprié. Des critères d'évaluation quantitative ont été développés afin de mesurer la fiabilité de l'approche que nous avons développée.

## II. Mise en correspondance stéréoscopique

### 1. Avant-propos

Le but d'un système de vision est de simuler le mieux possible le système de vision humain qui est capable de voir et de reconnaître son entourage. La mise en correspondance stéréoscopique est l'une des techniques qui visent à enrichir les fonctionnalités d'un système de vision artificielle et reproduire les fonctions de la vision naturelle. Cela se résume dans la fonction de description d'une scène dans un milieu naturel qui est généralement en 3D. L'idée de base est de déterminer la position d'un point dans l'espace à partir d'une ou plusieurs images fournies par le système d'acquisition.

La vision 3D vise à calculer la profondeur des différents points de la scène à partir du plan image ; cette action constitue un des objectifs de la vision intelligente, et c'est dans ce sens que la stéréoscopie intervient comme une technique utilisée pour reconstruire une scène 3D à partir des images prises sous des angles de vue différents.

### 2. Problématique de la mise en correspondance

Apparier les indices visuels extraits des images stéréoscopiques est une étape très importante dans le processus de la vision 3D. Ce processus réunit les primitives extraites d'une image de gauche et une image de droite pour réaliser une reconstruction 3D d'une scène. La reconstruction 3D est basée sur l'estimation d'une grandeur mesurant la distance entre les positions des primitives appariées et la construction d'une carte de disparités montrant la scène en mettant en évidence la profondeur des objets de la scène. La carte de disparités générée par le processus de mise en correspondance est utilisée pour estimer la carte de profondeur et reproduire une vue 3D de la scène.

La problématique a été traitée dans plusieurs travaux de recherches sans parvenir à une méthode générale [10, 11, 12]. Plusieurs approches existent donc dans le domaine et qui exploitent généralement les primitives des images stéréoscopiques telles que les points d'intérêt, les contours ou les régions. Toutefois, les différents types de primitives sont exploités par l'intermédiaire des mêmes caractéristiques qui sont soit photométrique, soit géométrique ou morphologique.

Les techniques de mise en correspondance restent difficiles en raison de la multiplicité des paramètres essentiels, principalement la pertinence des attributs d'appariement considérés.

Dans le processus de mise en correspondance, la difficulté de trouver les bonnes primitives de correspondance est due à des problèmes potentiels dans un processus de vision 3D tels que les occultations, les contractions, les variations de luminance... Tous les algorithmes proposés sont adaptés pour contrôler au moins une partie de ces problèmes.

En stéréovision, il y en a deux grandes classes d'algorithmes. La première est dite classe des algorithmes globaux, et qui prennent la totalité de l'image et traitent le problème d'estimation de la correspondance comme un problème de minimisation d'une fonction de cout globale représentant les erreurs de mise en correspondance. La seconde est dite classe des algorithmes locaux, et ils sont reclassés en une catégorie des algorithmes basés sur les régions des images et une autre catégorie basée sur les primitives des images. Les algorithmes basés sur les régions utilisent la corrélation de deux fonctions images sur les régions définies localement, alors que les algorithmes basés sur les caractéristiques établissent la correspondance entre les primitives des images stéréoscopiques.

### 3. Nos contributions

Nos contributions dans ce sens s'inscrivent dans une nouvelle classe d'algorithmes [12,46], et ils sont basés sur le formalisme mathématique des ondelettes. En effet, la mise en correspondance est le résultat d'une exploitation multi-échelle des modules maximaux des ondelettes et des orientations de ces modules calculés à partir des coefficients d'ondelettes. Cette classe d'algorithme constitue une solution intermédiaire entre les solutions locales et les solutions globales, et elle combine les techniques des approches locales et des approches globales. Dans la méthode proposée, nous utilisons le concept d'analyse multi-résolution pour apparier les primitives des deux images en exploitant leurs décompositions dans différents niveaux d'échelles.

Nous présentons durant cet axe notre méthode de mise en correspondance basée sur les modules maximaux des ondelettes. Cela est fait en exploitant les chaînes maximaux calculées à partir de ces modules maximaux et qui nous ont permis d'affiner la recherche des correspondances dans les deux images. En effet, à base de la transformée en ondelettes des images stéréoscopiques, nous construisons des cartes de modules des coefficients d'ondelettes et des cartes de phases (orientations) dans les différentes échelles à base desquelles on extrait les modules maximaux. À partir des modules maximaux qui persistent à travers les échelles on construit les chaînes maximaux multi-échelles. De ce fait, les points formant ces chaînes

maximas dans le plan espace-échelle vont être considérés comme des points d'intérêt dans le processus d'appariement. Les informations multi-échelles disponibles, vont permettre de chercher sous certaines contraintes géométriques, pour chaque point d'intérêt de l'image gauche son meilleur correspondant parmi les points constituant les chaînes maximales de l'image droite. Les résultats expérimentaux ont montré que le nombre de correspondance diminue au fur et à mesure que l'échelle augmente. Dans plusieurs essais, on obtient le correspondant unique par exploration des chaînes maximales de l'échelle la plus fine vers l'échelle la plus grossière avec un temps de calcul raisonnable.

Nous nous sommes basés sur le chaînage des modules maximales (construction des chaînes maximales) des ondelettes à travers les échelles et nous réalisons un appariement des différents points formant les chaînes pour trouver les bons correspondants dans l'image droite pour chaque extrémité de chaîne dans l'image gauche. La mise en correspondance est effectuée par estimation d'un test de similarité entre les points de chaque chaîne de gauche et les points des chaînes droite satisfaisant la contrainte géométrique épipolaire et d'orientation. La carte de disparités générée a un caractère multi-échelle et elle est construite selon un traitement simultané en espace et en échelle. L'aspect multi-échelle a permis de perfectionner la carte de disparités en permettant l'amélioration de la sélection des candidats potentiels à travers un critère multi-échelle basé sur une corrélation normalisée et une distinction géométrique. *Mallat et al* [47, 48,49] montrent que les extrémités des chaînes maximales coïncident avec les singularités de l'image caractérisées par leurs coefficients de Lipchitz, et avec un bon choix de la méthode de construction des chaînes maximales, l'algorithme d'appariement proposé prévoit non seulement la précision des résultats, mais il converge aussi rapidement.

### III. Conclusion

Nous nous proposons dans le présent travail de donner des réponses aux difficultés rencontrées lors de l'identification des panneaux de signalisation routiers dans une scène. Ces difficultés se ramènent essentiellement à un problème de luminance et de chrominance au niveau de certaines régions d'intérêt dans l'image. L'idée exploitée est d'utiliser les histogrammes couleurs 2D de l'image pour mettre en évidence l'information couleur. Nous justifions le choix d'un espace de couleurs adéquat pour notre méthode de détection et nous montrons qu'un choix judicieux favorise la mise en évidence au niveau de l'image des deux informations de chrominance et de luminance. Les histogrammes couleurs 2D générés par

notre approche, peuvent être considérés comme des histogrammes d'une image 2D sur laquelle une analyse par la théorie des ondelettes sera faite.

Nous avons entrepris une analyse multi-échelle de l'image 2D, basée sur l'utilisation de la transformée en ondelettes, qui nous a permis d'utiliser les coefficients des ondelettes pour extraire les modes (extremums) des histogrammes. Cela nous a permis de développer une méthode de détection de panneaux routiers dont les résultats sont très satisfaisants.

Nous avons aussi, dans le cadre de ce travail, utilisé l'analyse multi-échelle de la théorie des ondelettes pour développer une technique de mise en correspondance entre deux images permettant une reconstitution de la vision 3D. Cette technique est basée sur la représentation des images en modules maximaux des coefficients des ondelettes et exploite le caractère multi-échelle des contours couleur pour caractériser des points d'intérêts et ensuite, réaliser une mise en correspondance. En effet, nous exploitons les modules maximaux des ondelettes comme des points caractéristiques et nous incluons l'information multi-échelle de ces modules maximaux dans le processus d'appariement des pixels.

# **CHAPITRE 2 :**

# **SYNTHESE DES OUTILS**

# **UTILISES**



## I. Introduction

Les outils de traitement et d'analyse des images numériques sont très nombreux et variés. Les différentes approches mathématiques ont conduit à une variété de méthodes et outils en analyse et traitement des images numériques. Le principal objectif de ces méthodes est de donner une interprétation sémantique du contenu de l'image [13,14,15].

Ces méthodes de traitement des images numériques sont des solutions qui ont été développées pour des applications particulières. Il en découle que le choix de la méthode à utiliser pour analyser une image dépend de plusieurs facteurs : le champ d'application, la nature des images traitées, la performance des algorithmes de calcul,...

Le traitement des images numériques est une suite d'étapes. Le processus préliminaire consiste d'abord à préparer les images en améliorant leur qualité en se débarrassant du bruit provenant souvent des conditions d'acquisition. Après ce prétraitement, l'image sera segmentée afin d'identifier les objets nécessaires à l'interprétation de son contenu. La dernière étape consiste à extraire les objets de l'image, et préciser son contenu sémantique.[13]

Dans le domaine de la vision par ordinateur, on trouve plusieurs travaux réalisés pour le développement des chaînes de traitement des images numériques et qui traitent essentiellement la segmentation. La segmentation d'images est une étape essentielle car elle conditionne la qualité de l'interprétation future des éléments composant l'image. C'est un partitionnement de l'image en régions correspondant aux objets d'intérêt dans l'image. En fait, le système visuel humain effectue cette tâche naturellement. Automatiser la segmentation en vue de simuler l'œil humain est une tâche extrêmement complexe qui constitue un véritable challenge dans le domaine d'analyse des images. Plusieurs approches existent pour résoudre ce problème et la majorité des algorithmes sont basés sur deux principes : recherche de régions homogènes selon un critère dit critère d'homogénéité (couleur, texture,..) ou recherche de transitions dans le plan image pour calculer les frontières des objets.[13, 14]

Les méthodes de segmentation basées sur une recherche de régions homogènes se différencient par le fait qu'on traite le problème soit dans le plan image soit dans l'espace couleur. La couleur est souvent utilisée comme critère d'homogénéité privilégié. Le problème qui se pose pour ces méthodes de segmentation basées sur la couleur est la dimension 3D de la couleur. A. TRÉMEAU et als [27] divisent ces approches en trois classes selon la dimension couleur utilisée :

- L'approche dite marginale qui traite les composantes couleurs une par une puis fusionne les résultats.
- L'approche scalaire qui fusionne les composantes couleurs en une seule composante scalaire.
- L'approche vectorielle qui traite directement les composantes couleurs dans la forme vectorielle.

L'avantage de l'approche scalaire est la simplicité algorithmique. On se ramène par cette approche à un problème de segmentation des images en niveau de gris. L'inconvénient majeur de cette méthode est le non prise en considération des corrélations des composantes couleurs.

Certains auteurs proposent une approche intermédiaire dite bi-marginale [24] qui considère les composantes couleurs deux par deux. L'idée de cette approche est de conserver l'aspect marginal en préservant partiellement la corrélation entre les composantes. L'analyse est faite sur les trois combinaisons bi-composantes possibles, chacune de ces combinaisons est traitée suivant un schéma marginal de classification et de simplification. Les segmentations bi-marginales sont ensuite fusionnées par intersection, et le résultat est simplifié par regroupement de régions utilisant les informations bi-composantes comme trois sources d'informations indépendantes.

Les approches de segmentation par recherches de régions homogènes dans les espaces couleur peuvent utiliser des approches métriques (k-means par exemple), ou utiliser des approches probabilistes basées sur les histogrammes couleurs comme outil de classification, ou combiner les deux. L'atout des approches probabilistes est qu'elles ne nécessitent pas forcément une connaissance préalable sur les images à segmenter. En effet, ce sont des approches globales qui analysent la distribution statistique de la couleur dans l'intégralité de l'image en exploitant les histogrammes couleurs. La majorité des approches qui existent exploitent les histogrammes monodimensionnels (1D) au lieu d'exploiter les histogrammes tridimensionnels (3D). Bien que les utilisations de l'histogramme 1D bénéficient de la simplicité des calculs, une zone de couleurs présente dans l'histogramme 3D n'est pas toujours présente dans chacun des histogrammes 1D des composantes couleurs. Les méthodes qui exploitent directement un histogramme 3D ont le problème de la répartition spatiale des données, qui peuvent être très dispersées, et ont le problème de la complexité des algorithmes d'analyse [22, 24, 23].

En s'attachant à une approche d'analyse bi-marginale des images couleurs, une alternative intéressante réside dans l'exploitation des histogrammes 2D représentant les composantes couleurs deux par deux [23]. Ces histogrammes 2D réduisent le caractère très dispersé des points couleurs dans les histogrammes 3D, et simplifient la recherche des classes de pixels en préservant partiellement les interactions colorimétriques.

Un autre problème inhérent aux images couleurs est celui de la représentation colorimétrique des pixels. En effet, une image couleur peut être représentée dans différentes espaces couleurs, et chaque espace couleur présente ces propriétés particulières. Plusieurs auteurs ont montré l'importance d'un choix judicieux de l'espace couleur pour une opération de segmentation. D'ailleurs, ce choix important de l'espace couleur pourra améliorer ou dégrader les résultats [3, 22, 19].

Nous présentons dans la suite de ce chapitre une synthèse des outils de traitement des images couleurs, et on insiste particulièrement sur la segmentation, la colorimétrie et le calcul des histogrammes.

Les images représentent des signaux bidimensionnels et le formalisme des ondelettes offre des outils pour obtenir des transformées très adaptées pour l'exploration d'une image. La théorie des ondelettes s'est développée dès le début des années 80 [37, 38, 39] et son application à l'image a permis de mettre en valeur les structures les plus significatives de l'image. Dans le présent rapport, la théorie des ondelettes est notre outil d'analyse utilisé que ce soit au niveau de la vision 3D, ou au niveau du processus de traitement et d'analyse des images couleurs. Ce chapitre présente la théorie des ondelettes et son exploitation pour le traitement des images numériques, nous insisterons plus particulièrement sur les différentes formulations et techniques qui sont à la base des travaux présentés.

Les coefficients d'ondelettes calculés sur une image renseignent sur la variation locale des intensités autour d'un pixel de l'image, et cela grâce à la représentation des variations autour d'un pixel dans le sens vertical, horizontal et diagonal. Dans ce chapitre, nous allons présenter une autre formulation de la transformée en ondelettes en gardant seulement deux directions. Il s'agit d'une méthode issue du détecteur introduit dans les travaux de Canny [50] et développée par Mallat dans [51]. Nous obtenons ainsi une transformée en ondelettes adaptée, non seulement pour la détection des contours et surtout les points à fortes singularités [52],

mais elle caractérise aussi certaines propriétés de ces contours à savoir leur orientation géométrique et leur comportement multi-échelle.

La stéréoscopie est un processus de reconstruction de la profondeur d'une scène à partir de deux images de la même scène. En effet, nous nous basons sur le fait que la profondeur de la scène crée une disparité géométrique sur les deux images acquises selon deux angles différentes. Dans ce chapitre, nous présentons une introduction à la stéréoscopie en montrant l'importance de la mise en correspondance dans ce processus. Ces présentations sont très utiles pour la bonne compréhension de la méthode de mise en correspondance proposée et pour l'évaluation des résultats obtenus.

## II. Images numériques couleurs

Le terme image désigne des significations qui peuvent aller de la reproduction de la représentation d'un être ou une chose, à la représentation mentale que l'on fait de quelque chose ou à un concept plus physique lié à la reproduction d'un objet matériel par un matériel optique.

D'un point de vue mathématique, une image peut être regardée comme une fonction de deux variables réelles  $I$ . Avec  $I(x, y)$  l'amplitude de l'image (luminance par exemple), à la position définie par les coordonnées réels  $(x, y)$ .

La réalité continue des images analogiques, avec des supports  $D \subset \mathbb{R}^2$ , n'est pas adaptée à une utilisation informatique vue l'infinité des points de ce support. De ce fait, L'utilisation numérique des images nécessite un échantillonnage et une quantification (passage à un support  $D \subset \mathbb{Z}^2$ ) pour une approximation discrète de celles-ci. Une image discrète de support  $\mathbb{Z}^2$  ou sous ensemble de  $\mathbb{Z}^2$  peut être représentée par une grille rectangulaire dont les données peuvent être stockées dans une matrice de  $M$  lignes et  $N$  colonnes. La plupart des capteurs produisent de telles images, et des techniques d'échantillonnage et d'interpolation sont utilisées dans ces capteurs pour mapper une image continue dans sa forme discrète et vice versa.

La numérisation des images mène à une structure 2D parfaitement appropriée pour un traitement informatique. C'est un pavage du plan de l'image en éléments appelés pixels dont chacun porte une valeur scalaire dans le cas des images en niveau de gris ou un vecteur de valeurs dans le cas des images couleurs.

En liaison avec la théorie tri-colorimétrique qui stipule qu'une couleur peut être représentée par un système de trois coordonnées (espace couleur), et, de ce fait, une image couleur est une fonction qui assigne trois valeurs (trois coordonnées) à chaque point du support de l'image. Chaque point est ainsi caractérisé par un vecteur de trois dimensions. On peut penser à une image couleur comme trois images monochromatiques représentant chacune une primitive couleur [95]. La figure 4 montre la décomposition d'une image couleur en ses trois plans de primitives couleurs dans l'espace couleur *RGB*.

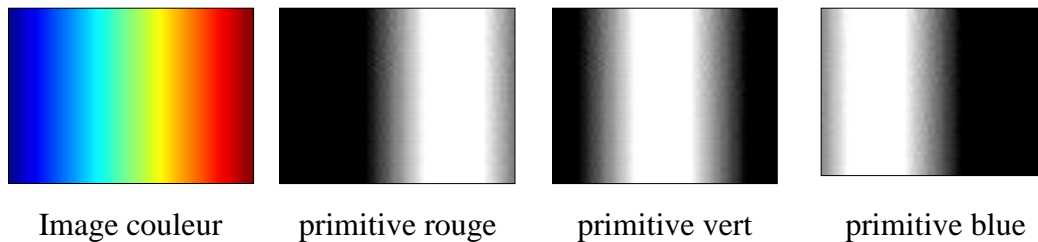


Figure 4 : Images couleur avec ces trois plans dans l'espace *RGB*

## 1. Les systèmes colorimétriques

L'étude des systèmes de représentation de la couleur dans le cadre d'un processus de segmentation des images couleurs a pour but d'améliorer les résultats. D'une part, dans la littérature, les méthodes de segmentation basées sur la couleur tentent de choisir l'espace couleur le mieux adapté pour leurs approches de traiter le problème de segmentation [28]. D'autre part, les méthodes de segmentation des images couleurs basées sur la classification des pixels analysent la distribution des couleurs dans l'espace couleur choisi, et supposent que les régions homogènes de l'image donnent naissance à des zones de forte concentration de valeurs dans cet espace.

### 1.1 La vision humaine

Plusieurs facteurs se réunissent pour donner la couleur d'un objet. Ces facteurs dépendent essentiellement des propriétés de la lumière incidente et des propriétés physiques de l'objet observé. En fait, les propriétés physiques de la lumière ont un impact sur la représentation informatique de la couleur.

### 1.2 La colorimétrie

Pour les systèmes de vision, utiliser la couleur avec son aspect spectrale (c'est-à-dire avec le choix de la description physique de la lumière) peut générer des algorithmes coûteux et avec de faible efficacité [28]. En effet, utiliser un spectre exige d'échantillonner l'intervalle [380nm-830nm] des longueurs d'ondes visibles et utiliser surtout des filtres pour connaître sa décomposition spectrale. La complexité et l'inefficacité sont bien liées au fait que les transitions dans les spectres sont douces et produisent un effet de fondu.

La colorimétrie donne une solution pratique alternative pour représenter la couleur numériquement. Au lieu d'utiliser des spectres complexes, générant beaucoup d'erreurs pour simuler le comportement humain envers la couleur, la colorimétrie utilise des mélanges de couleurs basées sur trois couleurs primaires tout comme les rétines de l'œil humain. La colorimétrie ne simule pas exactement le comportement du cerveau envers la couleur, mais elle s'attache à observer les effets de certaines stimulations lumineuses sur celui-ci et d'en tirer partie pour la définition des représentations colorimétriques.

La perception de la couleur est résultat de la répartition spectrale d'un stimulus. En effet, l'utilisation d'une série de sources lumineuses couvrant chacune un intervalle spectral permet de synthétiser n'importe quelle couleur. L'expérience a prouvé qu'on peut réduire le nombre de sources en supprimant certaines bandes spectrales et recréer un stimulus visuel équivalent à celui produit par le spectre continu, seulement en utilisant trois stimuli de longueurs d'ondes particulières convenablement choisies [29].

En 1931, la Commission Internationale de l'Eclairage (CIE) a adopté des illuminants et des observateurs standards afin de représenter l'ensemble des couleurs [30, 29]. La CIE a prouvé que pour décrire presque toutes les couleurs existantes, il suffit d'utiliser trois variables indépendantes représentant chacune une primitive couleur standard (rouge, vert et le bleu). De ce fait, les espaces de représentation de la couleur sont des systèmes de coordonnées tridimensionnels.

T. YOUNG [83,28] établit la théorie trichromatique de la lumière qui fut par la suite améliorée par H.V. HELMOLTZ partant de l'idée qu'on peut recréer presque toutes les couleurs du spectre visible à partir des primaires rouge, vert et bleu. C'est le modèle dite tri-stimulus très répondu et qui a été utilisé dans le codage de l'espace couleur *RGB*.

Un autre modèle portant le nom de modèle de couleurs opposées fut établi par E.HERING [28] au XIX siècle et qui est basé sur le fait que toutes les couleurs peuvent être décrites

comme contenant du rouge et du vert mais pas les deux en même temps. Ce modèle a organisé la couleur selon deux composantes chromatiques (correspondant respectivement à l'axe rouge-vert et l'axe bleu-jaune) et une composante achromatique (correspondant à l'axe blanc-noir représentant la luminosité).

Faugeras [31] a instauré un nouveau modèle (modèle de Faugeras) résultat d'une combinaison du modèle de tri-stimulus et des couleurs opposées. PRATT dans ces travaux [32] souligne que même si le modèle de tri-stimulus est très utilisé, il reste difficile de lier la perception humaine d'une couleur avec les composantes de ce modèle et propose de parler de teinte, de saturation et de luminance pour décrire une couleur.

### 1.3 Les espaces de représentation de la couleur

Les modèles de représentation de la couleur ont conduit à des systèmes de représentation tridimensionnels. Les systèmes de représentation de la couleur appelés espaces couleurs sont classés selon les modèles de représentation sous jacents. On trouve les espaces de couleurs primaires, les espaces couleurs luminance chrominance, les systèmes perceptuels uniformes et les systèmes d'axes indépendants.

#### 1.3.1 Les systèmes primaires

Ces systèmes se basent sur le modèle de tri-stimulus, ils sont le résultat des expériences réalisées avec un observateur standard adopté par la CIE et qui a été obtenu à partir d'une centaine d'expériences réalisés avec différents observateurs. La base de ces expériences, sont les fonctions colorimétriques standards qui correspondent au rouge, bleu et jaune. La figure 5 montre les résultats obtenus et présente les trois courbes correspondantes aux intensités des primaires RGB nécessaires pour synthétiser une couleur.

L'espace colorimétrique de base associé au modèle de tri-stimulus est le système *RGB*. Ce système est l'une des premières tentatives de quantification de la couleur. L'espace couleur *RGB* est l'espace le plus utilisé dans l'acquisition et le traitement des images couleurs. C'est un espace qui utilise un système de coordonnées cartésiennes et comme le montre la figure 6, la diagonale allant du blanc (1, 1,1) au noir (0, 0,0) (pour des valeurs de pixels comprise entre 0 et 1) est appelé axe de gris ou axe achromatique.

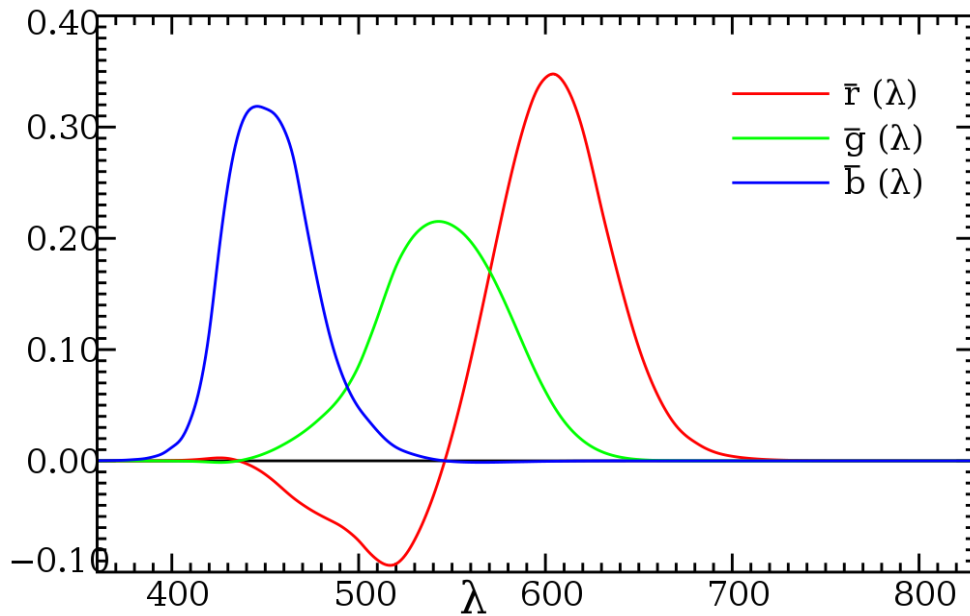


Figure 5 : variation du tri-stimuli en fonction des longueurs d'ondes correspondantes aux primaires *RGB*

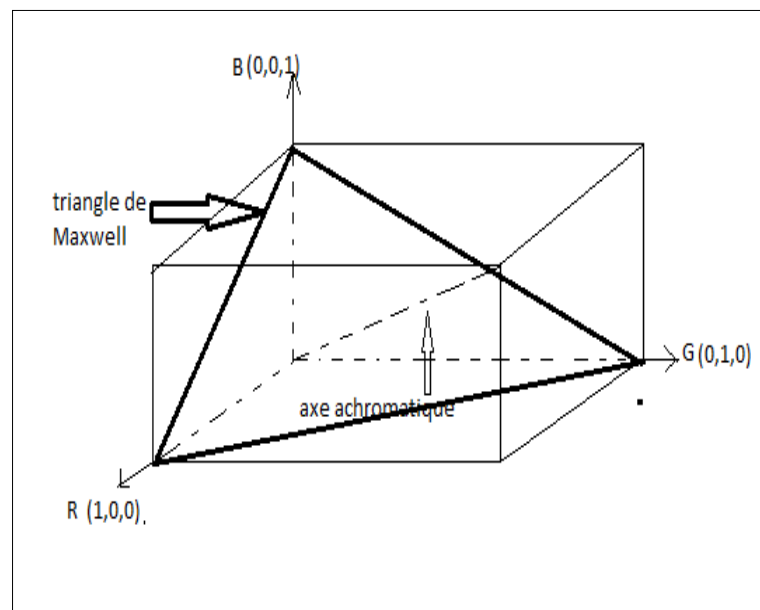


Figure 6 : L'espace de couleur *RGB*

L'utilisation répandue du système *RGB* en imagerie numérique est dû surtout à sa simplicité, il reste toujours un système de référence même s'il présente certains inconvénients. Ces inconvénients sont dus au fait que les valeurs correspondants au rouge peuvent avoir des valeurs négatives. Pour remédier à ces inconvénients, la CIE a créé un système de représentation de la couleur dite espace *XYZ* présentant les composantes *X*, *Y* et *Z* virtuelles obtenues par une simple transformation linéaire.



L'espace *RGB* est généralement la base des acquisitions des images, mais il n'est pas forcément le mieux adapté pour l'analyse. Toutefois, un certain nombre de transformations de cet espace vers d'autres espaces de représentation est utilisé pour une meilleure mise en évidence de l'information couleur.

### 1.3.2 Les systèmes luminance-chrominance

Les espaces luminance/chrominance sont le résultat de l'application d'une théorie dite théorie des couleurs antagoniste dans laquelle on tente de dissocier l'information d'intensité (ou luminance) de celle de chrominance. Faugeras [31] a proposé en 1976 un modèle afin de modéliser le fait que la réponse des cônes à un stimulus de couleur n'est pas proportionnelle à son intensité. Le résultat de ces travaux est le système *AC1C2* à partir duquel on déduit les composantes Saturation, Teinte et Luminance. L'un des espaces couleurs représentant cette famille de système est l'espace HSV (cf. figure 8).

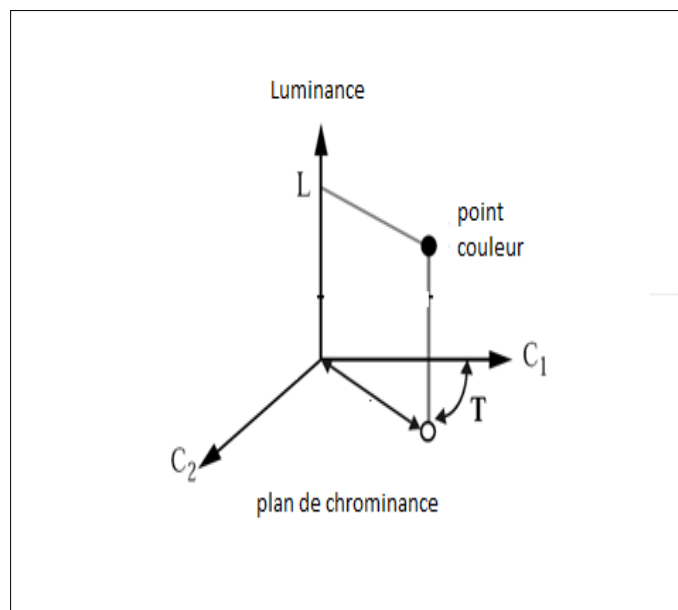


Figure 7 : Les composantes *S*, *T*, *L*

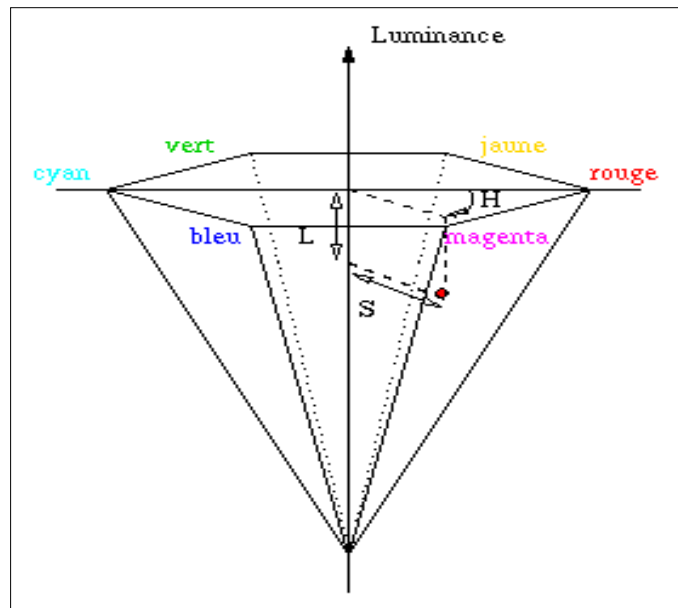


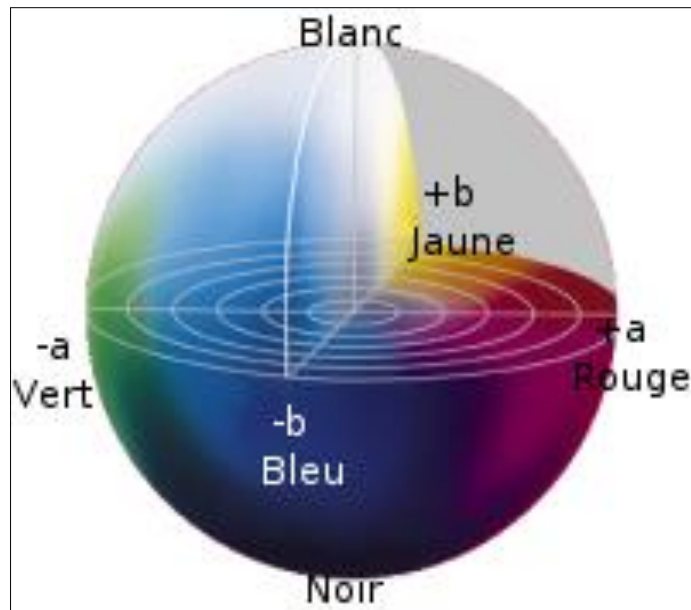
Figure 8 : L'espace couleur HSV

L'intérêt essentiel de cette représentation est qu'elle assure une dé-corrélation de l'information de chrominance de la luminance. Par exemple un rouge et un rose ne se distinguent que sur la composante  $S$  ou  $C$ . Au contraire, dans l'espace  $RGB$ , cette distinction se retrouve sur les trois composantes de base.

### 1.3.3 Les systèmes uniformes

Ces systèmes ont été produits pour uniformiser la représentation de la couleur par rapport à la perception humaine. De ce fait, chaque changement élémentaire en luminance et en chrominance est uniformément perceptible par l'œil humain. Ça permet d'avoir une harmonisation perceptuelle des différences entre les couleurs et la distance euclidienne entre deux couleurs à une valeur proche de la différence perçue par l'œil humain. Deux principaux espaces permettent de décrire ce type de systèmes : le système  $Lu^*v^*$  et le système  $La^*b^*$ [28].

Pour le système  $La^*b^*$ (cf. figure 9), la composante  $L$  décrit la luminance, la composante  $a^*$  décrit l'opposition vert-rouge, et la composante  $b^*$  décrit l'opposition bleu-jaune. Ce système a été conçu pour inclure toutes les couleurs perceptibles, et pour être un système de référence indépendant. Ces propriétés intéressantes de l'espace  $La^*b^*$  le favorisent comme espace d'analyse et de segmentation des images couleurs.

Figure 9 : Espace couleur  $L^*a^*b^*$ 

### 1.3.4 Les systèmes d'axes indépendants

L'un des problèmes de traitement et de la segmentation des images couleurs vient du fait que les espaces couleurs présentent des composantes corrélés. Une première solution dans le cadre d'une méthode de segmentation basée sur la couleur est de choisir l'espace couleur le mieux adapté en étudiant l'effet de la corrélation sur les résultats de la segmentation. En effet, les composantes corrélées et le degré de la corrélation entre ces composantes varient d'un espace couleur à l'autre, et on peut trouver un espace couleur adapté pour certains types d'applications. Une deuxième solution a été proposée par des auteurs, qui ont calculé des composantes indépendantes adaptées aux images traitées en appliquant des algorithmes dites algorithmes d'analyse en composantes principales (ACP). Ces algorithmes se basent sur une transformation dite de KARHUNEN-LOÉVE [33].

*Otha et alts*[34] ont introduit un nouvel espace de couleur  $I_1I_2I_3$  dont le but est de produire une segmentation comme celle basée sur une transformation de KARHUNEN- LOÉVE. Il a réussi à extraire des caractéristiques couleurs par application de la transformation dans des opérations de segmentation. Les caractéristiques les plus pertinentes ont été retenues et représentées sous forme de trois axes de couleur  $I_1$ ,  $I_2$ , et  $I_3$ .

A chacun de ces systèmes de représentation de la couleur correspond un espace tridimensionnel couleur. Un stimulus de couleur est représenté par les valeurs de ses

coordonnées dans cet espace. De ce fait, les valeurs des pixels d'une image numérique sont exprimées selon le système de représentation de la couleur choisi. D'ailleurs, une image numérique couleur peut être représentée dans plusieurs espaces couleurs et une simple transformation matricielle permet le passage d'un espace colorimétrique à un autre.

## 2. Les histogrammes couleurs

### 2.1 Calcul des histogrammes

Notre méthode de segmentation des images couleurs s'inscrit dans le cadre des méthodes de classification probabilistes. Ces méthodes exploitent l'histogramme couleur pour étudier les couleurs présentes dans l'image et calculer les classes des pixels. En effet, pour une image  $I$ , l'histogramme couleur  $H(I)$  est une estimation de la loi de densité de probabilité d'apparition des couleurs dans l'image. C'est une structure tridimensionnelle qui associe pour chaque point couleur dans l'espace couleur considéré le nombre de pixels ayant cette couleur dans l'image [22]. On calcule simplement les cellules de  $H$  avec la formule :

$$H_I(c) = \frac{n_c}{M \times N} \quad (3.1)$$

Avec  $n_c$  : les occurrences d'apparition de  $c$  et  $M \times N$  la taille de l'image.

La figure 10 fait apparaître les résultats de calcul des histogrammes couleurs 3D sur l'image de synthèse  $I$ . Nous avons calculé les histogrammes de l'image dans les espaces couleurs RGB, HSV et La\*b\*. Les cinq régions de l'image forment des nuages de points identifiables dans les histogrammes pour les différents espaces de couleurs.

La figure 11 fait apparaître les résultats de calcul des histogrammes 3D sur une image réelle  $J$ . on voit clairement les nuages de point représentant les régions de l'image et spécialement ceux correspondants aux panneaux routiers.

Les figures 10 et 11 montrent que l'analyse dans l'espace couleur repose sur la présence de nuages de points couleurs bien identifiés lorsqu'il s'agit de régions bien identifiés et séparés dans le plan image. Toutefois, les régions qui ont des couleurs proches donnent naissance à des nuages de points qui se chevauchent, ce qui rend la discrimination de ces régions difficile.

