

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université 20 Août 1955 -Skikda
Faculté des Sciences
Département d'Informatique
Ref :.....



جامعة 20 أوت 1955 -سكيكدة
كلية العلوم
قسم: الإعلام الآلي
المرجع:.....

Thèse

En vue de l'obtention du diplôme de

Doctorat de 3^o cycle (LMD) en Informatique

Option : *Computation et Cognition des Systèmes Informatiques*

Apport de la texture à la caractérisation d'images : Application à la recherche de l'image par le contenu visuel (CBIR)

Présentée par :

Salah Bougueroua

Soutenue publiquement le 18 mai 2017

Devant le Jury :

M. Mohamed REDJIMI,	Professeur,	Université 20 Août 1955 -Skikda,	Président
M. Mohamed BATOUCHE,	Professeur,	Université A.Mehri Constantine 2,	Examineur
M. Ramdane MAAMRI,	Professeur,	Université A.Mehri Constantine 2,	Examineur
M. Smaine MAZOUZI,	MC 'A',	Université 20 Août 1955 -Skikda,	Examineur
M. Bachir BOUCHEHAM,	Professeur,	Université 20 Août 1955 -Skikda,	Directeur de thèse

Année Universitaire : 2016/2017

Remerciements

Je tiens à exprimer toute ma gratitude au directeur de ma thèse, Pr. Bachir Boucheham, pour la liberté qu'il m'a donnée dans mon travail, sa grande disponibilité et surtout pour ses conseils judicieux.

Je remercie aussi chaleureusement M. Mohamed Redjimi, Professeur à l'Université 20 août 1955 de Skikda qui m'a fait l'honneur d'accepter de présider le jury de ma thèse. Je remercie également et vivement M. Mohamed Batouche, Professeur à l'Université Constantine 2, M. Ramdane Maamri, Professeur à l'Université Constantine 2, M. Smaine Mazouzi, Maître de Conférences A à l'Université 20 août 1955 de Skikda, de bien avoir voulu accepter de juger mon travail.

Je remercie sincèrement tous les enseignants au sein du département d'informatique de l'Université de Skikda, ainsi que mon ami et collègue Dr. Mohamed Belaoued pour ses conseils.

Enfin, mes vifs remerciements vont à mes parents pour tout ce qu'ils ont fait pour moi, sans eux, je ne serais jamais arrivé à cette étape.

Dédicaces

Je dédie cette thèse

À mes très chers parents et à toute la famille sans exception

À mes chers amis : Bilel Chebbour, Hamza Boubraïka, Haroune Zighed, Hichem Mekadem, Mohamed Belaoued, Mohamed Brahimí, Mehdi Boukehla, Nassim Aibeche, Nassreddine Chenniki, Noureddine Seddari, Slimane Benloucif,...

ملخص

تندرج هذه الاطروحة في سياق البحث عن الصور باستعمال مقارنة البحث عن طريق المحتوى. هذه المقاربة البديلة للبحث عن الصور بواسطة الكلمات المفتاحية، تركز على خصائص الصورة المرئية كاللون، النسيج، الشكل..... إلخ للبحث و استرجاع الصور ذات صلة بالاستعلام المقدم. لكن، حتى مقارنة البحث بالمحتوى تعاني هي الأخرى من بعض المشاكل التي لا تسمح لها بان تكون بالفعالية المرجوة كالاتباع عن إدراك المعنى الحقيقي للصورة.

من بين السمات المذكورة أعلاه، النسيج هو النقطة المحورية في بحثنا. هذه السمة جد معروفة ومستخدمة على نطاق واسع في حياتنا اليومية، ولكن النسيج هو من بين أصعب الخصائص للتوصيف نظرا لعدم وجود تعريف رسمي، دقيق وتوافقي، على الرغم من أن تحليل الملمس هو حقل بحث نشط مضى عليه أكثر من خمسة عقود. إذن في إطار هذه الرسالة نقوم بدراسة حالة الفن، ثم نقوم بتقديم عدة طرق عالية الأداء تركز على الخصائص ذات المستوى المنخفض.

أساسا، مساهمتنا الأولى تتمثل في تحسين الطريقة الشهيرة " أنماط الثنائية المحلية ". حيث نقترح دمج تدريجي للأنماط الثنائية المحلية، من خلال مخطط جديد لتحديد المناطق، والذي يختلف عن نظيره الذي يستند إلى الكتل. في المساهمة الثانية، نركز على الموجات. في هذا الصدد، سنقوم بالتبيان التجريبي للفرق بين توزيع قيم المعدلات والطاقات، ثم تأثير أخذ بعين الاعتبار الطاقة في المعادلات اللاحقة. المساهمة الرئيسية الثالثة هي مساهمة في مستوى مرحلة المقارنة. الفكرة تقوم على أساس الأخذ بعين الاعتبار للصور المسترجعة في لحظة ما للبحث عن الصورة التالية.

الكلمات المفتاحية: البحث عن الصور عن طريق المحتوى، النسيج، أنماط الثنائية المحلية، الموجات، استخلاص الخصائص، دمج تدريجي للمعلومة المحلية

Résumé

Dans ce travail, nous nous intéressons à la recherche d'images par le contenu. Cette intéressante approche, plus connue par son acronyme anglais CBIR (*Content Based Image Retrieval*), a vu le jour afin de remédier aux problèmes engendrés par son prédécesseur, basé sur le texte (TBIR : *Text based image retrieval*). Principalement, le TBIR souffre de deux inconvénients. Le premier consiste dans la difficulté de la tâche d'attribution des mots clés aux images, due à son aspect manuel. En outre, le deuxième inconvénient réside dans la subjectivité de la tâche. Afin de parer à ces inconvénients, l'approche alternative CBIR utilise les caractéristiques visuelles de l'image, telles que *la couleur, la texture, la forme...* etc. Cependant, le fait de se baser sur de telles caractéristiques de bas niveau pour représenter et exprimer le contenu d'une image, généralement très complexe, engendre un écart significativement important avec les sémantiques et les significations de celle-ci. Cet écart est communément connu par *le fossé sémantique*.

Particulièrement, l'attribut *texture* est très répandu et très utilisé, vu que les êtres humains sont capables de distinguer entre différentes manifestations de cet attribut. Néanmoins, sur le plan computationnel, la texture demeure un attribut difficile à caractériser. Dans le présent travail, la texture représente le point de focalisation. Par conséquent, un ensemble de méthodes basées textures seront abordées ainsi que la proposition de plusieurs contributions.

Principalement, notre première contribution est une amélioration de la célèbre méthode *LBP* (*Local Binary Pattern, motif binaire local*), où nous proposons d'intégrer graduellement l'information locale véhiculée par les motifs binaires locaux. Ceci est obtenu à travers un nouveau schéma de délimitation des régions, qui diffère de celui basé sur les blocs (régions rectangulaires classiques). Par conséquent, notre méthode est baptisée *GLIBP* (*Gradual Locality Integration of binary patterns*). Le succès de cette instance du *Framework* proposé *GLI* (*Gradual Locality Integration*) nous a poussés à investiguer la faisabilité de l'approche pour l'extraction des caractéristiques de la *couleur*. Dans ce contexte, la méthode *GLI-Color* a été proposée ainsi que deux de ses variantes. La première opère sur l'espace de couleur HSV, d'où son nom : *GLI-Color^(HSV)*. Quant à la deuxième variante, elle utilise les moments d'où son nom *M-GLI-Color*. Dans la seconde contribution, nous nous intéressons aux ondelettes dans le contexte de la caractérisation du contenu textural de l'image. Dans cette optique, la littérature existante montre que des auteurs affirment que la moyenne des sous-bandes d'ondelettes est égale à (ou proche de) zéro. Cette affirmation nous a paru digne d'intérêt, vu que les mesures calculées des différentes sous-bandes sont généralement les *moments statistiques*; or la moyenne est utilisée lors du calcul de ces derniers. Donc, comme première étape, nous avons vérifié expérimentalement la différence entre la distribution des moyennes des sous-bandes d'ondelettes et celle des énergies des sous-bandes. Les résultats obtenus montrent une différence intéressante entre les deux distributions. Effectivement, par opposition aux valeurs des moyennes qui se concentrent autour de l'axe de zéro, celles des énergies sont plus

dispersées. En se basant sur ces résultats, le calcul des moments statistiques autour de l'énergie au lieu de la moyenne, semble une modification efficace afin d'améliorer les capacités discriminatives de ces mesures. Cette hypothèse a fait l'objet de la deuxième étape où nous avons montré l'apport de la considération de l'énergie dans les formules des moments statistiques sur la performance. Notre troisième contribution principale agit quant à elle au niveau de l'étape de comparaison des systèmes CBIR. Elle repose sur une *stratégie gloutonne* basée sur une mesure d'homogénéité importée du domaine du *biclustering des Expressions de Gènes*, appelée *Mean Squared Residue* (MSR). Notre méthode est distinguée par le fait qu'elle profite de toutes les images déjà jugées pertinentes par le système CBIR, pour la sélection de l'image suivante. Ceci est complètement différent de l'approche classique des systèmes CBIR, où seulement la requête est utilisée lors de la recherche des images, puisque le résultat de la requête est alors traité comme un cluster homogène dont la croissance est réalisée de façon incrémentale.

L'ensemble des résultats obtenus par les différents algorithmes et techniques proposés ont montré des performances intéressantes, voire supérieures, par rapport à plusieurs travaux publiés. En particulier, la méthode GLIBP a montré des performances supérieures à celles des méthodes : *LBP* de [Ojala et al. \(1996\)](#), *DLEP* de [Murala et al. \(2012a\)](#), '*Block-based LBP*' de [Takala et al. \(2005\)](#), entre autres.

D'un autre côté, les résultats obtenus nous encouragent à continuer sur cet axe de recherche et aussi nous inspirent des perspectives à explorer dans le futur, particulièrement, l'exploitation d'autres techniques de biclustering des expressions génétiques dans le domaine de la recherche d'images par le contenu.

Mots-clés : CBIR, recherche d'images par le contenu, texture, LBP, ondelette, biclustering, extraction des caractéristiques, intégration graduelle de l'information locale

Abstract

In this work, we are interested in the retrieval of images by the content. This interesting alternative approach which is known more by its English acronym CBIR (*Content Based Image Retrieval*) was proposed in order to remedy to the problems caused by its text-based predecessor (TBIR: *Text based image retrieval*). Mainly, TBIR suffers from two disadvantages. The first one is the difficulty of assigning keywords to images due to its manual appearance. On the other hand, the second disadvantage is the subjectivity of the task. In order to counter these disadvantages, the alternative approach CBIR uses the visual characteristics of the image, such as *color, texture, shape...* etc., for satisfying the user query. However, relying on such low-level characteristics to represent and express the content of an image, generally very complex, creates a significant gap with the semantics and the meanings of the image content. This gap is commonly known by the *semantic gap*.

Particularly, “texture” is a very widespread and widely used characteristic, since human beings are able to distinguish between different manifestations of this attribute. Nevertheless, on the computational level, this attribute remains difficult to characterize. In the present work, texture represents the focal point. Therefore, a set of texture-based methods were addressed and several contributions were proposed.