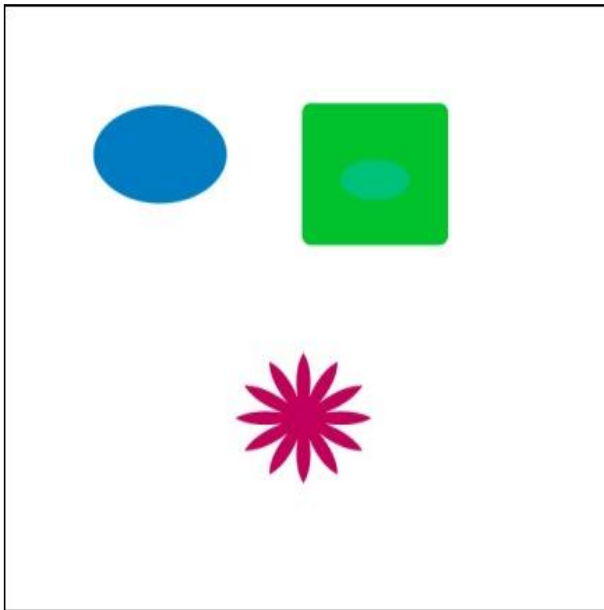
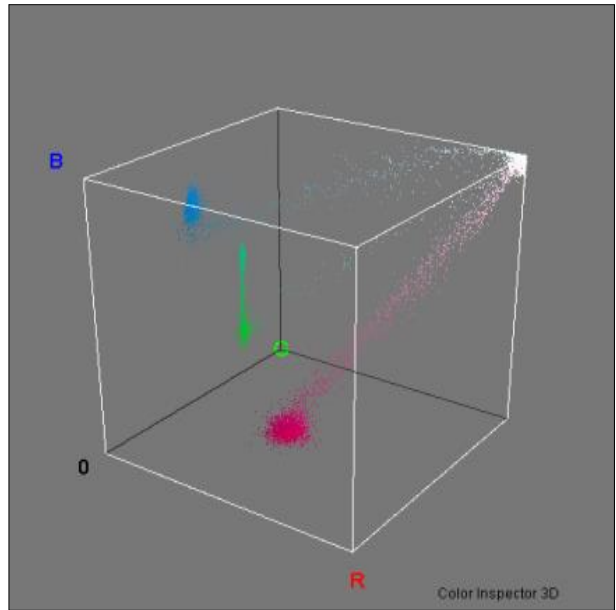
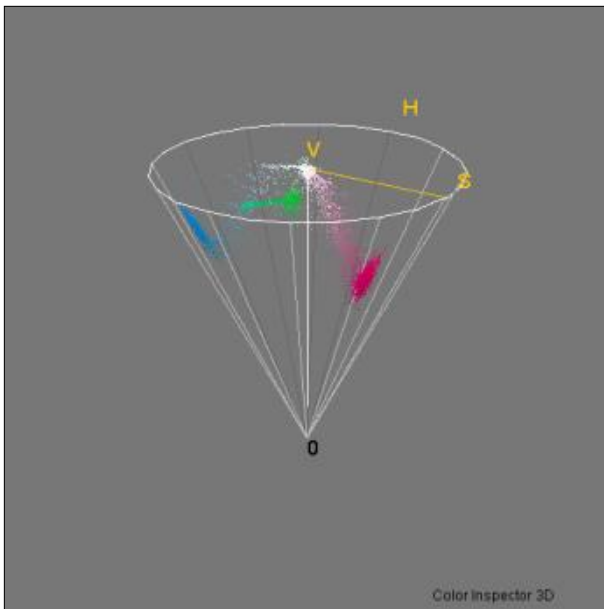


Image I



H(I) dans l'espace RGB



H(I) dans l'espace HSV

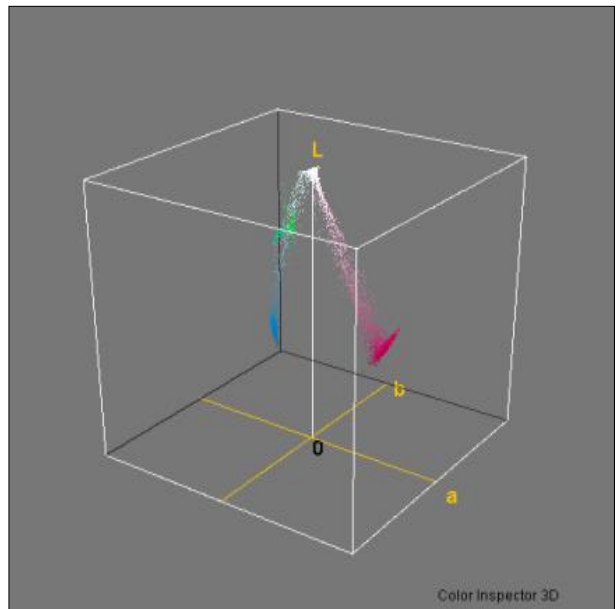
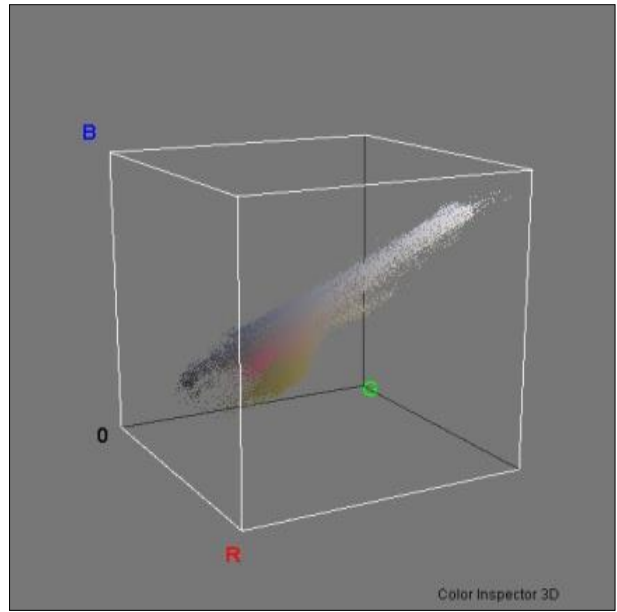
H(I) dans l'espace  $La^*b^*$ 

Figure 10: Image avec quatre régions de différentes couleurs et ces histogrammes 3D dans les espace  $RGB$ ,  $HSV$  et  $La^*b^*$

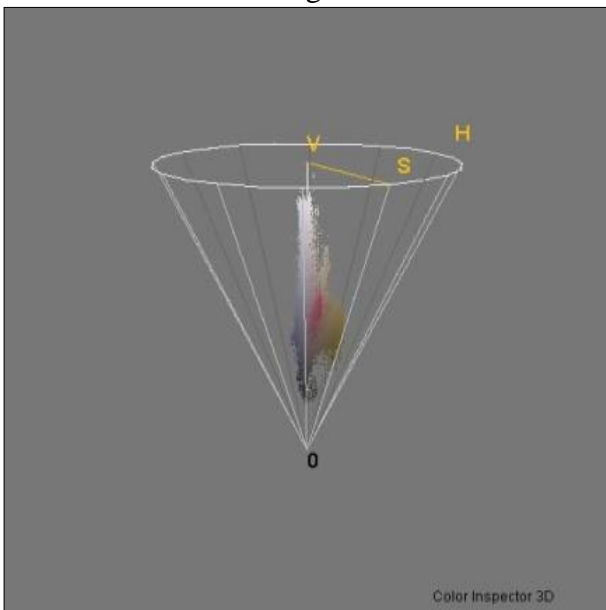
L'histogramme couleur mis en évidence les domaines de couleurs décrivant des régions de l'image. Chaque région donne naissance à une zone de forte concentration de points couleurs caractérisée par un sommet (mode). Il en découle qu'une construction des classes de pixels revient à détecter les modes présents dans l'histogramme. Cette technique de construction des classes de pixels donne de bons résultats lorsque les régions de l'image sont bien séparées et équiprobables.



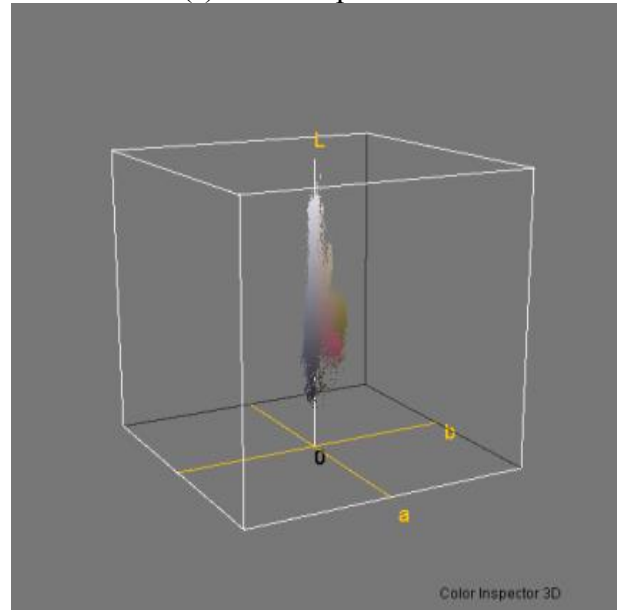
Image J



H(J) dans l'espace  $RGB$



H(J) dans l'espace  $HSV$



H(J) dans l'espace  $La^*b^*$

Figure 11: Image réelle avec ces histogrammes 3D dans les espace  $RGB$ ,  $HSV$  et  $La^*b^*$

La figure 12 montre une image synthèse avec une région de petite taille (région rouge). Le nuage de points couleur généré dans l'histogramme couleur correspondant est difficile à identifier par rapport aux autres nuages de points représentant les autres régions. De ce fait, les régions de l'image de petite taille sont moins représentées par les histogrammes et les modes auxquels elles donnent naissance sont difficiles à détecter.

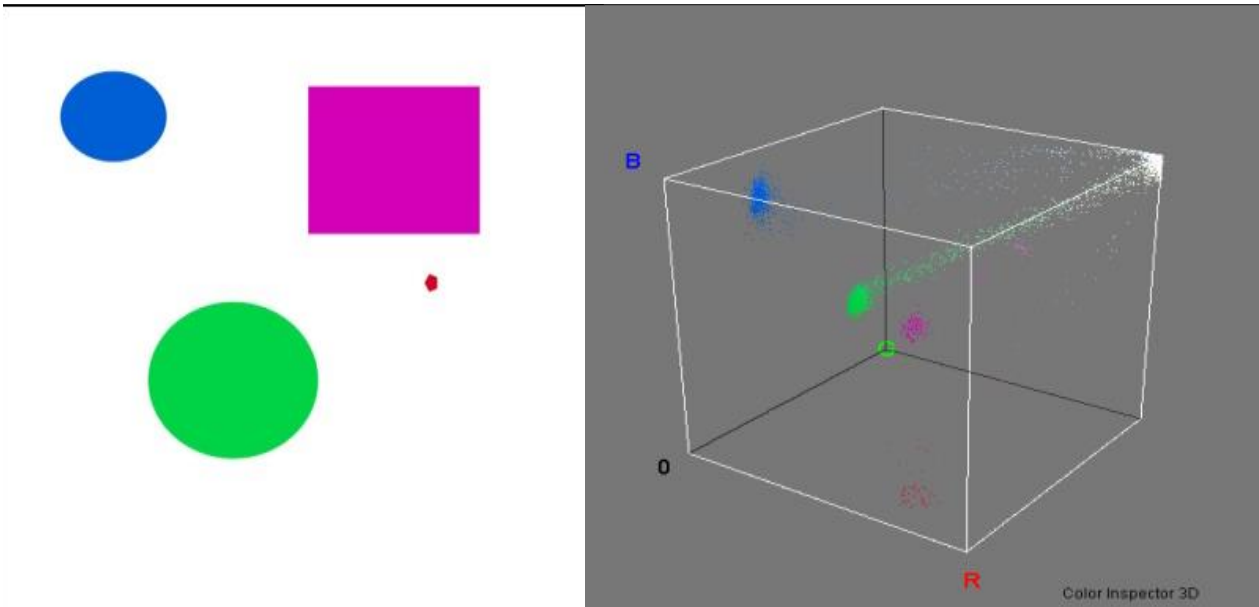


Figure 12 : Une image de régions et son histogramme couleur *RGB*

Les histogrammes tridimensionnels sont utilisés dans le cadre d'une approche de segmentation vectorielles tenant compte les trois primaires de l'espace couleur considéré. Pourtant, l'histogramme 3D est une structure complexe nécessitant des structures de données et des algorithmes de calcul complexes. De plus, les histogrammes 3D souffrent, dans certains cas, de la répartition spatiale des données qui peuvent être très dispersées.

Les algorithmes qui exploitent directement l'histogramme 3D dans des processus de segmentation sont rares. *J.G. Postaire et alts* [35] ont utilisé des techniques de la morphologie mathématique pour rechercher les modes de l'histogramme 3D. La plupart des algorithmes existants exploitent des histogrammes 1D en projetant l'histogramme 3D sur les trois composantes couleurs de l'espace couleur choisi afin de simplifier les calculs.

## 2.2 Projection de l'histogramme

Dans une projection de l'histogramme 3D sur les trois composantes couleurs (*cf. figure13*), l'analyse se réduit à une recherche de pics dans les trois histogrammes 1D de chacune des composantes couleurs (approche marginale). En effet, des méthodes proposent d'exploiter directement les histogrammes 1D en calculant le produit cartésien des intervalles modaux définies par les pics de ces histogrammes [22]. Le produit cartésien des intervalles modaux donne des domaines modaux dans l'espace couleur d'analyse. La classification des pixels se base sur une recherche des pixels qui appartiennent à ces domaines. Une autre

approche pour exploiter les histogrammes 1D d'une image couleur existe et qui propose de segmenter séparément les trois plans correspondants aux primitives couleurs de l'image de l'espace couleur choisi en se basant sur l'histogramme 1D [24]. De ce fait, chaque opération de segmentation tient compte d'une seule composante couleur et donne un plan couleur segmenté. Ensuite, on fusionne les trois plans couleurs segmentés pour avoir l'image entièrement segmentée.

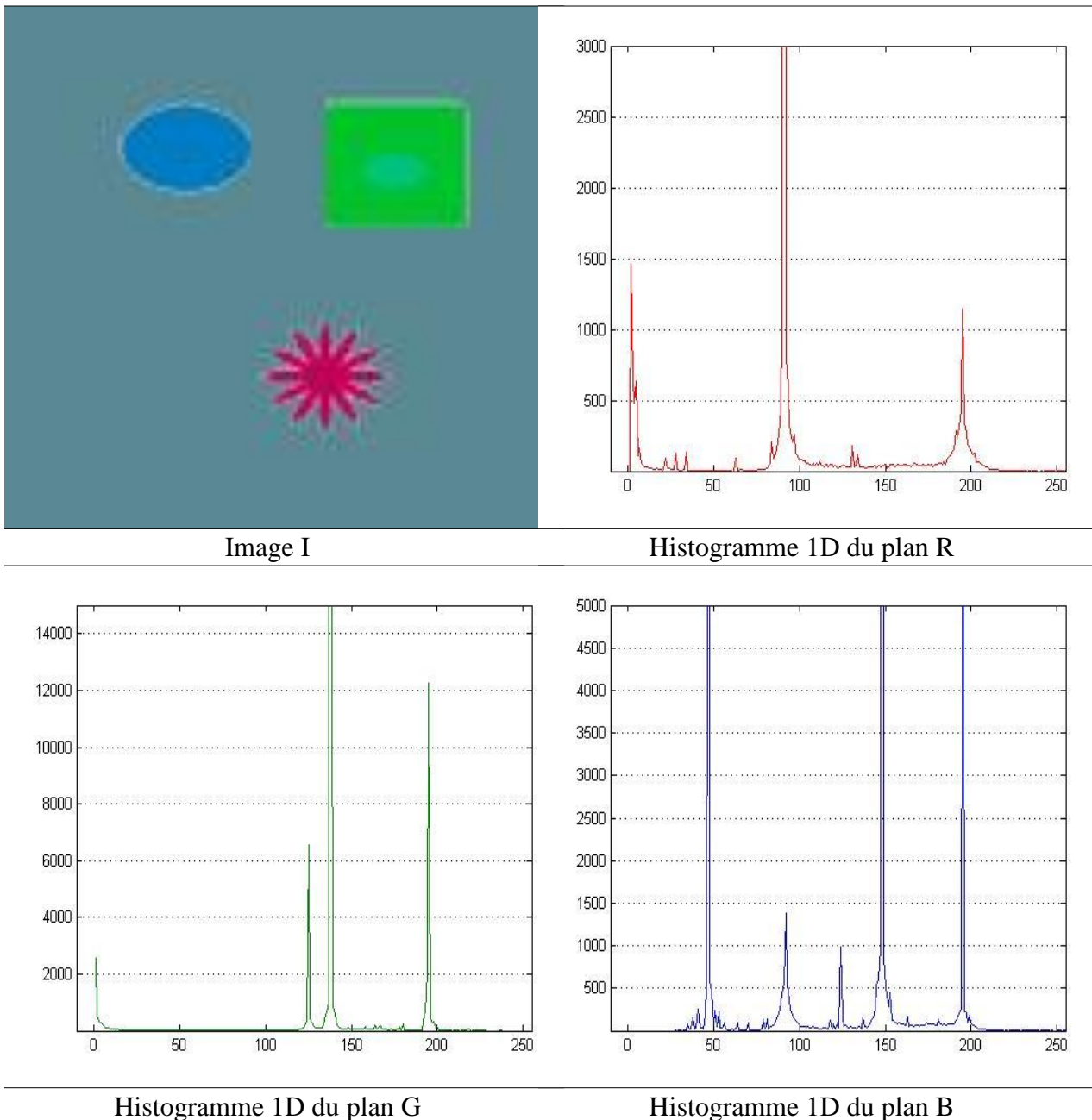


Figure 13 : Histogrammes monodimensionnels dans l'espace *RGB*



La figure 14 présente le schéma d'un algorithme utilisant une approche marginale de segmentation des images couleurs utilisant les histogrammes 1D. L'approche consiste en l'analyse séparée de chaque histogramme en produisant ainsi une segmentation partielle de l'image. Chaque analyse génère une classification du plan image correspondant et la fusion des plans image segmentés produit une segmentation entière de l'image.

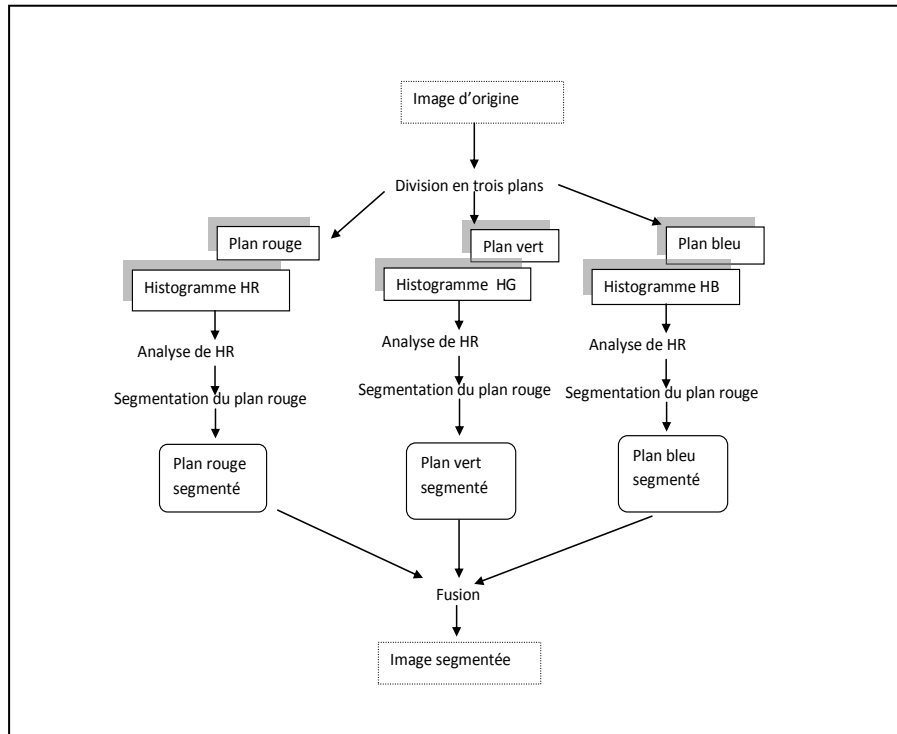


Figure 14 : Approche de segmentation marginale d'images couleur

### 2.3 Les histogrammes bidimensionnels

Les histogrammes 2D peuvent être calculés par projection des histogrammes 3D sur les plans formés par les composantes prises deux à deux. Par exemple, pour l'espace couleur *RGB*, on peut avoir des histogrammes 2D représentant les plans *RB*, *RG* ou *GB*.

De ce fait, on dispose de trois histogrammes 2D qui représentent la corrélation entre les composantes couleurs deux à deux. Toutefois, il suffit d'exploiter deux histogrammes 2D et pour certains types d'images un seul histogramme 2D suffit pour avoir les résultats escomptés. L'avantage d'utiliser les histogrammes 2D est qu'on réduit l'aspect trop dispersé des couleurs dans l'histogramme 3D et on simplifie surtout les calculs pendant la recherche des classes de pixels. Dans [24] une méthode de segmentation basée sur les histogrammes 2D

a été proposée et se base sur la quantification des histogrammes par des classifieurs morphologiques.

La méthode la plus évidente pour calculer les histogrammes 2D est mettre à zéros une composante et calculer le nombre de pixels représentant les deux composantes tenues en considération. Cette méthode donne une bonne représentation de la distribution des couleurs, notamment pour des composantes couleurs avec de faible corrélation. De ce fait, l'utilisation de cette méthode doit se faire avec une analyse dans un espace couleur présentant des corrélations faibles entre ces composantes couleurs.

La deuxième méthode est une projection de l'histogramme 3D sur le plan de deux des trois composantes couleurs. Si on note  $(X, Y, Z)$  le système couleur et  $H(x, y, z)$  l'histogramme 3D d'une image  $I$  de taille  $(N, N)$ , on estime l'histogramme  $H_{2D}(x, y)$ , relativement au plan  $(X, Y)$ , comme la somme des  $H((x, y, z_0))$  pour  $z_0 = 0 \dots N - 1$ , en introduisant si nécessaire comme paramètre de normalisation  $\max_z H(x, y, z)$ .

*Clement et als [36]* proposent une troisième méthode qui décrit l'histogramme 2D d'une image couleur comme une image de résolution spatiale de  $N \times N$ . En effet, les valeurs des pixels de l'image histogramme  $H_{2D}$  sont calculées par une réduction linéaire dynamique de l'histogramme entre 1 et  $\min(P_{max}, N)$ ;  $N$  est le nombre de niveaux d'intensités colorimétriques présentes dans l'image  $I$  et il est supposé le même pour les deux composantes.

Sachant que  $p(i, j)$  indique le nombre de pixels présentant les composants colorimétriques  $i$  et  $j$ , et  $P_{max}$  la valeur maximale de  $p$  et  $P_{min}$  la valeur minimale de  $p$  :

$$H_{2D}(i, j) = \text{round} \left( \left[ \frac{(\min(P_{max}, 255) - 1) \times p(i, j) - \min(P_{max}, 255) \times P_{min} + P_{max}}{P_{max} - P_{min}} \right] \right) \quad (3.2)$$

La figure 15 illustre les résultats de calculs des histogrammes 2D par projection bi-composante de l'histogramme 3D de l'image synthèse  $I$  dans l'espace couleur  $RGB$ . Les trois projections ne contiennent pas évidemment les mêmes informations. En effet, l'histogramme  $RB$  représente les informations de corrélation entre le rouge et le bleu en ignorant totalement la composante vert et, de ce fait, il a identifié clairement les nuages de points correspondant aux régions avec des quantités importantes de rouge et/ou de bleu. Cela reste valable aussi pour le histogramme  $RG$  et l'histogramme  $GB$ .

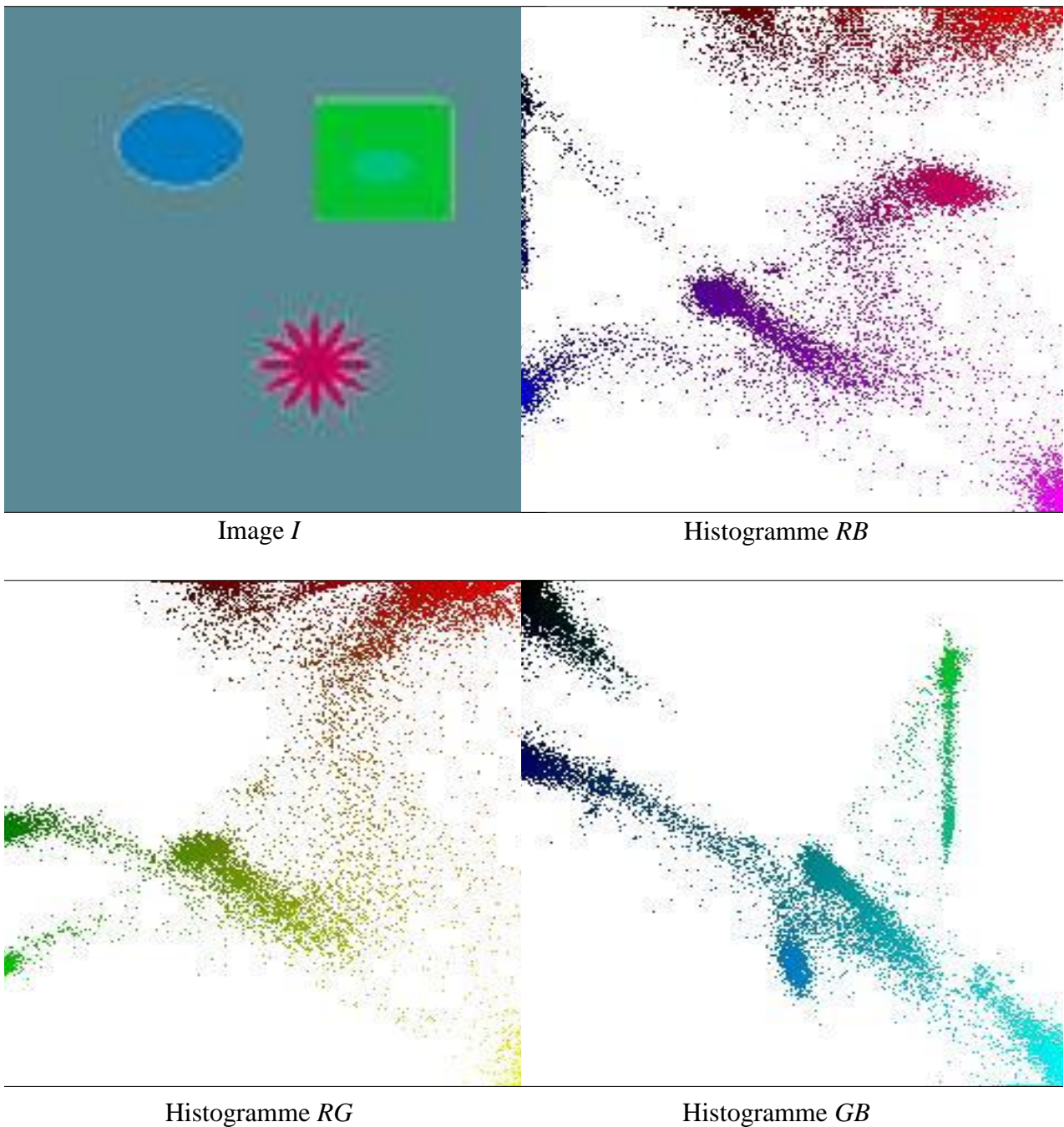


Figure 15 : Image synthèse est ces histogrammes 2D dans l'espace *RGB*

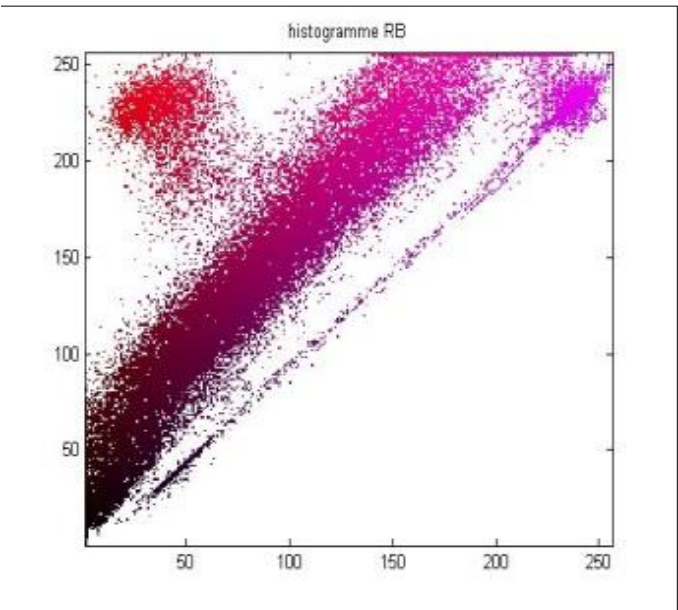
Nous remarquons la présence de régions de points à forte concentration et d'autres à faible concentration dans l'histogramme couleur 2D. Les zones de forte concentration (zones modales) observées correspondent aux régions des objets présents dans l'image.

La figure 16 présente les histogrammes 2D de deux images réelles qui présentent des scènes routières. Pour chaque image nous présentons son histogramme RB dans l'espace RGB et ses histogrammes  $L^*a^*$  et  $L^*b^*$  dans l'espace  $L^*a^*b^*$ . On distingue clairement la zone correspondant au panneau dans l'image et on remarque que cette distinction est d'autant plus

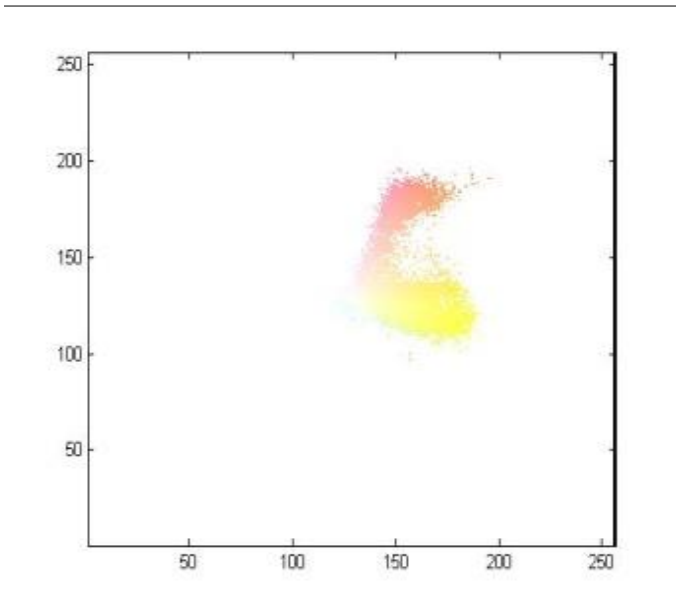
claire en passant du système *RGB* au système *La\*b\**.



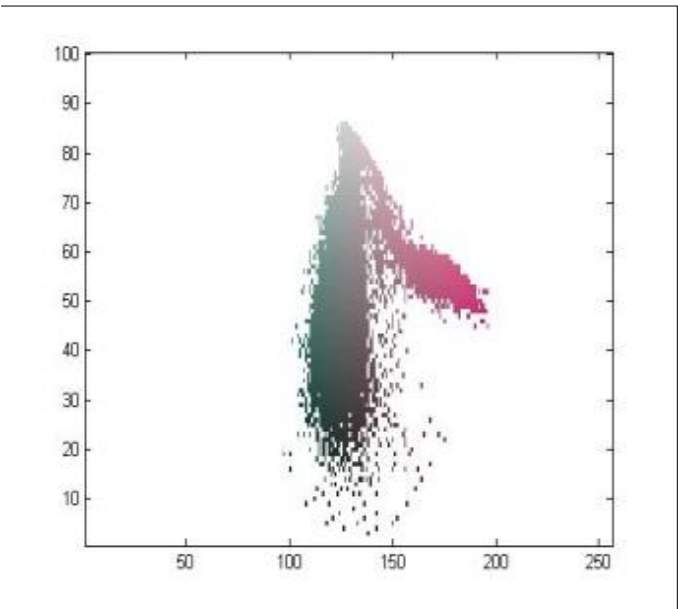
Image 1



Histogramme RB



Histogramme a\*b\*



Histogramme La\*

Figure 16: Zones panneaux dans l'histogramme 2D

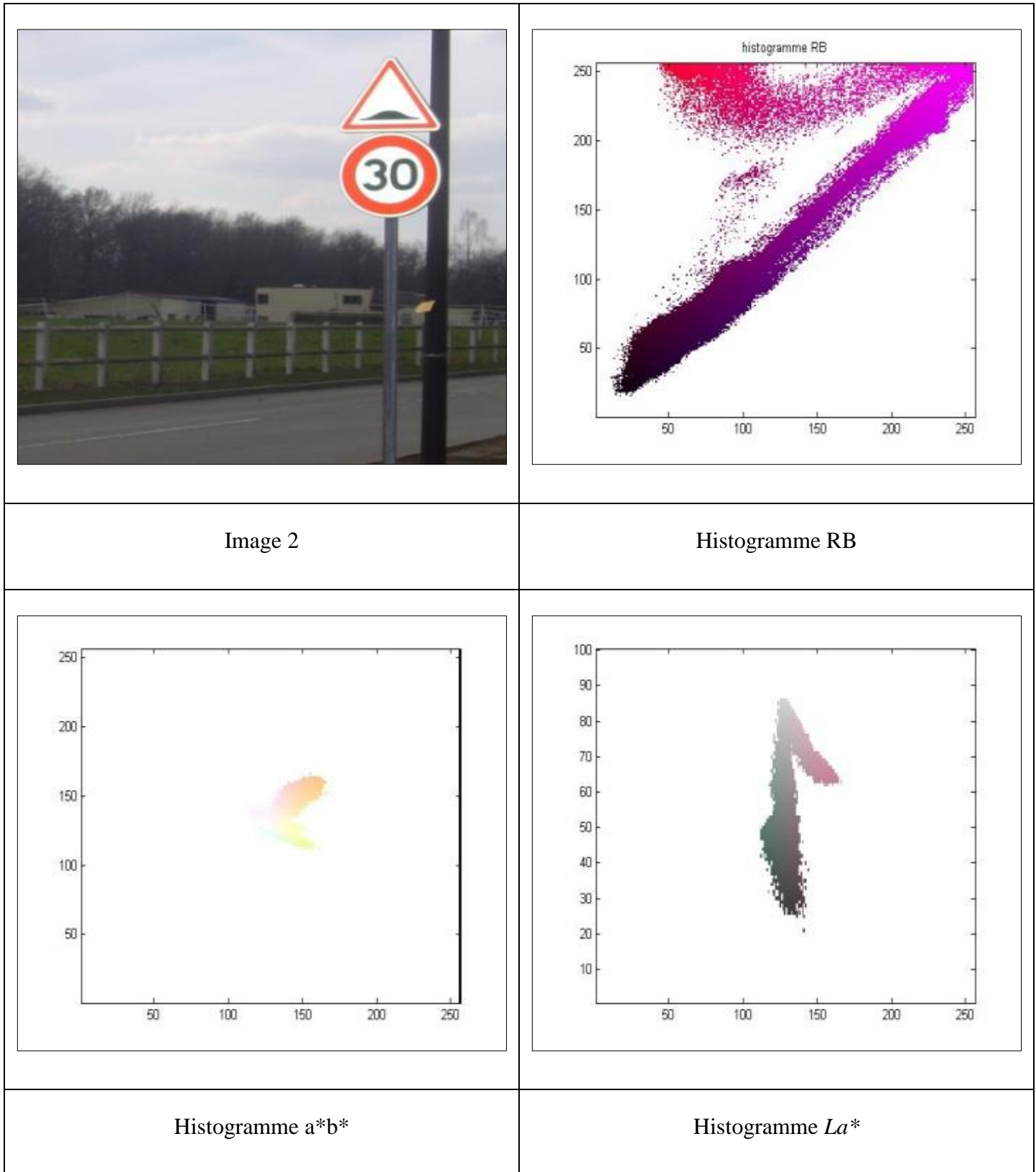


Figure 16: Zones panneaux dans l’histogramme 2D

Les figures illustrent les difficultés liées à l’exploitation des histogrammes couleurs. La présence de plusieurs panneaux ayant des valeurs chromatiques proches se traduit par des

zones de points qui se chevauchent et qu'on ne peut pas facilement distinguer. En effet, Dans l'image 2 de la figure 16, les nuages de points dans l'histogramme 2D correspondant aux deux panneaux contenant la même couleur, sont plus au moins confondus. Une deuxième difficulté illustrée est due au fait que les régions qui correspondent aux panneaux routiers donnent naissance à des petites zones de points dans l'histogramme couleur, alors que les régions de tailles plus importantes sont mieux représentées par l'histogramme et donnent naissance à des zones de points plus claires.

Ces difficultés sont d'autant plus présentes que la variation des paramètres chrominance-luminance des panneaux est forte à cause de l'exposition aux facteurs naturels telle le soleil, les intempéries ou l'éclairage nocturne.

### **III. La vision 3D**

#### **1. Introduction**

La vision est un système complexe et difficile à imiter. De ce fait, plusieurs études et recherches biologiques ont montré l'architecture de ce géant système qui donne le plaisir de voir et d'apercevoir le voisinage. Le développement des machines de calcul et de traitement de l'information a favorisé la production de nombreux travaux scientifiques traitants le problème d'un point de vue quantitatif.

Ainsi, on trouve différents modèles de la vision artificielle ou la vision par ordinateur capables de reproduire des fonctions de la vision naturelle en se basant sur la conception de systèmes de vision, souvent autonomes et capables d'effectuer des tâches telles que la reconnaissance ou la navigation. Il est devenu courant, dans les laboratoires de recherche de s'appliquer à la réalisation d'un robot capable de se déplacer de manière autonome dans le monde réel et plus loin de cela, la conception d'un véhicule intelligent capable de discriminer son entourage.

Ces robots, et ces véhicules disposent en général d'un système de vision dont la performance est due d'une façon commune à son système d'acquisition, son processus de traitement d'image et ses fonctions de reconnaissance. La fonction principale de telles machines peut se résumer dans la fonction de description de la scène dans son milieu naturel, et qui est généralement en 3D. Ceci passe par la détermination de la position de points dans

l'espace à partir d'une ou plusieurs images fournies par le système d'acquisition. Cette action constitue un des objectifs de la vision artificielle, et c'est dans ce commun but que la stéréoscopie intervient comme une technique utilisée pour reconstruire une scène 3D à partir des images prises sous des angles de vue différents.

## 2. Principe de la stéréovision

Stéréoscopie, Stéréovision, Vision stéréoscopique, tous ces termes notifient la même notion. Il s'agit globalement de l'ensemble des techniques et des procédés utilisées afin qu'on puisse réaliser et imiter une perception du relief 3D à partir de deux images. Lorsqu'on utilise deux images dans la stéréoscopie, on parle principalement de la stéréoscopie binoculaire. En effet, la vision en relief est naturelle, c'est l'image 2D qui demande une interprétation et une explication de la scène, car elle n'est qu'une projection perspective de la scène 3D dans un plan. Ce paragraphe traite les notions élémentaires de la vision stéréoscopique, ainsi il provoque les différentes contraintes et caractéristiques ayant une relation directe avec la mise en œuvre de ce système visuel artificiel.

La vision stéréoscopique, avant toute chose, est une faculté innée fondamentale de l'être humain et à l'insu de nous-mêmes nous la pratiquons de manière perpétuelle. Pourtant, peu de gens s'attardent à comprendre le fonctionnement de cette faculté si merveilleuse. En effet, notre corps est composé de deux yeux nous permettant de percevoir l'environnement qui nous entoure. Cependant, l'espace visuel périphérique vu par l'œil s'étend normalement de 60° en haut, 70° en bas et 90° environ latéralement ce qui correspond à un objectif photographique "grand angle". Chaque œil reçoit une image différente, cela est dû à sa position qui est différente de l'autre œil. Nos yeux sont éloignés de plus au moins de 65mm, distance inter pupille. De ce fait, l'image reçue par l'œil gauche est donc légèrement décalée par rapport à celle reçue par l'œil droit. Plus précisément c'est la rétine visuelle ou nerveuse qui est le lieu de traduction du message lumineux venant de l'extérieur vers un signal nerveux transmit au cerveau, via le nerf optique, qui traite et analyse les deux images. Le cerveau superposant ces deux images, et nous amène à un effet de relief. C'est la zone des champs de vision commune –vision binoculaire- des deux yeux qui permet de voir en 3D. La vue monoculaire nous permet aussi de voir en relief mais c'est très imprécis, nous conserverons principalement la vision binoculaire dans la suite de ce document.

La difficulté majeure de la stéréoscopie consiste à déterminer le décalage entre les images

prises, et cela conduit systématiquement à une opération de mise en correspondance, difficile à mettre en place, entre les primitives-images (pixels, régions, contours, etc.). Pour ce faire de manière simple et similaire au système visuel humain, le système stéréoscopique binoculaire n'utilise que deux images prises par deux capteurs légèrement décalés pour créer l'image gauche et l'image droite qui correspondent respectivement à l'image vue par l'œil gauche et l'image vue par l'œil droit.

En réalité, le processus de la stéréovision artificielle se réalise sur plusieurs étapes à savoir : Calibrage de la géométrie et de l'orientation des caméras, rectification et transformation pour simplifier la mise en correspondance, mise en correspondance et repérage des paires de points correspondantes dans chaque image et enfin reconstruction 3D par processus de triangulation. Nous donnons un bref aperçu de chacune de ses étapes dans le paragraphe ci après, mais pour plus de détails nous conseillons les références suivantes [[46](#), [53](#), [54](#)].

### **3. Aspects géométriques de la stéréovision**

Après avoir mis le point sur le principe de vision stéréoscopique chez l'être humain, nous abordons, dans cette partie, les aspects et les étapes nécessaires pour mettre en œuvre la stéréoscopie binoculaire artificielle. Autrement dit, à partir de deux caméras positionnées devant une scène, légèrement décalées, donnant ainsi les deux images gauche et droite de la scène.

Les caméras constituant le système de prise de vue de la vision stéréoscopie sont conçues pour qu'elles aient une convergence vers la même scène. Ce qui veut dire une même vue de la scène, et c'est ce réglage qui fait que nous obtenons des images gauche et droite, ainsi stipulant l'image vue par l'œil gauche et l'image vue par l'œil droit.



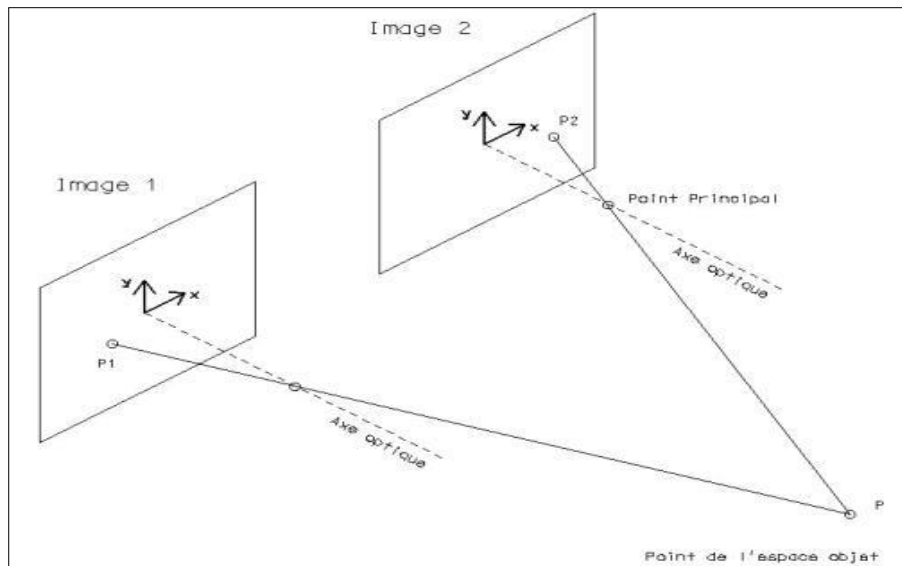


Figure 17 : Configuration générale du système stéréoscopique

### 3.1 Calibrage stéréoscopique

Le calibrage est un processus qui permet de déterminer la relation entre les coordonnées d'un point dans la scène, suivant un repère de référence, et ses coordonnées dans les deux images. Cette relation dépend principalement des paramètres intrinsèques et extrinsèques du système. En effet, les paramètres intrinsèques rassemblent l'ensemble des propriétés propres à chacune des caméras : la distance focale des objectifs des caméras et la taille et la résolution des capteurs. En revanche, les paramètres extrinsèques décrivent l'association des caméras dans le système : orientation et position des caméras par rapport au repère de référence.

### 3.2 Rectification

Cette opération permet de réduire l'espace de recherche des paires de points dans l'opération d'appariement. Pratiquement, l'utilisation des droites épipolaires suscite ce principe. Une droite épipolaire définit l'ensemble des points d'une image pouvant correspondre à un même point dans l'autre image, autrement dit, pour chaque point dans une image il y a une droite épipolaire dans l'autre. Ces droites épipolaires convergent vers un point nommé épipôle comme montré dans la figure 18 ci dessus. En fait, l'épipôle est défini comme étant la projection du centre optique de la caméra 1, par exemple, sur l'image 2 et vice-versa.

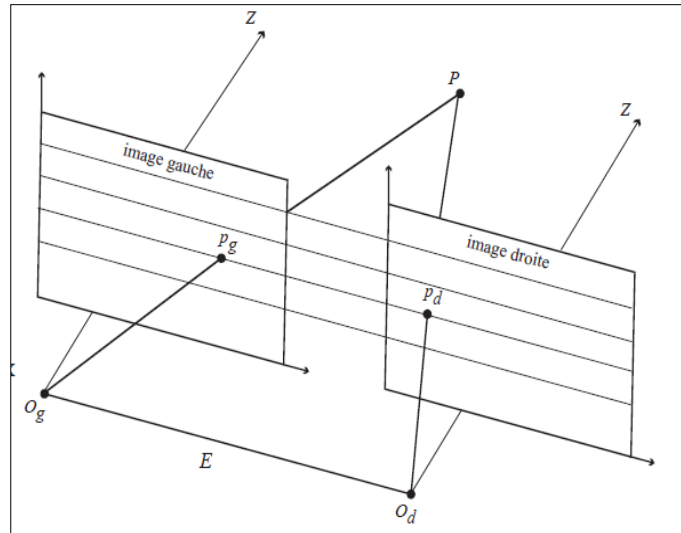


Figure 18 : Mise en œuvre des épipôles

En outre, il existe une configuration parfaite qui fait que les droites épipolaires seront parallèles et confondues avec les lignes de l'image. Cette configuration est obtenue quand les axes des deux caméras sont parallèles et quand la droite reliant les centres optiques est confondue avec les axes horizontaux des deux caméras (cf. figure 19).

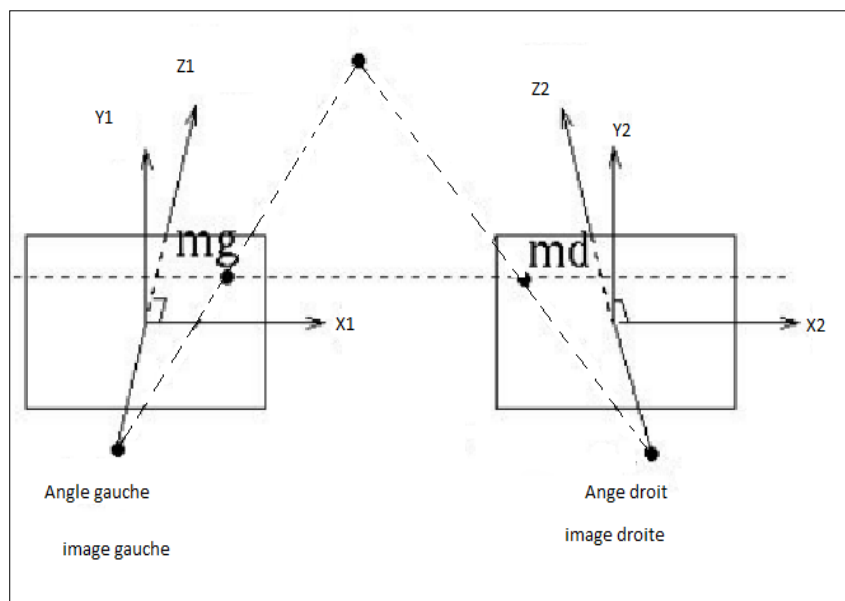


Figure 19 : Mise en œuvre de la rectification géométrique

Par ailleurs, il existe des méthodes analytiques pour mettre en place ce type de configuration [46], en trouvant l'équation des droites épipolaires. L'inconvénient de cette procédure est le temps supplémentaire de calcul.

## 4. Mise en correspondance

Dans cette partie nous traitons le processus de l'appariement des pixels dans la stéréovision de manière simple et précise. En fait, la mise en correspondance consiste à faire concorder des primitives dans l'image gauche à celles dans l'image droite, ces primitives définissent la nature des entités à correspondre. Notamment, les régions d'intérêt, les pixels, les contours, etc. En somme toute caractéristique robuste de l'image peut être le sujet de la mise en correspondance [46].

La détection des caractéristiques d'image ou les primitives peut être effectuée manuellement par les experts d'une application donnée ou automatiquement selon des méthodes et des algorithmes à définir selon l'application.

L'implémentation de la mise en correspondance peut se faire selon plusieurs approches de calcul : Une approche neuronale à l'aide des réseaux biologiques, une approche s'appuyant sur le calcul de la corrélation entre les pixels et qui suscite une prise en compte d'un ensemble de contraintes dans l'algorithme de calcul, ou une approche probabiliste [55, 56, 57].

Le but essentiel de l'appariement est, en premier temps, mettre en relation les primitives en question avec une pertinence élevée et un nombre élevé des entités corrélées, et en deuxième temps, avoir un nombre faible d'appariées erronées. Toutefois, il est difficile d'idéaliser cette opération vu le nombre des contraintes qui entrent en action pour la mettre en évidence.

Nous mentionnons quelques contraintes essentielles pour la mise en correspondance, qui se réunissent dans deux catégories à savoir :

- Les contraintes locales : Ces contraintes permettent de donner plus de précision au choix de couples et ainsi accorder de l'aide au calcul et à la recherche, desquelles on peut citer : La contrainte épi-polaire, la contrainte d'orientation, la contrainte de disparité maximale et/ou minimale.
- Les contraintes globales : Ces contraintes sont appliquées à l'ordre total des appariements plausible, les plus connues sont : La contrainte d'unicité, la contrainte d'ordre, la contrainte de continuité de la disparité, la contrainte de continuité figurale.

Toutes ces contraintes et ces ambiguïtés peuvent être mises en pratique d'une manière ou d'une autre selon des critères liées à l'application concernée [46].

## 5. Triangulation

Dans la dernière étape du processus de la stéréovision, il s'agit essentiellement de mettre en évidence la troisième dimension. Cette étape se base sur la carte de disparités obtenue dans l'étape de la mise en correspondance et elle calcule le décalage entre les paires souhaitées, autrement dit la différence entre les deux positions de chacun des points du couple mis en correspondance. En effet, ce décalage est inversement proportionnel à la profondeur avec un facteur qui dépend de la distance focale des caméras et de la distance entre les deux centres optiques des caméras. Le calcul de la profondeur revient au calcul des coordonnées  $(X, Y, Z)$  de chaque point  $P$  dans l'espace 3D en fonction des coordonnées  $(x, y)$  dans le plan image et qui sont données par la formule

$$Z = \frac{f * E}{d}, X = \frac{x * Z}{f} \text{ et } Y = \frac{y * Z}{f}.$$

Avec :

$f$  : distance focale,  $E$  : distance entre les centres optiques des caméras et  $d$  : disparité.

Plusieurs chercheurs ont montré leur intérêt à ce processus de mise en correspondance en proposant des solutions et des approches plus au moins fiables et chacun de sa manière de percevoir le problème [43, 55, 56,57, 58].

Nous avons ainsi mis le doigt sur la problématique fondamentale de la vision stéréoscopique, il s'agit de la mise en correspondance ou l'appariement, qui constitue l'un des thèmes de nos contributions dans ce rapport. Nous en donnerons les détails dans les chapitres suivants.

## IV. Conclusion

Ce chapitre présente une introduction aux images numériques et les méthodes de traitement utilisées, nous avons mis l'accent sur les techniques et les théories sur les quelles nous nous sommes basées pour introduire nos contributions. A ce niveau, nous avons mentionné une brève présentation des méthodes de segmentation pour situer notre approche et comprendre l'intérêt des résultats que nous allons présenter dans les chapitres suivants. Notre

approche utilise la couleur comme paramètre de segmentation, d'où l'intérêt de parler de la théorie colorimétrique et les espaces de couleur.

Les méthodes de segmentation des images numériques sont diverses, nous avons présenté les principales catégories de méthodes de segmentation, les méthodes globales basées sur la classification de pixels et les méthodes locales basées sur des critères de voisinage spatial. Dans la littérature, il existe aussi des méthodes basées sur la morphologie mathématique et des méthodes hybrides combinant les deux ou les trois approches. En général, les méthodes existantes dans la littérature sont utilisées en fonction de l'application désirée et fonctionnent souvent sous certaines hypothèses.

Une brève étude des systèmes de représentation de la couleur a été présentée. En effet, plusieurs systèmes de représentation de la couleur existent, et chacun de ces systèmes à son contexte et ses caractéristiques particulières qui peuvent améliorer ou réduire la qualité d'un processus de segmentation des images couleurs. Les espaces couleurs présentant des composants dé-corrélés sont très favorisés dans le cadre d'une méthode de segmentation basée sur la couleur, spécialement, une méthode probabiliste utilisant les histogrammes couleurs. Dans le domaine de la conduite intelligente, les images de la route sont prises dans des conditions temporelles et climatiques différentes, un espace couleur qui sépare le mieux les composantes chromatiques de la luminance sera privilégié pour l'analyse de telles images.

L'utilisation des histogrammes couleur est étroitement liée aux espaces couleurs. Les régions de l'image forment des zones de forte concentration dans les histogrammes, et ces zones sont caractérisées par des modes (extremum). Les espaces couleurs sont des systèmes de représentation tridimensionnels, et souvent ils font intervenir trois primitives pour composer une couleur quelconque. Pour représenter la distribution des primitives dans une image selon un espace couleur, les structures adéquates sont les histogrammes 3D. Vu les inconvénients liés à la complexité de l'utilisation de telles structures, les auteurs ont cherché des solutions pour la simplifier. La première approche de simplification dite approche marginale consiste en le traitement séparé des composantes couleurs à travers des histogrammes monodimensionnels et qui a le problème majeur de ne pas traiter l'aspect vectoriel de l'information couleur. La deuxième approche dite bi marginale et constitue un excellent compromis entre l'aspect vectoriel et l'aspect marginal, elle permet une simplification de l'analyse des histogrammes 3D tout en préservant la corrélation des composantes.

Nous sommes en mesure de développer un algorithme de segmentation globale et quasi-automatique. Les histogrammes couleurs restent l'outil adéquat pour réaliser de tels algorithmes relativement à notre approche de détection. Nous avons présenté les types des histogrammes qui existent dans la littérature et les méthodes qui existent pour les calculer. La raison principale pour laquelle nous avons choisi d'utiliser les histogrammes 2D dans la suite réside dans la simplicité d'analyse en préservant partiellement la corrélation entre les composantes couleurs. Une deuxième raison d'utiliser les histogrammes 2D est l'exploitation des ondelettes comme outil d'analyse. En effet, l'analyse avec le formalisme des ondelettes s'adapte très bien à une étude des signaux bidimensionnels.

La difficulté centrale de la stéréoscopie consiste à déterminer le décalage entre les images prises, et cela conduit systématiquement à une opération de mise en correspondance, souvent difficile à mettre en place, entre les pixels. Plusieurs méthodes de mise en correspondance existent, certaines traitent l'ensemble de l'image et construisent les cartes de disparité en réalisant un appariement pixel par pixel, d'autres méthodes cherchent à appairer d'abord les points caractéristiques (points d'intérêts) et utilisent le résultat pour reconstruire des cartes de correspondances.

Nous avons conçu une méthode de mise en correspondance qui exploite les modules maximaux des ondelettes comme des points caractéristiques. La particularité de notre approche réside dans l'utilisation des informations multi-échelles dans la sélection des candidats à appairer et dans le processus d'appariement.

Nous n'avons présenté que l'essentiel de nos outils et théories que nous avons exploité pour réaliser notre travail. Le chapitre quatre présente notre méthode de détection basée sur la couleur comme critère de segmentation et le chapitre cinq présente notre approche de mise en correspondance stéréoscopique.

# **CHAPITRE 3 :**

# **APPROCHE**

# **MATHEMATIQUE**

# I. Ondelettes comme outil d'analyse

## 1. Introduction

Les transformations mathématiques dans le domaine de traitement de signal sont des outils puissants pour étudier les phénomènes modélisés par des signaux. Ils servent à analyser et interpréter l'information en vue de leur exploitation. Les images représentent des signaux bidimensionnels et le formalisme des ondelettes offre des outils pour obtenir des transformées très adaptées pour l'exploration d'une image. La théorie des ondelettes s'est développée dès le début des années 80 [37, 38, 39] et son application à l'image a permis de mettre en valeur les structures les plus significatives de l'image.

## 2. Transformée de Fourier

Il est consistant de caractériser un signal par son contenu fréquentiel que par son contenu temporel (spectre). L'analyse de Fourier propose une approche d'analyse globale d'un signal. En effet, c'est une intégration de moins infini à plus infini dans laquelle toute notion de localisation temporelle ou spatiale disparaît.

La transformée de Fourier d'un signal s'exprime comme une somme infinie des fonctions trigonométriques de toutes fréquences. C'est une extension de la définition de série de Fourier pour les signaux périodiques, une telle sommation se présente sous forme d'intégrale [40].

$$\forall \omega \in \mathbb{R}, \text{ on a } TFf(\omega) = \int_{-\infty}^{+\infty} f(t)e^{-j\omega t} dt \quad (3.3)$$

Les images sont des signaux à deux dimensions (représentation des intensités lumineuses dans le plan), et la transformée de Fourier d'une image donnera une idée sur la distribution fréquentielle de l'image (spectre de Fourier).

Pour les images numériques, une des applications de la transformée de Fourier est la recherche des hautes fréquences (bruits) dans les images. Le passage au domaine fréquentiel donne une idée sur les fréquences présentes dans l'image et permettra de sélectionner les hautes fréquences et les annuler. Une transformation inverse permettra de retrouver l'image d'origine sans les hautes variations correspondant au bruit.

La transformée de Fourier d'un signal renseigne sur la distribution spectrale sans aucune information sur la localisation temporelle de changement des fréquences. Ce problème se



traduit par des limitations dans la détection des discontinuités et la détection des faibles variations du signal.

Le problème de la détection des discontinuités est l'un des soucis de l'analyse du signal vu l'existence capitale des signaux présentant des troncatures et des discontinuités. L'utilisation de la transformée de Fourier comme outil d'analyse n'est pas fiable dans les zones isolées. Cela a été étudié et montré pour la transformée de Fourier d'un signal tronqué qui suscite des fluctuations brusques autour des zones de troncature connu sous le nom de phénomène de Gibbs [40, 25]. En somme, la TF n'est pas capable de donner les bonnes composantes fréquentielles dans ces régions du signal.

Le deuxième inconvénient est la résolution, cela veut dire la précision, en effet la TF permet de traiter efficacement les signaux qui ne varient pas rapidement au cours du temps, autrement dit les signaux stationnaires, cependant le spectre obtenu ne donne aucune information sur la variation du temps dans le domaine temporel, ainsi la TF ne peut pas traiter les petites variations du signal dans lesquelles se trouvent généralement l'information utile.

Une solution pour combler les problèmes d'analyse avec la transformée de Fourier a été proposée. Cette solution consiste en multiplier le signal par une fenêtre glissante  $h(t - b)$ , c'est un découpage du signal en morceaux dans le but d'y localiser les fréquences.

Cette solution est connue sous le nom de Transformée de Fourier à fenêtre glissante :

$$W_f: \mathbb{R}^2 \rightarrow \mathbb{C} \quad (3.4)$$

$$(b, \omega) \rightarrow W_f(b, \omega) = \int_{\mathbb{R}} f(t) h(t-b) e^{-j\omega t} dt$$

Le fait de limiter l'analyse de Fourier à un voisinage temporel permettra de prendre en considération le facteur temporel dans l'analyse du signal, plusieurs types de fenêtres peuvent être utilisées dans ce sens : triangulaire, Gauss, Hamming, ...

La transformée de Fourier à fenêtre glissante permet l'analyse du signal dans des fenêtres temps-fréquence fixes, dont l'échelle de l'analyse correspond à la taille des fenêtres choisies. Cependant, pour les signaux transitoires ou il y a des variations de signal sur de courtes durées, cette transformation présente les mêmes limitations que la transformée de Fourier ordinaire.

La puissance de la transformée en ondelettes revient à sa capacité d'analyser un signal en adaptant la taille de la fenêtre selon la fréquence. C'est une transformation qui renseignera sur

le contenu fréquentiel en préservant la localisation, en permettant une représentation temps-fréquence ou espace-échelle du signal.

### 3. Transformée en ondelettes

#### 3.1 Présentation

La transformée en ondelette est une transformation adaptative, c'est une analyse de signal à l'aide d'une fonction qui a un début et une fin, de moyenne nulle, appelée ondelette. L'analyse consiste en une translation accompagnée d'une dilatation de l'ondelette sur tout le signal. Les ondelettes offrent une possibilité d'analyse (extraire des éléments remarquables, caractéristiques) et de représentation (approximation) d'un signal. Les principes mathématiques de la théorie des ondelettes sont détaillés davantage dans les références [25, 34, 38, 41].

Pour une fonction  $f(t)$  réelle à variable réelle, la transformée en ondelette est définie par :

$$\forall x \in \mathbb{R}, \forall a > 0 \text{ on a } TWf(x, a) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \overline{\psi_{x,a}}(t) dt \quad (3.5)$$

La fonction  $\overline{\psi_{x,a}}(t)$  est une translation et dilatation d'une fonction particulière appelée ondelette mère :

$$\psi_{x,a}(t) = \Psi\left(\frac{t-x}{a}\right) \quad (3.6)$$

L'ensemble  $\left\{ \psi_{x,a}(t) = \Psi\left(\frac{t-x}{a}\right) \right\}$ ,  $x \in \mathbb{R}$  et  $a \in \mathbb{R}^{+*}$  constitue la famille d'ondelettes filles générées par les dilatés et translatés de l'ondelette  $\Psi$ .

La fonction  $\Psi$  doit vérifier la condition dite d'admissibilité, cela s'exprime par :

$$\int_{-\infty}^{+\infty} \Psi(t) dt = 0 \quad (3.7)$$

La transformée d'ondelette  $TWf(x, a)$  de  $f$  à deux paramètres : le paramètre de position  $x$  qui représente le point autour duquel est centré l'analyse et le facteur d'échelle  $a$  qui indique la finesse de l'analyse. Donc plus  $a$  est grand plus la dilatation est grande est l'étude se fera sur une zone plus large (cf. figure 17).

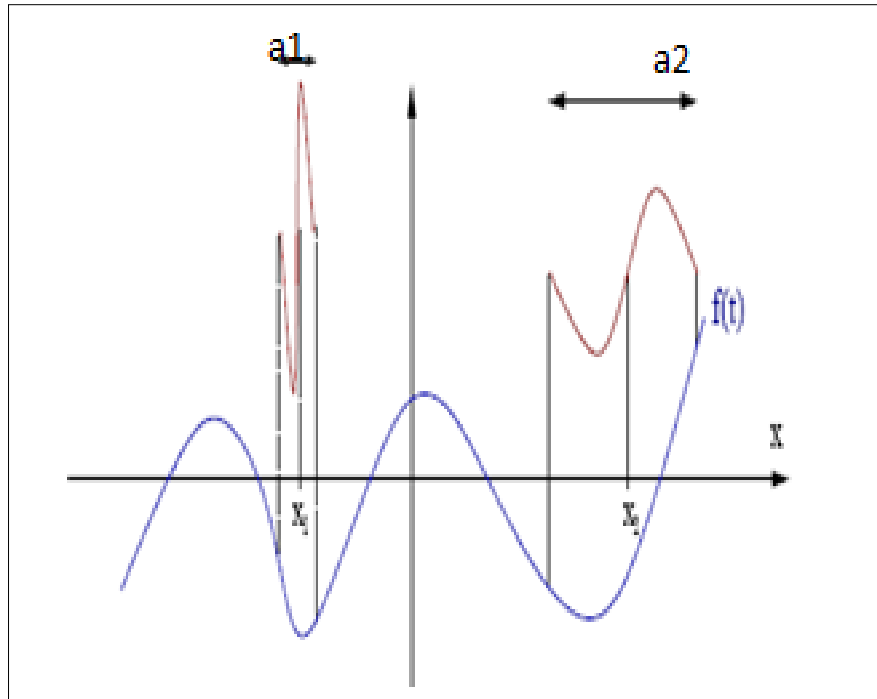


Figure 20 : Représentation de la transformée en ondelettes ( $\psi_{x_1, a_1}$  et  $\psi_{x_2, a_2}$ )

### 3.2 Exemples d'ondelettes

- Dérivées de la gaussienne :** La fonction  $g(x) = e^{-\pi x^2}$  est une Gaussienne. La moyenne de cette fonction sur  $\mathbb{R}$  n'est pas nulle, donc elle ne vérifie pas la condition d'admissibilité. Cependant, les dérivées de  $g$  possèdent des moyennes nulles sur  $\mathbb{R}$ . Une dérivée gaussienne d'ordre  $n$  vérifie la condition d'admissibilité et permet de profiter aussi de certaines propriétés des fonctions gaussiennes en particulier l'approximation numérique. La dérivée seconde d'une gaussienne est connue sous le nom du « chapeau mexicain ». Les coefficients d'ondelettes relatifs à la dérivée de la gaussienne s'obtiennent par simple dérivation de la convolution avec un noyau gaussien et on a un compromis entre la diffusion de l'information dans l'espace et sa diffusion fréquence. C'est l'une des ondelettes que nous avons utilisé dans nos travaux (cf. figure 21).

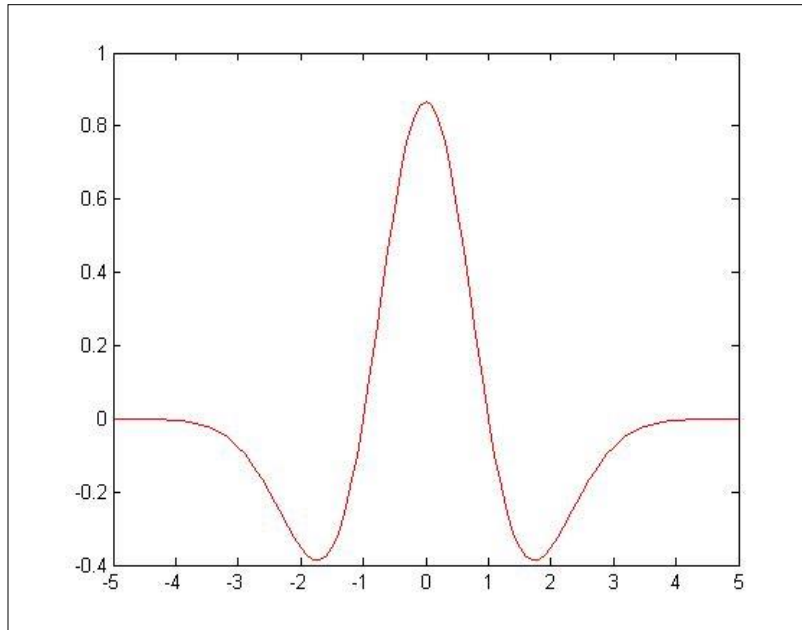


Figure 21 : Dérivée seconde d'une gaussienne, chapeau mexicain

- **L'ondelette de Haar** : C'est la plus populaire et la plus simple des ondelettes, elle est définie sur l'intervalle  $[0,1]$  ou  $[-\frac{1}{2},\frac{1}{2}]$  comme la fonction constante par morceaux qui vaut 1 sur  $[0,\frac{1}{2}[$  et  $-1$  sur  $]\frac{1}{2},1]$ . C'est une ondelette très simple à manipuler et très convenable pour les calculs numériques, elle présente en plus l'avantage d'être bien localisée, du fait que son support est compact. Par contre il s'agit d'une fonction discontinue et n'ayant qu'un seul moment nul, sa transformée de Fourier n'est pas une fonction fortement oscillante et on perd la localisation en fréquence (cf. figure 22).

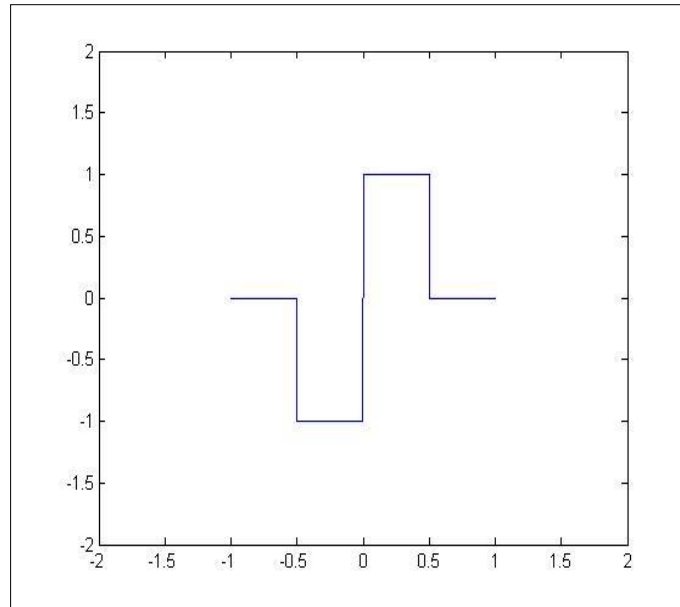


Figure 22: ondelette de Haar

### 3.3 Transformée en ondelettes bidimensionnelle

Les définitions et les formules utilisées dans le cas unidimensionnel peuvent être généralisées dans le cas 2D, il suffit dans ce cas de tenir en compte en plus des deux paramètres de position (translation) et d'échelle (dilatation) un troisième paramètre pour représenter l'orientation (direction). En 2D il y'en a le problème de la direction qui va se poser car la régularité ne peut être perçue de la même manière selon une direction ou une autre. Pour illustrer ce problème, si on parcourt verticalement une image, la régularité d'un objet isolé dont la forme est fine et verticale n'est pas perçue de la même manière que si l'image est parcourue horizontalement, car dans ce dernier cas une forte singularité serait nettement distinguée. Pour remédier à cette problématique on introduit la direction dans la définition de la transformée d'ondelettes par le biais d'un angle  $\theta \in [0, 2\pi]$  et on parle de transformée en ondelettes directionnelle [42]. Dans nos travaux, nous nous sommes appuyés sur ce principe en utilisant seulement les directions horizontales et verticales.

### 3.4 Transformée en ondelette dyadique

Pour implémenter les formules mathématiques continues liées aux ondelettes, il faut les discrétiser. Il faut penser surtout à un support discret pour les échelles et aussi à un ensemble de paramètres de translation discrets. Plusieurs approches de discrétisation des échelles peuvent être utilisées dans ce sens. En revanche, la discrétisation dyadique est la plus

populaire car elle permet un bon parcours des fréquences du signal et une implémentation simple et efficace  $a = 2^j, j = 1..n$ . Lors d'une discrétisation dyadique nous parlerons de transformée en ondelette dyadique [26].

Dans le cas discret, la transformée en ondelette génère des coefficients de calcul contenant l'information résultat de l'analyse, cette information se concentre sur une position et peut changer avec le changement d'échelle. L'important c'est que l'information recherchée correspond à un voisinage qui devient plus grand chaque fois l'échelle devient plus grand aussi. Le choix d'une discrétisation dans le domaine spatial qui dépend des échelles dyadiques sera très utile puisque ca va rendre la transformée en ondelette plus compact et plus performant en terme de calculs. Cette approche de discrétisation est largement adoptée et elle a une correspondance très intéressante avec la théorie d'analyse multi-résolution, ce qui a conduit a des algorithmes de calcul rapides et simples favorisant ainsi le succès des analyses avec les ondelettes.

#### 4. Analyse multi résolution et ondelettes

L'AMR (analyse multi résolution) consiste en une suite de sous espaces fermés et emboîtés  $V_j \in L^2(\mathbb{R}), j \in \mathbb{Z}$  vérifiant les propriétés suivantes :

- i-  $\forall j \in \mathbb{Z}, V_j \subset V_{j+1}$
- ii-  $\forall j \in \mathbb{Z}, f(t) \in V_j \Leftrightarrow f(2t) \in V_{j+1}$
- iii-  $\forall k \in \mathbb{Z}, f(t) \in V_0 \Leftrightarrow f(t-k) \in V_0$
- iv-  $\bigcap_{j=-\infty}^{+\infty} V_j = \{0\}$
- v-  $\bigcup_{j=-\infty}^{+\infty} V_j = L^2\{\mathbb{R}\}$
- vi- Il existe  $\phi \in V_0$  tel que  $\{\phi(t-n)\}_{n \in \mathbb{Z}}$  soit une base de Riesz de  $V_0$

Lors de l'approximation d'une fonction  $f$  dans un espace  $V_j$ , on dit que  $j$  est le niveau d'analyse de  $f$  (résolution). Cette approximation est plus fine que celle dans  $V_{j-1}$  et moins bonne que l'approximation dans  $V_{j+1}$ .  $V_0$  Soit l'espace qui donne l'approximation la plus grossière, et à partir des fonctions  $\phi$  (appelées fonctions d'échelle) on peut obtenir une base orthonormée de  $V_j$  pour tout  $j \in \mathbb{Z}$  :

$$\{\phi_{j,k}\}_{k \in \mathbb{Z}} = \{2^j \phi(2^j t - k)\}_{k \in \mathbb{Z}} \quad (3.7)$$

Ces bases sont suffisantes lorsqu'il s'agit d'une simple approximation de la fonction à analyser dans différentes espaces  $V_j$ , mais si on s'intéresse aux détails perdus lors du passage d'un niveau  $j$  à un niveau  $j-1$ , ces bases ne possèdent pas des propriétés directes permettant de calculer ces détails. Pour les signaux et les images, les détails résultat d'un passage d'un niveau à l'autre contiennent des informations sur les hautes fréquences, et on peut les utiliser pour déceler les singularités de la fonction analysée.