Mainly, our first contribution is an improvement of the famous LBP method (*Local Binary Patterns*), where we propose to integrate gradually the local information conveyed by the binary patterns. This is achieved through a new region delimitation scheme, which is different from that based on blocks (classic rectangular regions). Hence, our method is called *GLIBP (Gradual Locality Integration of binary patterns)*. The success of this instance of the proposed *Framework, GLI (Gradual Locality Integration)* encouraged us to investigate its feasibility for extracting the characteristics of *color* attribute. In this context, the *GLI-Color* method was proposed as well as two of its variants. The first operates on the HSV color space, hence its name: *GLI-Color^(HSV)*. As for the second variant, it uses the moments, hence its name *M-GLI-Color*. In the second contribution, we were interested in Wavelets in the context of the characterization of image textural content. From this perspective, the existing literature shows that authors claim that the means of the wavelet sub-bands is equal to (or close to) zero. This assertion seemed to us worthy of interest, since the calculated measurements of the different sub-bands are generally the *statistical moments* and the mean is used when calculating the latter. Thus, as a first step, we experimentally verified the difference between the distribution of the wavelet sub-bands means and that of the wavelet sub-bands energies. The obtained results show an interesting difference between the two distributions. Indeed, to the contrary of the means values, which concentrate around the axis of zero, those of the energies are more dispersed. Based on these results, the calculation of the statistical moments around the energy instead of the mean seemed to us to be an effective modification in order to improve the discriminative capacities of these measures. This hypothesis was the subject of the second step, where we

showed the effect of the consideration of energy in the statistical moment's formulas on the effectiveness. As for our third main contribution, it acts on the comparison step of the CBIR systems. This contribution relies on a *greedy strategy* based on a homogeneity measurement imported from the *biclustering of Gene's Expressions* domain, called *Mean Squared Residue (MSR)*. The proposed method is distinguished by the fact that it takes advantage of all the images already selected as pertinent for the selection of the following image. This is completely different from the traditional approach of CBIR systems, where only the query is used when searching for images. Indeed, in the proposed approach, the results corresponding to a given query are returned as a cluster, grown incrementally.

All the results obtained by the various proposed algorithms and techniques have shown interesting or even better performances compared to several published works. In particular, the GLIBP method showed performances superior to those of the methods: LBP of [Ojala et al. \(1996\)](#), DLEP of [Murala et al. \(2012a\)](#), 'Block-based LBP' of [Takala et al. \(2005\)](#), among others.

On the other hand, the results obtained encourage us to continue on this field of research and inspire us with perspectives to be explored in the future, particularly the exploitation of other biclustering techniques of genetic expressions in the field of content based image retrieval.

Key words: CBIR, image retrieval by the content, texture, LBP, wavelet, biclustering, features extraction, gradual locality integration

Table des matières

REMERCIEMENTS	III
ملخص	VII
RESUME	IX
ABSTRACT	XI
TABLE DES MATIERES	XIII
TABLE DES FIGURES	XVII
Liste des tableaux	XIX
CHAPITRE I. INTRODUCTION GENERALE.....	1
I.1 CONTEXTE DE TRAVAIL	1
I.2 PROBLEMATIQUE	3
I.3 OBJECTIFS DU TRAVAIL	3
I.4 CONTRIBUTIONS	4
I.5 ORGANISATION DU MANUSCRIT	5
CHAPITRE II. RECHERCHE D'IMAGES PAR LE CONTENU	7
II.1 SYSTEME DE RECHERCHE D'IMAGES PAR LE CONTENU.....	7
II.1.1 <i>Extraction des caractéristiques</i>	8
II.1.2 <i>Comparaison</i>	9
II.2 EVALUATION	11
II.3 BENCHMARKING	12
II.4 CONCLUSION.....	16
CHAPITRE III. DESCRIPTEURS VISUELS DE L'IMAGE.....	17
III.1 LA COULEUR.....	18
III.1.1 <i>Les espaces de couleurs</i>	19
III.1.1.1 Espace RGB.....	20
III.1.1.2 Espace TSV(HSV).....	20
III.1.1.3 Espace CIE XYZ.....	21
III.1.1.4 Espace CIE L*a*b*	22
III.1.1.5 Espace CIE L*u*v	22
III.1.2 <i>Les descripteurs de couleurs</i>	23
III.1.2.1 Descripteurs globaux.....	23
III.1.2.2 Descripteurs locaux.....	23
III.2 LA TEXTURE	26
III.2.1 <i>Discussions</i>	27
III.2.1.1 De quoi résulte la texture ?	27
III.2.1.2 Texture stationnaire et non-stationnaire	28
III.2.1.3 Quelques termes décrivant les primitives de texture	28
III.2.1.4 Quelques définitions de la texture	30
III.2.2 <i>Applications de la texture</i>	30
III.3 LA FORME.....	32

III.3.1	<i>Descripteurs basés-Contour</i>	34
III.3.1.1	Code de Freeman (Chain coded string).....	34
III.3.1.2	Descripteur de Fourier.....	35
III.3.2	<i>Descripteurs basés-Région</i>	36
III.3.2.1	Les moments invariants.....	36
III.3.2.2	Les moments de Zernike.....	36
III.4	LES POINTS D'INTERET.....	37
III.4.1	<i>Détecteurs</i>	38
III.4.1.1	Détecteur de Harris.....	38
III.4.1.2	Détecteur Hessien.....	39
III.4.1.3	Laplacien de Gaussien.....	39
III.4.1.4	Différence de Gaussien.....	40
III.4.2	<i>Descripteurs</i>	40
III.4.2.1	SIFT.....	40
III.4.2.2	SURF.....	40
III.4.2.3	GLOH.....	41
III.4.2.4	LPDF.....	41
III.5	COMBINAISON DES DESCRIPTEURS DE TEXTURE ET DE COULEUR.....	42
III.5.1	<i>Les méthodes parallèles</i>	43
III.5.2	<i>Les méthodes séquentielles</i>	43
III.5.3	<i>Les méthodes intégratives</i>	44
III.5.3.1	Monocanal (single-channel).....	44
III.5.3.2	Multicanal (multi-channel).....	45
III.6	CONCLUSION.....	45
CHAPITRE IV.	DESCRIPTEURS DE TEXTURE	47
IV.1	LES DESCRIPTEURS.....	48
IV.1.1	<i>Les matrices de cooccurrence de Haralick</i>	48
IV.1.2	<i>Les descripteurs de Tamura</i>	52
IV.1.2.1	La grossièreté.....	52
IV.1.2.2	Le contraste.....	53
IV.1.2.3	La directivité.....	53
IV.1.3	<i>Les motifs binaires locaux</i>	54
IV.1.3.1	Variantes du LBP.....	57
IV.1.3.2	Applications.....	69
IV.1.4	<i>Les ondelettes</i>	70
IV.1.4.1	Présentation des ondelettes.....	70
IV.1.4.2	Les ondelettes et les images.....	72
IV.1.4.3	Applications et descripteurs des ondelettes pour la caractérisation des images.....	73
IV.1.5	<i>Les filtres de Gabor</i>	74
IV.1.5.1	Présentation.....	74
IV.1.5.2	Applications.....	77
IV.2	CLASSIFICATION DES DESCRIPTEURS.....	81
IV.2.1.1	Approches statistiques.....	81
IV.2.1.2	Approches structurelles.....	83
IV.2.1.3	Approches basées sur les modèles.....	84
IV.2.1.4	Approches basées sur les transformations.....	86
IV.3	CONCLUSION.....	88
CHAPITRE V.	VERS DES DESCRIPTEURS VISUELS PLUS PERFORMANTS POUR LA RECHERCHE DE	
L'IMAGE	89	
V.1	INTEGRATION GRADUELLE DE L'INFORMATION LOCALE.....	90
V.1.1	<i>Motivation</i>	90

V.1.2	<i>Gradual Locality Integration of Binary Patterns (GLIBP)</i>	91
V.1.2.1	La rotation d'ellipses	92
V.1.2.2	Mise en échelle et l'enchaînement d'ellipses	92
V.1.2.3	La comparaison	93
V.1.2.4	Expérimentation.....	94
V.1.2.5	Résultats et discussion	94
V.1.3	<i>GLI pour la couleur</i>	98
V.1.3.1	GLI-Color et ses deux variantes GLI-Color ^(HSV) et M-GLI-Color.....	98
V.1.3.2	Résultats et discussion	99
V.2	EXPLOITATION DE L'ÉNERGIE DES SOUS-BANDES D'ONDELETTES	102
V.2.1	<i>Motivation</i>	102
V.2.2	<i>L'énergie et la moyenne de sous-bande d'ondelettes</i>	102
V.2.2.1	Remarques et discussion.....	104
V.2.3	<i>L'impact de l'intégration de l'énergie dans les moments d'ondelettes</i>	104
V.2.4	<i>Expérimentation</i>	105
V.2.5	<i>Résultats et discussion</i>	105
V.3	CONCLUSION.....	108
CHAPITRE VI. UNE NOUVELLE METHODE GLOUTONNE POUR LA RECHERCHE D'IMAGES BASEE MSR (MEAN SQUARED RESIDUE)		111
VI.1	MOTIVATION.....	111
VI.2	RAPPEL	112
VI.2.1	<i>Biclustering</i>	112
VI.2.2	<i>Mean Squared Residue (MSR)</i>	112
VI.3	LE MSR POUR LA RECHERCHE D'IMAGES	113
VI.3.1	<i>Exemple</i>	114
VI.3.2	<i>Expérimentation</i>	115
VI.3.3	<i>Résultats et discussion</i>	115
VI.4	CONCLUSION.....	117
CHAPITRE VII. CONCLUSION GENERALE		119
VII.1	PERSPECTIVES.....	120
BIBLIOGRAPHIE		121
ANNEXE (BIBLIOGRAPHIE DU CANDIDAT)		137

Table des figures

FIGURE II-1 SCHEMA ILLUSTRATIF D'UN SYSTEME CBIR.	8
FIGURE II-2 QUELQUES IMAGES DE LA BASE COREL 1k (WANG).	14
FIGURE II-3 QUELQUES IMAGES DE LA BASE MIRFLICKR	14
FIGURE II-4 LES DIFFERENTES TEXTURES DE LA BASE KYLBERG.	14
FIGURE II-5 QUELQUES TEXTURES DE LA BASE BRODATZ.	15
FIGURE II-6 QUELQUES TEXTURES DE LA BASE VISTEX.	15
FIGURE II-7 LES TEXTURES DE LA BASE OUTEX.	16
FIGURE III-1 DIAGRAMME SCHEMATIQUE DE L'ŒIL HUMAINE.	19
FIGURE III-2 L'ESPACE DE COULEURS RGB.	20
FIGURE III-3 L'ESPACE DE COULEURS HSV.	21
FIGURE III-4 QUELQUES TEXTURES DE DIFFERENTES BASES DE DONNEES.	26
FIGURE III-5 ROLE DE LA TEXTURE.	27
FIGURE III-6 IMPACT DE LA DISTANCE SUR L'APPARENCE DE TEXTURE (UMD HR DATASET).	28
FIGURE III-7 DEUX IMAGES DE TEXTURES, CHACUNE EST CONSTITUEE DE DEUX REGIONES ONT LES MEMES STATISTIQUES DU 2EME ORDRE. (A) : LES DEUX REGIONES NE SONT PAS DISCRIMINABLES PAR LE SVH PREATTENTIF, (B) : LES DEUX REGIONES SONT FACILEMENT DISCRIMINABLES. (TUCERYAN ET JAIN, 2011)	29
FIGURE III-8 EXEMPLES D'IMAGES AERIENNES DE LA BASE USC-SIPI.	31
FIGURE III-9 QUELQUES FORMES REPRESENTANT DIFFERENTS OBJETS (IMAGES DE LA BASE MPEG400).	32
FIGURE III-10 EXEMPLES D'IMAGES DE LOGOS (LOGO DATASET).	33
FIGURE III-11 CLASSIFICATION DES REPRESENTATIONS DE LA FORME ET LES TECHNIQUES DE DESCRIPTION (ZHANG ET LU, 2004).	34
FIGURE III-12 GRILLE POLAIRE, (A): 4 BINS SUR LA DIRECTION RADIALE ET 16 SUR LA DIRECTION ANGULAIRE; (B) : 3 BINS SUR LA DIRECTION RADIALE ET 8 SUR LA DIRECTION ANGULAIRE (SANS DIVISION DU BIN CENTRAL)	41
FIGURE III-13 APPROCHE PARALLELE.	43
FIGURE III-14 APPROCHE SEQUENTIELLE.	43
FIGURE III-15 APPROCHE INTEGRATIVE. (A) : SINGLE CHANNEL, (B) : MULTI CHANNEL.	44
FIGURE IV-1 ILLUSTRATION DE DIFFERENTES DIRECTIONS DES MCNG.	49
FIGURE IV-2 EXEMPLE DE CALCUL DES MCNG.	49
FIGURE IV-3 LBP DE BASE. EXEMPLE DE CALCUL.	55
FIGURE IV-4 EXEMPLES D'IMAGE LBP, (A GAUCHE) : DEUX IMAGES EN COULEUR, (A DROITE) : LES IMAGES LBP CORRESPONDANTES	55
FIGURE IV-5 LBP GENERALE.	56
FIGURE IV-6 LTP, UPPER ET LOWER PATTERNS.	59
FIGURE IV-7 VOISINAGE DU LDP.	60
FIGURE IV-8 OPPONENT COLOR LBP.	62
FIGURE IV-9 LA METHODE BASEE SUR LES BLOCS.	62
FIGURE IV-10 LA METHODE BASEE SUR LES BLOCS PRIMITIFS.	63
FIGURE IV-11 LTRIDP, EXEMPLE.	65
FIGURE IV-12 (A) : nLBP, (B) : dLBP.	67
FIGURE IV-13 SS-3D-LTP.	68
FIGURE IV-14 ENHANCED LBP.	69
FIGURE IV-15 LES ESPACES D'APPROXIMATION.	71
FIGURE IV-16 LES ONDELETTES. LA DECOMPOSITION PYRAMIDALE.	73
FIGURE IV-17 LA COURBE D'UNE FONCTION GAUSSIENNE, ET CELLE DE LA FONCTION DE GABOR (PARTIE REELLE).	75
FIGURE IV-18 UNE FONCTION GAUSSIENNE (2D).	76
FIGURE IV-19 PLOT DE LA FONCTION DE GABOR (PARTIE REELLE).	76

FIGURE IV-20 PLOT DE LA FONCTION DE GABOR (PARTIE IMAGINAIRE).....	77
FIGURE IV-21 SIX MOTIFS SUR UNE GRILLE 2x2.	82
FIGURE IV-22 LES VOISINAGES DE DIFFERENTS ORDRES-MRF.	85
FIGURE IV-23 FOURIER LOCAL. (A) : UNE SEQUENCE DE 8 PIXELS, (B) : 4 SEQUENCES DE 4 PIXELS.	87
FIGURE V-1 POSITIONNEMENT D'ELLIPSES (S=3, N=7).	91
FIGURE V-2 ILLUSTRATION DES REGIONS ELLIPTIQUES SUR L'IMAGE LBP (S=3, N=7). À GAUCHE : L'IMAGE ORIGINALE, A DROITE : LES REGIONS ELLIPTIQUES SUR L'IMAGE LBP.	92
FIGURE V-3 ENCHAÎNEMENT DES ELLIPSES.	93
FIGURE V-4 ILLUSTRATION DE LA COMPARAISON DE LA METHODE GLIBP, (A,B,U,N,S)=(50,25,1.4,10,2).....	94
FIGURE V-5 LA METHODE GLI-COLOR, ILLUSTRATION DES REGIONS ELLIPTIQUES SUR LES TROIS CANAUX DE L'ESPACE RGB.....	98
FIGURE V-6 ILLUSTRATION DES REGIONS ELLIPTIQUES SUR LES TROIS CANAUX : HUE, SATURATION ET VALUE DE L'ESPACE DE COULEUR HSV.	99
FIGURE V-7 COMPARAISON DES TROIS VARIANTES EN TERMES DE TEMPS DE COMPARAISON MOYENNE (SECONDS).	101
FIGURE V-8 LES MOYENNES ET LES ECARTS TYPE DES MOYENNES DES CANAUX POUR CHACUNE DES CLASSES DE LA BASE OUTEX_TC_00000.	103
FIGURE V-9 LES MOYENNES ET LES ECARTS TYPE DES ENERGIES DES CANAUX POUR CHACUNE DES CLASSES DE LA BASE OUTEX_TC_00000.	103
FIGURE V-10 COMPARAISON DE PERFORMANCES DES DIFFERENTES METHODES EN TERMES DE LA PRECISION MOYENNE, (OUTEX_TC_00000).	107
FIGURE V-11 COMPARAISON DE PERFORMANCES DES DIFFERENTES METHODES EN TERMES DE LA PRECISION MOYENNE, (BRODATZ).	108
FIGURE VI-1 LES ETAPES D'ELARGISSEMENT DE L'ENSEMBLE I.	114

Liste des Tableaux

TABLEAU II-1 QUELQUES BASES DE DONNEES DE REFERENCE UTILISEES DANS LA LITTERATURE.	13
TABLEAU III-1 TABLEAU RECAPITULATIF DES DESCRIPTEURS DE COULEUR DISCUTES.....	25
TABLEAU IV-1 QUELQUES FILTRES PASSE-HAUT DE DECOMPOSITION POUR QUELQUES TYPES D'ONDELETTES	73
TABLEAU IV-2 TABLEAU RECAPITULATIF DES VARIANTES LBP DISCUTEES.	78
TABLEAU IV-3 TABLEAU RECAPITULATIF DES DESCRIPTEURS DE TEXTURE DISCUTES-(MCNG).	79
TABLEAU IV-4 TABLEAU RECAPITULATIF DES DESCRIPTEURS DE TEXTURE DISCUTES-(TAMURA).....	80
TABLEAU IV-5 TABLEAU RECAPITULATIF DES TRAVAUX EXPLOITANTS LES ONDELETTES.	80
TABLEAU V-1 LES RESULTATS DE LA METHODE GLIBP EN TERMES DE PRECISION MOYENNE PONDEREE, AVEC LES PARAMETRES (A,B,U,S)=(50,25,1.4,2) ET DIFFERENTES VALEURS POUR N, EN UTILISANT LA METRIQUE DE DISTANCE MANHATTAN ET D1 (w=20).....	95
TABLEAU V-2 LES RESULTATS DE LA METHODE GLIBP EN TERMES DE PRECISION MOYENNE PONDEREE, AVEC LES PARAMETRES (A,B,U,S)=(30,15,2,3) ET DIFFERENTES VALEURS POUR N, EN UTILISANT LA METRIQUE DE DISTANCE MANHATTAN ET D1 (w=20).....	95
TABLEAU V-3 COMPARAISON DE LA METHODE PROPOSEE AVEC QUELQUES METHODES EXISTANTES EN TERMES DE PRECISION MOYENNE PONDEREE EN UTILISANT LA METRIQUE DE DISTANCE MANHATTAN, (A,B,U,N,S)= (50,25,1.4,10,2).	96
TABLEAU V-4 COMPARAISON DE LA METHODE PROPOSEE AVEC QUELQUES METHODES EXISTANTES EN TERMES DE PRECISION MOYENNE, (A,B,U,N,S)= (50,25,1.4,10,2) (w=9).....	97
TABLEAU V-5 COMPARAISON DE LA METHODE PROPOSEE AVEC QUELQUES METHODES EXISTANTES EN TERMES DE PRECISION MOYENNE, (A,B,U,N,S)= (50,25,1.4,10,2).....	97
TABLEAU V-6 LA PRECISION MOYENNE PONDEREE ET LA PRECISION MOYENNE DE GLI-COLOR POUR w=10 ET w=20, AVEC DIFFERENTS ENSEMBLES DE PARAMETRES (A-B-U-N-S), EN UTILISANT LES METRIQUES DE DISTANCES MANHATTAN ET D1.....	100
TABLEAU V-7 COMPARAISON DE GLI-COLOR(HSV) AVEC GLI-COLOR EN TERMES DE LA PRECISION MOYENNE PONDEREE ET LA PRECISION MOYENNE POUR w=10 ET w=20, EN UTILISANT DIFFERENTS ENSEMBLES DE PARAMETRES (A-B-U-N-S), ET LA METRIQUE DE DISTANCE MANHATTAN.	100
TABLEAU V-8 COMPARAISON DE LA METHODE PROPOSEE GLI-COLOR AVEC QUELQUES METHODES EXISTANTES EN TERMES DE PRECISION MOYENNE (w=10).....	101
TABLEAU V-9 COMPARAISON EN TERMES DE TEMPS DE COMPARAISON (SECONDS).	102
TABLEAU V-10 L'ECART TYPE DES MOYENNES DE MOYENNES ET DES MOYENNES D'ENERGIES SUR TOUTE LA BASE (INTER-CLASS). ...	104
TABLEAU V-11 COMPARAISON DE LA METHODE PROPOSEE AVEC QUELQUES COMBINAISONS DE CARACTERISTIQUES EN UTILISANT 2 NIVEAUX DE DECOMPOSITIONS, (OUTEX_TC_00000, ET K=20).....	106
TABLEAU V-12 COMPARAISON DE LA METHODE PROPOSEE AVEC QUELQUES COMBINAISONS DE CARACTERISTIQUES EN UTILISANT 3 NIVEAUX DE DECOMPOSITIONS (OUTEX_TC_00000, ET K=20).	106
TABLEAU V-13 COMPARAISON DE LA METHODE PROPOSEE AVEC QUELQUES COMBINAISONS DE CARACTERISTIQUES EN UTILISANT 2 NIVEAUX DE DECOMPOSITIONS, (BRODATZ, AND K=16).	107
TABLEAU V-14 COMPARAISON DE LA METHODE PROPOSEE AVEC QUELQUES COMBINAISONS DE CARACTERISTIQUES EN UTILISANT 3 NIVEAUX DE DECOMPOSITIONS, (BRODATZ, AND K=16).	108
TABLEAU VI-1 COMPARAISON DE LA METHODE PROPOSEE AVEC LA DISTANCE EUCLIDIENNE EN TERMES DE PRECISION MOYENNE (LA BASE OUTEX_TC_00000, WINDOW=20).....	115
TABLEAU VI-2 COMPARAISON DE LA METHODE PROPOSEE AVEC LA DISTANCE EUCLIDIENNE EN TERMES DE PRECISION MOYENNE (LA BASE BRODATZ (16 SUB-IMAGES×112)).	116

Chapitre I. Introduction

Générale

« Une image vaut mille mots. »
Confucius

Sommaire

I.1	CONTEXTE DE TRAVAIL	1
I.2	PROBLEMATIQUE	3
I.3	OBJECTIFS DU TRAVAIL	3
I.4	CONTRIBUTIONS	4
I.5	ORGANISATION DU MANUSCRIT	5

I.1 Contexte de travail

La démocratisation des appareils photo numériques et les différents outils d'acquisition, ainsi que l'amélioration d'autres facteurs tels que les outils de stockage et d'archivage, et aussi l'augmentation de la bande passante d'internet a contribué à l'accumulation d'un nombre gigantesque de données multimédias. En particulier, les images sont très répandues et très utilisées dans notre vie personnelle, pour mémoriser nos souvenirs avec la famille et les amis, ou professionnelle pour décrire, par exemple, des anomalies, des situations ou des évènements dans des domaines très utiles, parfois même sensibles, tels que le domaine médical et le domaine militaire. Véritablement, la popularité et l'utilité des images sont liées aux avantages de ces dernières par rapport aux autres médias, tels que la rapidité d'acquisition et les ressources requises réduites, mais avant tout leurs grandes capacités expressives.

Quelques statistiques sur le nombre d'images du grand public, chargées par les différents services et applications disponibles sur internet, montrent clairement la popularité de ce type de média. Par exemple, le nombre de photos partagées dans une seconde sur Instagram¹ est 743 photos². De son côté, Flickr³ a enregistré en 2015, environ 2 millions de photos publiques

¹ <https://www.instagram.com/>

² <http://www.internetlivestats.com/one-second/#instagram-band> (visité le 23/09/2016-00 : 20)

³ <https://www.flickr.com/>

chargées par jour⁴. Quant à Facebook, selon son fondateur Mark Zuckerberg, 2 716 000 photos sont déposées toutes les 20 minutes⁵.