Pour représenter les détails d'une approximation lors d'un passage d'un niveau de résolution  $j$  à un niveau de résolution  $j-1$ , on choisit un espace complémentaire orthogonale de  $V_{j-1}$  souvent noté  $W_{j-1}$  et qu'on appelle espace des détails :

$$V_j = V_{j-1} \oplus W_{j-1} \text{ avec } V_{j-1} \perp W_{j-1} \quad (3.8)$$

Le fait de chercher des bases orthonormées de ces espaces permettra de représenter le détail d'une approximation [26] [43].

La fonction  $\phi$  vérifie l'équation dite d'échelle :

$$\exists h = (h_k)_{k \in \mathbb{Z}} \text{ tel que } \phi(t) = \sum_k h_k \phi(2t - k) \quad (3.9)$$

Cette relation permettra de construire une ondelette  $\psi$  tel que  $\{\psi(t - n)\}_{n \in \mathbb{Z}}$  une base orthonormée de  $W_0$  et la dilatation de celle-ci permettra de construire des bases orthonormées des autres espaces de détails :

$$\psi(t) = \sqrt{2} \sum_k (-1)^{1-k} h_{1-k} \phi(2t - k) \text{ ou } g_k = (-1)^{1-k} h_{1-k} \quad (3.10)$$

Les suites  $\{h_k\}_{k \in \mathbb{Z}}$  et  $\{g_k\}_{k \in \mathbb{Z}}$  sont deux filtres respectivement passe bas et passe haut, construits soigneusement à partir de la fonction d'échelle  $\phi$ .

Les signaux et les images à manipuler étant de résolution finie. De ce fait, on considère leur projection dans l'espace d'approximation le plus grossier  $V_J$  où  $J$  est lié à la résolution du signal ou l'image. Supposons que le signal est de taille  $N$  avec  $N = 2^J$ , on représente donc le signal par le vecteur d'approximation  $C_J = (C_{J,k})_{k=0..2^J-1}$  de coefficients dans la base de  $V_J$ . La transformée en ondelettes orthogonale du signal est obtenue donc par sa décomposition dans  $V_J = W_{J-1} \oplus W_{J-2} \oplus \dots \oplus W_0 \oplus V_0$  et ce en faisant, à chaque échelle  $j$  une convolution des coefficients d'approximations  $C^j$  avec les filtres  $h$  et  $g$  et en ne gardant que les coefficients d'indices pairs, c'est l'étape de décimation. Ceci est traduit par l'algorithme,

dit d'analyse ou de convolution-décimation suivant :

$$C_{j-1}[k] = (C_j * \bar{h})[2k] \forall k = 0, \dots, 2^{j-1} - 1 \quad (3.11)$$

$$d_{j-1}[k] = (C_j * \bar{g})[2k] \forall k = 0, \dots, 2^{j-1} - 1 \quad (3.12)$$

$$\text{Avec } \bar{x}[n] = x[-n]$$

La transformée en ondelette s'obtient alors par itération de deux opérations de convolution des données par les deux filtres  $h$  et  $g$  (cf. Figure 23).  $h$  est un filtre passe bas qui produit un lissage des données en gardant que les basses fréquences et  $g$  un filtre passe haut qui sélectionnera les hautes fréquences (détails). La prochaine approximation s'applique de nouveau sur les coefficients  $C_{j-1}$  obtenus.

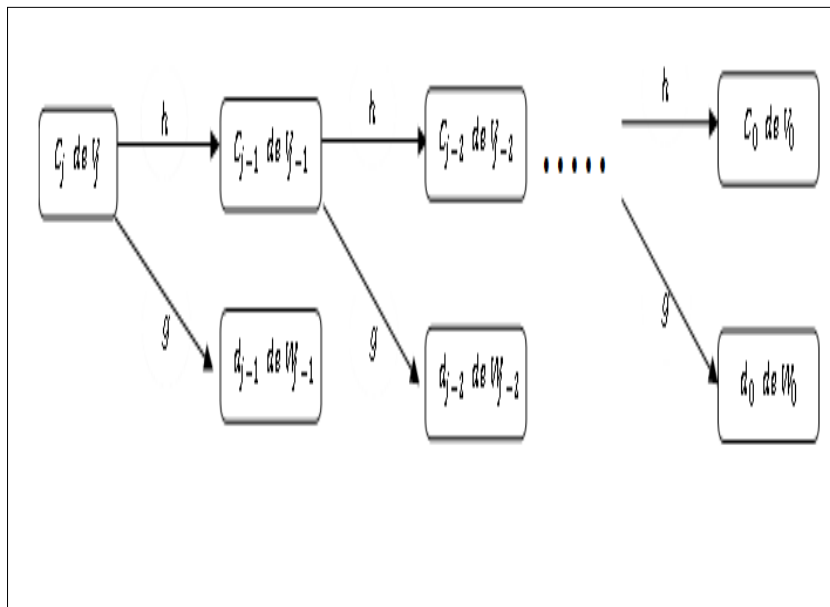


Figure 23 : Analyse à l'aide d'un banc de filtres.

#### 4.1 Exemple : ondelettes de Haar

Une analyse multi résolution de Haar se base sur la fonction d'échelle suivante :

$$\phi(x) = \begin{cases} 1 & \text{si } 0 \leq x < 1 \\ 0 & \text{si non} \end{cases} \quad (3.13)$$

Cette fonction vérifie les hypothèses d'une analyse multi résolution, et nous avons :

$$\phi_{j,k}(x) = 2^j \phi(2^j x - k) \quad (3.14)$$



Donc :

$$\phi_{j,k}(x) = \begin{cases} 1 & \text{si } \frac{k}{2^j} \leq x < \frac{k+1}{2^j} \\ 0 & \text{si non} \end{cases} \quad (3.15)$$

- $V_0$  C'est l'espace vectoriel des fonctions constantes sur  $[0,1[$ .
- $V_1$  c'est l'espace vectoriel des fonctions constantes sur  $[0, \frac{1}{2}[$  et  $[\frac{1}{2}, 1[$ .
- ...
- Généralement :  $V_j$  constituera l'espace vectoriel des fonctions constantes par morceaux sur  $\{[\frac{k}{2^j}, \frac{k+1}{2^j}[ ]_{k=0..2^j-1}$

Pour une fonction  $f(x)$  définie sur l'intervalle  $[0,1[$ , un échantillonnage dyadique de cette fonction en  $2^J$  points permet d'écrire :

$$f(x) = \sum_{p=0}^{2^J-1} f(\frac{p}{2^J}) \phi_{J,p}(x) \quad (3.15)$$

Donc  $2^J$  définit l'échelle la plus fine de la décomposition avec les coefficients  $c_{J,p} = f(\frac{p}{2^J})$ . L'approximation de  $f$  dans un espace  $V_j$   $j < J$  consiste à écrire la fonction dans la base  $\{\phi_{j,k}(x)\}, k = 0..2^j - 1$  engendrant cet espace. C'est une base orthogonale pouvant être orthonormée en multipliant  $\phi_{j,k}$  par  $\sqrt{2}$ .

$$\langle \phi_{j,k}(x), \phi_{j,k}(x) \rangle = 1/2^j \text{ et } \langle \phi_{j,k}(x), \phi_{j,k'}(x) \rangle = 0 \text{ pour } k \neq k'$$

L'obtention des espaces complémentaires lors du passage d'une résolution à l'autre, s'obtient par la fonction :

$$\psi(x) \begin{cases} 1, & 0 \leq x < 1/2 \\ -1, & \frac{1}{2} \leq x < 1 \\ 0, & \text{ailleurs} \end{cases} \quad (3.16)$$

Les fonctions  $\psi$  vérifient la relation :  $\psi_{j,k}(x) = 2^j \psi(2^j x - k)$  avec :

$$\psi_{j,k}(x) \begin{cases} 1, & \frac{k}{2^j} \leq x < \frac{k}{2^j} + \frac{1}{2^{j+1}} \\ -1, & \frac{k}{2^j} + \frac{1}{2^{j+1}} \leq x < \frac{k+1}{2^j} \\ 0, & \text{ailleurs} \end{cases} \quad (3.17)$$

La famille  $\{\psi_{j,k}(x)\}, k = 0..2^j - 1$  forme une base orthonormée de l'espace vectoriel  $W_j$  avec  $V_{j+1} = V_j \oplus W_j$ .

Donc  $\phi$  représente la fonction d'échelle de haar et  $\psi$  représente l'ondelette de haar.

Approximation de la fonction  $f(x)$  :

$$f(x) = \sum_{p=0}^{2^{j-1}-1} c_{j-1,p} \phi_{j-1,2p}(x) + \sum_{p=0}^{2^{j-1}-1} d_{j-1,p} \psi_{j-1,2p}(x) \quad (3.18)$$

On peut calculer les coefficients  $c_{j-1,p}$  et  $d_{j-1,p}$  par récurrence à partir des  $c_{j,p}$  :

$$c_{j-1,p} = \frac{c_{j,2p} + c_{j,2p+1}}{2} \text{ et } d_{j-1,p} = \frac{c_{j,2p} - c_{j,2p+1}}{2}$$

## 4.2 L'analyse multi résolution dans le cas 2D

L'AMR dans un espace  $L^2(\mathbb{R}^2)$  se fait de la même manière que dans le cas monodimensionnel. Et le produit tensoriel fournit un moyen simple pour la définir.

$\{v_{j+1} = V_{j+1} \otimes V_{j+1}\}, j \in \mathbb{Z}$  Un espace  $L^2(\mathbb{R}^2)$  et  $V_j$  espace de  $L^2(\mathbb{R})$ .

Donc :

$$v_{j+1} = (V_j \otimes V_j) \oplus (V_j \otimes W_j) \oplus (W_j \otimes V_j) \oplus (W_j \otimes W_j)$$

Donc on peut écrire :

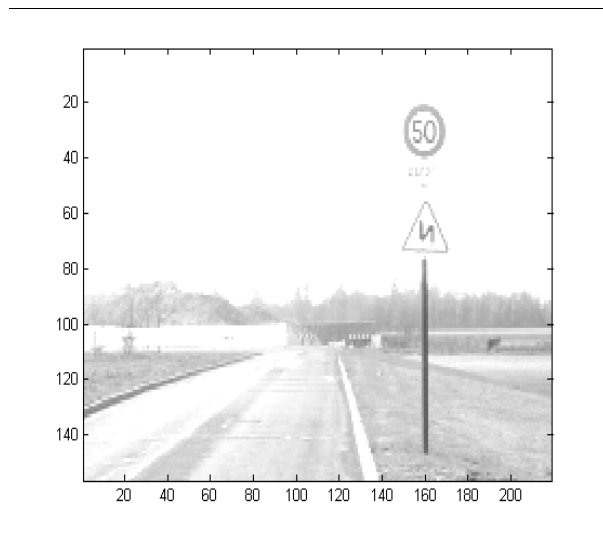
$$v_{j+1} = v_j \oplus w_j^1 \oplus w_j^2 \oplus w_j^3 \quad (3.19)$$

La base de  $w_j^1$  sera écrite en fonction de  $\phi$  et  $\psi$  :  $\{\psi_{j,k} \phi_{j,p}\}, k, p \in \mathbb{Z}$

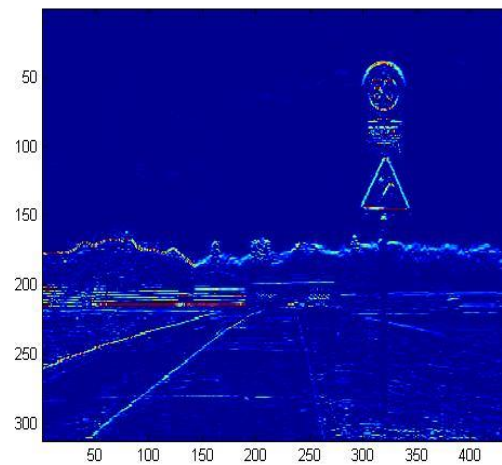
La décomposition consiste en la décomposition de chaque espace d'approximation en un sous espace d'approximation plus grossier et trois sous espaces de détails. Donc, l'application d'une décomposition basée sur une analyse multi-résolution d'un signal bidimensionnel conduit à une approximation de ce signal dans l'espace d'approximation, les détails de cette approximation seront décrits dans les espaces de détails correspondants. L'application de ce type de décomposition aux images numériques, fournit : Une approximation de l'image, Les détails horizontaux de l'approximation (informations sur les arrêts horizontaux), Les détails verticaux de l'approximation (informations sur les arrêts verticaux) et les détails diagonaux (cf. figure 24).



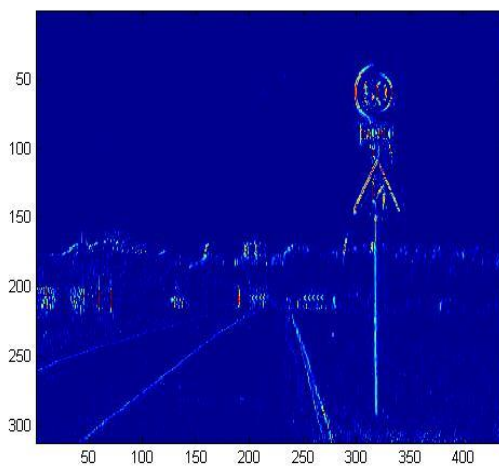
Image (A)



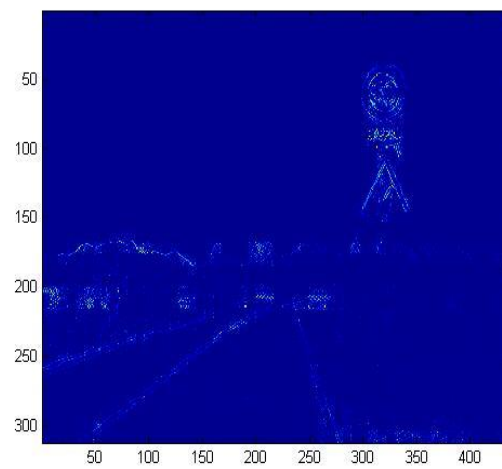
Approximation de l'image



Détails horizontaux



Détails verticaux



Détails diagonaux

Figure 24 : Transformée en ondelettes orthogonale 2D

## II. Transformé en ondelettes et contours multi-échelles

### 1. Modules maxima de la transformée en ondelettes

L'étude des modules maxima de la transformée en ondelettes d'une fonction  $f$  est intéressante dans la mesure où ces modules maxima sont liés aux singularités de la fonction. Ces modules maxima ont de grandes valeurs aux points où la fonction  $f$  est irrégulière et de faibles valeurs là où la fonction  $f$  est régulière [26] [59]. Les modules maxima sont définis dans le cadre d'une analyse multi-échelle avec les ondelettes. Chaque module maxima dans l'espace est lié à une échelle particulière  $s$ .

Généralement, un maximum local d'une fonction  $f$  est un point  $x$  dans lequel  $f$  a la plus grande valeur dans le voisinage de  $x$ . Ce voisinage sera un intervalle de valeurs dans le cas 1D ou une boule dans le cas 2D. Avec l'ajout d'une direction privilégiée dans le cas 2D nous aurons de même un intervalle comme dans le cas 1D.

Soit  $f : I \rightarrow \mathbb{R}$  une fonction dérivable.

Pour  $x_0 \in I$  nous avons :  $f(x_0)$  est un maximum ou minimum (extremum) local implique que :  $f'(x_0) = 0$  le signe de  $f'(x_0)$  varie autour de  $x_0$  et selon à ce que  $f(x_0)$  est un maximum ou un minimum.

Un module maxima d'ondelettes est défini comme un point  $(x_0, s_0) \in L \times \mathbb{R}_+^*$  où la fonction  $|TWf(x, s_0)|$  est localement maximum.  $TWf(x, s_0)$  est la transformée en ondelettes de la fonction  $f$  à l'échelle  $s_0$  et  $|TWf(x, s_0)|$  sont module.

*Mallat et alts* dans [25] ont une proposition qui exprime le lien des modules maxima avec les singularités d'une fonction. Pour  $f$  prise dans un intervalle  $[a, b]$ , s'il existe une échelle  $s_0$  tel que  $|TWf(x, s_0)|$  n'a aucun maximum local dans cette intervalle pour tous les échelles inférieurs à  $s_0$ , alors  $f$  est uniformément Lipschitz  $-n$  sur tout intervalle  $[a + \varepsilon, b - \varepsilon]$  avec  $\varepsilon > 0$ .

La valeur de  $|TWf(x_0, s_0)|$  est d'autant plus importante pour un module maxima que la singularité est marquée. En outre, une singularité conduit à un certain nombre de maximum d'autant plus grand que l'ondelette est oscillante.

Pour caractériser la régularité en un point particulier, on étudie le comportement des coefficients d'ondelettes autour de ce point. De ce fait, on doit traiter précisément les variations de ces coefficients dans l'ensemble des points  $(u, s)$  du plan espace échelle ou les coefficients sont influencés par ce point. Ces coefficients constituent ce qu'on appelle un cône d'influence dans le plan espace échelle (*cf. figure 25*).

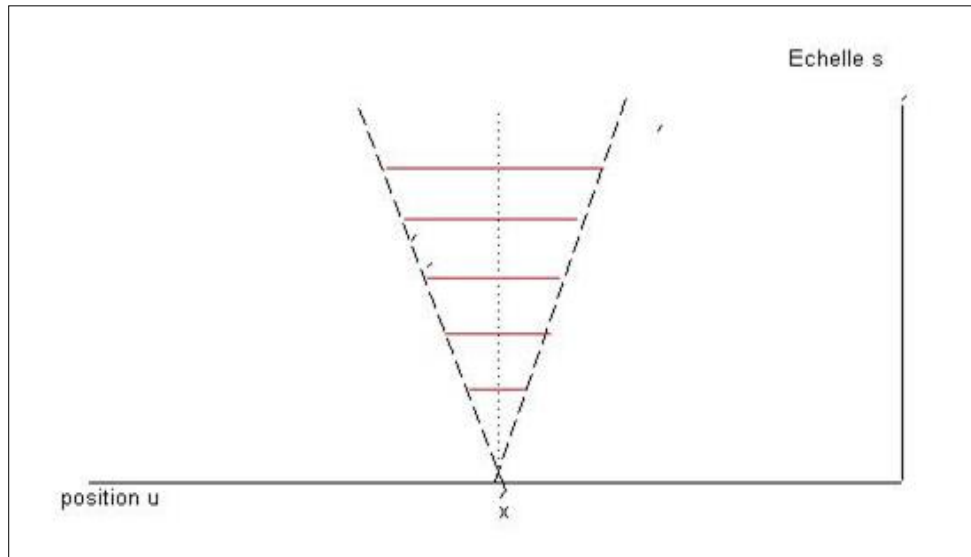


Figure 25 : Cône d'influence d'un point  $x$

Mallat suggère de suivre seulement les modules maxima à travers les échelles dans le cône d'influence au lieu d'analyser tous les coefficients, et il a montré que le comportement d'un module maxima à travers les échelles définit une courbe (une ligne maxima) dans le plan espace-échelle.

Généralement, une fonction  $f$  ne peut comporter des singularités qu'aux points où les modules de sa décomposition en ondelette sont maximums. Certaines décompositions en ondelettes garantissent que les maxima qui apparaissent dans un niveau d'échelle sont liés à des maxima dans les niveaux en dessous de cette échelle formant ainsi des lignes maximas. La détection et le calcul des singularités d'une fonction correspondent donc à la recherche des lignes maximas persistants à travers les échelles.

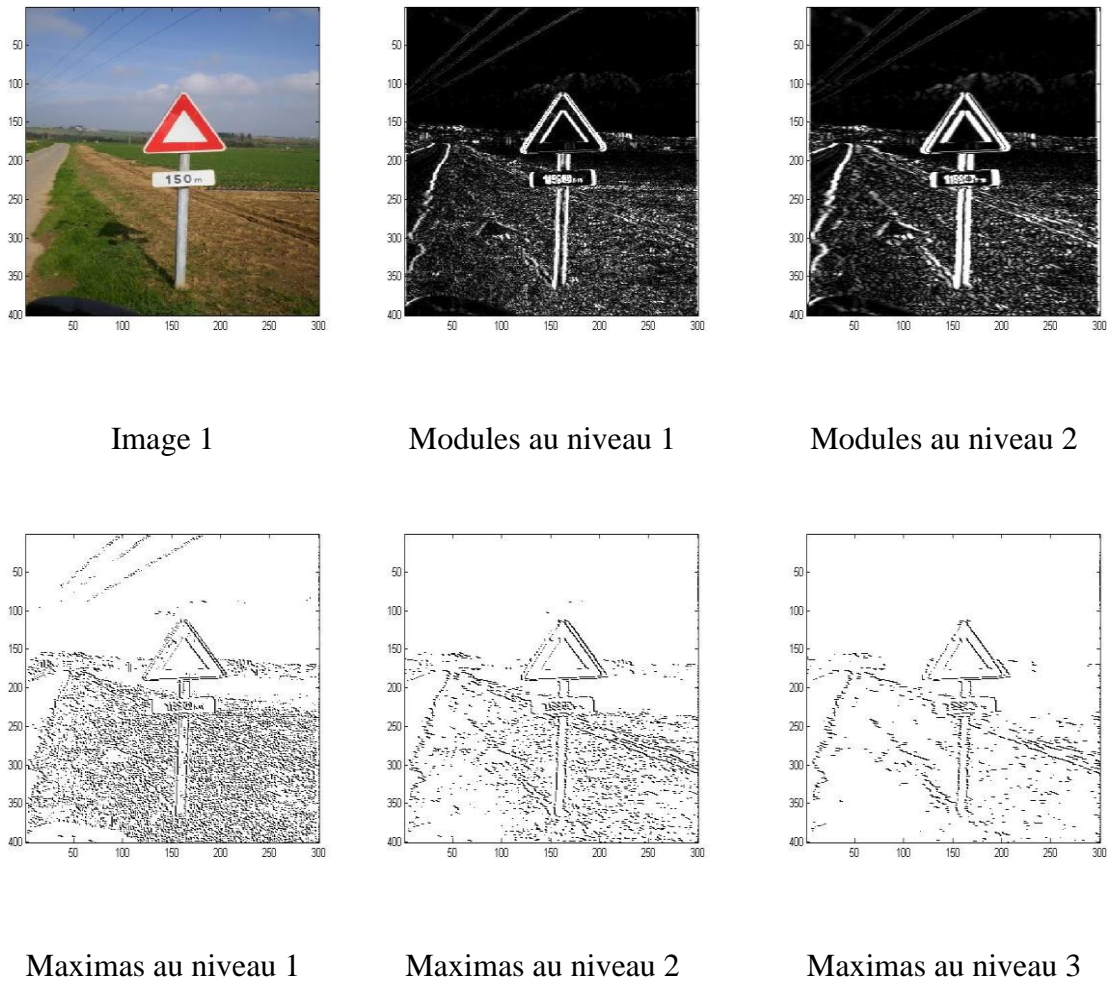


Figure 26 : Image 1, ses cartes de modules aux niveaux 1,2 ses cartes de maximas aux niveaux 1,2 et 3

Les singularités d'un signal comportent l'information la plus pertinente dans l'analyse de celui-ci. En effet, et dans le cas des images numériques ces singularités décrivent des points de forte changement d'intensité (points d'intérêt, contours, bruit,..).

L'utilisation des modules maximas s'avère très intéressante pour extraire les caractéristiques d'une image, notons que ces modules maximas permettent de mettre en évidence certaines structures locales dans les grandes échelles, les détails sont importantes dans les échelles les plus fines et ils se sont atténués et seule une tendance globale persiste (cf. figure 26).

## 2. Modules maximas et détecteur de Canny

L'analyse multi-échelle avec les ondelettes appliquée sur une image donne des coefficients

d'ondelette qui renseignent sur la variation locale des intensités autour des pixels. Ces variations autour d'un pixel sont exprimées dans le sens vertical, horizontal et diagonal. *Canny et als* [50] proposent une méthode issue du détecteur de Canny et développée par [51], et qui permet une autre formulation de la transformée en gardant seulement deux directions. Cette transformée nommée « transformée en ondelette au sens de Canny » est très adaptée pour la détection des contours et les points à fortes singularités [52]. Aussi, elle caractérise certaines propriétés géométriques des contours détectés et permet d'étudier leurs comportements multi-échelles.

Etant donné une image  $I$ , on considère deux ondelettes mères  $\psi^x$  et  $\psi^y$ , il s'agit, généralement, de dérivées suivant  $x$  et  $y$  d'un noyau de lissage, et on calcule les transformées en ondelettes, détectant les singularités respectivement horizontales et verticales, au pixel  $P$  et à l'échelle  $j$  par :

$$\begin{cases} D^1 I(P, j) = \left\langle I(X) \left| \frac{1}{2^j} \psi^x \left( \frac{X-P}{2^j} \right) \right. \right\rangle = I * \tilde{\psi}_j^x(P) \\ D^2 I(P, j) = \left\langle I(X) \left| \frac{1}{2^j} \psi^y \left( \frac{X-P}{2^j} \right) \right. \right\rangle = I * \tilde{\psi}_j^y(P) \end{cases}$$

où  $\tilde{\psi}_j^i(u, v) = \frac{1}{2^j} \psi_j^i(-u, -v) = \frac{1}{2^j} \psi^i \left( \frac{-u}{2^j}, \frac{-v}{2^j} \right)$  pour  $i = x$  et  $i = y$ .

Si on considère  $\psi^x = \frac{\partial \theta}{\partial x}$  et  $\psi^y = \frac{\partial \theta}{\partial y}$  où  $\theta$  est un noyau de lissage, les transformées d'ondelettes s'écrivent donc :

$$\begin{cases} D^1 I(P, j) = I * \tilde{\psi}_j^x(P) = I * \frac{\partial \tilde{\theta}_j}{\partial x}(P) \\ D^2 I(P, j) = I * \tilde{\psi}_j^y(P) = I * \frac{\partial \tilde{\theta}_j}{\partial y}(P) \end{cases} \quad (3.20)$$

Et grâce à la commutativité entre la convolution et la dérivation on a :

$$\begin{aligned} D^1 I(P, j) &= 2^j \frac{\partial}{\partial x} (I * \tilde{\theta}_j)(P) \\ D^2 I(P, j) &= 2^j \frac{\partial}{\partial y} (I * \tilde{\theta}_j)(P) \end{aligned} \quad (3.21)$$

Où  $\tilde{\theta}_j(u, v) = \frac{1}{2^j} \theta \left( \frac{-u}{2^j}, \frac{-v}{2^j} \right)$ . Ainsi, la transformée en ondelettes s'exprime sous sa forme vectorielle par :

$$DI(P, j) = \langle D^1 I(P, j), D^2 I(P, j) \rangle = 2^j \vec{\nabla} (I * \tilde{\theta}_j)(P) \quad (3.22)$$

Et on conclut que cette transformée n'est rien d'autre que le gradient de l'image lissée par un noyau de convolution.

Généralement, les maximas d'ondelettes au sens de Canny sont définis à partir du gradient de l'image, et plus précisément par son module et son orientation. L'intérêt majeur de ces maximas est qu'ils sont bien localisés sur les contours d'une image avec une sensibilité faible aux bruits.

### 3. Maximas des ondelettes et contours

L'un des problèmes majeurs dans le traitement de l'image est la détection de contours [37, 42, 59, 60]. En effet, et grâce à une caractérisation aussi précise que possible des contours, on peut avoir une information sur les différents objets présents dans l'image [61]. Plusieurs applications se sont appuyées sur les contours que ce soit pour la compression, la reconnaissance ou la prise de décision dans le domaine de l'imagerie numérique [60]. Le formalisme des ondelettes a contribué énormément à l'avancée dans ce domaine, et on trouve plusieurs travaux qui proposent des bases d'ondelettes exploitant la géométrie de l'image, on peut citer les ridgelets, les beamlets, les curvelets [50, 51, 62, 63, 64, 65] et d'autres bases. Toutes ces bases sont utilisées pour une meilleure caractérisation des contours multi-échelles dans une image numérique.

Si on revient à la définition d'un contour dans une image, on se trouve face à plusieurs façons possibles pour la donner. Nous allons choisir dans ce rapport la version la plus logique qui consiste à définir un contour comme un ensemble de pixels qui se distinguent par une forte singularité par rapport à leurs voisins. On dira qu'on a une singularité en un pixel si le gradient de l'intensité lumineuse est maximum dans la direction du gradient et c'est le cas lorsque la valeur du gradient au pixel en question est fortement supérieure à celles de tous ses voisins. Cela nécessite bien sûr une définition d'un seuil de comparaison et une topologie pour définir les voisins d'un pixel.

Cette définition suppose que l'image soit dérivable, c'est à dire régulière pour qu'on puisse disposer d'une image gradient exploitable, ce qui n'est pas souvent le cas pour les images réelles. Pour contourner cette difficulté on lisse l'image, pour la rendre régulière, par convolution avec un noyau de lissage puis on calcule le gradient par la suite. C'est le cas des travaux de Canny dans [50] exploités par Mallat et Zhong dans [47] pour donner une nouvelle notion de contours multi-échelles. Cette technique est basée sur l'exploitation de la



transformée en ondelettes, en établissant une image-contours à chaque échelle. Les points de cette image-contours, à l'échelle  $2^j$ , ne sont rien d'autres que les maxima locaux de la transformée en ondelettes à l'échelle  $2^j$ . *Mallat et als* [48, 51] ont montré que seule la donnée des images-contours, pour toutes les échelles, est suffisante pour reconstruire l'image d'une manière très satisfaisante.

Etant donnée qu'un contour dans une image peut être vue soit d'une manière discontinue si l'on observe dans la direction de son gradient, soit d'une manière continue et régulière si on l'observe dans la direction de sa tangente. Et puisque la transformée en ondelettes, comme nous l'avions expliqué, se ramène au calcul du gradient de l'image lissée, on va conclure que le contour est défini là où le module du gradient est maximum, d'où l'utilisation de la représentation polaire de la transformée en ondelettes. L'écriture polaire du vecteur  $DI(P, j)$  en calculant son module  $MDI(P, j)$  et son angle d'orientation  $AI(P, j)$ , est:

$$MDI(P, j) = \sqrt{(D^1I(P, j))^2 + (D^2I(P, j))^2}$$

$$AI(P, j) = \begin{cases} \tan^{-1} \frac{D^2I(P, j)}{D^1I(P, j)} \text{ si } D^1I(P, j) > 0 \\ \pi - \tan^{-1} \frac{D^2I(P, j)}{D^1I(P, j)} \text{ si } D^1I(P, j) < 0 \end{cases} \quad (3.23)$$

Le module de la transformée en ondelette est aussi le module du gradient de l'image lissée, et l'angle  $AI(P, j)$  est son orientation. L'image-contour, à l'échelle  $j$ , est obtenue en cherchant les maxima des modules coefficients d'ondelettes  $MDI(P, j)$ , dans la direction  $AI(P, j)$ .

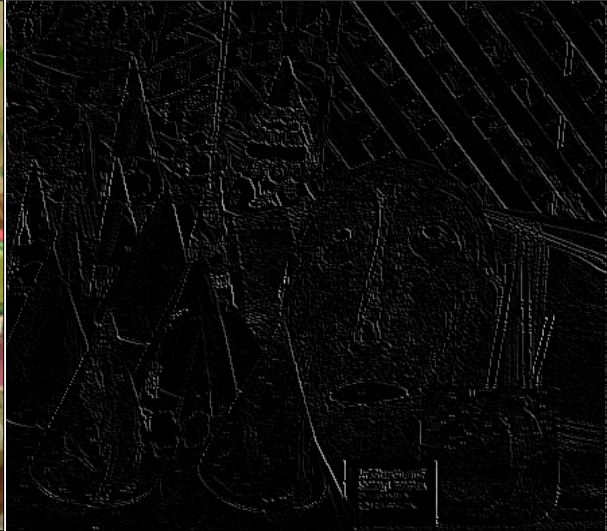


Image cones

Maximas niveau=1



Maximas niveau=2

Maximas niveau=3

Figure 27: Images contours dans les échelles : 1,2 et 3

### III. Conclusion

Nous avons donné un intérêt particulier aux ondelettes, vu le succès de la théorie des ondelettes dans plusieurs domaines et spécialement dans le domaine de traitement des images numériques. L'étude que nous avons présentée dans ce sens, concernant la théorie des ondelettes explique notre intérêt. Notre réflexion au sujet de l'analyse et l'interprétation des histogrammes 2D a été conclue par la décision d'adopter les ondelettes comme outil d'analyse et d'approximation. Ce choix des ondelettes est lié à leurs caractéristiques intéressantes et à la possibilité d'avoir des approximations multi-échelles des fonctions bidimensionnelles qui mettent l'accent sur les différentes structures tout en protégeant les détails.

Les points contours et les points de fortes transitions dans les images représentent l'information la plus pertinente. En effet, ces points décrivent les frontières des objets et renseignent sur les principales structures contenues dans l'image. Les coefficients d'ondelettes calculés par analyse multi-échelle permettent de mieux localiser ces points et suivre leurs comportement à travers les échelles. Ces caractéristiques intéressantes de l'analyse avec les ondelettes nous ont menés à l'exploiter dans un processus de mise en correspondance dans le cadre de la vision stéréoscopique.

Le premier chapitre et le deuxième chapitre ont mis l'accent sur les outils nécessaires pour que nous abordions la problématique de notre thématique de recherche. Il comprend une présentation des approches et outils utilisés communément par la communauté scientifique pour le traitement des images numériques. Nous donnons un intérêt particulier aux techniques de calcul des histogrammes couleur et à la théorie des ondelettes.

# **CHAPITRE 4:**

## **ANALYSE MILTIECHELLE DES HISTOGRAMMES 2D**

## I. Introduction

Le travail que nous proposons s'intègre dans un processus de traitement automatique des images couleurs issues d'un système de vision embarqué pour l'automobile. Toutefois, les images utilisées dans cette partie sont, soit issues du système de vision du laboratoire, soit artificielles et fabriquées sur mesure. Nous supposons que chacune de ces images est préparée en faisant appel à plusieurs fonctions de traitement de bas niveau, souvent traduites par un filtrage de l'image pour minimiser la présence de l'information inutile.

Ces fonctions sont liées soit au système d'acquisition, tel que le choix de l'espace couleur et qui peut être changé par un simple algorithme de conversion, soit aux connaissances à priori. Dans notre application, par exemple, l'une de ces fonctions se traduit par le fait de considérer que les panneaux routiers utiles se positionnent géométriquement en haut et à droite de l'image et que leurs couleurs sont souvent distinguées dans l'image (rouge, bleu, jaune).

L'image subit ensuite nos fonctions de segmentation pour reconstruire les régions qui correspondent aux panneaux routiers présents dans l'image. Nous exigeons de ces méthodes, seulement une confirmation de la présence des panneaux routiers. Nous supposons que seules les régions homogènes de l'image correspondent aux objets d'une scène, et cela convient aux images liées à la sécurité routière. Les panneaux routiers dans l'image peuvent être interprétés comme des régions très étroites et à forte énergie (les valeurs de la luminance et la chrominance sont maximums sur ces régions).

Nous présentons dans ce chapitre nos solutions pour surmonter les contraintes liés à la détection des panneaux routiers en utilisant une approche couleur. On se propose une approche de segmentation simple basée sur une classification des pixels de l'image via l'analyse des histogrammes 2D. Les décompositions des histogrammes 2D avec les transformées en ondelettes ont calculé des approximations de ces histogrammes dans plusieurs degrés de résolution. Ces approximations permettent d'identifier au mieux les modes significatifs des histogrammes à base desquels nous calculons les classes des pixels de l'image de départ. On note que les histogrammes 2D sont une simplification de la représentation statistique de la distribution de la couleur dans l'image, qu'on peut grandement améliorer avec un choix judicieux de l'espace couleur. Le fait de travailler dans un cadre 2D

et d'appliquer des analyses avec les transformées en ondelettes bidimensionnelles sur les histogrammes 2D a un impact sur la performance de l'approche et sur le temps de traitement.

Notons aussi que la propriété multi-échelle des transformées en ondelettes nous a permis de contrôler au mieux la différence des valeurs des pixels avec une précision maîtrisée. Cela a permis de différencier deux objets présents dans l'image même quand ils ont deux couleurs proches et en assurer la présence dans l'image.

## **II. Analyse des histogrammes et détection des panneaux**

### **1. Analyse des histogrammes**

Rappelons que les histogrammes 2D représentent la distribution des couleurs dans l'image selon des composantes prises deux par deux dans l'espace couleur choisi. L'histogramme *RB* par exemple donne la distribution du rouge et du bleu prise ensemble. De ce fait, les régions qui contiennent des quantités importantes de rouge ou de bleu ou les deux peuvent être clairement identifiées. Cependant, nous n'avons aucune information pour la composante verte.

L'utilisation des histogrammes 2D convient parfaitement dans un problème de segmentation mettant l'accent sur des composantes couleurs particulières de l'espace couleur choisi, et nous pouvons fusionner les résultats d'analyse de deux histogrammes pour perfectionner d'avantage la segmentation. Les panneaux de signalisation routiers ont des couleurs et des formes bien distinguées par rapport aux autres objets de la scène routière. De ce fait, l'utilisation des histogrammes 2D s'adapte à notre problème de détection des panneaux routiers.