Devant ce nombre immense d'images, les techniques de gestion de bases d'images doivent être en mesure de les gérer. En particulier, la recherche de l'image (*image retrieval*) est une opération cruciale dans ce contexte. Principalement, deux approches existent dans la littérature pour la recherche de l'image. La première approche remonte aux années 1970s. Elle est basée sur l'utilisation *des mots clés* pour la recherche et la récupération d'images, d'où le nom de l'approche TBIR (*Text Based Image Retrieval*). Cette approche est encore utilisée par les différents moteurs de recherche. Peut-être, le meilleur exemple à citer ici est le moteur de recherche *Google image*, sur lequel les mots clés sont utilisés pour récupérer des images répondant à une requête. L'utilisation d'une telle approche nécessite que les images soient étiquetées par des mots clés qui résument et reflètent leurs contenus. Cependant devant un nombre énorme d'images, la tâche d'association des mots clés pour chaque image est clairement longue et délicate. En outre, l'annotation manuelle est subjective et relative à la perception de la personne qui s'en charge, ses sentiments, ses cultures... etc. ([Sethi et al., 2001](#))

La deuxième approche se base sur le contenu visuel de l'image. Elle a commencé d'émerger et de s'imposer depuis les années 90s. Cette nouvelle approche est appelée CBIR (*Content Based Images Retrieval*). Le paradigme CBIR extrait les caractéristiques de quelques attributs de l'image, tels que la couleur, la texture, la forme... etc., sous forme de vecteur, appelé *vecteur de caractéristiques* (*signature*, en anglais : *feature vector*). L'avantage de la signature par rapport à la représentation en pixels est la compression significative de la représentation de l'image. Cependant, la motivation la plus importante est le gain en corrélation entre la représentation de l'image et sa sémantique ([Wang et al., 2001](#)).

L'étape d'extraction joue un rôle primordial dans les systèmes CBIR. Car, les résultats finaux dépendent fortement de cette étape. De ce fait, les caractéristiques extraites doivent être autant que possible discriminantes et réduites pour une meilleure efficacité et une meilleure performance. Dans l'étape suivante, la similarité entre la signature représentant le contenu de l'image requête et celles représentant les images de la base, est mesurée. À la fin, le système retourne les images les plus similaires à l'utilisateur.

Plusieurs systèmes CBIR ont été développés au cours des années, tels que : QBIC (Query By Image Content) ([Niblack et al., 1993](#)), Blobworld ([Carson et al., 1999](#)), NETRA ([Ma et Manjunath, 1999](#)), Photobook ([Pentland et al., 1996](#)), SIMPLicity (Semantics-sensitive Integrated Matching for Picture Libraries) ([Wang et al., 2001](#)), Qatris iManager ([Barrena et al., 2015](#)), pour ne citer que quelques-uns. La description d'un ensemble de systèmes CBIR peut être trouvée dans ([Veltkamp et Tanase, 2002](#)).

Parmi les attributs cités ci-dessus, la texture a été exploitée dans beaucoup de travaux pour la recherche de l'image par le contenu, tels que ([Huang et Dai, 2003](#); [Manjunath et Ma, 1996](#);

⁴ <https://www.flickr.com/photos/franckmichel/6855169886> (visité le 23/09/2016-00 : 50)

⁵ <http://www.planetoscope.com/Internet-/1217-nombre-de-photos-deposees-sur-facebook.html> (visité le 27/12/2016-19 : 45)

[Murala et al., 2012b](#); [Verma et Raman, 2016](#)). En effet, dans les années 1990s, la texture était déjà connue et plusieurs pas ont été faits avec cet attribut dans différents problèmes tels que la classification, la segmentation... etc. Plus précisément, les premiers travaux qui s'y sont intéressés remontent aux années 1960s avec les travaux de [Julesz \(1962\)](#), puis de [Haralick et al. \(1973\)](#) et aussi de [Tamura et al. \(1978\)](#). À noter que la texture reste encore sans définition consensuelle précise. Néanmoins, la recherche sur cet attribut est encore très active.

I.2 Problématique

En réalité, l'inconvénient majeur des systèmes CBIR réside dans ce qui est appelé : *le fossé sémantique* (*semantic gap*) qui se trouve entre les caractéristiques extraites qui sont de bas niveaux et les sémantiques d'images qui sont de hauts niveaux. Le fossé sémantique est défini par [Smeulders et al. \(2000\)](#) comme suit :

*"The semantic gap is the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data have for a user in a given situation."*⁶

Selon [Liu et al. \(2007\)](#), [Eakins et Graham \(1999\)](#) ont classifié les requêtes en CBIR en trois niveaux :

- *Niveau 1* : la requête typique à ce niveau est la requête par *image-exemple*, 'trouver les images similaires à celle-ci'. La recherche est basée sur les caractéristiques de différents attributs de bas niveau.
- *Niveau 2* : la recherche à ce niveau se base sur les objets. À cet effet, les objets sont identifiés par des caractéristiques avec un certain degré d'inférence logique. Par exemple, 'trouver les images de voitures'.
- *Niveau 3* : la recherche à ce niveau est faite par des attributs abstraits. Ce qui laisse la possibilité à une importante quantité de raisonnement de haut niveau sur le but des objets et les scènes décrites. Par exemple, 'trouver les images des fêtes d'anniversaire'.

Les niveaux 2 et 3 représentent la recherche sémantique, alors que le fossé qui se trouve entre les niveaux 1 et 2 est le fossé sémantique ([Eakins et Graham, 1999](#)).

I.3 Objectifs du travail

Vu la problématique présentée ci-dessus, nous nous sommes focalisés tout au long des années de préparation de cette thèse sur la réduction du fossé sémantique. En fait, bien qu'il existe plusieurs techniques de haut niveau dans la littérature pour réduire le fossé sémantique (principalement catégorisées par [Liu et al. \(2007\)](#) en cinq grandes catégories), certainement l'amélioration des systèmes CBIR à travers l'utilisation des techniques de haut niveau dépend

⁶ Le fossé sémantique est le manque de coïncidence entre les informations que l'on peut extraire des données visuelles et l'interprétation que les mêmes données ont pour un utilisateur dans une situation donnée.

également de la qualité des caractéristiques extraites. Autrement dit, la qualité des sorties dépend de la qualité des entrées, ‘*garbage in garbage out*’. D’où l’un des objectifs de cette thèse est la proposition de signatures de bas niveau plus sophistiquées.

En bref, les objectifs de ce travail sont :

- ✓ Dresser un état de l’art sur les différents descripteurs de différents attributs en général, et sur ceux de texture en particulier, afin de mettre en évidence la texture et son apport à la caractérisation de l’image.
- ✓ Mettre en évidence les méthodes LBP et les ondelettes.
- ✓ Proposer des méthodes plus sophistiquées pour l’extraction des caractéristiques de texture en particulier (étape d’extraction des caractéristiques), en les appliquant sur la recherche des images par le contenu. Ce qui permet de montrer, une deuxième fois, l’apport de la texture à la caractérisation d’images. Toutefois, cela ne nous empêche pas de faire des contributions qui agissent au niveau de l’étape de comparaison.

I.4 Contributions

Dans le cadre de cette thèse, nous avons apporté des contributions originales concernant l’attribut ‘texture’ auquel nous nous sommes intéressés. En général, nos contributions peuvent être vues selon trois axes :

- ✓ **L’intégration graduelle de l’information locale (*Gradual Locality Integration, GLI*) :**

Notre première idée se préoccupe de l’amélioration de la capture de l’information locale et sa bonne utilisation. Pour cela, nous proposons un schéma non trivial de délimitation de régions de forme elliptique, complètement différent de celui basé sur les blocs. Étant donné que les régions sélectionnées par le framework proposé sont délimitées par des ellipses de différentes tailles, ceci permet d’intégrer l’information locale caractérisant cette image graduellement et, ainsi, d’avoir une meilleure localisation du contenu discriminatif de l’image. Ce schéma a été inspiré de la forme des filtres de Gabor. Il est à noter que, dans cette étude, les histogrammes sont utilisés pour représenter les caractéristiques de différentes régions elliptiques en se basant sur l’attribut de texture. Principalement, cette idée est utilisée pour capturer l’information locale de motifs binaires (Local binary patterns, LBP). D’où le nom de notre contribution principale (*Gradual Locality Integration of binary patterns, GLIBP*) ([Bougueroua et Boucheham, 2014, 2016a](#)).

Se basant sur le succès du framework GLI (*Gradual Locality Integration*), la même idée fut appliquée avec succès aussi sur la couleur. Il s’en est dégagé alors deux techniques de caractérisation de l’image basées sur l’attribut de couleur : GLI-Color ([Bougueroua et Boucheham, 2015](#)), GLI-Color^(HSV) et M-GLI-Color.

✓ **Les sous-bandes d'ondelettes et l'énergie :**

Lors de nos lectures bibliographiques concernant les ondelettes dans le contexte de la caractérisation de la texture, nous avons constaté que quelques auteurs déclarent que la moyenne des coefficients d'une sous-bande est égale à (ou proche de) zéro. En partant de cette constatation, nous avons commencé par la vérification expérimentale de la différence entre la distribution des moyennes des sous-bandes d'ondelettes et celle de leurs énergies. Les résultats obtenus ont effectivement montré une intéressante différence entre les distributions. Étant donné que les mesures calculées des sous-bandes d'ondelettes sont généralement les moments statistiques, le calcul de ces derniers autour de l'énergie au lieu de la moyenne semble une modification efficace afin d'améliorer les capacités discriminatives de ces mesures. Pour cela, nous avons montré l'impact de cette modification sur la performance ([Bougueroua et Boucheham; Bougueroua et al., 2015](#)).

✓ **Une méthode gloutonne basée *Mean Squared Residue* (MSR) rendant des résultats de CBIR plus homogènes :**

Notre 3^e idée est basée sur notre observation que dans beaucoup de systèmes CBIR, la comparaison est faite *requête-image de la base*, puis le classement des images de la base selon leurs similarités par rapport à la requête. Pour cela, nous avons proposé de rechercher l'image de rang i en prenant compte toutes les images déjà trouvées (rang inférieur), y compris la requête ([Bougueroua et Boucheham, 2016b](#)). Une autre originalité de la méthode est l'utilisation du *Mean Squared Residue* (MSR) dans le CBIR. Sachant que cette mesure est originalement proposée pour le biclustering des expressions génétiques.

Nous rapportons que toutes les méthodes proposées ont été implantées et comparées à des méthodes existantes très performantes. Il s'est avéré que toutes nos propositions sont soit plus performantes soit au moins aussi performantes que les méthodes existantes, comme il va être illustré dans la suite du manuscrit.

I.5 Organisation du manuscrit

Le reste du présent manuscrit est organisé comme suit :

Le second chapitre est consacré à la présentation du schéma général des systèmes CBIR.

Dans le chapitre trois, nous abordons les différents descripteurs visuels, la couleur, la forme et les points d'intérêt, ainsi qu'un ensemble de travaux qui y s'intéressent. En ce qui concerne la texture, nous nous limitons à sa présentation, discussion, ainsi qu'à la mise en évidence de différents domaines d'application. Nous mettons l'accent également sur les différentes approches de combinaison de descripteurs de couleur et de texture.

Dans le quatrième chapitre, nous dresserons l'état de l'art de l'attribut visuel 'Texture'. Un état de l'art relatif aux méthodes de texture les plus significatives, de notre point de vue, est d'abord dressé. Dans la deuxième partie de ce chapitre, nous discutons également la classification des descripteurs selon les approches suivies par ces derniers.

Par le chapitre cinq, nous entamons la partie des contributions personnelles. Dans ce cinquième chapitre, nous présentons nos contributions qui ont été proposées pour améliorer l'extraction des caractéristiques de texture. Principalement, nos contributions seront présentées selon les deux grandes idées, à savoir : *intégration graduelle de l'information locale*, puis *l'impact de l'intégration de l'énergie dans les moments d'ondelettes*.

Le sixième chapitre de ce manuscrit, nous l'avons consacré à la présentation de notre nouvelle approche, utilisant une nouvelle stratégie pour la recherche d'images par le contenu. Cette dernière est une contribution qui agit au niveau de l'étape de comparaison, en contraste des précédentes qui se situent à l'étape d'extraction.

Chapitre II. Recherche d'images par le contenu

« Celui qui déplace la montagne, c'est celui
qui commence à enlever les petites pierres. »

Confucius

Sommaire

II.1	SYSTEME DE RECHERCHE D'IMAGES PAR LE CONTENU.....	7
II.1.1	Extraction des caractéristiques.....	8
II.1.2	Comparaison	9
II.2	EVALUATION	11
II.3	BENCHMARKING	12
II.4	CONCLUSION.....	16

A travers ce chapitre, nous allons présenter les grandes lignes de la recherche d'images par le contenu. Pour cela, nous allons présenter et discuter quelques notions en relation étroite avec la recherche d'images par le contenu. Ces notions seront discutées au fur et à mesure que les différentes sections de ce chapitre sont développées. Principalement, nous nous intéressons dans le présent chapitre à la présentation des différentes étapes des systèmes CBIR. Nous présenterons également quelques méthodes d'évaluation des performances de ces systèmes. Enfin, quelques bases de données (base d'images) seront présentées.

II.1 Système de recherche d'images par le contenu

Comme il a été avancé dans l'introduction générale, l'approche de la recherche d'images par le contenu est apparue dans les années 90s pour remédier aux problèmes engendrés par l'approche textuelle. Cette nouvelle approche est plus connue par l'acronyme anglais CBIR qui réfère à *Content Based Image Retrieval*. CBIR a été défini par [Wang et al. \(2001\)](#) comme étant *l'ensemble des techniques permettant de récupérer des images sémantiquement pertinentes à partir d'une base de données d'images en fonction des caractéristiques d'image automatiquement dérivées.*⁷

Principalement, les systèmes CBIR sont constitués de deux étapes (Figure II-1) : *Extraction des caractéristiques* (feature extraction), et *la comparaison* (matching). Cependant, un système

⁷ "Content-based image retrieval (CBIR) is the set of techniques for retrieving semantically-relevant images from an image database based on automatically derived image features."

CBIR peut intégrer d'autres étapes ou modules, par exemple une étape de *prétraitement*, ou encore un module de *contrôle de pertinence* (Relevance Feedback)... etc.

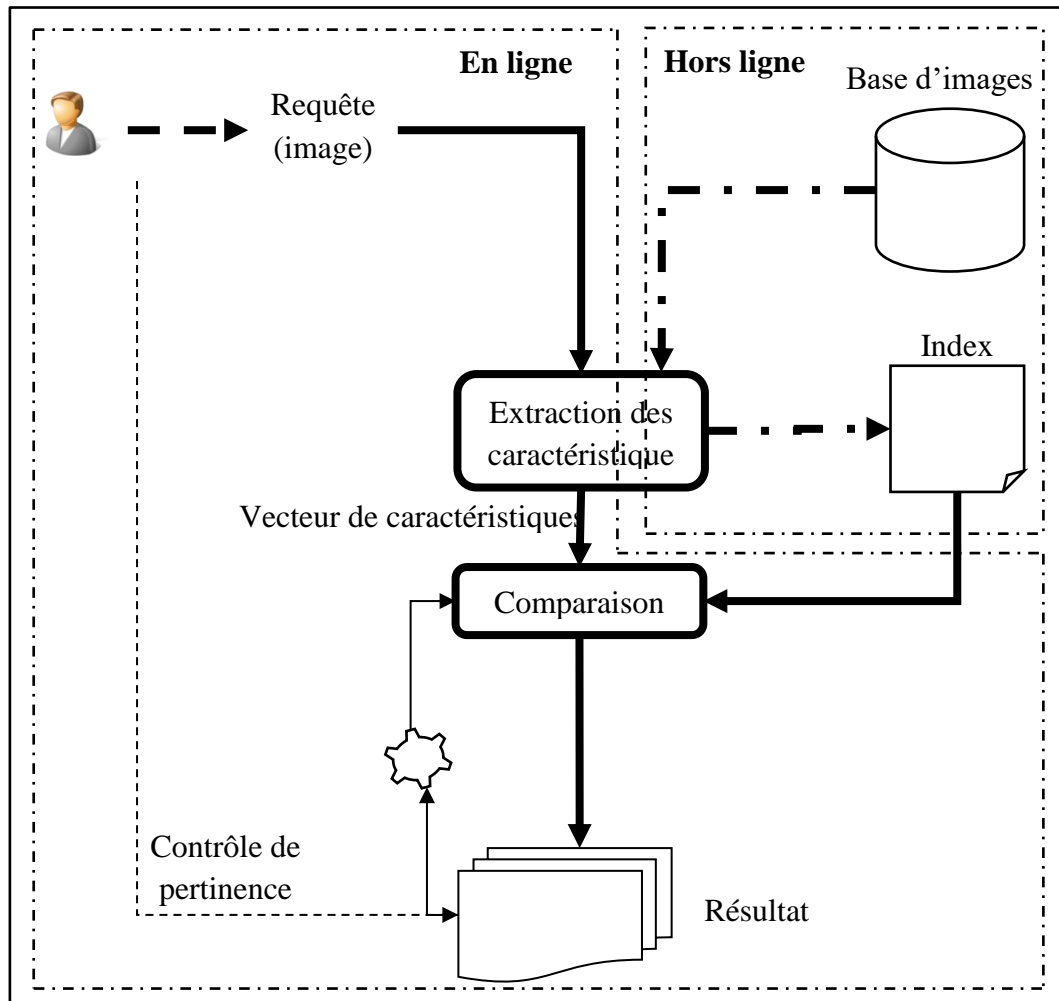


Figure II-1 Schéma illustratif d'un système CBIR.

II.1.1 Extraction des caractéristiques

Les méthodes utilisées dans cette étape doivent être suffisamment performantes et aussi efficaces afin qu'elles puissent extraire des caractéristiques *discriminantes* de tailles raisonnables à partir des attributs considérés. Ces nécessités sont attendues puisque toutes les étapes subséquentes reposeront sur ces caractéristiques. Le résultat de cette étape est une représentation numérique de l'image brute, généralement sous forme de vecteur, appelée *vecteur de caractéristiques* ou encore *signature* (en anglais, *feature vector*).

L'avantage de la signature par rapport à la représentation en pixels est la compression significative de la représentation de l'image. Cependant, la raison la plus importante est de gagner de meilleure corrélation entre la représentation de l'image et la sémantique ([Wang et al., 2001](#)).

Le terme *feature* réfère à une caractéristique distinctive des données qui signifie quelque chose pour quelqu'un⁸ ([MPEG Requirements Group, 1999](#)). En réalité, le terme '*descriptor*'

⁸ "A Feature is a distinctive characteristic of the data which signifies something to somebody"

(descripteur) est aussi largement utilisé dans la littérature. Ce terme a été défini également dans ([MPEG Requirements Group, 1999](#)) comme étant la représentation d'une caractéristique, qui définit la syntaxe et la sémantique de la représentation de caractéristique⁹. Dans ([Smith, 2002](#)), le terme '*descriptor*' est considéré comme étant une quantité numérique qui décrit une caractéristique (feature).

Les caractéristiques extraites peuvent être globales si elles sont extraites de l'image entière, ou locales si elles sont extraites des parties de l'image. En effet, la meilleure manière pour extraire les caractéristiques locales est de segmenter l'image (segmentation forte) afin de séparer les différents objets de l'image. Malheureusement, une méthode performante qui fonctionne sur toutes les images et sous différentes conditions n'existe pas. Une alternative est d'utiliser une segmentation faible (weak segmentation) qui segmente l'image en se basant sur un critère donné sans que les différentes régions soient nécessairement couvrantes l'objet entier ([Smeulders et al., 2000](#)).

II.1.2 Comparaison

Une fois, les caractéristiques sont extraites de l'image requête sous forme de vecteur de caractéristiques, ce dernier doit être comparé avec ceux des images indexées (les images de la base) afin de trouver les images similaires. À cet effet, une métrique de distance ou une mesure de similarité est utilisée. Les images de la base sont ensuite réordonnées par ordre croissant (ou décroissant) selon qu'une distance ou mesure de similarité est utilisée. Enfin, les tops k images sont retournées à l'utilisateur.

Le rôle de cette étape est également important et décisif vu que les caractéristiques extraites peuvent être de différents types. Plusieurs études comparatives ont montré clairement l'impact des mesures de comparaison sur les performances des systèmes ([Malik et Baharudin, 2013a](#); [Patil et Talbar, 2012](#); [Rubner et al., 2001](#)). En effet, les mesures utilisées dans cette étape varient de simples mesures à des mesures plus sophistiquées. [Rubner \(2000\)](#) distingue entre deux types de distances : *Bin-By-Bin Dissimilarity Measures*, et *Cross-Bin Dissimilarity Measures*. Les mesures de distance du premier type comparent chaque bin d'un histogramme avec son correspondant dans le deuxième histogramme (de même indice). D'où le premier inconvénient de ces mesures, vu que l'information *across bins* n'est pas utilisée. Le deuxième inconvénient, toujours selon ([Rubner, 2000](#)), est la sensibilité à la taille des bins. Parmi les mesures de distance basées sur ce principe, nous citons la distance *Euclidienne*, *Manhattan*, *Canberra*, et aussi D_1 ... etc. D'autres mesures ont été proposées spécifiquement pour manipuler des histogrammes, telles que *Histogram Intersection* ([Swain et Ballard, 1991](#)). D'autres mesures issues de la théorie de l'information sont également exploitées, telle *Kullback-Leibler*. Ces distances sont définies, respectivement, comme suit (H et K étant deux vecteurs de mesure de même taille, notamment des histogrammes):

⁹ "A Descriptor (*D*) is a representation of a Feature. A Descriptor defines the syntax and the semantics of the Feature representation."

$$D_{Euc}(H, K) = \sqrt{\sum_i (H[i] - K[i])^2} \quad (\text{II.1})$$

$$D_{man}(H, K) = \sum_i |H[i] - K[i]| \quad (\text{II.2})$$

$$D_{can}(H, K) = \sum_i \frac{|H[i] - K[i]|}{|H[i]| + |K[i]|} \quad (\text{II.3})$$

$$D_1(H, K) = \sum_i \left| \frac{H[i] - K[i]}{1 + H[i] + K[i]} \right| \quad (\text{II.4})$$

$$D_{HI}(H, K) = 1 - \frac{\sum_i \min(H[i], K[i])}{\sum_i K[i]} \quad (\text{II.5})$$

$$D_{KL}(H, K) = \sum_i H[i] \log \frac{H[i]}{K[i]} \quad (\text{II.6})$$

En revanche, les mesures de la deuxième catégorie distinguée par [Rubner \(2000\)](#), *Cross-Bin Dissimilarity Measures*, comparent également les bins non-correspondants. Parmi les mesures de ce type, nous citons la *distance quadratique*. Cette dernière a été exploitée dans le contexte de recherche d'images par le contenu par [Hafner et al. \(1995\)](#). Elle est définie comme suit :

$$D_{Quad}(H, K) = \sqrt{(\mathbf{H} - \mathbf{K})^T \mathbf{A} (\mathbf{H} - \mathbf{K})} \quad (\text{II.7})$$

Où \mathbf{H} et \mathbf{K} sont des vecteurs listant les entrées de H et K . $\mathbf{A} = a_{ij}$ est une matrice de similarité, où a_{ij} dénote la similarité entre le bin i et j .