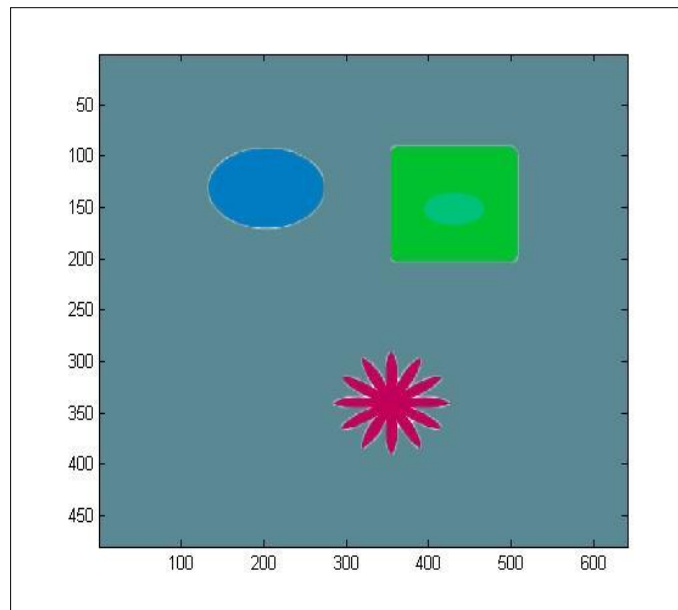
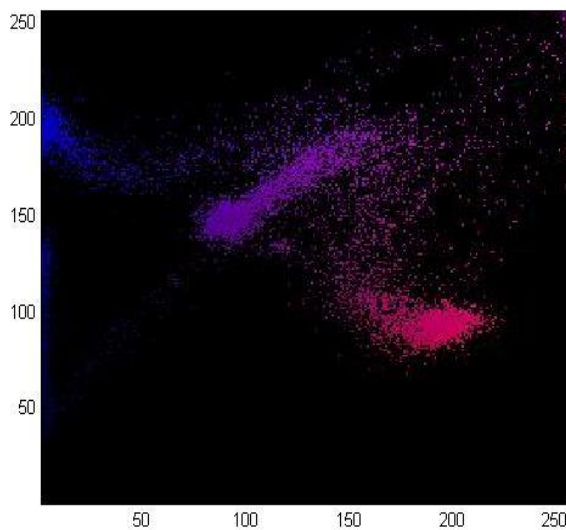
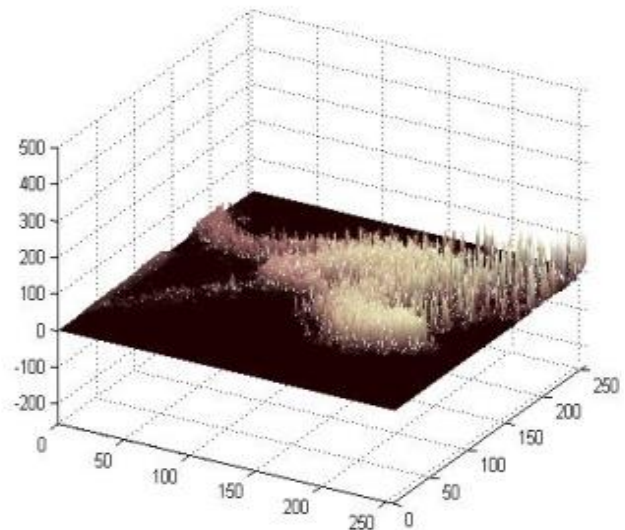


Image 1



histogramme RB



Vue 3D

Figure 28 : Vue 3D des histogrammes 2D

La figure 28 présente l'histogramme couleur RB d'une image synthèse (image1). Nous remarquons que les points couleurs présentés dans l'histogramme *RB* en haut ont un caractère très dispersé, cela a induit des nuages de points non compacts avec beaucoup de pics. Nous avons appliqué une analyse multi-résolution sur les histogrammes 2D, et nous avons obtenu dans chaque niveau de résolution une approximation qui représente une version plus lisse de l'histogramme avec des zones de couleurs plus compactes.

Les axes des histogrammes 2D dans l'espace couleur *RGB* sont tous échantillonnés sur un intervalle de valeurs  $[0,255]$ . Cependant, l'échantillonnage des axes des histogrammes dans l'espace couleur *La\*b\** varie d'une composantes à l'autre. L'axe *a\** contient des valeurs dans l'intervalle  $[-127,128]$  représentant l'opposition rouge-vert, l'axe *b\** contient des valeurs dans le même intervalle représentant l'opposition bleu-jaune, et l'axe *L* contient des valeurs dans l'intervalle  $[0,100]$  représentant le niveau de la luminance.

L'histogramme *a\*b\** décrit la distribution des couleurs dans le plan chromatique *a\*b\**. Le plan *a\*b\** contient les informations de corrélation entre les composantes *a\** et *b\** dans l'espace couleur *La\*b\** en ignorant la composante achromatique *L* représentant le niveau de la luminance dans les pixels. De ce fait, le résultat d'une segmentation basée seulement sur l'histogramme *a\*b\** sera théoriquement indépendant de la composante luminance. Le choix de l'espace *La\*b\** peut favoriser l'étude du facteur luminance dans une opération de segmentation des images couleurs, de plus, il permettra de contrôler mieux l'impact des changements de luminance pour les scènes réelles.

On présente dans la figure 29 l'histogramme *RB* de l'espace *RGB* et l'histogramme *a\*b\** de l'espace *La\*b\** de la même image couleur. Les vues 3D visualisées pour les deux histogrammes montrent que les régions couleurs apparaissent plus claires et plus compacts dans l'espace *La\*b\** que dans l'espace *RGB*. Ceci est dû aux propriétés d'uniformité de l'espace couleur *La\*b\**, et de la corrélation moindre entre les composantes couleurs relativement à l'espace *RGB*.

La méthode de segmentation proposée se base sur une analyse multi échelle avec les ondelettes des histogrammes 2D de l'image étudiée. Le but de l'analyse est d'extraire les pics de l'histogramme identifiant des régions de couleurs homogènes dans le plan image. Généralement, une analyse multi échelle dans un cadre bidimensionnel calcule une décomposition de l'espace de représentation en un sous espace d'approximation et trois sous espaces de détails. De ce fait, et au départ, les valeurs de l'histogramme 2D sont les coefficients de représentation dans un espace qu'on note  $V_0$ , est cela représente le niveau 0 d'analyse. L'analyse de l'histogramme consiste en une décomposition de celui-ci dans le sous espace d'approximation  $V_1$  et dans les trois espaces de détails  $W_1^1, W_1^2, W_1^3$  représentant ainsi le niveau d'analyse 1. Généralement, chaque niveau d'analyse *j* donne une décomposition de l'histogramme 2D dans un espace  $V_j$  dit espace d'approximation, et dans des espaces de détails  $W_j^1, W_j^2, W_j^3$ . Les coefficients de l'histogramme dans l'espace d'approximation



décrivent une version plus lisse de l'histogramme avec atténuation des détails. Les coefficients de l'histogramme dans les espaces de détails contiennent surtout les informations qu'on perd lors d'un passage d'un niveau de décomposition à l'autre et qui coïncident avec les hautes fréquences dans l'histogramme.

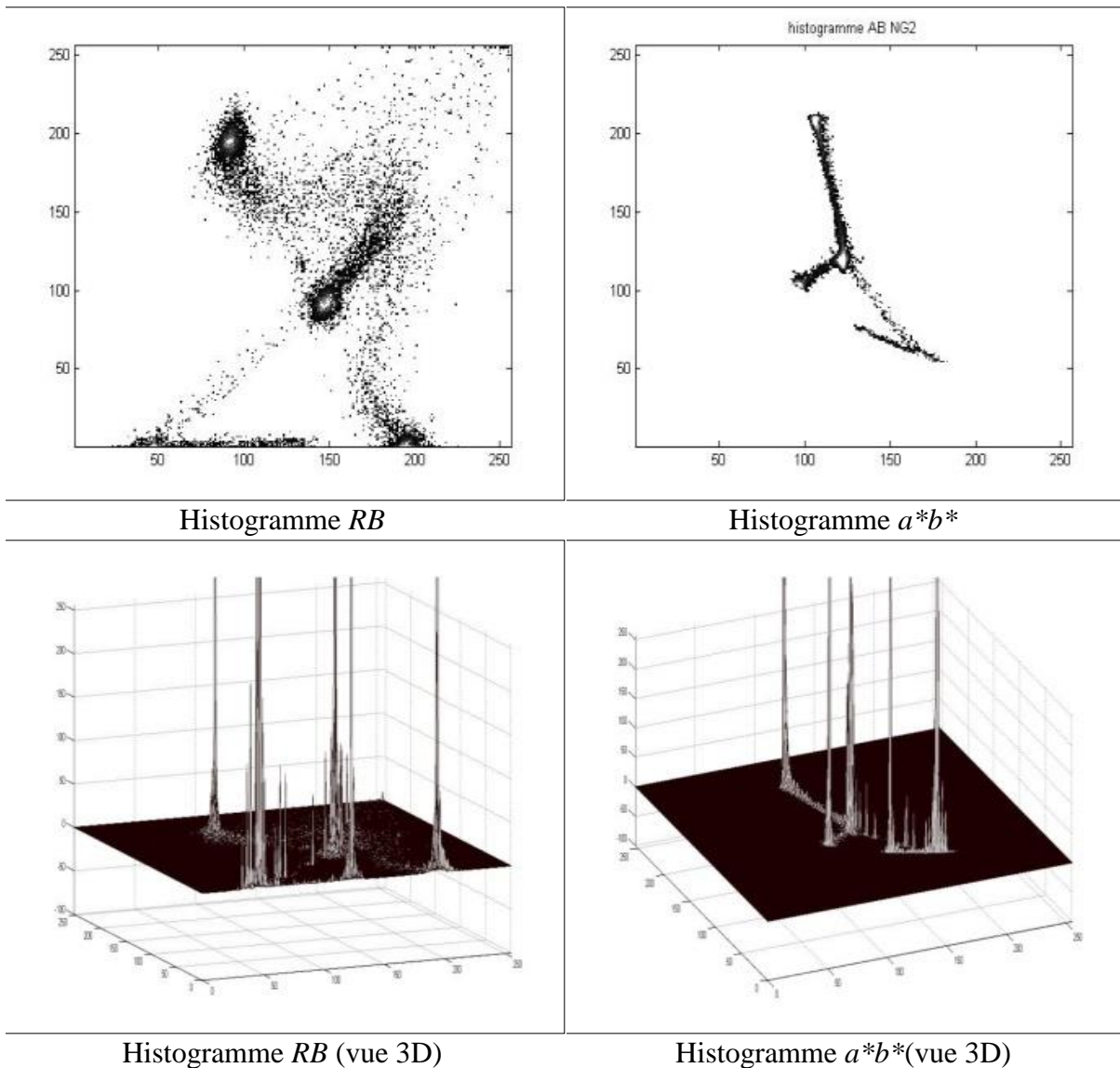


Figure 29 : Les histogrammes 2D dans les espaces  $RGB$  et  $La^*b^*$

La figure 30 illustre les décompositions de l'histogramme de l'image synthèse dans les niveaux 1,2 et 5. Avec l'augmentation de l'échelle d'analyse, nous avons une approximation représentant une version de l'histogramme plus lisse que les versions dans les échelles

antérieures. Ces versions présentent des zones plus compactes avec beaucoup moins de détails du à une forte dispersion des points. Les zones compactes définissent des domaines modaux qui se concentrent sur des valeurs pics caractérisant ainsi des régions homogènes dans l'image source. La classification à base des histogrammes repose sur l'identification des valeurs pics afin de regrouper les pixels définissant les domaines modaux correspondants dans une même classe. Toutefois, après calcul de l'histogramme, plusieurs valeurs pics y apparaissent et qui représentent des points isolés dans l'image, ou qui sont dues à une faible régularité de la distribution des couleurs au sien d'une même région de l'image. Les approximations multi échelles des histogrammes 2D réduisent non seulement les points isolés au fur et à mesure que l'échelle augmente, mais aussi elles régularisent d'avantage les domaines modaux.

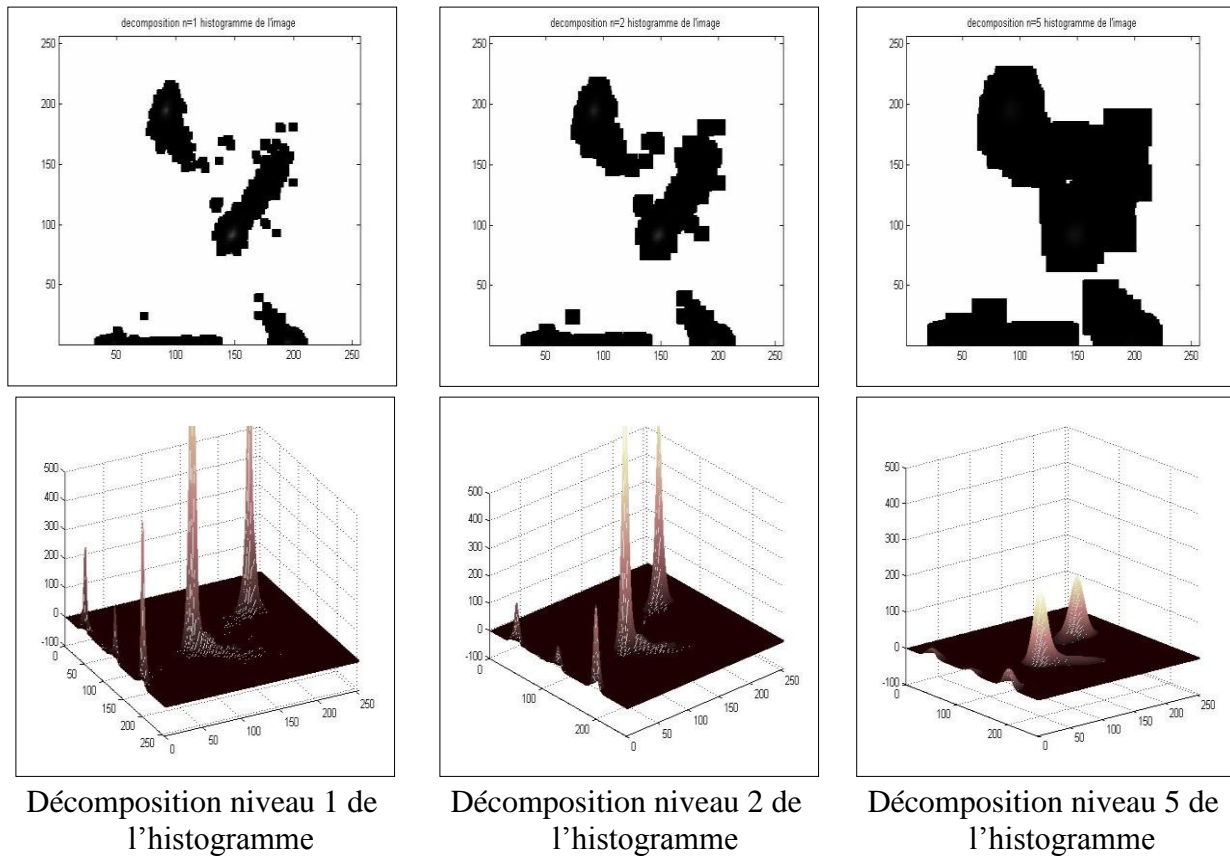


Figure 30 : Approximation de l'histogramme *RB* dans différents niveaux d'échelles

## 2. Approche de détection

Alors que les approximations de l'histogramme avec une analyse multi-échelle donnent des versions de l'histogramme avec moins de points isolés et avec moins de pics insignifiants, les détails de ces approximations modélisent la perte d'information lors d'un passage d'un niveau de résolution à un autre niveau plus grossier. De ce fait, l'évolution de ces détails peut être utilisée pour estimer la régularité de l'histogramme à chaque niveau de résolution. Un histogramme régulier manifeste des pics représentant les principales régions de l'image, et le détail de l'analyse du aux hautes fréquences devient moins important.

L'analyse multi échelle que nous avons appliquée sur l'histogramme a favorisé la discrimination des régions ayant des couleurs proches. Ces régions peuvent générer des domaines modaux qui se chevauchent dans l'histogramme, et la distinction des pics caractérisant ces domaines devient difficile. L'analyse effectuée en passant d'une échelle à l'autre a permis de mettre l'accent sur deux valeurs pics de couleurs suffisamment proches et les distinguer.

Le tableau 1 présente le résultat de l'analyse multi échelle avec les ondelettes appliquée sur l'image de synthèse *I*. Nous avons calculé dans chaque niveau d'analyse le nombre de pics présents dans l'approximation, et nous avons modélisé le détail généré résultat de cette approximation en calculant le nombre de modules maximaux présents. Le tableau récapitule l'évolution des pics et l'évolution des modules maximaux à travers les échelles.

Niveau d'analyse	Histogramme RB N=0	N=1	N=2	N=6	N=7	N=8	N=9
Nombre de pics	87	39	24	9	6	5	5
Nombre de maximaux		83	74	49	47	46	46

Tableau 1 : Le nombre de modes et de modules maximaux de l'histogramme de l'image *I* de niveau 0 au niveau 9.

Ce résultat montre que le nombre de pics et de modules maximaux de l'échelle 6 à l'échelle 9 commence de devenir assez stable. Au fur et à mesure que l'échelle augmente, les

approximations de l’histogramme manifestent des domaines modaux plus compacts et le degré de détail diminue.

Nous cherchons à obtenir des zones de points réguliers dans les approximations, ces zones décrivent les domaines modaux représentatifs, et traduisent les différentes régions de couleurs qui existent dans les images sources. De ce fait, les niveaux de résolution où les approximations de l’histogramme deviennent relativement régulières et uniformes constituent les niveaux de résolution adéquats pour extraire les modes significatifs. La classification des pixels à base de ces modes conduit à une meilleure caractérisation des régions homogènes de l’image, tout en évitant les éventuelles sur-classifications dues aux pics isolés dans l’histogramme.

Il faut noter qu’il est important de pouvoir calculer les modes de l’histogramme d’une façon quasi automatique pour une méthode de classification qui se veut être globale, et sans connaissances préalables. La plupart des méthodes de classification qui existent dans la littérature sont des méthodes supervisées. En effet, ces méthodes exigent des paramètres comme les centres des classes ou au minimum le nombre de classes pour classifier les pixels.

Notre méthode qui exploite l’analyse multi échelle de l’histogramme permet de calculer les modes de l’histogramme qui fournissent les classes de pixels consistant, et qui correspondent à des régions de forte homogénéité dans l’image originale.

Une illustration est donnée dans la figure 31, nous avons réalisé des classifications des pixels de l’image  $I$  à partir des modes extraits dans les approximations de l’histogramme bidimensionnel dans les niveaux d’échelles de 6 à 9 et nous présentons le résultat de la classification dans les niveaux 6,7, et 8. Pendant l’analyse, les niveaux de résolution fins inférieurs à 6 ont généré beaucoup de détails, et le nombre de pics dans ces niveaux est très important. C’est à partir du niveau 6 que le détail des approximations change peu pendant la décomposition, les grandes structures apparaissent plus claires dans l’histogramme et les approximations commencent à se stabiliser.

Nous avons calculé neuf modes dans l’approximation de niveau d’analyse 6, et les régions correspondant à une classification à base de ces pics induit une sur classification des pixels. En effet, les régions résultat de la sur classification correspondent surtout à des variations importantes de la chrominance au sein d’une même région ou dans les frontières de ces régions.

Le niveau d'analyse 7 fait apparaître six modes, et c'est à partir de ce niveau que nous avons un faible changement des détails de la décomposition de l'histogramme. Dans ce niveau d'analyse nous avons l'apparition d'une région de plus résultat de la sur classification (cf. figure 31), et la zone correspondante apparaît sur les frontières de la région bleu. Pour l'approximation dans le niveau d'analyse 8, et qui a généré cinq modes, nous identifions exactement les cinq régions de l'image. Nous avons une classification correcte des pixels de l'image, résultat de la bonne sélection des modes représentatifs de l'histogramme pour le niveau d'analyse 8. On note aussi qu'à partir de ce niveau, les modules maxima calculés deviennent relativement stables.

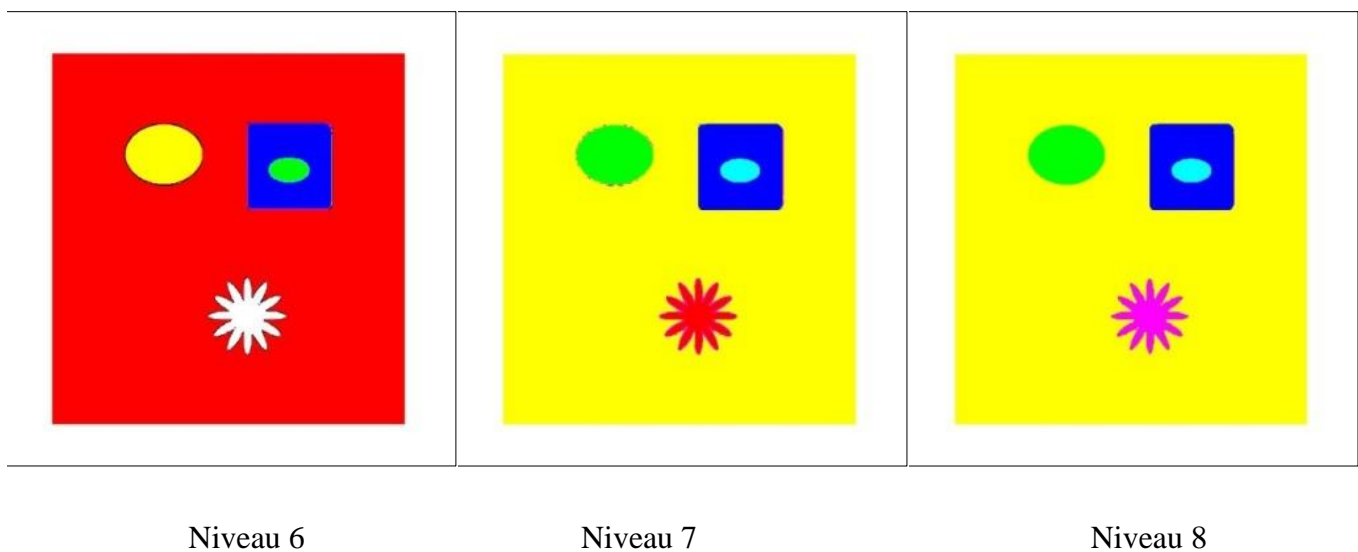


Figure 31 : Classification dans les niveaux de résolution 6,7 et 8

Nous avons résumé dans la figure 32 les étapes de l'algorithme de segmentation proposé. L'étape de calcul des histogrammes 2D doit être faite après le choix de l'espace couleur pour l'analyse et après le choix des composantes couleurs à étudier. L'étape d'analyse et le cœur de ce processus de segmentation, c'est une suite de décompositions itératives de l'histogramme 2D avec la transformée en ondelettes. De ce fait, chaque itération produit une approximation de l'histogramme plus une carte des modules maxima calculés à partir des détails résultats de l'approximation. Le résultat de l'étape d'analyse est la sélection des modes significatifs de l'histogramme 2D. Les modes calculés sont utilisés dans un processus de classification des pixels de l'image pour identifier les régions homogènes de l'image.

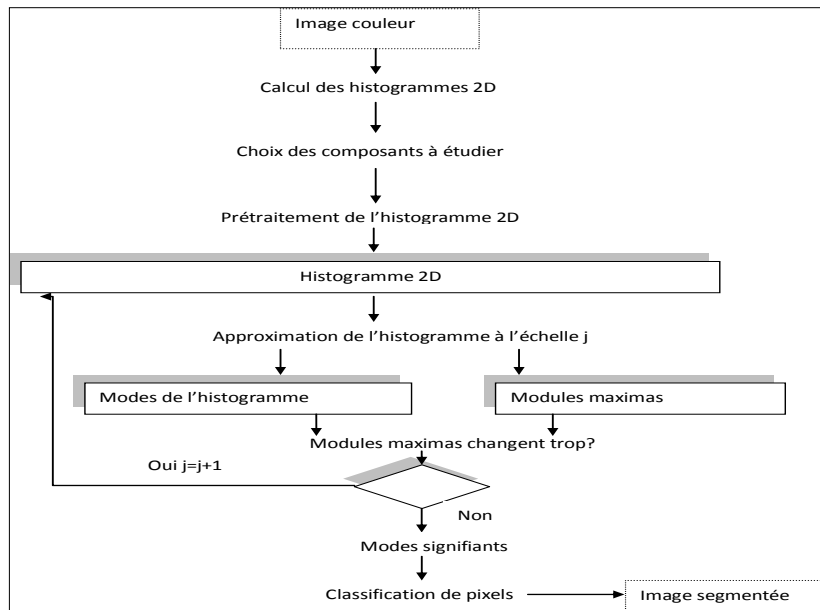


Figure 32 : Processus de segmentation à base d'analyse des histogrammes 2D

### III. Résultats de l'approche

Pour évaluer la qualité de l'approche proposée, nous nous sommes basés sur les techniques d'évaluation existantes et sur la comparaison des résultats avec les résultats obtenus avec d'autres algorithmes de segmentation des images qui utilisent la couleur comme critère de segmentation. Nous présentons également une comparaison des résultats de la segmentation dans l'espace  $RGB$ , avec les résultats de la segmentation dans l'espace  $La^*b^*$ .

Généralement, et en plus d'une évaluation visuelle (qualitative) utiliser pour mesurer l'efficacité d'une méthode de segmentation, les auteurs utilisent des critères d'évaluation quantitatifs pour mesurer la fiabilité des résultats. Ces critères d'évaluation sont utilisés pour comparer une approche de segmentation nouvelle à des approches existantes, ou bien pour tester ou paramétrer un algorithme en fonction de leurs applications. Malheureusement, nous n'avons pas un critère d'évaluation robuste pour évaluer les résultats d'une segmentation, mais nous pouvons utiliser plusieurs et combiner leurs résultats.

Plusieurs critères d'évaluation existent dans la littérature [44][45], et qu'on regroupe généralement en deux catégories : les critères d'évaluation supervisée basé sur une segmentation de référence appelée « vérité de terrain » et les critères d'évaluation non supervisée basée sur l'homogénéité des régions obtenues.

Une technique d'évaluation supervisée postule qu'on possède des connaissances à priori sur le domaine. Dans ce cas, on calcule la qualité d'une segmentation en comparant les résultats obtenus avec une vérité de terrain (segmentation de référence) représentant le résultat idéal. Ces critères d'évaluation supervisée mesurent au mieux la divergence des résultats de la segmentation à évaluer par rapport à l'image de référence.

L'avantage d'une méthode d'évaluation supervisée est qu'elle fournit des résultats fiables indépendamment de la nature des régions des images. Cependant, le problème de telle méthode est qu'elle demande une vérité de terrain qui n'est pas toujours disponible et difficile à construire.

Une technique d'évaluation non supervisée a l'avantage d'être automatique. En effet, c'est une quantification de la qualité des résultats de la segmentation sans connaissances à priori, et on se base sur des statistiques calculées pour chaque région détectée.

Nous avons évalué les résultats obtenus par notre approche de détection en utilisant des critères d'évaluation non supervisés. Cela est fait pour approuver la qualité de notre approche et la comparer avec l'approche marginale qui utilise la fusion des segmentations obtenues par des analyses des histogrammes 1D. Nous avons choisi de reprendre quelques critères proposés dans la littérature pour comparer les résultats.

## 1. Evaluation qualitative

Pour mieux décrire les résultats et approuver la qualité de notre approche de segmentation, nous avons choisi des images représentant des scènes réelles. L'image 1 de la figure 33 représente une scène routière. La région panneau (en rouge) de l'image est de taille petite par rapport aux régions constituant le fond de l'image. De ce fait, les régions de fond génèrent des nuages importants au niveau de l'histogramme, et la région qui correspond au panneau semble moins représentée. Les approximations multi-échelles de l'histogramme nous a permis d'identifier les petits nuages de points couleurs dans telles situations et de mettre en évidence les nuages de points liées aux pixels des panneaux routiers.

### Résultats de l'image 1 :

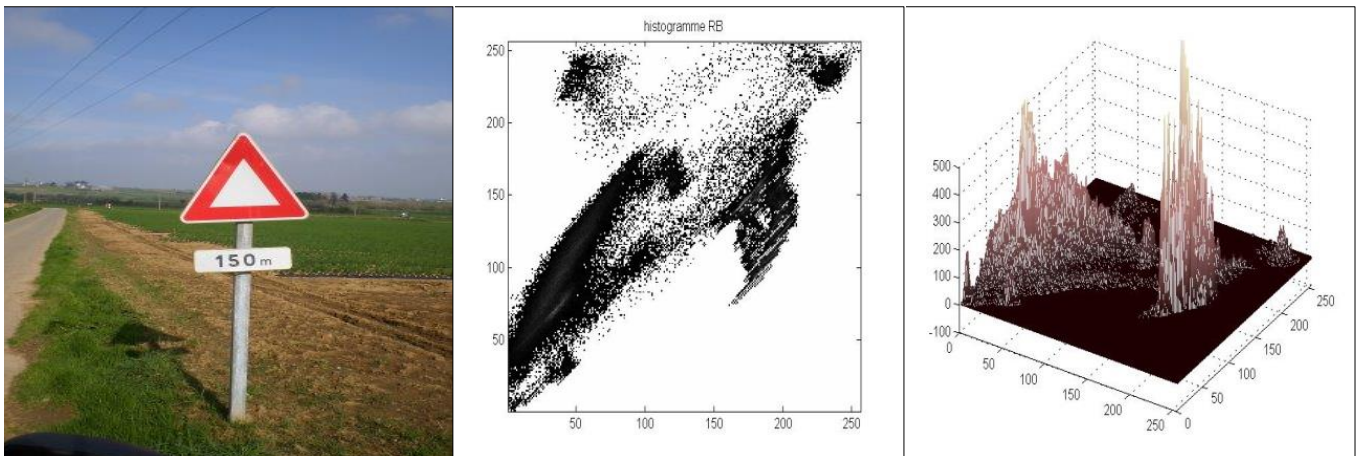


Image 1

Histogramme *RB*Histogramme *RB* vue 3D

Figure 33 :L'image 1, son histogramme *RB*, une vue 3D de l'histogramme

Pendant l'analyse des histogrammes de l'image 1, le niveau d'analyse  $N=9$  est le niveau le plus approprié qui a permis de détecter toutes les régions significatives de l'image, même ceux avec de petites tailles. L'analyse dans le niveau huit induit une sur classification résultat d'une approximation non encore uniforme. Cependant, le niveau dix génère un nombre de classes inférieur au nombre de classes généré par les niveaux neuf et huit.

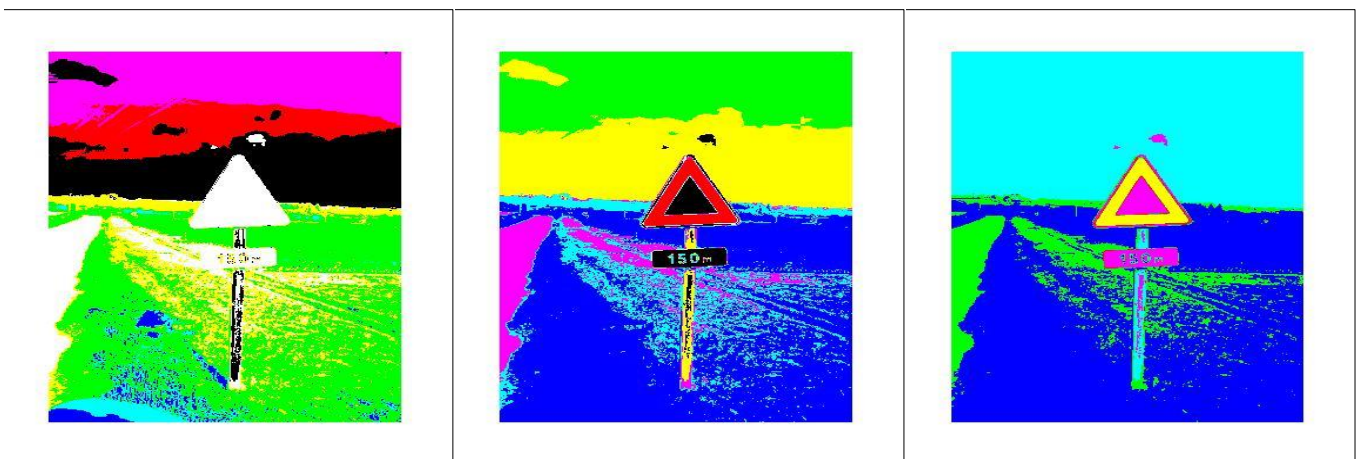
Image 1 :  $N=8$ Image 1 :  $N=9$ Image 1 :  $N=10$ 

Figure 34 : Des segmentations de l'image 1, pour les niveaux d'analyse 8,9 et 10



## Résultats pour l'image 2 :

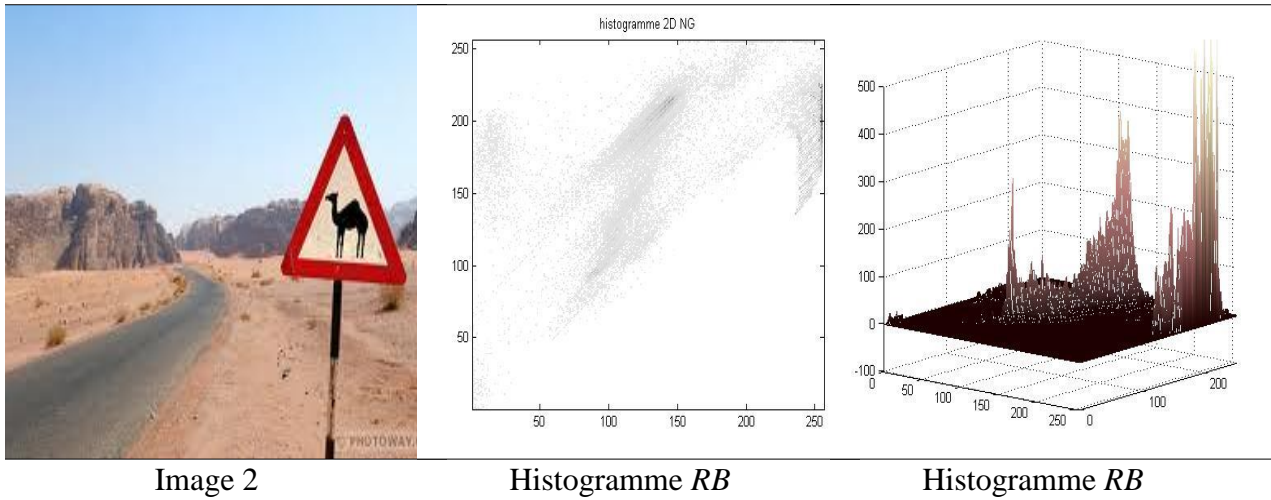


Image 2 segmentée : N=6



Image 2 segmentée : N=7

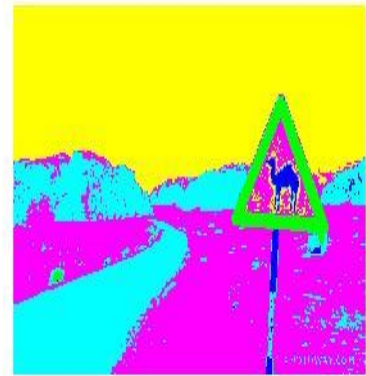


Image 2 segmentée: N=10

Figure 35 : Segmentations de l'image 2 dans les niveaux d'analyse 6,7 et 10

## 2. Evaluation quantitative

Les critères d'évaluation quantitative est une mesure d'homogénéité des régions ou de contrastes entre régions. Le principe de ces critères est simple, et il est basé sur la variabilité inter-région ou sur l'uniformité de l'intensité ou la couleur à l'intérieur des régions.

### Exemple de critères :

Critère de Levine et Nazif :  $CLN = \sum_{i=1}^N \frac{\sigma_i^2}{\sigma_{max}^2}$  avec  $N$  le nombre de régions et  $\sigma_i^2$  la variance de la région  $i$ .

Critère de Liu et Yang :  $CLY = \frac{1}{1000 \times A} \sqrt{N} \sum_{i=1}^N \frac{E_i^2}{\sqrt{A_i}}$  avec  $A_i$  la surface de la région  $i$ ,  $N$  le nombre de régions et  $E_i$  la somme des distances euclidiennes entre les couleurs des pixels de la région  $i$  et la couleur attribuer à cette région.

Le tableau suivant présente une liste des valeurs de quelques critères d'évaluation mesurés pour notre approche de segmentation (ligne 1) et pour l'approche de segmentation marginale (ligne 2) basée sur les histogrammes  $ID$  appliquées sur l'image 1 :

Levin et Nazif intra-region	Liu et Yang	Borsotti	Levin et Nazif inter-region
0.1202	0.0014649	0.9954	0.2161
0.2296	0.012656	0,18063	0.2046

Tableau 2 : Valeurs des critères d'évaluation: ligne 1(approche de segmentation proposée) ; ligne2 (approche de segmentation marginale)

Les valeurs des critères d'évaluation confirment la qualité de l'approche de détection proposée et appuient notre évaluation visuelle des segmentations obtenues. Généralement, les petites valeurs pour les critères de Levin et Nazif et Liu et Yang indiquent que les résultats d'une segmentation sont bons.

### 3. Résultats dans l'espace $La^*b^*$

Rappelons que l'espace  $La^*b^*$  fait partie des espaces couleurs uniformes adopté et standardisé par la commission internationale de l'éclairage CIE. Dans l'espace  $La^*b^*$  la couleur est caractérisée par un paramètre d'intensité  $L$  et deux paramètres de chrominance  $a^*$  et  $b^*$ . Nous avons repris les images traitées précédemment et nous avons appliqué la procédure de segmentation dans l'espace  $La^*b^*$ .

L'étude des modes calculés à partir des approximations multi-échelles et le suivi des modules maxima à travers les échelles montrent que le niveau six est le niveau d'analyse le plus approprié pour l'image synthèse (cf. Figure 36). La classification basée sur les modes, extraits de l'approximation de l'histogramme au niveau six, donne de bons résultats. L'évolution des modules maxima à travers les échelles confirme le choix de ce niveau d'analyse.