L'inconvénient commun entre les mesures ci-dessus, est qu'elles ne comparent que des histogrammes de même taille (même des histogrammes irréguliers). Or, il est très difficile de calculer des histogrammes de taille fixe pour toutes les images tout en gardant un meilleur compromis entre l'expressivité et l'efficacité. Pour cela, [Rubner \(2000\)](#) a proposé une mesure très intéressante pour la recherche d'image, appelée *Earth Mover's Distance* (EMD), qui manipule *des signatures* au lieu des histogrammes. L'EMD s'est basée sur la solution du problème de transportation, qui consiste à trouver le coût minimal qui doit être payé pour transformer une distribution à une autre.

Soit deux signatures P et Q , tel que $P = \{(\mathbf{p}_1, w_{p_1}), \dots, (\mathbf{p}_m, w_{p_m})\}$ et $Q = \{(\mathbf{q}_1, w_{q_1}), \dots, (\mathbf{q}_n, w_{q_n})\}$ décrivant m et n clusters respectivement, où \mathbf{p}_i et \mathbf{q}_i sont les représentants des clusters. w_{p_i} et w_{q_i} sont les poids des clusters correspondants. Soit également $\mathbf{D} = d_{ij}$ la matrice des distances

entre les caractéristiques respectives (*ground distance*¹⁰), où d_{ij} représente le *ground distance* entre les clusters \mathbf{p}_i et \mathbf{q}_j .

Donc, pour calculer la distance entre les deux signatures, on calcule en premier, le flow $\mathbf{F} = [f_{ij}]$, qui minimise le coût spécifié par l'équation suivante :

$$Work(P, Q, \mathbf{F}) = \sum_{i=1}^m \sum_{j=1}^n d_{ij} f_{ij} \quad (\text{II.8})$$

Sujet des contraintes suivantes :

$$f_{ij} \geq 0 \quad 1 \leq i \leq m, 1 \leq j \leq n \quad (\text{II.9})$$

$$\sum_{j=1}^n f_{ij} \leq w_{\mathbf{p}_i} \quad 1 \leq i \leq m \quad (\text{II.10})$$

$$\sum_{i=1}^m f_{ij} \leq w_{\mathbf{q}_j} \quad 1 \leq j \leq n \quad (\text{II.11})$$

$$\sum_{i=1}^m \sum_{j=1}^n f_{ij} = \min \left(\sum_{i=1}^m w_{\mathbf{p}_i}, \sum_{j=1}^n w_{\mathbf{q}_j} \right) \quad (\text{II.12})$$

La distance EMD est alors la normalisation du *Work* résultant par le flow total :

$$EMD(P, Q) = \frac{\sum_{i=1}^m \sum_{j=1}^n d_{ij} f_{ij}}{\sum_{i=1}^m \sum_{j=1}^n f_{ij}} \quad (\text{II.13})$$

Pour des systèmes plus performants, d'autres auteurs ont choisi la combinaison de mesures de similarité ([Arevalillo-Herráez et al., 2008](#)). Alors que d'autres auteurs sont partis au-delà de la combinaison, en proposant la sélection des distances comme il a été fait par [Mosbah et Boucheham \(2016\)](#).

Dans ce contexte, nous avons également proposé une contribution originale, qui consiste à apporter et adapter une mesure originellement proposée pour le biclustering des expressions génétiques au domaine de la recherche d'images par le contenu ([Bougueroua et Boucheham, 2016b](#)).

II.2 Evaluation

Avoir une méthode d'évaluation ou une mesure qui reflète les capacités des systèmes est une nécessité absolue pour n'importe quel domaine. Les mesures d'évaluation permettent la

¹⁰ "A *ground distance* is defined as a distance between individual feature vectors in the underlying feature space." ([Rubner et al., 2001](#))

quantification des performances de différents systèmes et méthodes. Ceci permettra de faire des comparaisons objectives entre ces derniers. Dans le contexte de la recherche d'images par le contenu, plusieurs mesures ont été considérées dans la littérature. Parmi les mesures les plus utilisées en CBIR, nous recensons la *précision* et le *rappel* (*precision-recall*). Ces deux mesures ont été empruntées de la recherche de l'information (IR : *Information Retrieval*). La précision est définie comme étant la fraction des images pertinentes récupérées par le système par rapport au nombre total des images récupérées. D'un autre côté, le rappel (recall) est la fraction des images pertinentes récupérées par le système par rapport au nombre total des images pertinentes dans la base.

Formellement, la précision et le rappel peuvent être présentés comme suit ([Smeulders et al., 2000](#)) : étant donné un ensemble de données D et une requête q , alors l'ensemble de données D peut être divisé en deux ensembles par rapport à la requête q : un ensemble des images pertinentes $R(q)$ et son complément $\bar{R}(q)$. En supposant que l'ensemble des images récupérées par le système pour la requête q est $A(q)$. Alors, les équations de la précision (p) et du rappel (r) peuvent être données comme suit :

$$p = \frac{|A(q) \cap R(q)|}{|A(q)|} \quad (\text{II.14})$$

$$r = \frac{|A(q) \cap R(q)|}{|R(q)|} \quad (\text{II.15})$$

En fait, la précision souffre d'un inconvénient dû à la considération de l'ensemble des images récupérées et elle n'est pas affectée par l'ordre des images pertinentes dans cet ensemble ([Datta et al., 2008](#)). La précision pondérée (WP : *Weighted Precision*) ([Wang et al., 2001](#)) est l'une des solutions à ce problème. Cette mesure prend en compte le rang des images pertinentes en calculant la précision moyenne dans un ensemble de K -images retournées :

$$WP = \frac{1}{K} \sum_{k=1}^K \frac{|A(q)_k \cap R(q)|}{|A(q)_k|} \quad (\text{II.16})$$

Où $A(q)_k$, représente l'ensemble des k premières images récupérées par le système pour la requête q .

D'autres mesures existent également telles que : *Average Normalized Modified Retrieval Rank* (ANMRR), dont la description peut être trouvée dans ([Manjunath et al., 2001](#)). En plus de la détermination de l'existence des résultats pertinents, ANMRR prend en compte, également, le rang des résultats ([Shrivastava et Tyagi, 2014](#)).

II.3 Benchmarking

Les bases de données (data sets) ou encore les bancs d'essai sont indispensables dans le processus de développement de méthodes, afin d'établir des comparaisons entre les méthodes

de différentes générations d'une manière unifiée et équitable. La littérature dispose de plusieurs bases d'images de différents types, obtenues par différents moyens d'acquisition sous différentes conditions. Elles (les bases d'images) sont de plusieurs types : des bases des images de scènes, d'objets, de textures... etc.

La base '*Corel dataset*' est sans doute l'une des bases les plus populaires dans le domaine de CBIR, grâce à la diversité des images qu'elle contient (véhicule, animaux, fleurs...), ce qui la rend une base très représentative des situations réelles. En effet, un sous ensemble de 1000 images a été créé de cette base, connu par la base de Wang ou encore '*Corel 1K dataset*' (Figure II-2).

Quant aux bases des images de textures, plusieurs bases existent également. Parmi les bases de textures les plus utilisées dans le domaine de CBIR, nous citons la base *Brodatz*, *Outex*, *VisTex*. Un nombre important de ces bases ont été considérées dans les deux intéressantes études compréhensives suivantes ([Bianconi et Fernández, 2014](#); [Hossain et Serikawa, 2013](#)). Par exemple, dans ([Bianconi et Fernández, 2014](#)) les auteurs ont différencié entre trois types de bases de textures : des bases dont les images de textures sont issues de différents domaines biomédicaux ; des bases de textures naturelles, généralement composées de classes de textures hétérogènes dominées principalement par des scènes générales en plein air comme des bâtiments, plantes... etc. ; et enfin, des bases de matériaux réels tels : le bois, le granite, la céramique... etc.

Nous rapportons dans le Tableau II-1, ci-dessous, quelques informations sur quelques bases d'images de référence utilisées dans la littérature. Des exemples d'images contenues dans ces bases sont illustrés dans les figures ci-dessous (Figure II-2 -Figure II-7).

Tableau II-1 Quelques bases de données de référence utilisées dans la littérature.

Nom de la base	Taille d'images	Nbr Images/classe (nbr classes)	Niveau de gris ou Couleur	Lien
Corel 1k (Wang database)	256 × 384 (384 × 256)	100 (10)	Couleur	http://wang.ist.psu.edu/docs/related/
MIRFLICKR -1M	N/A	1 000 000	Couleur	http://press.liacs.nl/mirflickr/mirdownload.html
Brodatz	640×640	1 (112)	Niveau de gris	http://multibandtexture.recherche.usherbrooke.ca/original_brodatz.html
OUTex (TC_00000)	128×128	20 (24)	Niveau de gris	http://www.outex.oulu.fi/index.php?page=test_suites
VisTex (Reference Textures)	128×128, 512×512	167	Couleur	http://vismod.media.mit.edu/vismod/imagery/VisionTexture/vistex.html
Kylberg	576×576	160 (28)	Niveau de gris	http://www.cb.uu.se/~gustaf/texture/



Figure II-2 Quelques images de la base Corel 1k (Wang).



Figure II-3 Quelques images de la base MIRFLICKR

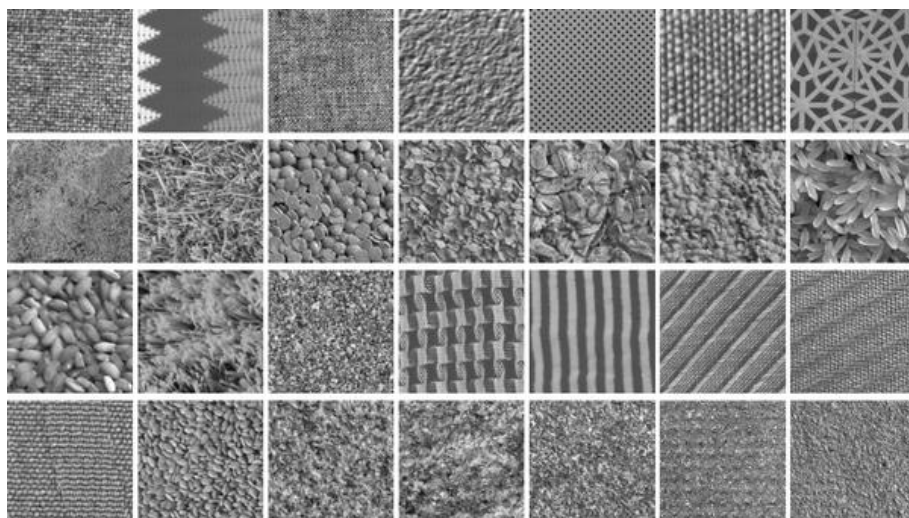


Figure II-4 Les différentes textures de la base Kylberg.

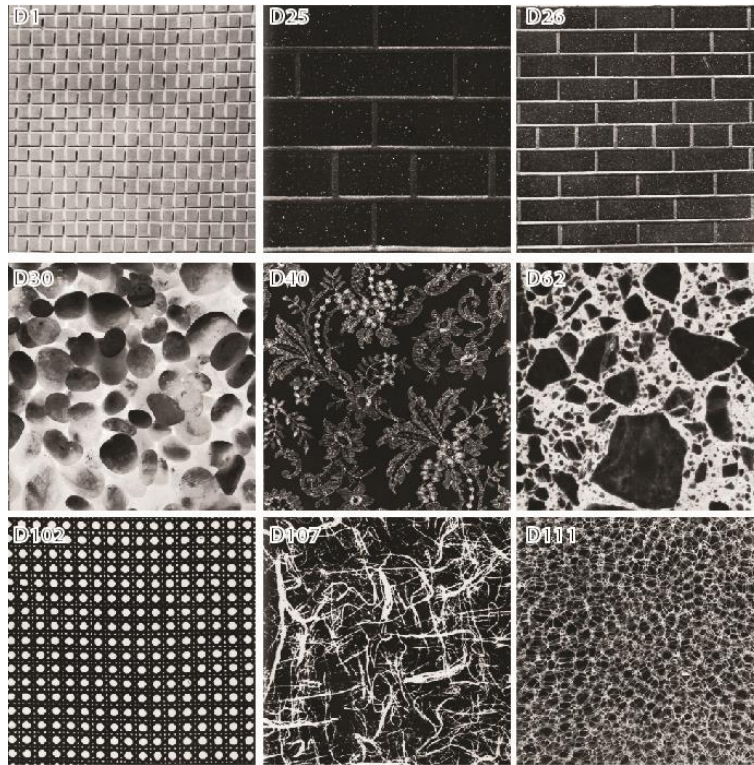


Figure II-5 Quelques textures de la base Brodatz.



Figure II-6 Quelques textures de la base VisTex.

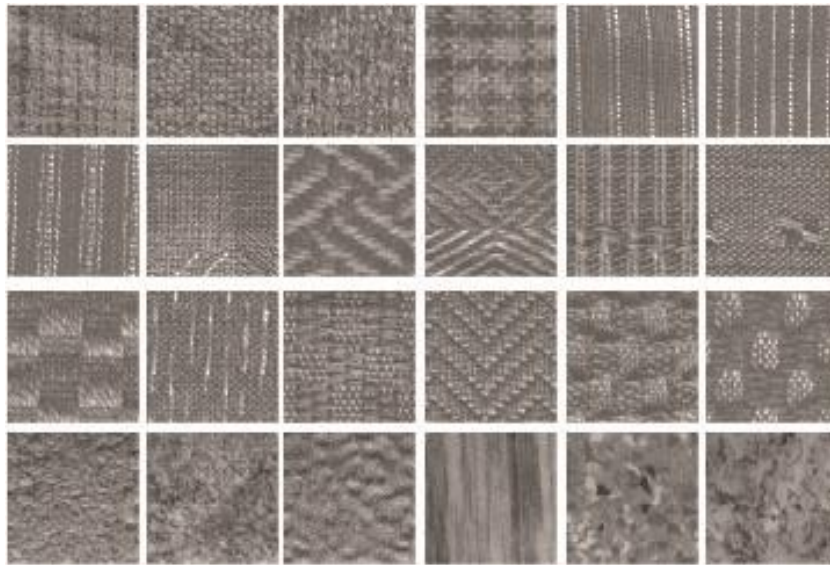


Figure II-7 Les textures de la base Outex.

II.4 Conclusion

Dans ce chapitre, nous avons présenté et discuté les étapes principales des systèmes CBIR, à savoir, l'étape d'extraction des caractéristiques et l'étape de comparaison. Nous avons aussi présenté quelques mesures permettant l'évaluation des performances de ces systèmes. Nous avons également avancé quelques bases de données qui ont été considérées dans la littérature pour l'évaluation des systèmes CBIR.

Le chapitre suivant est consacré à l'étape d'extraction des caractéristiques.

Chapitre III. *Descripteurs visuels de l'image*

« Les couleurs sont des actions de la lumière... »
Goethe

Sommaire

III.1	LA COULEUR.....	18
III.1.1	<i>Les espaces de couleurs</i>	19
III.1.1.1	Espace RGB	20
III.1.1.2	Espace TSV(HSV)	20
III.1.1.3	Espace CIE XYZ	21
III.1.1.4	Espace CIE L*a*b*	22
III.1.1.5	Espace CIE L*u*v.....	22
III.1.2	<i>Les descripteurs de couleurs</i>	23
III.1.2.1	Descripteurs globaux	23
III.1.2.2	Descripteurs locaux.....	23
III.2	LA TEXTURE	26
III.2.1	<i>Discussions</i>	27
III.2.1.1	De quoi résulte la texture ?.....	27
III.2.1.2	Texture stationnaire et non-stationnaire.....	28
III.2.1.3	Quelques termes décrivant les primitives de texture	28
III.2.1.4	Quelques définitions de la texture.....	30
III.2.2	<i>Applications de la texture</i>	30
III.3	LA FORME.....	32
III.3.1	<i>Descripteurs basés-Contour</i>	34
III.3.1.1	Code de Freeman (Chain coded string).....	34
III.3.1.2	Descripteur de Fourier	35
III.3.2	<i>Descripteurs basés-Région</i>	36
III.3.2.1	Les moments invariants	36
III.3.2.2	Les moments de Zernike.....	36
III.4	LES POINTS D'INTERET	37
III.4.1	<i>Détecteurs</i>	38
III.4.1.1	Détecteur de Harris.....	38
III.4.1.2	Détecteur Hessien.....	39
III.4.1.3	Laplacien de Gaussien.....	39
III.4.1.4	Différence de Gaussien	40
III.4.2	<i>Descripteurs</i>	40
III.4.2.1	SIFT	40
III.4.2.2	SURF.....	40
III.4.2.3	GLOH.....	41
III.4.2.4	LPDF	41
III.5	COMBINAISON DES DESCRIPTEURS DE TEXTURE ET DE COULEUR	42
III.5.1	<i>Les méthodes parallèles</i>	43
III.5.2	<i>Les méthodes séquentielles</i>	43
III.5.3	<i>Les méthodes intégratives</i>	44

III.5.3.1	Monocanal (single-channel).....	44
III.5.3.2	Multicanal (multi-channel)	45
III.6	CONCLUSION.....	45

Le but de ce chapitre est de présenter quelques descripteurs de différents attributs de l'image (*couleur, texture, forme, point d'intérêt*). Le premier attribut discuté dans ce chapitre est la couleur. Pour cela, quelques espaces de couleur seront présentés, puis nous passerons immédiatement aux différents descripteurs de couleur. En ce qui concerne le deuxième attribut 'texture', nous nous limiterons dans le présent chapitre à sa présentation et à sa discussion seulement, ainsi qu'à la mise en évidence des différents domaines d'application, alors que ses descripteurs sont reportés au prochain chapitre. Dans ce chapitre, nous discuterons également la forme, ainsi que les points d'intérêt. Pour terminer à la fin par les différentes approches de combinaison de la couleur et de la texture.

III.1 La couleur

La couleur est l'un des attributs les plus importants dans la vie quotidienne de l'être humain, dû à son symbolisme, ses révélations sur la personnalité et ses effets psychologiques, ainsi qu'à ses capacités discriminatives. En fait, le traitement de l'information commence au niveau de l'œil dès que la lumière réfléchiée y pénètre en traversant ses différentes parties : la cornée, l'humeur aqueuse, le cristallin, l'humeur vitrée pour arriver à la rétine sur laquelle l'image est formée (Figure III-1¹¹). La rétine est composée de deux types de photorécepteurs ; les cônes qui sont responsables de la vision diurne et colorée, et les bâtonnets qui sont plus nombreux, responsables de vision crépusculaire et nocturne en noir et blanc¹². Selon le type de photorécepteurs actifs, trois types de vision sont distingués : la vision *photopique*, *scotopique*, et *mésopique*, lorsque seulement les cônes, les bâtonnets, ou les deux types sont actifs, respectivement (Fairchild, 2005, p. 8). Également, trois types de cônes sont différenciés : *L*, *M* et *S*, selon les longueurs d'ondes auxquelles ils sont sensibles. En plus de cette différence, les cônes *S* sont relativement peu peuplés à travers la rétine et totalement absents dans la zone la plus centrale de la fovéa. Ainsi il y a beaucoup plus de cônes *L* et *M* que de cônes *S*, avec des populations relatives approximatives des cônes *L* : *M* : *S* estimées à 40 : 20 : 1.

¹¹ https://fr.wikipedia.org/wiki/%C5%92il_humain (visitée le 19/12/2016 à 19 : 15)

¹² <https://fr.wikipedia.org/wiki/Rétine>

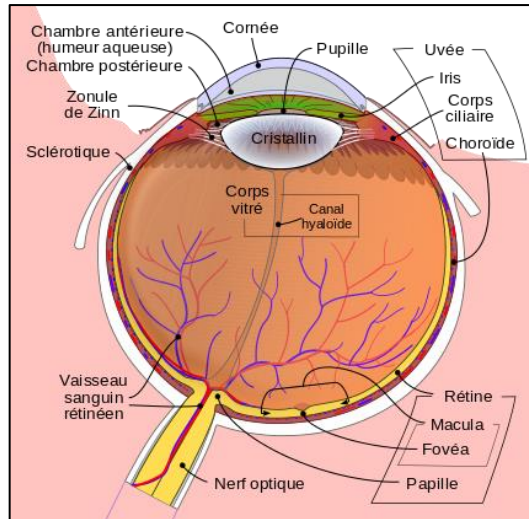


Figure III-1 Diagramme schématisique de l'œil humain.
(Source : Wikipédia)

Le rôle de la couleur est également important dans le royaume de la vision par ordinateur. En effet, la couleur est parmi les premiers attributs utilisés pour l'indexation et aussi la recherche d'images par le contenu, par exemple : ([Stricker et Orengo, 1995](#); [Swain et Ballard, 1991](#)). Donc, afin d'utiliser la couleur, ses caractéristiques doivent être extraites. Pour cela, plusieurs méthodes ont été proposées durant les années, essayant de capturer le plus ample d'informations pour que les différents systèmes soient plus performants.

Dans la sous-section (III.1.2), nous survolerons plusieurs méthodes de l'état de l'art. Néanmoins, vu l'impact de l'espace de couleur sur les performances des méthodes, nous en présenterons, en premier, quelques-uns dans la sous-section qui suit.

III.1.1 Les espaces de couleurs

Un espace de couleur est un espace géométrique tridimensionnel, dont les axes sont définis d'une manière appropriée afin que les symboles pour toutes les perceptions de couleurs possibles de l'homme ou d'autres animaux s'y intègrent dans un ordre correspondant à l'ordre psychologique. Dans cet espace, chaque perception de couleur est représentée par un point ([Kuehni, 2003](#)). Ainsi, un espace de couleur peut être défini comme une méthode par laquelle une couleur est spécifiée, créée et visualisée ([Ford et Roberts, 1998\(b\)](#)). Ces trois axes peuvent être vus comme une imitation des trois types de récepteurs (cônes) qui se trouvent dans la rétine de l'œil humain, chaque partie du spectre de lumière (visuel) stimule un type spécifique de ces récepteurs. En effet, il existe plusieurs espaces de couleur développés au cours des années et adaptés pour un objectif ou un autre. Par exemple, certains espaces sont plus adaptés aux programmes de dessin et de l'art (machine), alors que d'autres sont convenables pour la perception. Il est à noter que nous nous sommes basés pour le développement des prochaines sous-sections, sur les références suivantes ([Ebner, 2007](#); [Fairchild, 2005](#); [Westland et Ripamonti, 2004](#)).

III.1.1.1 Espace RGB

Cet espace de couleur est très utilisé en CBIR. Il est dépendant du périphérique (matériel), non linéaire par rapport à la perception humaine, ainsi la spécification des couleurs dans cet espace est semi-intuitive (Ford et Roberts, 1998). Le gamut de couleur est représenté par un cube (Figure III-2), dont les valeurs des axes (rouge, vert, bleu) se varient de 0 à 255 (ou de 0.0 à 1.0 pour le RGB normalisé). Le point de coordonnées (R, G, B)= (0, 0, 0) représente la couleur noire, alors le blanc est situé au corner opposé (255, 255, 255). Les six corners restants du cube sont étiquetés par : rouge, vert, bleu, jaune, magenta et cyan.