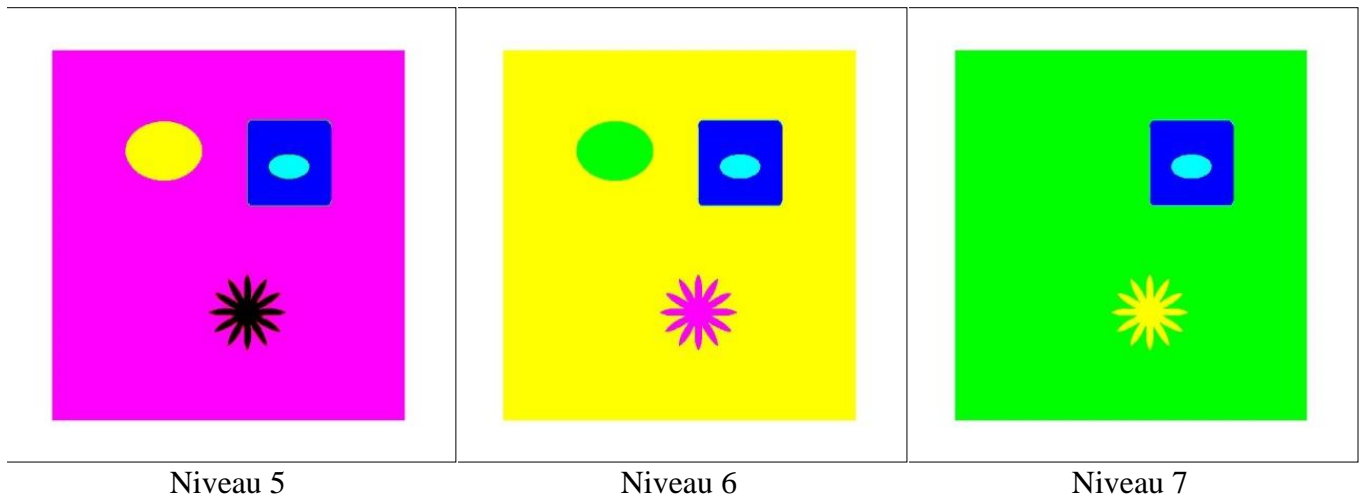


Figure 36 : Segmentation de l'image synthèse par analyse de l'histogramme  $a^*b^*$  dans les niveaux 5,6 et 7

Les résultats de la méthode de segmentation dans l'espace  $La^*b^*$  sont mieux que ceux obtenus dans l'espace  $RGB$  (cf. figures 36, 37,38). En effet, le nombre d'approximations de l'histogramme  $a^*b^*$  ou  $La^*$  est inférieur à celui de l'histogramme  $RB$ . Cela réduit considérablement le temps de traitement, et il est dû essentiellement à la dispersion réduite des nuages de points couleur dans les histogrammes  $La^*b^*$  par rapport à ceux des histogrammes calculés dans l'espace  $RGB$ .

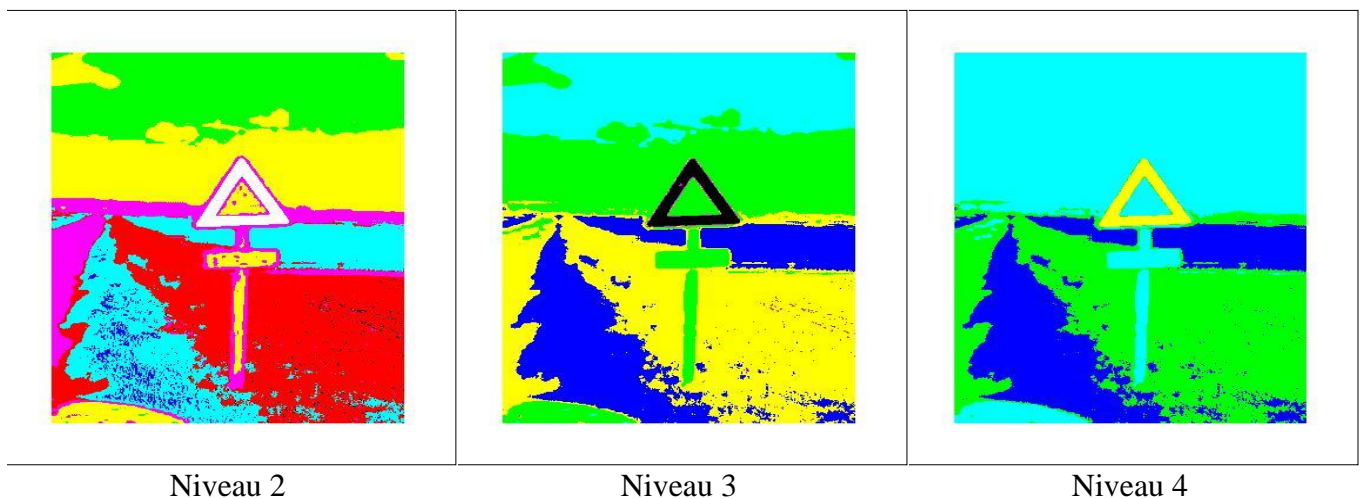


Figure 37: Segmentations de l'image 1 par analyse de l'histogramme  $a^*b^*$  dans les niveaux 2,3 et 4.

Ces résultats se justifient par l'uniformité de l'espace couleur  $La^*b^*$  et la faible corrélation entre les composantes couleurs par rapport à l'espace  $RGB$ . Nous avons aussi constaté que l'espace  $La^*b^*$  favorise l'identification des régions uniformes, et cela semble intéressant dans un post traitement cherchant les objets dans la scène en se basant sur leurs couleurs et leurs formes.

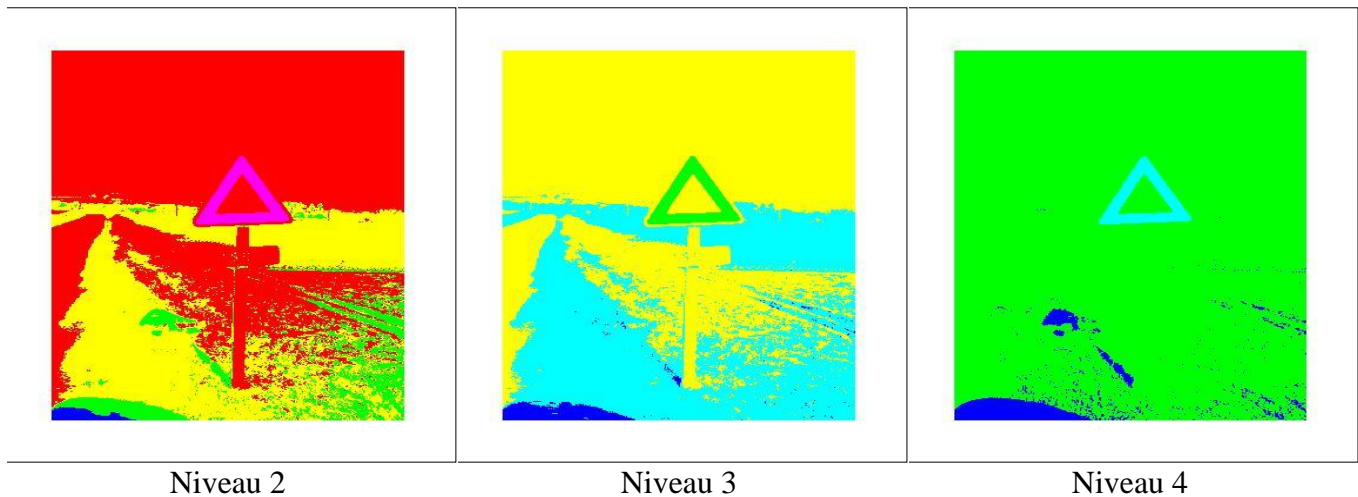


Figure 38 : Segmentations de l'image 1 par analyse de l'histogramme  $La^*$  dans les niveaux 2,3 et 4.

### Comparaison des résultats des deux espaces :

Images	Espace $RGB$				
	Levin et Nazif intra-region	Liu et Yang	Borsotti	Levin et Nazif inter-region	Nombre d'approximations
Image 1	0.1202	0.0014649	0.9954	0.2161	9
Image de synthèse	0.0193	0.73164	0.8973	0.94 72	11

Tableau 3 : Valeurs des critères d'évaluation dans l'espace  $RGB$

Images	Espace $La^*b^*$				
	Levin et Nazif intra-region	Liu et Yang	Borsotti	Levin et Nazif inter-region	Nombre d'approximations
Image 1	0.1879	0,00018275	0,8484	0.1642	3
Image de synthèse	0.0189	0.82032	0.9973	0.9844	6

Tableau 4 : Valeurs des critères d'évaluation dans l'espace  $La^*b^*$ 

L'espace  $La^*b^*$  manifeste de bonnes valeurs pour les critères d'évaluation quantitatifs sur nos images de tests. Aussi, en utilisant l'espace  $La^*b^*$ , nous pouvons mieux contrôler les problèmes de variation d'éclairage liés aux conditions d'acquisition ou aux conditions climatiques. Toutefois, nous avons toujours le problème de choix des composants à étudier dans l'histogramme 2D, ce problème nécessite d'étudier davantage les résultats de l'approche dans d'autres espaces couleurs notamment les espace couleurs hybrides.

#### IV. Conclusion

Dans ce chapitre, nous avons présenté les résultats de notre approche de détection sur quelques images de tests. Nous avons étudié les résultats de l'analyse des histogrammes 2D dans l'espace  $RGB$  et dans l'espace  $La^*b^*$ . Nous avons présenté les résultats pour une évaluation visuelle, et nous avons présenté les valeurs de quelques critères d'évaluation quantitatives.

L'exploitation des histogrammes 2D dans un processus de segmentation est beaucoup plus simple que l'exploitation des histogrammes 1D ou des histogrammes 3D pour les images couleurs. L'utilisation des histogrammes 2D dans un processus de classification est plus adaptée pour certains types d'applications. En effet, et pour des applications qui cherchent des objets d'intérêts particuliers (des couleurs particulières), on peut se restreindre à étudier seulement la projection sur deux composantes particulières et négliger les deux autres projections. Le choix des composantes couleurs à traiter doit se combiner à un choix adéquat

de l'espace couleur dans lequel l'analyse est faite, et cela afin d'améliorer d'avantage les résultats de la segmentation.

Pour justifier la qualité de notre approche de segmentation, nous avons réalisé une comparaison avec une approche marginale basée sur les histogrammes monodimensionnels. Cette approche marginale nécessite l'analyse de trois histogrammes 1D et fusionne ensuite les résultats en ignorant totalement la corrélation entre les composantes couleurs. Le choix des histogrammes 2D nous a permis de réduire le temps de calcul, de minimiser les erreurs liés à la corrélation entre les composantes couleurs, et réduire les éventuelles sur classification résultat de la combinaison.

Nous avons présenté les résultats de l'approche dans l'espace  $La*b^*$  pour montrer que les résultats d'une segmentation peuvent changer d'un espace couleur à un autre. En effet, chacun des espaces couleurs présente ces caractéristiques particulières. L'espace  $La*b^*$  est un espace de référence dans la colorimétrie, et qui présente deux propriétés intéressantes :

- Espace uniforme
- Il sépare l'information chrominance de l'information luminance.

Nous avons calculé des histogrammes 2D des images couleurs dans l'espace  $La*b^*$ . Ces histogrammes génèrent des nuages de points couleurs moins dispersés que les nuages des histogrammes calculés dans l'espace  $RGB$ . Par conséquent, cela a réduit le nombre d'approximations pendant l'analyse multi-échelles et a permis d'identifier très tôt les régions uniformes dans l'image.

Ce chapitre représente le premier axe de recherche que nous avons abordé et nos contributions pour réaliser une méthode de détection des panneaux routiers à base de la couleur. L'étude faite dans les chapitres précédents est essentielle afin de situer notre approche et de présenter les théories et outils sur lesquels nous nous sommes appuyés pour réaliser nos contributions.

Nous nous sommes basés dans notre travail, pour mettre en évidence la présence des panneaux, sur l'exploitation des histogrammes 2D auxquels nous appliquons une analyse en ondelettes. Cela nous a permis d'extraire les caractéristiques quantitatives des images à travers les histogrammes 2D. En effet, et dans un cadre bidimensionnel, l'analyse avec les ondelettes permet d'approximer un signal 2D (image) dans différents niveaux de résolution.

Ces approximations permettent de mettre en évidence certaines structures au fur et à mesure que l'échelle augmente.

Une étude des approximations itératives des histogrammes 2D a permis de mettre en évidence les pics décrivant les régions des images. Les modules maxima calculés pendant l'analyse décrivent les hautes fréquences et ils sont liés à des points de forte transition (généralement ça coïncide avec un fort changement de l'intensité). Pour les structures plus au moins réguliers au niveau des histogrammes, ces maxima connaissent un faible changement en passant d'un niveau d'échelle à un autre.

La technique choisie et les outils utilisés sont le résultat d'une importante réflexion pour surmonter les défis que présente une méthode de détection basée sur la couleur. Ces défis sont de plus en plus importants lorsqu'il s'agit d'une méthode de détection des panneaux routiers dans le cadre d'une conduite assistée. L'analyse multi-échelle nous a permis d'une part de mettre en évidence un panneau (une région petite dans l'image) présent dans une scène contenant des objets de tailles importantes, et d'autre part de distinguer la présence de deux objets de couleurs presque identiques.

Les résultats dans l'espace  $La*b*$  sont plus intéressants, cela confirme l'importance de l'espace couleur choisi pour l'analyse. Nous avons fixé comme perspectives, l'utilisation des systèmes d'axes indépendantes tel que le système d'Ohta ou même calculer un système de représentation adapté au type d'images que nous traitons. Cet axe de recherche est très important pour l'amélioration de la qualité des résultats et pour améliorer le temps de traitement.

L'évaluation des résultats de la segmentation par évaluation visuelle ou par les critères quantitatifs confirment nos choix. Nos résultats donnent de bonnes valeurs pour la majorité des critères et cela en comparaison avec une segmentation marginale avec les histogrammes 1D.

Au cours de ce travail, nous avons développé plusieurs fonctionnalités pour le test des différents algorithmes et techniques proposées. Cette problématique nous a servi de disposer de plusieurs codes propres écrits soit en Java, soit en Matlab enrichis par d'autres bouts de code propriétés de Matlab et d'autres bibliothèques publiques référencées. Nous avons fixé comme perspectives à court terme d'intégrer les différents algorithmes développés dans un Framework de détection des panneaux routiers en mettant en place des plateformes de test.

Généralement, les algorithmes conçus sont simples et ne sont pas coûteux en calcul mais ils sont à optimiser pour les adapter au temps réel.

Le travail présenté traite seulement la partie détection, l'algorithme est en mesure d'informer si la scène contient ou non un panneau routier. La partie détection constitue la première étape dans un processus de reconnaissance réalisé avec différentes techniques. Nous comptons étudier d'avantage les transformée en ondelettes et la caractérisation des singularités dans le sens de les exploiter dans une approche combinant la couleur et la forme en vue d'extraire des caractéristiques facilitant la reconnaissance des panneaux.

Le caractère multi-échelle des ondelettes et leur puissance dans la détection des points de transition dans les signaux, et particulièrement dans les images numériques, nous a conduit à les exploiter dans une méthode de mise en correspondance pour une vision 3D, que nous avons présenté dans le chapitre cinq de ce manuscrit.



# **CHAPITRE 5 :**

# **MISE EN**

# **CORRESPONDANCE**

# **MULTIECHELLE**

# I. INTRODUCTION

## 1. Avant-propos

La mise en correspondance stéréoscopique est l'une des techniques qui vise à enrichir les fonctionnalités d'un système de vision artificielle et reproduire les fonctions de la vision naturelle. Cela se résume dans la fonction de description d'une scène dans un milieu naturel qui est généralement en 3D. L'idée de base est de déterminer la position d'un point dans l'espace à partir d'une ou plusieurs images fournies par le système d'acquisition.

La vision 3D vise à calculer la profondeur des différents points de la scène à partir du plan image, cette action constitue un des objectifs de la vision intelligente, et c'est dans ce sens que la stéréoscopie intervient comme une technique utilisée pour reconstruire une scène 3D à partir des images prises sous des angles de vue différents.

Nous présentons dans cette partie une méthode de mise en correspondance basée sur les modules maxima des ondelettes. L'exploitation des chaînes maxima nous a permis d'affiner la recherche des correspondances. En effet, à base de la transformée en ondelettes des images stéréoscopiques, nous construisons des cartes de modules des coefficients d'ondelettes et des cartes de phases dans les différentes échelles à base desquelles on extrait les modules maxima et on construit les chaînes maxima. De ce fait, les points formant des chaînes maxima dans le plan espace-échelle vont être considérés comme des points d'intérêt dans le processus d'appariement. Les informations multi-échelles disponibles, vont permettre de chercher sous certaines contraintes géométriques, pour chaque point d'intérêt de l'image gauche son meilleur correspondant parmi les points constituant les chaînes maxima de l'image droite. Les résultats expérimentaux ont montré que le nombre de correspondance diminue au fur et à mesure que l'échelle augmente. Dans plusieurs essais, on obtient le correspondant unique par exploration des chaînes maxima de l'échelle la plus fine vers l'échelle la plus grossière avec un temps de calcul raisonnable.

## 2. Problématique de la mise en correspondance

Apparier les indices visuels extraits des images stéréoscopiques, est une étape très importante dans le processus de la vision 3D. Ce processus réunit les primitives extraites d'une image de gauche et une image de droite pour réaliser une reconstruction 3D d'une scène. La reconstruction 3D est basée sur l'estimation d'une grandeur mesurant la distance

entre les positions des primitives appariées et la construction d'une carte de disparités montrant la scène en mettant en évidence la profondeur des objets de la scène. La carte de disparités générée par le processus de mise en correspondance est utilisée pour estimer la carte de profondeur et reproduire une vue 3D de la scène.

La problématique a été traitée par plusieurs recherches sans parvenir à une méthode générale [10, 11,12]. Plusieurs approches existent dans le domaine et qui exploitent généralement les primitives des images stéréoscopiques telles que les points d'intérêt, les contours ou les régions. Toutefois, les différents types de primitives sont exploités par l'intermédiaire des mêmes caractéristiques qui sont soit photométrique, soit géométrique ou morphologique.

Les techniques de mise en correspondance restent difficiles en raison de la multiplicité des paramètres essentiels, principalement la pertinence des attributs d'appariement considérés. Dans le processus de mise en correspondance, la difficulté de trouver les bonnes primitives de correspondance est due à des problèmes potentiels dans un processus de vision 3D tels que les occultations, les contractions, les variations de luminance.... Tous les algorithmes proposés sont adaptés pour contrôler au moins une partie de ces problèmes.

En stéréovision, il y en a deux grandes classes d'algorithmes. La première est dite classe des algorithmes globaux, et qui prennent la totalité de l'image et traitent le problème d'estimation de la correspondance comme un problème de minimisation d'une fonction de coût globale représentant les erreurs de mise en correspondance. La seconde est dite classe des algorithmes locaux, et ils sont reclassés en une catégorie des algorithmes basés sur les régions des images et une autre catégorie basée sur les primitives des images. Les algorithmes basés sur les régions utilisent la corrélation de deux fonctions images sur les régions définies localement, alors que les algorithmes basés sur les caractéristiques établissent la correspondance entre les primitives des images.

Nos contributions dans cette partie s'inscrivent dans une nouvelle classe d'algorithmes [12,46] basés sur le formalisme mathématique des ondelettes. En effet, la mise en correspondance est le résultat d'une exploitation multi-échelle des modules maximaux des ondelettes et des orientations de ces modules calculés à partir des coefficients d'ondelettes. Cette classe d'algorithme constitue une solution intermédiaire entre les solutions locales et les solutions globales, et elle combine les techniques des approches locales et des approches

globales. Dans la méthode proposée, nous utilisons le concept d'analyse multi-résolution pour appairer les primitives des deux images en exploitant leurs décompositions dans différentes échelles.

Nous nous sommes basés sur le chaînage des modules maxima (construction des chaînes maxima) des ondelettes à travers les échelles et nous réalisons un appariement des différents points formant les chaînes pour trouver les bons correspondants dans l'image droite pour chaque extrémité de chaîne dans l'image gauche. La mise en correspondance est effectuée par estimation d'un test de similarité entre les points de chaque chaîne de gauche et les points des chaînes droite satisfaisant la contrainte géométrique épipolaire et d'orientation. La carte de disparités générée a un caractère multi-échelle et elle est construite selon un traitement simultané en espace et en échelle. L'aspect multi-échelle a permis de perfectionner la carte de disparités en permettant l'amélioration de la sélection des candidats potentiels à travers un critère multi-échelle basé sur une corrélation normalisée et une distinction géométrique. *Mallat et als* [47, 48,49] montrent que les extrémités des chaînes maxima coïncident avec les singularités de l'image caractérisées par leurs coefficients de Lipschitz, et avec un bon choix de la méthode de construction des chaînes maxima, l'algorithme d'appariement proposé prévoit non seulement la précision des résultats, mais il converge aussi rapidement.

## II. MISE EN CORRESPONDANCE MULTIECHELLE

### 1. Introduction

Les coefficients d'ondelettes calculés sur une image renseignent sur la variation locale des intensités autour d'un pixel de l'image, et cela grâce à la représentation des variations autour d'un pixel dans le sens vertical, horizontal et diagonal. Dans ce chapitre, nous allons présenter une autre formulation de la transformée en ondelettes en gardant seulement deux directions. Il s'agit d'une méthode issue du détecteur introduit dans les travaux de Canny [50] et développé par Mallat dans [51]. Nous obtenons ainsi une transformée en ondelettes adaptée, non seulement pour la détection des contours et surtout les points à fortes singularités [52] mais elle caractérise aussi certaines propriétés de ces contours à savoir leur orientation géométrique et leur comportement multi-échelle.

Basé sur l'idée que dans le traitement de l'image, les données visuelles sont mises de façon hiérarchique par l'échelle, la transformée en ondelettes décompose une image en blocs élémentaires qui sont bien localisées à la fois en échelle et en espace. En conséquence de cette

décomposition, l'image peut être représentée par une approximation de résolution grossière et le détail de toutes les échelles intermédiaires.

Parmi les principaux intérêts de la représentation en ondelettes est que l'examen des coefficients d'ondelette d'une image donne des informations à propos de son analyse. En effet, dans [50, 66], il est prouvé que la régularité accentue la décroissance des coefficients d'ondelettes et nous parlons de la régularité locale liée au support d'analyse en ondelettes [67, 68].

Ainsi, si l'image est en permanence régulière partout sauf en quelques points isolés, les coefficients d'ondelettes estimées seront affectés que si le support de l'ondelette d'analyse contient ces points. Les représentations multi-échelles sont les mieux adaptées pour se focaliser sur l'information dans l'image dont la régularité n'est pas homogène vu que les coefficients d'ondelette au delà d'un certain seuil se concentrent seulement auprès des singularités (contours, points isolés, etc.).

Une des décompositions multi-échelles qui exploite bien les propriétés d'ondelettes et qui permettent de faire une étude, dans des niveaux différents, sur les points de discontinuité présents dans une image est la transformée en modules maxima d'ondelettes proposé par Mallat et Zhong [47, 69].

Le principe des modules maxima d'ondelettes est de donner une image des caractéristiques pour les identifier, ils transmettent les caractéristiques évidentes de l'image en fonction des modules bien définis accompagnées de directions bien déterminées.

Dans ce contexte, nous pouvons choisir deux ondelettes  $\psi^x = \frac{\partial \Theta}{\partial x}$  et  $\psi^y = \frac{\partial \Theta}{\partial y}$  où  $\Theta$  est un noyau de lissage. En effet, et pour ce travail nous avons choisi une ondelette Gaussien ( $\Theta = G$ ) pour assurer la valorisation des modules maxima et assurer la continuité des chaînes maxima [47, 50]. Ainsi la transformée en ondelette d'une image peut s'écrire comme un opérateur différentiel multi échelle.

$$\begin{pmatrix} W^x f(u,v,j) \\ W^y f(u,v,j) \end{pmatrix} = \begin{pmatrix} f * \psi_j^x(u,v) \\ f * \psi_j^y(u,v) \end{pmatrix} = 2^j \vec{\nabla} (f * G_j(u,v)) \quad (5.1)$$

Les modules de la transformée en ondelettes sont définies par :

$$Mf(u, v, j) = \sqrt{|W^x f(u, v, j)|^2 + |W^y f(u, v, j)|^2} \quad (5.2)$$

Sa direction est définie par :

$$\theta f(u, v, j) = \text{tang}^{-1}\left(\frac{W^x f(u, v, j)}{W^y f(u, v, j)}\right) \quad (5.3)$$

Le module maximal de la transformée en ondelette peut être calculé par résolution de l'équation:

$$\partial Mf(u, v, j) = 0$$

Autrement dit, et à une échelle donnée le point  $(u_0, v_0)$  tel que  $Mf(u, v)$  est localement maximum selon la norme de gradient dans le voisinage de  $(u_0, v_0)$  dans la direction  $\theta f(u, v, j)$ .

## 2. Approche de mise en correspondance

### 2.1 Construction des chaînes maximales

Rappelons que les travaux de Mallat et Zhong ont montré que les coefficients d'ondelettes sont influencés par les singularités résidant dans le support de l'ondelette analysante. On parle de cône d'influence pour exprimer l'ensemble de points où les coefficients d'ondelette peuvent être influencés par la singularité d'un point. Un point singulier génère un maximum local pour les coefficients d'ondelette à travers le cône d'influence jusqu'à une échelle particulière. De ce fait, pour caractériser les singularités, on observe le comportement des maximums locaux à travers les échelles. L'étude des singularités se fait par construction des lignes connectant les maximums locaux dans le plan espace-échelle, se sont des courbes qu'on appelle lignes maximales ou chaînes maximales.

En pratique, nous commençons avec une image de taille  $(N, N)$ ,  $N = 2^J$ . Nous obtiendrons une transformation en ondelette dyadique où le paramètre échelle est une puissance de 2. La transformation est non décimée, la méthode génère pour chaque niveau d'échelle une transformation en ondelettes avec la même taille que la taille de l'image initiale (le même nombre de points que l'image). Ensuite, on calcule les modules maximaux d'ondelettes et ses orientations. Les résultats de cela sont deux matrices, une matrice contient les valeurs des modules maximaux et une matrice contenant les directions. Les valeurs des maximaux sont des valeurs strictement positives, la matrice des maximaux aura des valeurs non nulles là où nous

avons des maxima et des valeurs nulles dans les autres points. Chaque échelle générera une matrice des maxima que nous allons indexer par sa valeur.

Nous construisons les chaînes maximales par chaînage d'un module maxima de location  $(x_j, y_j)$  dans la matrice  $MMI(\dots, j)$  des modules maxima à l'échelle  $j$ , avec son successeur  $(x_{j-1}, y_{j-1})$  dans la matrice  $MMI(\dots, j-1)$  des modules maxima à l'échelle  $j-1$  [68, 70]. La recherche des successeurs sera limitée au voisinage définie par le cône d'influence, et les chaînes peuvent être indexées par  $(x_j, y_j)_{j=J \dots m}$  et on note  $CMI(x_j^i, y_j^i, m)$ . L'entier  $m$  indique la longueur de la chaîne, et il est définie tel que le dernier point de la chaîne soit un module maxima à l'échelle  $J-m$ . notons que toute chaîne est composée seulement de maxima et caractérisée par une extrémité  $MMI(x_j, y_j, J)$  à l'échelle  $J$  qui coïncide avec un point singulier de l'image (cf. figure 39).

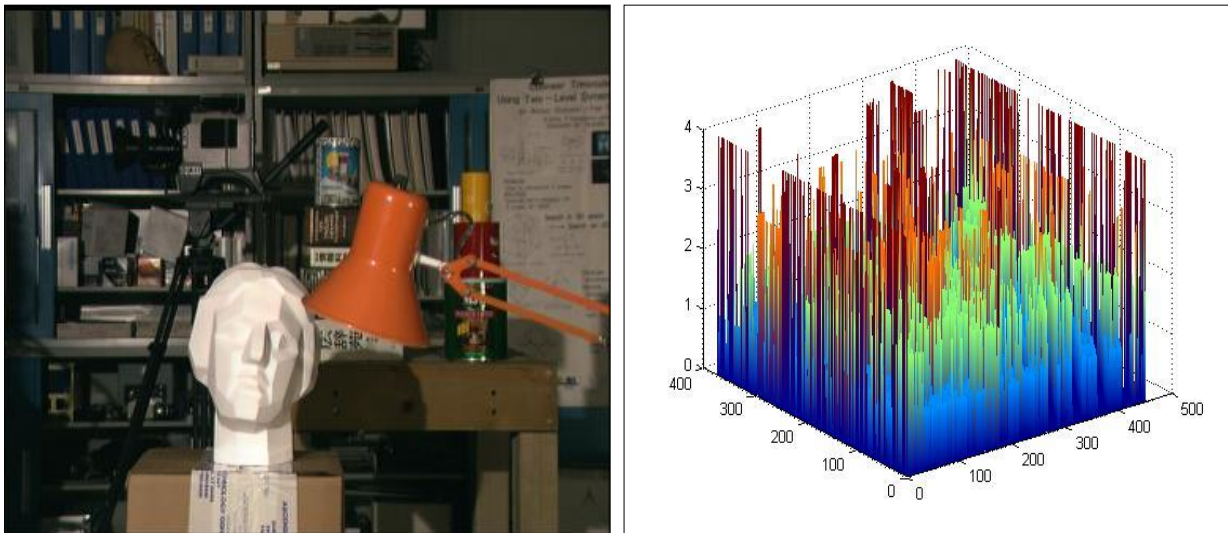


Image tsukuba

Chaînes maximales de l'échelle 1 à 4

Figure 39 : Evolution des maxima de l'image tsukuba de niveau 1 à 4 (avec seuillage)

## 2.2 Mise en correspondance multi échelle

Pour deux images couleurs rectifiées à partir d'un système de vision calibré. On associe à chaque image sa représentation en chaînes de modules maxima  $(x_j^i, y_j^i)_{j=J \dots m}$ ,  $i = g$  pour l'image gauche et  $i = d$  pour l'image droite. Avec les représentations définies au-dessus, nous avons trois approches d'appariement possibles. La première est d'apparier directement les

modules maximas  $MMI_i(\dots, j)$  des deux images en toute échelle  $j = J - 1 \dots L$ . La deuxième performe l'appariement des extrémités de chaque chaîne  $CMD(x_j^d, y_j^d, m')$  par un algorithme non linéaire. La troisième approche propose de considérer les chaînes maximas comme primitives.

Durant ce travail nous développons la deuxième approche et nous donnons les idées de traitement avec la troisième, la première semble approximativement au travail de Bhatti fait dans ce sens et on invite le lecteur à le découvrir dans [71].

Pour la chaîne  $CMG(x_j^g, y_j^g, m)$  de l'image gauche, on cherche dans toutes les chaînes de l'image droite telle que les extrémités respectent la contrainte épi-polaire avec le pixel  $(x_j^g, y_j^g)$ . De ce fait, tous les maximas de l'image gauche qui n'ont pas de correspondant dans l'image droite vérifiant la contrainte épi-polaire sont supprimés. Ainsi, on performe le test de similarité entre  $MMI(x_j^g, y_j^g, J)$  et toute module maximas dans les extrémités des chaînes droites sélectionnées. On note  $\{(x_{j,k}^d, y_{j,k}^d)_{k=1..c}\}$  l'ensemble des extrémités de la chaîne droite qui ont réalisé le test de similarité. Si  $c = 1$ , on conclut que  $(x_j^g, y_j^g)$  est mappé avec  $(x_{j,1}^d, y_{j,1}^d)$ . si  $c > 1$  on passe à l'échelle suivante, et on examine le test entre les successeurs, voir entre  $(x_{j-1}^g, y_{j-1}^g)$  et  $\{(x_{j-1,k}^d, y_{j-1,k}^d)_{k=1..c}\}$ . Cette étape sera itérée dans l'échelle  $j$ , jusqu'à la fin de la chaîne en assurant nécessairement la mise en correspondance de  $(x_j^g, y_j^g)$  et l'un des points de  $\{(x_{j,k}^d, y_{j,k}^d)_{k=1..c}\}$ .

La similarité est évaluée, sous la contrainte d'orientation, par utilisation de l'une des mesures de similarité disponibles en littérature. La majorité des mesures de similarité utilisées dans les méthodes de mise en correspondance sont des mesures de distorsion (SAD) et (SSD) et la corrélation croisée normalisée (NCC) [71]. Les mesures SAD et SSD sont connues plus rapide en calculs que le NCC qui est plus précis et qui nous permettra de raffiner le test de similarité. Dans les échelles fines, nous avons plusieurs points contours résultats d'un effet de bruit de l'image dans la représentation des modules maximas. La présence de ces faux maximas et comme il est connu que les mesures SSD et SAD sont justifiées lorsque nous avons un bruit additive Gaussien ou exponentiel, nous avons opté pour le choix d'une combinaison des tests de similarité NCC et SAD pour les échelles fines. A noter que pour les échelles grandes le processus de lissage supprime la majorité des faux maximas et une mesure NCC est suffisante.



Pour trouver la correspondance entre les pixels des maximas  $MMG(x_j^g, y_j^g, j)$  et  $MMD(x_j^d, y_j^d, j)$ , on maximise la corrélation ou on minimise la distorsion. Pour chaque point  $(x_j^g, y_j^g)$ , on cherche le point  $(x_j^d, y_j^d)$  dans la chaîne  $CMD(x_j^d, y_j^d, m')$ , d'un voisinage de taille  $s$  qui a le maximum de corrélation ou le minimum de distorsion avec le voisinage du point  $(x_j^g, y_j^g)$  dans la chaîne  $CMD(x_j^g, y_j^g, m)$ . selon la contrainte épi polaire  $y_j^d = y_j^g$  et la recherche est faite seulement selon  $(x_j^d, y_j^g)$  dans la ligne épi polaire.

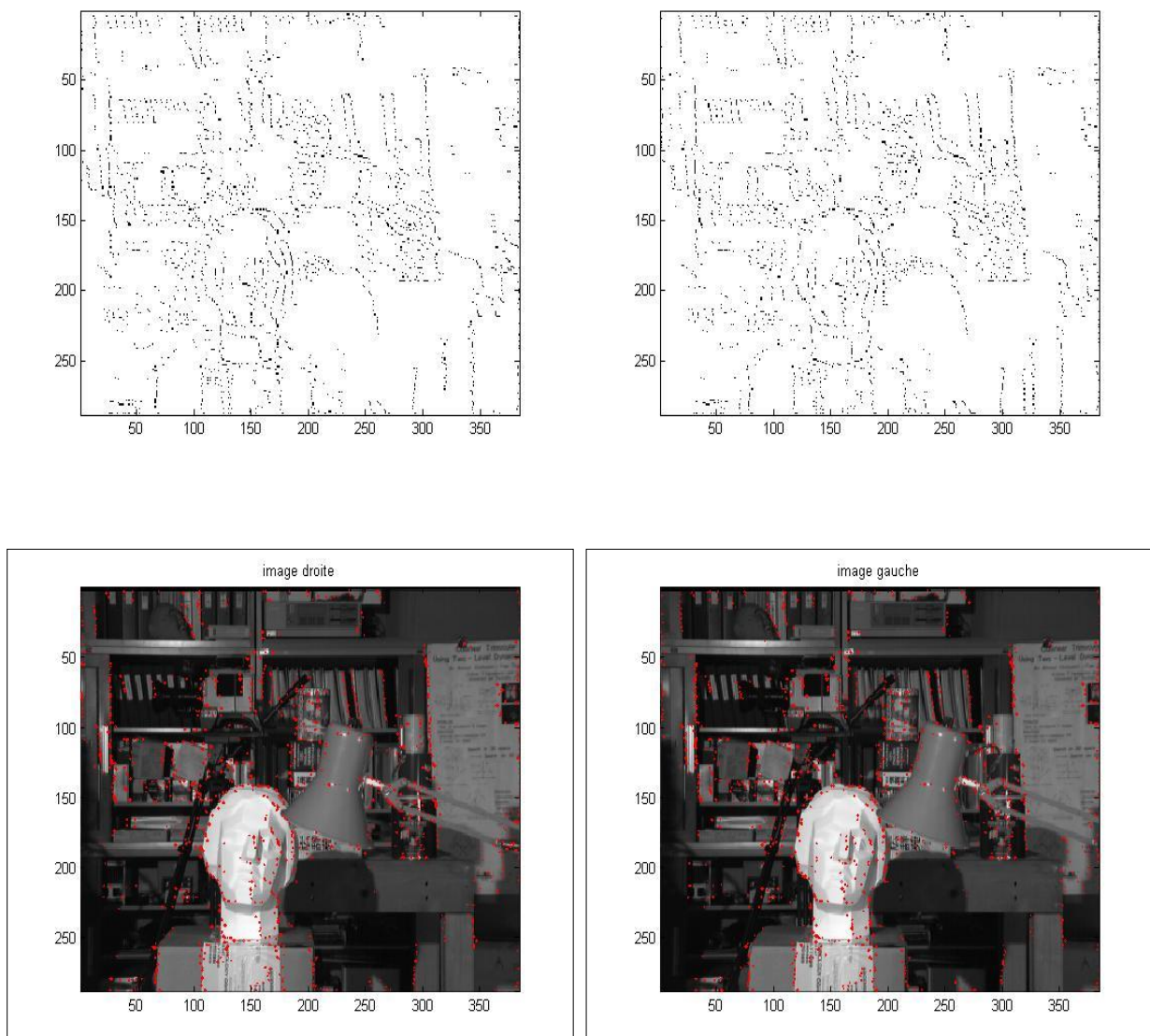


Figure 40 : Maximas appariés dans les images droite et gauche de l'image tsukuba

Dans la figure 40, nous avons calculé les chaînes maximales de l'image Tsukuba et nous avons réalisé un appariement multi échelle de ces chaînes maximales selon l'approche proposée. Les appariements validés sont présentés sur les deux images gauche et droite montrant ainsi la qualité des appariements et les points d'intérêt que présentent les points extrémités des chaînes maximales appariées.

### 3. Résultats de l'approche

Nous avons testé la méthode présentée dans ce chapitre sur des images téléchargées du site [72] [www.middlebury.edu/stereo/](http://www.middlebury.edu/stereo/). Ces images sont créées pour le test des algorithmes de mise en correspondance et l'évaluation des cartes de disparités calculées. Nous présentons les résultats sur quelques images avec la même luminance.

Nous présentons l'application de l'algorithme dans différents niveaux d'échelles  $j$  avec un suivi du comportement des maxima de la transformée en ondelettes à travers ces échelles. Les figures 41,42 et 43 montrent l'évolution des chaînes maximales calculées du niveau 1 au niveau 4, et l'appariement concerne seulement les maxima qui persistent à travers les échelles. Les figures présentent les cartes des maxima appariées et les points d'intérêts générés sur les images en question.

Les résultats expérimentaux sont réalisés en utilisant une ondelette gaussienne. Nous y présentons les points  $(x_j^d, y_j^d)$  de l'image droite qui ont validé le test de similarité avec le point  $(x_j^g, y_j^g)$  de l'image gauche. En effet, les points appariés définissent les profondeurs des primitives existantes dans les deux images. Cela est théoriquement suffisant, car les primitives contiennent l'information la plus pertinente dans l'image, et le chaînage des modules maximaux nous a permis d'affiner d'avantage l'appariement en restreignant l'appariement aux contours les plus significatifs des objets présents dans les images.

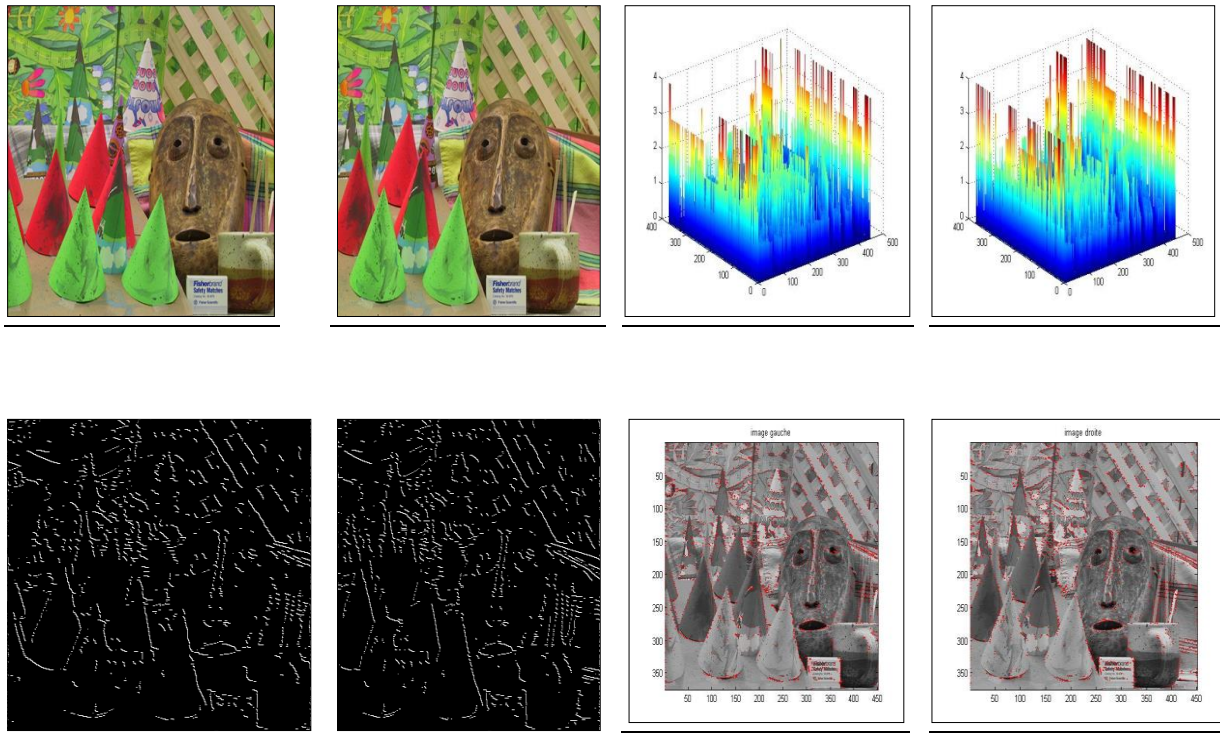


Figure 41 : Image cones (gauche, droite), les chaines maximales de niveau 1 à 4, les cartes des maximas appariées, les points d'intérêt correspondants

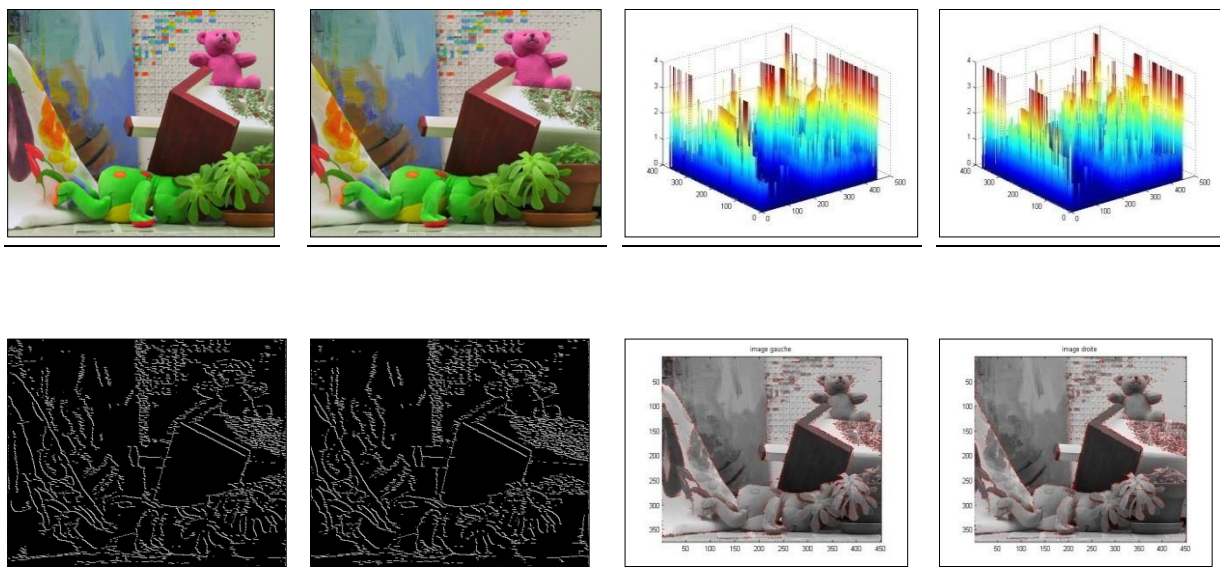


Figure 42 : Image teddy (gauche, droite), les chaines maximales de niveau 1 à 4, les cartes des maximas appariées, les points d'intérêt correspondants

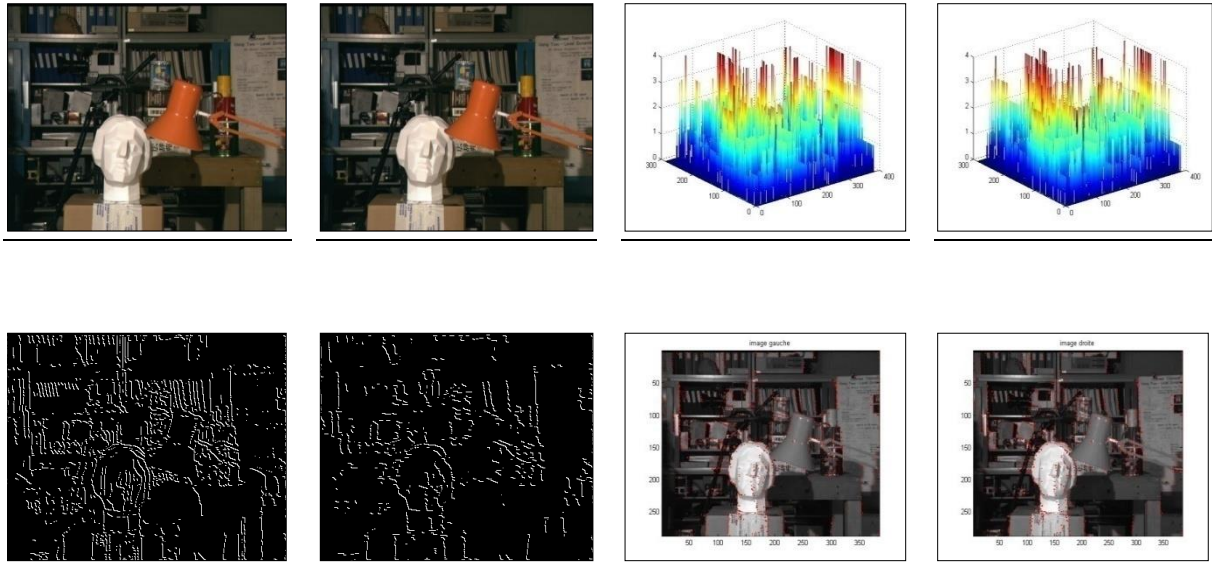


Figure 43 : Image tsukuba(droite, gauche), les chaines maximas de niveau 1 à 4, les cartes des maximas appariées, les points d'intérêt correspondants

Pour valider la fiabilité de nos appariements basés sur les primitifs maxims, nous avons utilisé les techniques de propagation de germes [96] pour construire des cartes de disparité pour l'ensemble de l'image (cf. figure 44). Les cartes présentent les profondeurs des objets dans nos images de test qu'on peut perfectionner avec des techniques de raffinement.

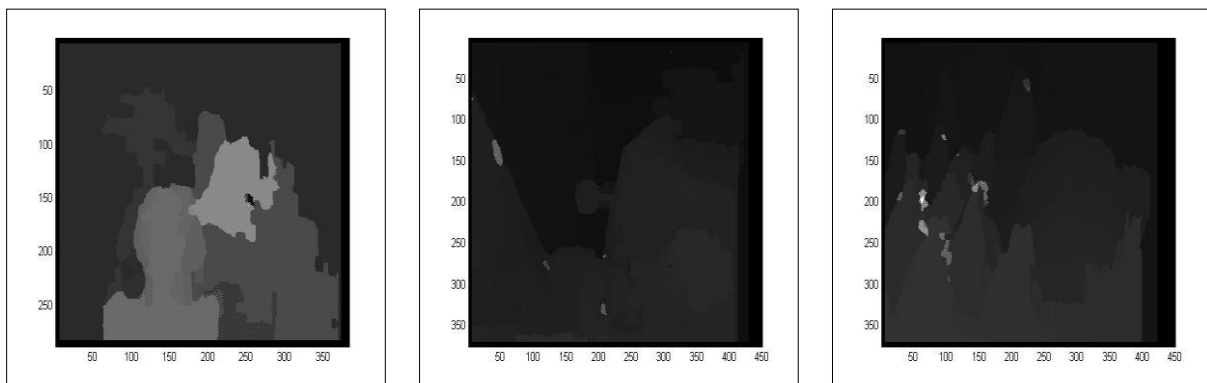


Figure 44 : Cartes de disparités des images tsukuba,teddy et cones sans post traitement

### III. Conclusion

Ce chapitre présente une introduction à un algorithme non linéaire de mise en correspondance multi-échelle basé sur les modules maxima des ondelettes comme primitives des images. Les chaînes des modules maxima sont utilisées pour trouver la meilleure correspondance pour un maxima de l'image gauche parmi les maxima candidats de l'image droite.

Le processus exploite intelligemment l'aspect multi-échelles des maxima et offre un algorithme précis et rapide. Notons aussi que les premiers résultats de construction des cartes de disparité sont très encourageants. La suite de ces travaux sera d'affiner les cartes de disparités et les comparer aux techniques référencées et publiées dans la base de données Middlebury.

La seconde perspective c'est d'intégrer la régularité Lipschitz des pixels [[52](#),[67](#), [73](#)], dans le processus de test de similarité comme une information additive. Il est prouvé qu'il existe une forte relation entre les maxima et la régularité d'une fonction. Ça nous permettra d'exploiter d'avantage la nature des modules maxima comme primitives caractérisant différentes singularités dans l'image.

# CONCLUSIONS ET PERSPECTIVES

Le système de vision humaine est un système difficile à reproduire. De ce fait, et en l'absence de solutions scientifiques générales pour réaliser une tâche de vision, les solutions existantes sont des solutions approximatives attachées à des domaines d'application particuliers et traitent le problème d'un point de vue spécifique. Les applications de la vision artificielle sont très nombreuses, et dans ce mémoire nous avons abordé des problèmes d'utilisation des systèmes de vision dans le domaine de la conduite assistée.

Les travaux présentés s'inscrivent dans le cadre d'un projet de développement d'un système de conduite intelligente adopté par le laboratoire de la recherche scientifique (LabTIC). Le principal objectif de ce projet est de développer un système de vision embarqué à base des outils d'analyse et de traitement des images numériques.

Généralement, il y en a deux principaux axes de préoccupations pour les concepteurs de tels systèmes : l'axe de la vision 3D et l'interprétation de l'entourage (détection des obstacles) et l'axe de reconnaissance des panneaux routiers.

Dans ce mémoire, nous avons apporté des contributions dans les deux axes. En partant de ce qui est déjà réalisé dans le domaine et en étudiant les outils et les techniques disponibles. Cette étude a orienté nos travaux et nos choix clés pour développer nos propres approches soit pour réaliser une méthode de reconnaissance des signaux routiers, soit pour traiter le problème de mise en correspondance dans le cadre d'une vision 3D.

Tout au long de ce mémoire, nous avons adopté les ondelettes comme outils d'analyse, ce choix stratégique a été appuyé par le succès de l'analyse en ondelette dans le domaine de l'imagerie et la simplicité des algorithmes de calculs surtout d'un point de vue implémentation. Ce choix s'est confirmé par les algorithmes développés et les résultats obtenus sur nos images de tests.

Dans la première partie, nous avons proposé une méthode de détection des panneaux routiers basée sur la couleur. Les approches qui utilisent la couleur comme paramètre de détection des panneaux routiers supposent que les couleurs des panneaux sont bien distinguées. Ce sont des approches globales caractérisées surtout par leur simplicité d'implémentation. En revanche, ces méthodes manifestent quelques problèmes et difficultés liés à l'utilisation de la couleur comme critère de détection, et qui s'ajoutent aux problèmes propres aux images de la route.

Les particularités des scènes routières sont dues au fait que les objets panneaux doivent être différenciés par rapport aux autres objets de la route et que les conditions d'acquisitions peuvent affecter la qualité des images. Deux principales questions aux quelles nous avons répondu : comment distinguer les objets panneaux qui constituent des petites régions des autres objets de tailles importantes, et comment discriminer deux objets qui ont des couleurs proches. La réponse à ces questions était le résultat des réflexions dans le cadre d'une approche couleur, tout en insistant sur la simplicité des solutions.

Nous avons opté pour le choix des histogrammes couleurs pour analyser la couleur des objets de l'image, et nous avons pensé à ce niveau à l'importance des espaces de couleurs en comparant les résultats obtenus dans deux espaces différents. Nous avons fixé comme perspective immédiate l'étude des résultats de la méthode dans un espace couleur présentant des composants couleurs dé-corrélés, comme les systèmes basés sur les axes indépendants.

L'étude des histogrammes couleurs a montré la multiplicité des approches exploitant de telles structures. Toutefois, une approche bi-marginale utilisant des histogrammes 2D donne de bons résultats en combinant la simplicité de l'approche marginale et partiellement l'aspect corrélation de l'approche vectorielle.

Le problème de discrimination des panneaux a été traité en utilisant des histogrammes 2D présentant une corrélation partielle des composantes couleurs. Une analyse multi-échelle de ces histogrammes nous a permis de mieux localiser les régions couleurs présents dans l'image, et spécialement les régions correspondantes aux panneaux routiers.

Les résultats présentés dans cette partie de ce rapport sont très encourageants, et nous avons confirmé la qualité de l'approche par des tests sur des images artificielles et des images réelles. La qualité des segmentations observées sur nos images synthèses et sur les images réelles a été approuvée par des mesures quantitatives d'un certain nombre de critères d'évaluation non supervisée.

Comme perspectives, nous envisageons de combiner cette approche avec une approche de reconnaissance de panneaux routiers, dans le but de construire un système fiable et robuste pour la détection et la reconnaissance des panneaux routiers dans une image. Une alternative intéressante pour remédier à cette problématique est la combinaison de l'approche couleur proposée avec une approche d'identification des formes.



Nous avons dédié la deuxième partie de ce mémoire pour présenter nos contributions dans la vision 3D. L'idée d'exploiter les primitives des images pour la mise en correspondance a fait l'objet de plusieurs recherches, certaines proposent d'utiliser les points d'intérêt des images, d'autres utilisent les contours ou même les régions.

L'originalité de notre approche de mise en correspondance se réside non seulement dans l'exploitation des points contours significatifs dans les images, mais utiliser les informations multi-échelles dans le processus d'appariement. Aux appariements traditionnels basés sur des mesures de corrélation entre voisinage des pixels, nous avons introduit des mesures multi-échelles qui utilisent les modules maxima d'ondelettes et leurs orientations. Cela se fait essentiellement par l'établissement des cartes de maxima multi-échelle et les cartes des orientations et l'exploration des chaînes maxima minimisant ainsi l'effort de recherche aux points les plus significatifs dans les deux images stéréoscopiques.

Ce processus présente également une simplicité de la représentation. En effet, une image sera analysée par ondelettes ce qui donne une décomposition de l'image dans des bases orthogonales obtenues par dilatation et translation de l'ondelette. L'analyse multi-échelle utilisée permet une représentation hiérarchisée à travers les échelles, et c'est une décomposition constituée d'une approximation dans une échelle grossière et des détails aux échelles intermédiaires permettant de la reconstruire. L'intérêt de telle représentation est que la taille des coefficients d'ondelette informe sur le contenu de l'image, et que nous les avons exploités pour réaliser la mise en correspondance stéréoscopique. Nous avons testé notre approche de mise en correspondance sur des images de la base middlebury, et nous avons établie des cartes de disparités dont les profondeurs sont logiques, réelles et claires.

Les résultats des différents travaux ont fait l'objet d'un certain nombre de communications et de publications acceptés dans les revues scientifiques. Nous avons montré, une fois de plus, l'efficacité des ondelettes dans des processus de traitement des images numériques.

Les codes de tests que nous avons développés soit en Matlab soit en Java forment le noyau d'un Framework de test intégrant une large panoplie de fonctionnalités de traitement des images numériques que nous sommes entrain de finaliser.

## BIBLIOGRAPHIE

- [1] F.A. Aly and A.E. Alaa. Detection, categorization and recognition of road signs for autonomous navigation. In Proceeding of Advanced Concepts for Intelligent Vision System, Brussels, Belgium, Aug 2004.
- [2] A. de la Escalera, J.M. Armingol, J.M. Pastor, and F.J. Rodriguez. Visual sign information extraction and identification by deformable models for intelligent vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 5(2) :57–68, 2004.
- [3] U. Zakir, A.N.J. Leonce, E.A. Edirisinghe, Road Sign Segmentation based on colour spaces: a comparative study, Proceeding of the 11th IASTED International Conference, Computer Graphics and Imaging (CGIM 2010).
- [4] M. Bénallal, J. Meunier, Real-time color segmentation of road signs, CCECE 2003- CCEGEI 2003, Montréal, Mai 2003.
- [5] C.F. Paulo, P.L. Correia, Automatic detection and classification of traffic signs, *Image Analysis for Multimedia Interactive Services*, 2007. WIAMIS '07. Eighth International Workshop on.
- [6] H. Fleyeh, Traffic Signs color detection and segmentation in poor light conditions, 2005. Conference on Machine Vision Applications, Tsukuba Science City, Japan.
- [7] Arturo de la Escalera, Luis E. Moreno, Miguel Angel Salichs, and José Maria Armingol. Road traffic sign detection and classification. *IEEE Transactions on Industrial Electronics*, 44(6) :848–859, Dec 1997
- [8] Paul L. Rosin. Measuring shape : ellipticity, rectangularity, and triangularity. *Machine Vision and Applications*, 14(3) :172–184, July 2003.
- [9] Timofte R., Zimmermann K., Gool L. V., « Multi-View Traffic Sign Detection, Recognition, and 3D Localisation », *IEEE Workshop on Applications of Computer Vision WACV'09*, Institute of Electrical and Electronics Engineers, Snowbird, Utah, USA, p. 69-76, 7-9 décembre, 2009.
- [10] P. B. Zadeh, C. V. Serdean “ A Hierarchical multiwavelet based stereo correspondence matching techniques “, *EURASIP'2011*, pp 393-397
- [11] G. Qu, D. Zhang and P. Yan “ Medical image fusion by wavelet transform modulus maxima”, *Optics Express Vol 9, No 4* pp 184-190, 2001.
- [12] D. Scharstein and R. Szeliski, [www.Middlebury.edu/stereo/](http://www.Middlebury.edu/stereo/),
- [13] J.-P. Cocquerez et S. Philipp, *Analyse d'Images: filtrage et segmentation*. Masson1995.
- [14] A. TRÉMEAU, C. FERNANDEZ-MALOIGNE & P. BOTTON. « Image numérique couleur : de l'acquisition au traitement ». Paris : Dunod, 2004.
- [15] L. MacQueen, *Some methods for classification and analysis of multivariate observations*. Dans *Proceeding 5th Berkeley Symp.*, pp 281–297 1967.
- [16] Sitao Feng, Evaluation of color Segmentation Algorithms in Red Color of traffic signs detection,2010. Thesis Computer Engineering, Nr:E4003D
- [17] C. Maaoui, H. Laurent, B. Emile, Reconnaissance et détection robuste d'objets couleur, 2005. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images.

- [18] CABANI I; TOULMINET G; BENSRAHAIR A " Color-based Detection of Vehicle Lights" IEEE Intelligent Vehicles Symposium, Las Vegas, USA Année:2005.
- [19] A.Ezzine & c.Achraf Une nouvelle méthode de segmentation basée sur les ondelettes: application pour la détection des feux de stop d'un véhicule. Proceeding Optique 2006, INPT Rabat.
- [20] Faïcel Chamroukhi(2012). « Méthodes de classification pour la segmentation d'image » .(USTV/LSIS). <http://chamroukhi.univ-tln.fr/>.
- [21] B. Soheilian ,A. Arlicot,N. Paparoditis."Extraction de panneaux de signalisation routière dans des images couleurs ». IGN, MATIS.France
- [22] Ludovic Macaire, « exploitation de la couleur pour la segmentation et l'analyse des images». Thèse. Laboratoire LAGIS. Université des sciences et technologie de Lille.
- [23] F. KURUGOLLU, B. SANKUR, and A. HARMANCI, Color image segmentation using histogram multithresholding and fusion. Image and Vision Computing, 19(13):915-928, 2001.
- [24] O. Lezoray, C. Charrier "Color image segmentation by unsupervised 2D histogram clustering and Dempster-Shafer region merging " LUSAC EA 2607.
- [25] F.G. Meyer, Coifman R.R. Brushlets · a tool for directional image analysis and image compression, Applied and Computational Harmonic Analysis (Academic Press), Vol 4, pp 147-187, 1997.
- [26] Olivier LE CADET: « Méthodes d'ondelettes pour la segmentation d'images. Applications à l'imagerie médicale et au tatouage d'images », thèse de Doctorat, Institut National Polytechnique de Grenoble, Septembre 2004.
- [27] A. TRÉMEAU, C. FERNANDEZ-MALOIGNE, and P.BONTON, Image numérique couleur. Dunod, 2004.
- [28] Cyril Meurie, « Segmentation d'images couleur par classification pixellaire et hiérarchie de partitions ». Thèse de doctorat, l'UNIVERSITE de CAEN, october 2005
- [29] T. Carron, « Segmentations d'images couleur dans la base Teinte-Luminance-Saturation: approche numérique et symbolique ». Thèse de doctorat. Université de savoie, 1995.
- [30] Commission Internationale de l'Eclairage, Proceedings of the International congress on Illumination, Cambridge, Cambridge University Press, 1931.
- [31] Faugeras O.D., « Digital Image Processing within the framework of a Human Visual Model », IEEE transactions on Acoustic, speech and signal processing, Vol. 27, n°7, pp. 380-393, 1979.
- [32] Pratt W.K., « Digital Image Processing », John Wiley & Sons, New-York, 1978.
- [33] M. SAVOJI & R. BURGE. « Note on different methods based on the Karhunen-Loeve expansion and used in image analysis ». Computer Vision, Graphics and Image Processing, vol. 29, pages 259–269, 1985.
- [34] Y. OHTA, T. KANADE & T. SAKAI. « Color information for region segmentation ». Dans Computer Graphics and Image Processing, vol. 13, pages 222–241, 1980.
- [35] J.G. Postaire, R.D. Zhang, et C. Lecocq-Boite, *Cluster analysis by binary morphology*. Dans IEEE Trans. On PAMI, vol.15, no.2, pp. 170-180 1993.
- [36] Clément, A. & Vigouroux, B.(2003). Unsupervised segmentation of scenes containing vegetation (Forsythia) and soil by hierarchical analysis of bi-dimensional histograms. *Patt. Recogn. Lett.*, Vol. 24,(2003) 1951–1957
- [37] I. Daubechies, Ten lectures on Wavelets, SIAM, Philadelphia, PA, 1992.
- [38] Y. Meyer, Wavelets, Ed. J.M. Combes et al., Springer Verlag, Berlin, p. 21, 1989.
- [39] J.P. Kahane, P.G. Lemarie, Fourier Series and Wavelet, Gordon and Breach, 1996.
- [40] C. Gasquet, P. Witomski, Analyse de Fourier et applications, Masson 1990
- [41] S. Mallat, a wavelet tour of signal processing, Academic Press, second edition, 1998.

- [42] A. Cohen Ondelettes, analyses multirésolutions et filtres miroir en quadrature, Ann. Inst. H. Poincaré, Anal. Non Linéaire, 7 :439-459, 1990.
- [43] P. Flandrin Temps-fréquence, Hermès science publications, 1998.
- [44] Sylvie Philipp-Foliguet, Laurent Guigues: "Evaluation de la segmentation d'images : état de l'art, nouveaux indices et comparaison" ; ETIS, UMR CNRS 8051
- [45] C. Rosenberger, *Mise en Oeuvre d'un Système Adaptatif de Segmentation d'Images*. Thèse de Doctorat, Université de Rennes 1, décembre 1999.
- [46] Y.RUICHEK. Perception de l'environnement par stéréovision Application à la sécurité dans les systèmes de transports terrestres. USTL. 2005.
- [47] S. Mallat and S. Zhong, " Characterization of signals from multiscale edges ", IEEE Trans. on PAMI, vol. 14, pp. 710-732, July 1992.
- [48] S. Mallat and W. Hwang, " Singularity detection and processing with wavelets ", IEEE Trans. Information Theory, vol. 38, no. 2, pp.617-643, 1992.
- [49] S. Meignen, S. Achard, Time Localization of Transients with Wavelet Maxima Lines, IEEE. Trans. sign. Proc., accepted, 2004.
- [50] J. Canny A computational approach to edge detection, IEEE Trans. Patt. Anal. and Mach. Intell., vol PAMI-8 no. 6, pp. 679-698, Novembre 1986.
- [51] S. Mallat, A theory for multiresolution signal decomposition : the wavelet representation, IEEE Trans. Patt. Anal. and Mach. Intell., 11(7) :674-693, Juillet 1989.
- [52] S. Jaffard, Exposants de Holder en des points données et coefficients d'ondelettes, Note au compte-rendu de l'Académie des Sciences, France, 308 ser. I, pp 79-81, 1989.
- [53] Scharstein, D. and Szeliski, R. (1998). Stereo matching with nonlinear diffusion. International Journal of Computer Vision, 28(2):155–174.
- [54] Ohta and Takeo (1985). Stereo by intra- and inter-scanline search using dynamic programming. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-7:139–154.
- [55] A. Benschraï, P. Miché, and R. Debrie (1996). Fast and automatic stereo vision matching algorithm based on dynamic programming method. Pattern Recognition Letters, 17:457–466.
- [56] Birchfield and Tomasi (1999). Depth discontinuities by pixel-to-pixel stereo. International Journal of Computer Vision, 35(3):269–293.
- [57] Kolesnik, M. I. (1993). Fast algorithm for the stereo pair matching with parallel computation. In Chetverikov, D. and Kropatsch, W. G., editors, 5th International Conference on Computer Analysis of Images and Patterns, pages 533–537, Budapest, Hungary. Springer-Verlag.
- [58] Medioni, G. and Nevatia, R. (1985). Segment-based stereo matching. Computer Vision, Graphics, and Image Processing, 31:2–18.
- [59] Christophe Damerval ; " Ondelettes pour la détection de caractéristiques en traitement d'images. Application à la détection de région d'intérêt.. "Mathematics. Université Joseph-Fourier - Grenoble I, 2008. France.
- [60] P. Flandrin Temps-fréquence, Hermès science publications, 1998.

- [61] K. Berkner, R.O. Wells A new hierarchical scheme for approximating the continuous wavelet transform with applications to edge detection IEEE Signal Processing Letters, vol 6 no. 8, pp. 148-153, Août 1999
- [62] E.J. Candes, D.L. Donoho, Curvelets : a surprisingly effective non-adaptative representation of objects with edges, Curves and Surfaces, L.L. Schumaer et. al. (eds), Vanderbilt University Press, Nashville, TN, 2000.
- [63] D.L. Donoho, I. M. Johnstone, Minimax estimation via wavelet shrinkage, Ann. Statist. 26, 1998.
- [64] D.L. Donoho, Wedgelets : Nearly Minimax Estimation of Edges, Annals of statistics, 27(3) pp859-897, 1999.
- [65] M.N. Do, M. Vetterli, Contourlets : a Directional Multiresolution Image Re- presentation, Proc. International Conference on Image Processing, Vol. 1, p.357-360, septembre 2002
- [66] A. Cohen and A. Ezzine "Quadratures singulières et fonctions d'echelle ", CRAS, t.323, Serie I, 829-834, 1996.
- [67] I. Daubechies, Ten lectures on Wavelets, SIAM, Philadelphia, PA, 1992.
- [68] O. Lezoray. Segmentation d'images couleur par morphologie mathématique et classification de données par réseaux de neurones : application à la classification de cellules en cytologie des séreuses. Thèse de doctorat, Université de Caen, janvier 2000.
- [69] S. Zhong, "Edge Representation from Wavelet Transform Maxima". PhD thesis, New York University, September 1990.
- [70] S. D. Zeno. "A note on the gradient of multi-image". Computer Vision Graphics and Image Processing, 33 :116-125.
- [71] L. Distefano, Stefano Mattocia, " Fast template matching using bounded partial correlation ", Machine vision and application, NO 13, pp. 213-221, 2003.
- [73] Y. Meyer, Ondelettes et opérateurs I : ondelettes, Paris, Herman, 1990
- [74] A. Bhatti,S. Nahavandi "Stereo Image matching Using Wavelet Scale-Space Representation", Proceeding of CGIV'06, pp 267-272, 2006
- [75] S.Mallat, W.L.Hwang Singularity detection and processing with Wavelet,IEEETrans.Info. Theory,38(2): 617-643, Mars 1992.
- [76] S. Mallat, " Wavelets for a vision ", Processing of IEEE, vol. 84, NO 4, pp. 604-614, April 1996
- [77] A. Bhatti,S. Nahavandi "Stereo Image matching Using Wavelet Scale-Space Representation", Proceeding of CGIV'06, pp 267-272, 2006.
- [78] M. Fontaine. Segmentation non supervisée d'images couleur par analyse de la connexité des pixels. Thèse de doctorat, Université Lille 1, decembre 2001.
- [79] O. Le Cadet, V. Perrier : Edge Detection and characterization, en cours de rédaction, 2003.
- [80] R. Ohlander, K. Price et D. R. Reddy. "Picture segmentation using a recursive region splitting method". Computer Graphics and Image Processing, 8 :313-333, 1978.
- [81] S. H. Park, I. D. Yun et S. U. Lee. "Color image segmentation based on 3D clustering : morphological approach". Pattern Recognition, 31(8) :1061-1076, 1998.

- [82] T. Uchiyama et M. A. Arbib. "Color image segmentation using competitive learning". *IEEE Trans. On Pattern Analysis and Machine Intelligence*, 16(12) :1197–1206, 1994.
- [83] T.YOUNG. « On the theory of light and colors ». *Philosophical Transactions of the Royal Society*, page 91, 1802.
- [84] S. Zucker, *Region growing: Childhood and adolescence*. *Computer Graphics and Image Processing* 5, pp. 382-399 1976.
- [85] R. Haralick et L. Shapiro, *Image segmentation techniques*. *Computer Vision Graphics Image Processing*, vol. 29, pp. 100-132 1985.
- [86] R. Gonzalez et R. Woods, *Digital image processing*. Dans Addison-Wesley Publishing Company, Reading, MA 1993.
- [87] S. Horowitz et T. Pavlidis, *Picture segmentation by a directed split-and-merge procedure*. Dans 2nd International Joint Conference on Pattern Recognition, pp. 424–433 1974.
- [88] Jorge S. : Notes for Image processing and Vision.
- [89] John Guild, « *The Colorimetric Properties of the Spectrum* », dans *The Royal Society*, 1<sup>er</sup> janvier 1993
- [90] C. GARBAY. « Modélisation de la couleur dans le cadre de l'analyse d'images et de son application à la cytologie automatique ». Thèse de doctorat, Institut National Polytechnique de Grenoble, Décembre 1979.
- [91] L. Shafarenko, M. Petrou et J. Kittler, *Histogram-based segmentation in a perceptually uniform color space*. Dans *IEEE Trans. On Image Processing*, vol.7, pp.1354-1358 1998.
- [92] J. Liu, Y.H. Yang, Multiresolution color image segmentation, *IEEE Trans. PAMI.*, July, No. 16, 1999. pp 689-700.
- [93] Ramana L. Rao and Lakshman Prasad, "Segmentation by Multiresolution Histogram Decomposition", [citeseer.ist.psu.edu/rao95segmentation.html](http://citeseer.ist.psu.edu/rao95segmentation.html).
- [94] G. Finlayson, S. Hordley et P. Hubel. "Illuminant estimation for object recognition". *Color Research and Application*, 27(4) :260–270, 2002.
- [95] Gabriel Peyré, « Le traitement numérique des images » — *Images des Mathématiques*, CNRS, 2011
- [96] Guillaume Gales "Pixel matching for binocular stereovision by propagation of feature points matches and region-based randomized voting scheme". Thèse de doctorat .IRIT de Toulouse. Juillet 2011.
- [97] A. Ezzine Abdelhak, Youssef Yazidi Alaoui, Achraf Cohen, « Adaptive multiscale Stereo images Matching based on wavelet transform modulus maxima », *International Journal of Image IJIP*, Vol 5, Issue 5 date October 2012.
- [98] YAZIDI ALAOUI Youssef, EZZINE Abdelhak, "Images segmentation method based on a multi scale analysis of 2D histograms" ,*International Journal of Emerging Trends in Engineering and Development, IJETED*, Issue3 volume 6 date November 2013.

## Publications

**Article 1:** International Journal of Image Processing (IJIP), Volume (6) : Issue (5) : 2012

### **Adaptive multiscale stereo images matching based on wavelet Transform Modulus Maxima**

**Abdelhak EZZINE**

*aezzine@uae.ac.ma*

*ENSAT/ SIC/LABTIC*

*Abdelmalek ESSAADI University*

*Tangier, 99000, Morocco*

**Youssef Yazidi Alaoui**

*yazidi17@gmail.com*

*ENSAT/ SIC/LABTIC*

*Abdelmalek ESSAADI University*

*Tangier, 99000, Morocco*

**Achraf Cohen**

[cohenachraf@gmail.com](mailto:cohenachraf@gmail.com)

*ENSAT/ SIC/LABTIC*

*Abdelmalek ESSAADI University*

*Tangier, 99000, Morocco*

---

#### **Abstract**

In this paper we propose a multiscale stereo correspondence matching method based on wavelets transform modulus maxima. Exploitation of maxima modulus chains has given us the opportunity to refine the search for corresponding. Based on the wavelet transform we construct maps of modules and phases for different scales, then extracted the maxima and then we build chains of maxima. Points constituents maxima modulus chains will be considered as points of interest in matching processes. The availability of all its multiscale information, allows searching under geometric constraints, for each point of interest in the left image corresponding one of the best points of constituent chains of the right image. The experiment results demonstrate that the number of corresponding has a very clear decrease when the scale increases

**Keywords:**Maxima modulus, stereo matching, wavelet transform , maxima modulus chains.

---

## 1. INTRODUCTION

Matching the extracted visual indexes of stereoscopic images is a very significant step within the process of 3D version. It is made up of bringing together the primitive extracts of right and left images so as to realize a 3D reconstruction of the scene. This is based on estimating the offset between the positions of matched primitives and building a disparity map showing the scene by highlighting the depth of scene objects. The disparity map generated from the correspondence matching process, along with a stereo camera are used to estimate the depth map and produce the 3D view of the scene.

This problematic is largely treated by a variety of researchers yet still there is no general method [6,8,10,26], many approaches exist in the field and exploit generally the primitives such as reference points, contours or areas but generally via the same characteristics which are either photometric, geometric or either morphologic.

This technique remains difficult due to the multiplicity of essential parameters, mainly the pertinence of considerable attributes, in the process of matching and the difficulty to find the good corresponding primitive is subject to a number of potential problems in vision process like occlusion, ambiguity, illuminative variations and radial distortion and all algorithms proposed are adapted to control at least some of these problems. We find many class of stereo vision algorithms, the first class is named global algorithms (GA) who deal with the correspondence estimation process as global cost function optimization problem and the second one is the local algorithms (LA) and are categorized into either area based or feature based algorithms [24]. The area-based algorithm used the correlation of the two image functions over locally defined regions and the feature-based algorithms establish correspondence between the selected primitives of the image.

Our method is one of the new class of algorithms [5,6], it is based on the concept of wavelets because the matching is the result of a multi-scale operation of the modulus and the phase of wavelet coefficients. This class is considered in the middle way between the local and global algorithm, these combine the best features of the LA and GA techniques and uses the multiresolution concept to involves the matching of two images at different scales.

This method is based on the wavelets modulus maxima chain (CMI) and involves matching of the different points forming the chain (CMI) to find the best corresponding in the right image for each extremity of one chain in left image. The search of the best one is done by estimating a similarity test between the points of each left chain and points of the right chains satisfying the geometric constraint like epipolar and orientation. The disparity map generated has a multiscale character as it is built based on the simultaneous treatment of space and scale. The multiscale character will permit the refinement of the disparity map and possibility to refine the selection of the potential candidates using a multiscale criterion based on normalized correlation and geometric refinement. As proved in Mallat [7,22,12], the extremity of the chain coincides with a singularity of the image characterized by its coefficient of Lipchitz, thus with a good choice for constructing maxima chains, the proposed matching algorithm not only provides the precise results but it also converges rapidly.



Firstly, we are going to recapitulate the main principles of wavelet transform modulus maxima (MMW). Then, we move to present the method and its particularities. Finally, we synthesis the major results of the method we intend to compare with other attributes.

## 2. Wavelets transform modulus maxima

Based on the idea that in the image processing, visual data is put in a hierarchic way via the scale, wavelets transform decompose image into elementary blocs that are well localized both on scale and space. As a result of this decomposition, the image can be represented by the approximation to the coarse resolution and detail to all intermediate scales.

Among the main interests of representation in the form of wavelets we state the coefficient wavelet length exam of an image can inform us about it analysis. Indeed, in [2, 9], it is proved that the regularity accentuates the decay of wavelet coefficients; we are talking about the local regularity related to the analyzing wavelet support [17, 14].

Thereby, if the image is continuously regular everywhere except at a few isolated points the estimated wavelet coefficients will be affected only if the support of the analyzing wavelet contains these points or edges. In this sense, multiscale representations are better suited to focus the information in an image whose regularity is not homogeneous because the wavelet coefficients above a certain threshold focus only near the singularities (edges, single points, etc.).

One of the multiscale decompositions that exploits well the wavelet properties which permit making a study, in a different levels, about the discontinuity points present in an image is wavelet transform modulus maximas proposed by Mallat and Zhong [1,7].

The principle of wavelet modulus maxima can give the image of characteristics to identify it, they transmit obvious features of the image depending on well defined models and well determined directions. For more details about wavelets modulus maxima we refer to [2, 13,14, 21] and present below the essential terms.

Suppose two wavelets  $\psi^x$  and  $\psi^y$  such that  $\psi^x = -\frac{\partial G}{\partial x}$  and  $\psi^y = -\frac{\partial G}{\partial y}$  and  $\int_{-\infty}^{+\infty} G(t)dt \neq 0$ . For this work we choose  $G$  as a Gaussian wavelet to ensure both a valorization of wavelet maxima and continuity of the wavelet maxima chains [4, 7]. Thus wavelet transform of an image  $I$  can be written as a multiscale differential operator

$$\begin{pmatrix} W^x I(u, v, j) \\ W^y I(u, v, j) \end{pmatrix} = \begin{pmatrix} I * \psi_j^x(-u, -v) \\ I * \psi_j^y(-u, -v) \end{pmatrix} = 2^j \vec{\nabla} (I * G_j(-u, -v)) \quad (1)$$

The polarized representation of this gradient vector offers the wavelet transform modulus

$$MI(u, v, j) = \sqrt{|W^x I(u, v, j)|^2 + |W^y I(u, v, j)|^2} \quad (2)$$

and its direction is defined by the angle  $\theta I(u, v, j) = \tan^{-1} \left( \frac{W^y I(u, v, j)}{W^x I(u, v, j)} \right)$ . (3)

Then the local maximum of the wavelet transform modulus  $MMI(u, v, j)$  can be found by solving  $\partial MI(u, v, j) = 0$ . Otherwise on some scale  $j$ , the points  $(u_0, v_0)$  such that  $MMI(u, v, j)$  is a local maximum, according the  $L_\infty$  norm, on a neighborhood of  $(u_0, v_0)$  in the direction  $\theta MI(u, v, j)$ . We note that for image color we consider the maxima according DiZenzo approach[11].

### 3. Multilevel matching method

In works of Mallat and Zhong [7] is prove that a wavelet coefficient can be influenced by a singularity resides within support of the analyzing wavelet; we talk about the cone of influence to express the set of points where the wavelet coefficients may be influenced by a singularity at a point  $(u_0, v_0)$ . This singularity will generate local maxima of wavelet coefficient upon the cone and till a particular scale. In order to characterize this singularity, we will observe the behavior of local maxima in terms of scale; this will be much easier to follow if the maxima, between two successive scales, are connected together to form a curve in the plane scale space so-called line or chain maxima.

In practice, we start with an image of size  $(N, N)$ ,  $N = 2^J$ . We obtain the wavelet transform according dyadic approach [3,11,18] in which scale parameter is expressed by the power of 2,  $2^j$ ,  $j = J \dots L$ ,  $L$  is the coarsest scale. This method generates on each scale, a wavelet transform with the same number of pixels as the image  $I$ ; in spite of that the support of the wavelet is increasingly large when scales grow. After that we calculate the wavelets transform modulus and directions by equations (2) and (3). The results are two matrix, the maxima modulus matrix  $MMI(\dots, j)$  and the angle matrix  $\theta MI(\dots, j)$  for each scale  $j = J - 1 \dots L$ . The image  $I$  is then represented by  $(MMI(\dots, j))_{j=J-1 \dots L}$ , positions and values of the wavelet coefficients at each scale  $j$  when the module is a local maximum in the direction  $\theta$ .

We construct the line or chain of maxima by a chaining  $(x_j, y_j)$ , the location of each element of  $MMI(\dots, j)$ , with  $(x_{j-1}, y_{j-1})$ , the location of its successor in  $MMI(\dots, j-1)$  [21,23]. The search for these successors is limited to the neighborhood defined by the cone of influence and every chain can be indexed by  $(x_j, y_j)_{j=J \dots m}$  and noted  $CMJ(x_j^J, y_j^J, m)$  [14,11]. The integer  $m$  indicate the length of the chain, and is defined such that the last point of the chain is ones of maxima of the wavelet transform at scale  $J-m$ . We note that each chain is composed of maxima and is characterized by its extremity  $MMI(x_j, y_j, J)$  at the level  $J$  that coincides with a singular point of the image.

#### 3.1 The matching approach

Given two rectified color images, from a vision system calibrated. we associate to each one its maxima modulus chains representation  $(x_j^i, y_j^i)_{j=J \dots m}$ ,  $i = g$  for the left image and  $i = d$  for the right one. With different representations defined above, we are faced with three possible approaches to mapping. The first is to match directly the maxima modulus transforms  $MMI_i(\dots, j)$  of the two images at each scale  $j = J - 1 \dots L$ . The second performs the matching of the extremity of each chain  $CMG(x_j^g, y_j^g, m)$  with an extremity of an chain  $CMG(x_j^d, y_j^d, m')$  by a non-linear algorithm. The third approach proposes to consider the maxima chains as primitives.

In this paper we develop the second approach and we will give an idea about the treatment of the third, we note that the first one looks approximately like the work of Bhatti and we invite the reader to discover in this work [5].

To chain  $CMG(x_j^g, y_j^g, m)$  of the left image, we will search for all Chain of the right image such that their extremities respect the epipolar constraint with the pixel  $(x_j^g, y_j^g)$ ; all maxima of the left that have no match in the right image verifying the epipolar constraint are dropped. We then perform a test of

similarity between  $MMI(x_j^g, y_j^g, I)$  and each one of the maxima modulus on extremities of the rights chains selected before. We note  $\{(x_{j,k}^d, y_{j,k}^d)_{k=1..c}\}$  the set of extremity of the right chains validating this test. If  $c = 1$ , we conclude that  $(x_j^g, y_j^g)$  is matched with  $(x_{j,1}^d, y_{j,1}^d)$ . If  $c > 1$ , we move to the next scale and we redo the estimating test between the successors, i.e.  $(x_{j-1}^g, y_{j-1}^g)$  and  $\{(x_{j-1,k}^d, y_{j-1,k}^d)_{k=1..c}\}$ . This step is repeated, recursively on the scale  $j$ , until the end of the chain and it will result necessarily matching of  $(x_j^g, y_j^g)$  and ones of  $\{(x_{j,k}^d, y_{j,k}^d)_{k=1..c}\}$  points.

The similarity is evaluated, under the orientation constraint, by using ones off different similarity measure available in the literature. Major similarity measures used in matching methods are measures of distortion (SAD) and (SSD) and the normalized cross correlation (NCC)[15]. It is known that SAD and SSD are computationally fast then NCC who is more accurate and it allowed us to refine the similarity test. At fine scales, there are many edge points created by the image noise in the wavelet transform maxima modulus representation. By reason of the presence of these wrong maxima and because it is well known that the SSD and SAD are justified when the additive noise distribution is Gaussian or exponential, we are led to combine the NCC and SAD to the similarity test for the fine scales. We note that at larges scales the smoothing process removes most of the wrong maxima and the NCC is sufficient.

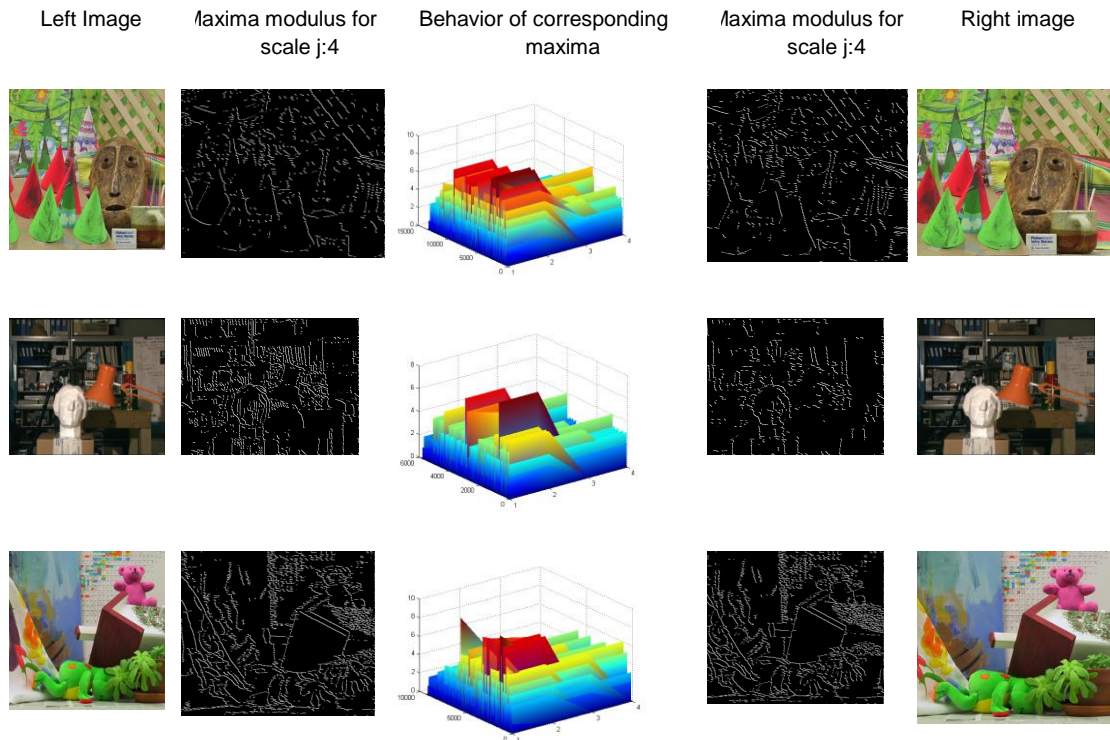
To find the correspondence between pixels of  $MMG(x_j^g, y_j^g, j)$  and  $MMD(x_j^d, y_j^d, i)$ , one can maximize a correlation or minimize the distortion measure. For each point  $(x_j^g, y_j^g)$ , we search for a point  $(x_j^d, y_j^d)$  in  $CMD(x_j^d, y_j^d, m')$ , whose a neighborhood of a size  $s$  has a maximum correlation or maximum distortion with the neighborhood of the point  $(x_j^g, y_j^g)$  in  $CMD(x_j^g, y_j^g, m)$ . Due to the epipolar constraint we have  $y_j^d = y_j^g$  and the search is reduced to points  $(x_j^d, y_j^g)$  on the epipolar line.

#### 4. Experiments results

The algorithm presented in this paper is tested on the images from web site [10,10b] [www.middlebury.edu/stereo/](http://www.middlebury.edu/stereo/). These images are created for the stereo matching algorithms we present the results obtained for four of them with the same illumination. In order to give visual results, the performance of the proposed nonlinear algorithm is presented for different scales  $j$  with the dyadic approach wavelet transform so we can follow the behavior of a maximum on different scales.

The experimental results were generated using the Gaussian wavelets. The figure 1 present the images cones, teddy and tsukuba, and there maxima modulus on different scale  $j$ : 5,4,3,2,1. In figure 1 are presented the number of points  $(x_j^d, y_j^g)$  of the right image validating the similarity test with  $(x_j^g, y_j^g)$  point of the left image. We present these results only for 50 points  $(x_j^g, y_j^g)$  represents the great maxima.

We observe that the number of corresponding has a very clear decrease when the scale increases.



**Figure 1:** in third column we observe the decreasing behavior of number of corresponding maxima in right image at scale  $j:4$  for everyone of maxima in left image (x axis :scale; y axis: maxima and z axis: number of corresponding)

## 5. Conclusion

This paper introduced a nonlinear algorithm matching based on wavelet transform maxima modulus. The maxima modulus chains are used to find the best corresponding maxima in the right image of each maximum in left image. This process exploits intelligently the multiscale character and offers an accurate and fast algorithm. We note also that the first results of the construction of disparity maps are very encouraging. Following this work is, firstly, affine the disparity maps and compare it to techniques benchmarked and published in the Middlebury database; and secondly integer the Lipschitz regularity of pixels [2,16,17]. in the similarity process as an additive information, because it was proven that there is a strong relationship between maxima and this regularity. This will allow us to exploit the nature of the maxima as edge points in the image and propose a non-linear feature-based algorithm.

## REFERENCES

- [1] S. Zhong, "Edge Representation from Wavelet Transform Maxima". PhD thesis, New York University, September 1990.
- [2] Y. Meyer, "Ondelettes et opérateurs ", vol. I. Paris: Hermann, 1990.
- [3] M. Shensa, "Wedding the à trous and Mallat algorithms, " IEEE Trans. Signal Processing, vol. 40, no. 10, pp. 2464-2482, 1992.
- [4] J. Canny, "computational approach to edge detection ", IEEE Trans. on PAMI, vol. 8, no. 6, pp. 679-698, 1986.
- [5] A. Bhatti,S. Nahavandi "Stereo Image matching Using Wavelet Scale-Space Representation", Proceeding of CGIV'06, pp 267-272, 2006

- [6] P. B. Zadeh, C. V. Serdean " A Hierarchical multiwavelet based stereo correspondence matching techniques ", EURASIP'2011, pp 393-397
- [7] S. Mallat and S. Zhong, " Characterization of signals from multiscale edges ", IEEE Trans. on PAMI, vol. 14, pp. 710-732, July 1992.
- [8] G. Qu, D. Zhang and P. Yan " Medical image fusion by wavelet transform modulus maxima", Optics Express Vol 9, No 4 pp 184-190, 2001.
- [9] A. Cohen and A. Ezzine "Quadratures singulières et fonctions d'échelle ", CRAS, t.323, Serie I, 829-834, 1996.
- [10] D. Scharstein and R. Szeliski, [www. Middlebury.edu/stereo/](http://www.Middlebury.edu/stereo/),
- [10b] D. Scharstein and R. Szeliski, Stereo matching with nonlinear diffusion. Int. J. of Computer vision, 28(2), 1998.
- [11] S. Dizenzo, " A note on the gradient of a multi-image ", CVGIP, vol. 33, pp. 116-125, 1986.
- [12] S. Mallat and W. Hwang, " Singularity detection and processing with wavelets ", IEEE Trans. Information Theory, vol. 38, no. 2, pp.617-643, 1992.
- [13] P. Carré, " Maxima d'ondelettes couleur, color wavelet maxima ", traitement du signal 2004 vol 21 special L'image numérique couleur 623-632.
- [14] S. Mallat, " Wavelets for a vision ", Processing of IEEE, vol. 84, NO 4, pp. 604-614, April 1996.
- [15] L. Distefano, Stefano Mattocia, " Fast template matching using bounded partial correlation ", Machine vision and application, NO 13, pp. 213-221, 2003.
- [16] S. Jaffard, " Exposants de Holder en des points données et coefficients d'ondelettes", CRAS France,308 ser I ,pp79-81,1989
- [17] I. Daubechies, Ten lectures on Wavelets, SIAM, Philadelphia, PA, 1992.
- [18] S. Mallat, a wavelet tour of signal processing, Academic Press, second edition, 1998.
- [19] J. Canny A computational approach to edge detection, IEEE Trans. Patt. Anal. and Mach. Intell., vol PAMI-8 no. 6, pp. 679-698, Novembre 1986.
- [20] M. Holschneider, R. Kronland-Martinet, J. Morlet, P. Tchamitchian Wavelets, Time-Frequency Methods, and Phase Space, Springer-Verlag, Berlin, 1989.
- [21] O. Lecadet Méthodes d'ondelettes pour la segmentation d'images. Applications à l'imagerie médicale et au tatouage d'images, these de doctorat, 2003.
- [22] S. Meignen, S. Achard, Time Localization of Transients with Wavelet Maxima Lines, IEEE. Trans. sign. Proc., accepted, 2004.
- [23] Y. Ruichek, Perception de l'environnement par stereovision , Application à la sécurité dans les systèmes de transports terrestres, Phdthesis, USTL, France 2006.

**Article 2:** *International Journal of Emerging Trends in Engineering and Development, IJETED, Issue3*  
volume 6 date November 2013.

# Images segmentation method based on a multi scale analysis of 2D histograms

YAZIDI AIAOUI Youssef, EZZINE Abdelhak

*LABTIC, ENSA-Tanger*

*BP 1818, Tanger Principale Tanger 9900*

*yazidi17@gmail.com*

*ezzine.abdelhak@gmail.com*

*abstract*

**In this paper we propose an image segmentation method based on multi-scale analysis using wavelets, this method has been developed and adapted to detect road signs (road signs and traffic lights). It is an approach to highlight objects in a road scene especially objects corresponding to a road signs. It relies on the use of 2D histograms highlighting the different modes and based on operation of multi-scale representations of an image with an analysis with wavelets. That said the identification of colors with weak presence in the image as well as the relevant discrimination of different colors in the image, which discriminates objects of the same color visually.**

**Key words:** wavelet transforms multi resolution, segmentation, 2D histogram.

## I. INTRODUCTION

Several works propose different segmentation algorithms for color images. This works are based on the distribution, often probabilistic, of information in the image or based on the knowledge in front of the image [1, 2]. The results of different approaches are the identification of regions of the image spatially coherent as well as their content: methods of growth regions.

One of these categories is based on a classification that groups the pixels of the image, regarded as a statistical distribution, in classes of individuals. The histogram of the image is thus used as a tool to identify several characteristics of these classes. For color images, the choice of the representation of the histogram becomes very important and according to one method or another we may remove or highlight primitives of the image (rare objects present in the image) and then promote their detection.

From a statistical point of view, the pixels with very similar colors are grouped in the same region which indicates a region of very high values in the histogram and therefore the presence of a mode. Different methods of classification, so deal this area by translating it into a homogeneous region not necessarily related and it is from here that the difficulties begin. Indeed, highlighting or discernment of objects having grayscale close but not identical in the image becomes very difficult. This is proven by experience to the images in the intelligent driving domain, essentially images containing road signs.

The use of methods based on the use of the histogram in the case of detecting road signs is ineffective as a road sign (which is a small region) located next to a large homogeneous region (vehicle, tree, road, building or advertising panel) would rarely be restored [3,4]. Such difficulties, which we are confronted is heightened by the confusion of a share of road signs with other objects of the road and on the other hand the variation of chrominance and luminance settings panel which is due to exposure to natural factors such as the sun, weather or variable light.

To overcome this problem, we propose a method based on the use of 2D histograms in order to take into account the color. Also the choice of a suitable color space will allow easily solve problems

highlighting both information chrominance and luminance. 2D histogram then considered as a 2D image processing by wavelet analysis, using essentially the detection of wavelet coefficients to distinguish the modes as much as possible. We managed to produce a method for detecting road signs and the first results are very satisfactory because the multi scale property of wavelet allows as controlling the difference of pixels values with a controlled precision and thus distinguishing between two objects in the image even when they have two colors close and ensure their presence in the image. Also, our approach has the advantage of considerably reducing the computation time results of the analysis of the histogram instead of analyzing the entire image.

## II. PRELIMINARIES

### A. Multi resolution and orthogonal wavelet transform

A multi-resolution analysis of  $L_2(\mathbb{R})$  consists of a sequence of nested closed subspaces  $(V_j)_{j = -\infty \dots +\infty}$  which satisfy the following properties:

$$\forall j \in \mathbb{Z}, V_j \subset V_{j+1} \quad (1)$$

$$\forall j \in \mathbb{Z}, f(t) \in V_j \Leftrightarrow f(2t) \in V_{j+1} \quad (2)$$

$$\forall k \in \mathbb{Z}, f(t) \in V_0 \Leftrightarrow f(t - k) \in V_0 \quad (3)$$

$$\lim_{j \rightarrow -\infty} V_j = \prod_{j=-\infty}^{+\infty} V_j = \{0\} \quad \lim_{j \rightarrow +\infty} V_j = \prod_{j=-\infty}^{+\infty} V_j = L^2\{\mathbb{R}\}$$

$$(4)$$

There exists  $\theta \in L^2(\mathbb{R})$  such that  $\{\theta(t - n)\}_{n \in \mathbb{Z}}$  is a base Riez of  $V_0$ .

$j$  is called the resolution and represents the level of analysis of the function  $f$ ; the approximation in  $V_j$  of  $f$  is twice as fine as in  $V_{j-1}$ , but less good, twice that in  $V_{j+1}$ .

#### 1) Detail space of wavelet :

The details are lost in passing a resolution  $j$  to a coarser resolution and more "smooth"  $j-1$  are high-frequency image components, we call  $W_{j-1}$  the space containing these details:

Let  $W_{j-1}$  the orthogonal complement of  $V_{j-1}$  in  $V_j$  :

$$V_j = V_{j-1} \oplus W_{j-1}, \text{ and } V_{j-1} \perp W_{j-1}$$

According to the definition of an AMR, we have  $V_0 \subset V_1$ . With  $\phi(t) \in V_0$ ,  $\phi(t)$  also belongs to  $V_1$ , there exists then a sequence  $h = (h_k)_{k \in \mathbb{Z}}$  such that:

$$\phi(t) = \sum_{k \in \mathbb{Z}} h_k \sqrt{2} \phi(2t - k)$$

$(\sqrt{2} \phi(2t - k))$  is an orthonormal base of  $V_1$ ) This relationship allows to build  $h_k$ .

On the other hand  $W_0 \subset V_1$ : if  $\psi(t)$  is a function of  $W_0$ , there is a sequence  $g = (g_k)_{k \in \mathbb{Z}}$  such as:

$$\psi(t) = \sum_{k \in \mathbb{Z}} g_k \sqrt{2} \phi(2t - k)$$

This relationship and the previous are called tow scales relationships.

The simplest example to illustrate this concept is the multi-resolution analysis of Haar, which is simply an approximation of  $f$  by piecewise constant functions on intervals of length  $2^j$ .

$$\psi(x) = \begin{cases} 1, & x \in [0, 1/2] \\ -1, & x \in [1/2, 1] \\ 0, & x \notin [0, 1] \end{cases}$$

$$\psi_j^i(x) = \psi(2^j x - i), i = 0, \dots, 2^j - 1$$

In practice, signals and manipulated images are necessarily finite resolution. We assume that the 1D signal  $f$  to be analyzed is defined on  $[0, 1]$  and includes  $N = 2^j$  points: it is represented by a vector  $c_j$  with  $c_j = \{c_{j,k}\}, k = 0..2^j - 1$  coefficients in the basis of scaling functions of  $V_j$  space.

In the final case, the wavelet transforms amounts to a filtering or convolution [5]:

$$a_j^k = \sum_{l \in K(j)} h_j^{k,l} a_{j+1}^l$$

$$d_j^m = \sum_{l \in M(j)} g_j^{m,l} a_{j+1}^l$$

The filter defined by  $h_i$  is a low-pass filter encoding signal approximations (low frequency).

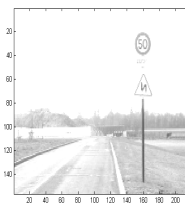
The filter  $g_i$  is defined by a high-pass filter encoding signal details (high frequencies).

### B. Multi resolution and tow-dimensional wavelet transform

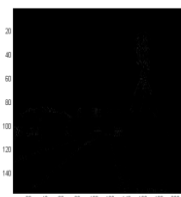
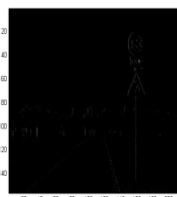
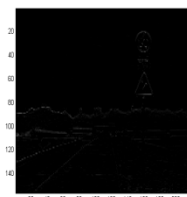
The general definition of a multi-resolution analysis of  $L_2(R_2)$  is similar to that given in the one-dimensional case. It suffices to consider functions with values in  $R_2$  and not in  $R$ .



original image



decomposition L=1





horizontal details L=1 vertical details L=1 diagonal details L=1

Fig 2 : Orthogonal wavelet transform of an image

Each approximation space  $V_j$  is decomposed into a coarser approximation space  $V_{j-1}$  and 3 spaces of details. Consider the case of a square image, so a function defined on a square domain. A digital image has a finite number of pixels. An image  $2^j \times 2^j$  will be considered as belonging to  $V_j$  and decomposed into the spaces  $(W_{1j}, W_{2j}, W_{3j})$   $j=L \dots J-1$  and in  $V_L$ .

### III. :CALCULATING HISTOGRAM 2D

Among the various methods of color image segmentation, there are methods on a statistical approach which consists to highlight areas of the image by analyzing the color histogram. We have several ways to use the histogram; the most natural of it is to consider a 3D histogram representing a cloud, in the coordinate system (XR, YG, ZB), such that each point of coordinates (xR, yG, zB) refers to the number of pixels of the image using the color (R, G, B).

Another way would be to consider three 1D histograms, for each color component. In this work, we use the concept of 2D histogram that expresses only two color components, which can be interpreted as a 2D projection of the 3D histogram in terms of two components. We have three 2D histograms, but we often operate only on two, for example in the (L, a\*, b\*) system, we can use H2D (L, a) and H2D (L, b) for the detection of the presence of signs.

Several ways are possible to estimate the 2D histogram, the first and most obvious is to take into account only the two present components, and set the third to zero. This approach is ideal if the three components are highly uncorrelated, but this is not the case of color systems generally available. This has motivated us to lead our research for the construction of a suitable color space that has this decorrelation property. The second technique is a projection of 3D histogram on the plan of 2 components such as: If we denote 3D Histogram H (r, g, b) we calculating H2D (r0, g0) as the sum of H (r0, g0, b) for b = 0, ..., L-1, if necessary by introducing as normalization parameter max (H (r0, g0, b)).

A 2D histogram of a color image I binds  $p(x_1, x_2)$ , the number of pixels in the image I which represent the color component (x1, x2). Since each color axis is quantized to 256 levels, the histogram can be represented by an image J with a spatial resolution of  $256 \times 256$  [9]. The  $p_n(x_1, x_2)$  value of the pixel of coordinates (x1, x2) in J is obtained by a linear dynamic contraction of the histogram between 1 and  $M = \min(P_{max}, 255)$ .

$$P_n(x_1, x_2) = \text{round} \left[ \frac{(M - 1)p(x_1, x_2) - MP_{min} + P_{max}}{P_{max} - P_{min}} \right]$$

With  $P_{max}$  and  $P_{min}$  are respectively the maximum and minimum values of P.

In the illustrations below, we notice the presence of areas with high concentration and other with low concentration. It comes the peaks separated by valleys. Modal areas correspond to objects in the image. We clearly distinguishes the area corresponding to the sign in the image and we note that this distinction is particularly clear from the RGB system to the L a\* b\* system. The cloud that we are interested in is in the northern part of the image histogram RG.

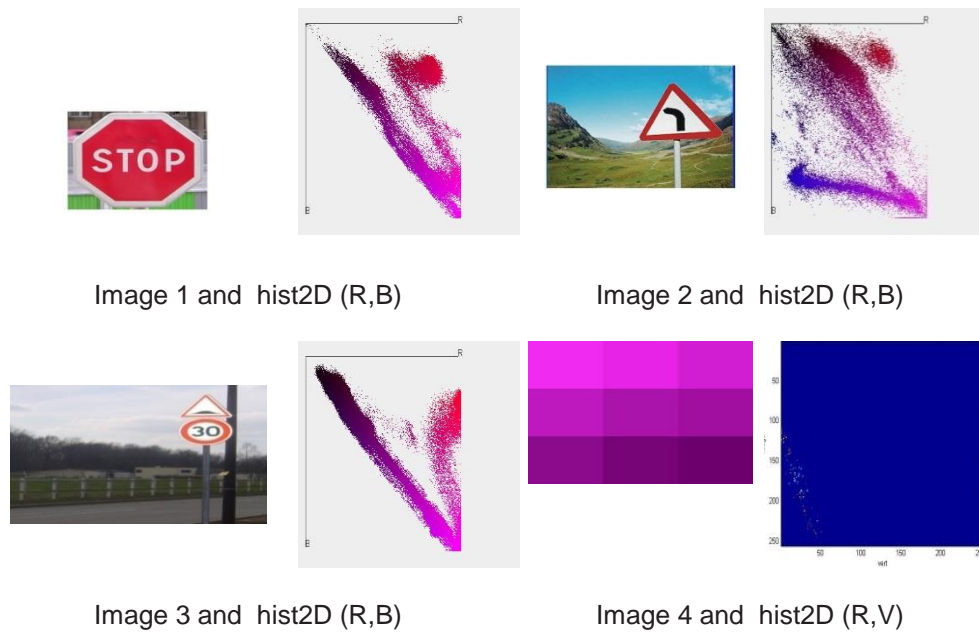


Fig 3 : images (1, 2, 3, 4) and their 2D histograms

#### IV. RESULTS

The first difficulties in the interpretation of the 2D histogram appear when there's presence of several identical panels as is the case of image 3; indeed the point clouds in the 2D histogram corresponding to the different panels containing the same color are more or less alike.

To work around this type of problem and highlight a difference between the panels we perform a multi-resolution analysis on the image of histogram by wavelet. Specifically we use the algorithm of Mallat, Zhong, [5] to be able to highlight the peaks in the 2D image histogram and have the opportunity to refine by changing scale.

The figure below shows the result of segmentation of the image (a) with peaks extracted from the 2D histogram according to multi-resolution decompositions in levels 3, 4 and 5.

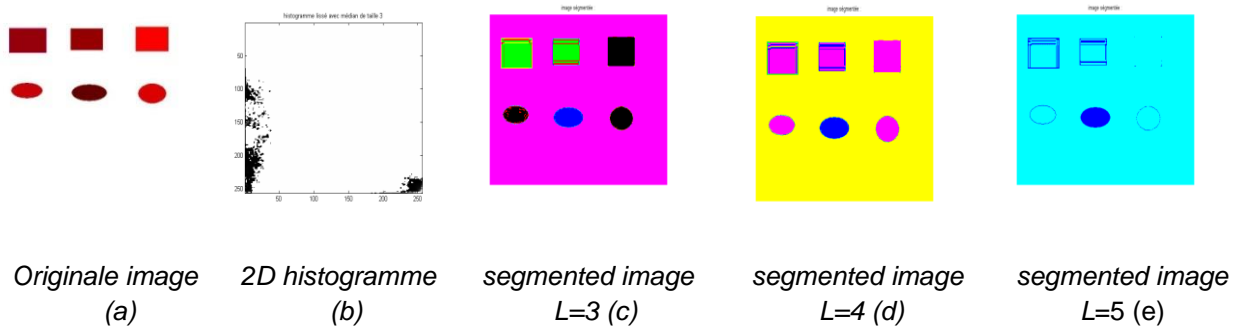


Fig 4 : Original image (a), the 2D histogram (b), (c, d, e) segmented images with levels decomposition (3, 4,5) respectively.

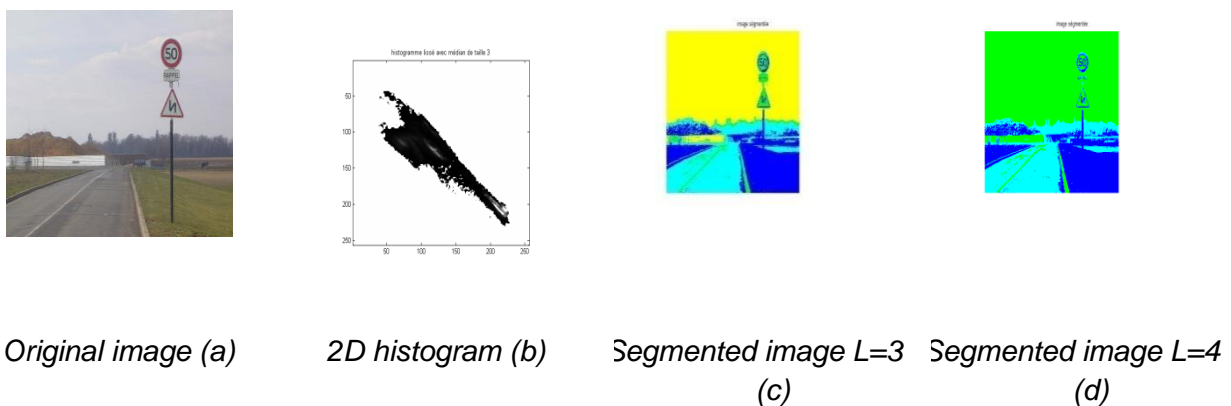


Fig 5 : Original image (a), the 2D histogram (b), (c, d, e) segmented images with levels decomposition (3, 4,5) respectively.

## V. CONCLUSIONS

We have developed an approach to segmentation of color images which is based on a multi-resolution analysis of the 2D histogram. The method was applied to a set of synthetic and real images to approve the performance of the segmentation algorithm. Also, the method has been particularly applied to the detection of road signs and the results are more satisfactory than those obtained using other segmentation algorithms especially those using 1D histograms.

## REFERENCES

- [1] J. Liu, Y.H. Yang, Multirésolution color image segmentation, IEEE Trans. PAMI., July, No. 16, 1999. pp 689-700.
- [2] Ramana L. Rao and Lakshman Prasad, "Segmentation by Multiresolution Histogram Decomposition", citeseer.ist.psu.edu/rao95segmentation.html.
- [3] CABANI I; TOULMINET G; BENSRAHAI A " Color-based Detection of Vehicle Lights" IEEE Intelligent Vehicles Symposium, Las Vegas, USA Année:2005
- [4] A.Ezzine & co Une nouvelle méthode de segmentation basée sur les ondelettes: application pour la détection des feux de stop d'un véhicule. Proceeding Optique 2006, INPT Rabat
- [5] S. Mallat and S. Zhong, Wavelet transform maxima and multiscale edges, in Wavelets and Their Applications, G. Beylkin, R. Coifman, I.Daubechies, S. Mallat, Y. Meyer, L. Raphael, M. Ruskai (eds.), Cambridge, MA,
- [6] Y. Meyer, Ondelettes et opérateurs I : ondelettes, Paris, Herman, 1990
- [7] A.Ezzine, A. Cohen, Quadratures singulières et fonctions d'échelle", CRAS, t.323, Serie I, 829-834, 1996.
- [8] I. Daubechies, "Ten lectures in wavelets", Cbms-Nsf Regional Conference Series in Applied Mathematics, No 61
- [9] Lezoray, O. & Cardot, H. (2003). Hybrid color image segmentation using 2D histogram clustering and region merging, Proc. Int. Conf. on Image and signal processing:ICISP'03, pp. 22-29, Agadir, 2003, Maroc