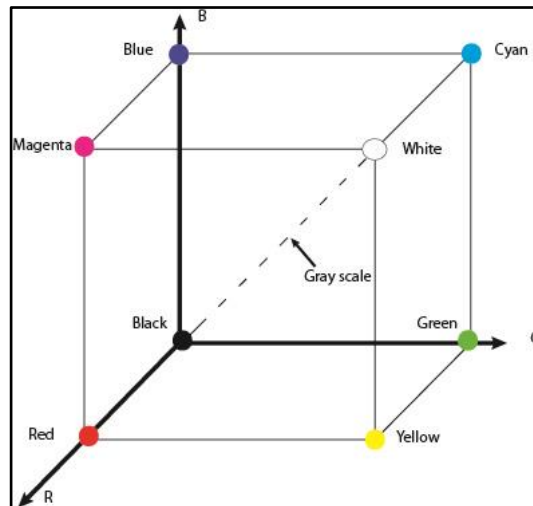


Figure III-2 L'espace de couleurs RGB.

III.1.1.2 Espace TSV(HSV)

L'espace TSV (Teinte-Saturation-Valeur) a été considéré dans beaucoup de travaux sur le CBIR (Jasmine et Kumar, 2014; ma, 2009; Sural et al., 2002; Wang et al., 2014; Youngeun et al., 2010). Il se ressemble au RGB quant à la dépendance au matériel et la non-linéarité par rapport à la perception humaine. En revanche, il a l'avantage lors la spécification des couleurs qui est une tâche très intuitive, ce qui l'a rendu très utile dans les programmes de dessin et de peinture (Ford et Roberts, 1998). Une couleur est déterminée en sélectionnant la teinte puis elle est ajustée en réglant sa saturation et son intensité. Cet espace de couleur prend la forme d'un hexcone (Figure III-3¹³). La transformation à cet espace à partir le RGB normalisé est faite comme suit (Ebner, 2007, p. 98) :

Soit :

$$max = \max\{r, g, b\} \tag{III.1}$$

$$min = \min\{r, g, b\} \tag{III.2}$$

Alors, les trois composants H, S et V vont prendre des valeurs dans l'intervalle [0, 1], ils sont calculés comme suit :

¹³ https://en.wikipedia.org/wiki/HSL_and_HSV

$$V = \max \quad (\text{III.3})$$

$$S = \frac{\max - \min}{\max} \quad (\text{III.4})$$

$$H = \left\{ \begin{array}{l} \frac{1}{6} \frac{g - b}{\max - \min} \quad \text{if } r = \max \\ \frac{1}{6} \left(2 + \frac{b - r}{\max - \min} \right) \quad \text{if } g = \max \\ \frac{1}{6} \left(4 + \frac{r - g}{\max - \min} \right) \quad \text{if } b = \max \end{array} \right\} \quad (\text{III.5})$$

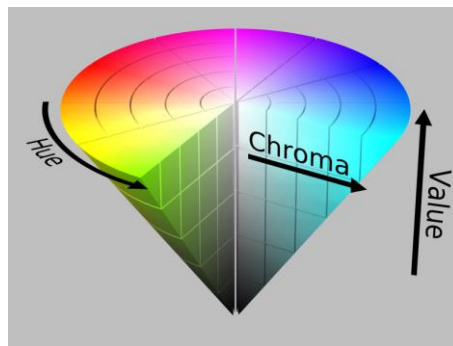


Figure III-3 L'espace de couleurs HSV.
(Source Wikipédia)

III.1.1.3 Espace CIE XYZ

Cet espace est défini par la *Commission Internationale d'Eclairage* (CIE). XYZ est un espace de couleur additif. Il décrit une couleur en termes de quantités nécessaires des trois primaires pour créer la lumière d'une telle couleur (Kerr, 2010). Il sert généralement à la transformation entre les autres espaces de couleur. C.-à-d., il joue le rôle d'un intermédiaire entre l'espace source et l'espace cible. Par exemple, pour transformer une couleur décrite dans l'espace RGB par le vecteur $[R, G, B]^T$ (les valeurs RGB affichées), nous devons utiliser une matrice 3×3 de transformation :

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = M * \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (\text{III.6})$$

Où :

$$M = \begin{bmatrix} X_r & X_g & X_b \\ Y_r & Y_g & Y_b \\ Z_r & Z_g & Z_b \end{bmatrix} \quad (\text{III.7})$$

M est une matrice de transformation de l'espace RGB vers l'espace XYZ. Ses coefficients dépendent du blanc de référence ([Ebner, 2007, p. 79](#)).

III.1.1.4 Espace CIE $L^*a^*b^*$

$L^*a^*b^*$ ou (CIELab) est un espace de couleur non linéaire (la transformation), perceptuellement uniforme, proposé par la Commission Internationale d'Eclairage en 1976. C'est un espace indépendant du matériel. Il se compose de 3 axes : L^* , a^* et b^* . Le signal achromatique est représenté par l'axe L^* alors que les signaux chromatiques sont représentés par les axes a^* et b^* qui varient de rouge à vert et de jaune à bleu, respectivement. La transformation à cet espace est faite à partir de l'espace XYZ en utilisant les équations (III.8-III.10) ([Westland et Ripamonti, 2004, p. 50](#)).

$$L^* = \begin{cases} 116 \left(\frac{Y}{Y_n} \right)^{\frac{1}{3}} - 16 & \text{if } \frac{Y}{Y_n} > 0.008856 \\ 903.3 \left(\frac{Y}{Y_n} \right) & \text{if } \frac{Y}{Y_n} \leq 0.008856 \end{cases} \quad (\text{III.8})$$

$$a^* = 500 \left(f \left(\frac{X}{X_n} \right) - f \left(\frac{Y}{Y_n} \right) \right) \quad (\text{III.9})$$

$$b^* = 200 \left(f \left(\frac{Y}{Y_n} \right) - f \left(\frac{Z}{Z_n} \right) \right) \quad (\text{III.10})$$

Où X_n , Y_n et Z_n représentent les tri-stimulus de la couleur achromatique 'Blanche', la fonction f est définie comme suit :

$$f(X) = \begin{cases} X^{\frac{1}{3}} & \text{if } X > 0.008856 \\ 7.787X + \frac{16}{116} & \text{if } X \leq 0.008856 \end{cases} \quad (\text{III.11})$$

III.1.1.5 Espace CIE $L^*u^*v^*$

$L^*u^*v^*$ (ou CIELuv) est un autre espace de couleur non linéaire, perceptuellement uniforme, défini (en 1976) par la Commission Internationale d'Eclairage. La transformation à cet espace est faite à partir de l'espace XYZ. Le composant L^* est calculé de la même manière que dans l'espace CIELab par l'équation ci-dessus (Eq.III.8). Quant aux deux composants u^* et v^* , sont calculés par les deux équations suivantes ([Westland et Ripamonti, 2004, p. 51](#)) :

$$u^* = 13L^*(u' - u'_n) \quad (\text{III.12})$$

$$v^* = 13L^*(v' - v'_n) \quad (\text{III.13})$$

Où v'_n et u'_n représentent les stimuli de la couleur d'un objet blanc. u' et v' sont calculé comme suit :

$$u' = \frac{4x}{x + 15y + 3z} \quad (\text{III.14})$$

$$v' = \frac{9x}{x + 15y + 3z} \quad (\text{III.15})$$

III.1.2 Les descripteurs de couleurs

Dans la présente section, nous survolons plusieurs descripteurs de couleur parmi les plus connus. Étant donné qu'un descripteur peut être global ou local selon qu'il est basé sur la description de l'image globalement ou localement, nous allons aborder les différents descripteurs selon cette catégorisation.

III.1.2.1 Descripteurs globaux

[Swain et Ballard \(1991\)](#) ont proposé le premier et le fameux Histogramme de couleur. Ce descripteur caractérise la distribution des couleurs de l'image, ce qui fait de lui une représentation grossière de cette dernière. Les histogrammes ont l'avantage d'être invariants aux translations et rotations autour de l'axe de vue, et ne varient que légèrement lorsque des changements surviennent sur le point de vue, l'échelle, ou l'occlusion ([Swain et Ballard, 1991](#)). Afin de calculer la similarité entre les histogrammes de deux images, les auteurs ont proposé la technique d'*intersection d'histogrammes* (cf. Eq.II.5).

Afin d'établir les histogrammes, une étape de quantification est très recommandée pour gagner de l'espace mémoire. Toutefois, le choix d'un bon paramètre pour la quantification est problématique. En plus, malgré une quantification optimale, des effets non préférés peuvent apparaître ([Stricker et Orengo, 1995](#)). D'où une nouvelle version d'histogrammes, plus robuste au paramètre de quantification a été proposée par [Stricker et Orengo \(1995\)](#), qui consiste à sauvegarder les histogrammes cumulatifs des couleurs. Dans le même travail, les auteurs proposent une deuxième approche basée sur les caractéristiques dominantes de l'histogramme. Ainsi, ils calculent les trois premiers moments pour chaque canal de couleur.

Une autre approche pour la caractérisation globale de l'image consiste à prendre en compte les couleurs dominantes. Le *descripteur de couleur dominante* ([Manjunath et al., 2001](#)) (DCD, *Dominant Color Descriptor*) est l'un des descripteurs les plus efficaces de MPEG-7. Afin d'aboutir à cette représentation, les couleurs de l'image (ou de la région d'intérêt) sont regroupées en clusters, ce qui en résulte un nombre réduit de couleurs. Puis, pour chaque couleur résultante, un pourcentage est calculé.

III.1.2.2 Descripteurs locaux

L'inconvénient majeur des descripteurs globaux est qu'ils n'intègrent pas de l'information spatiale. La plus simple manière d'en capturer est de diviser l'image en blocs, puis établir un histogramme pour chaque bloc. En effet, [Stricker et Dimai \(1996\)](#) ont divisé l'image en cinq régions floues, partiellement chevauchantes, chacune dans un coin est la cinquième au milieu. Ensuite, les auteurs établissent les trois premiers moments pour chaque région et canal de couleur. La même stratégie de partitionnement est suivie dans ([Appas et al., 1998](#)), cependant le troisième moment est calculé autour du deuxième sous prétexte que le deuxième moment est

moins sensible aux changements qui apparaissent à cause de petits changements de luminance. Ensuite, ils combinent les trois moments de chaque canal de couleur et région en une seule caractéristique. Récemment, [Bougueroua et Boucheham \(2015\)](#) ont appliqué leur approche d'intégration graduelle de l'information locale avec succès sur la couleur, d'où son nom : GLI-Color. La méthode sera vue en détail dans le Chapitre V. (cf. sec.V.1.3).

Une méthode basée sur le raffinement d'histogramme a été proposée par [Pass et Zabih \(1996\)](#), appelée *vecteur de couleur cohérent* (CCV, *Color Coherent Vector*). Cette méthode distingue entre les histogrammes des régions extraites ; *Cohérent* lorsque la région satisfait un nombre minimal de pixels (seuil), sinon *incohérent*. [Jing et al. \(1997\)](#) ont proposé une méthode promotrice dans l'état de l'art, baptisée *corrélogramme de couleur* (Color Correlogram). Un corrélogramme est un tableau, dans lequel la valeur de l'élément situé en (i, j) dans la $k^{\text{ème}}$ entrée représente la probabilité de trouver un pixel de couleur j à une distance k d'un autre de couleur i .

Un autre descripteur efficace et performant a été proposé par [Kasutani et Yamada \(2001\)](#), baptisé (CLD, *Color Layout Descriptor*). La construction de ce descripteur se passe à travers quatre étapes. En premier, l'image est partitionnée en blocs (8×8 blocks), puis une couleur dominante est sélectionnée pour chaque bloc. Dans la troisième étape, la transformation de chaque composant de l'espace de couleur (Y, Cb, Cr) en utilisant 8×8 DCT. Finalement, seulement quelques coefficients de basse fréquence sont extraits et quantifiés. Étant donné sa compacité et ses performances, ce descripteur a été également exploité pour la recherche de segments vidéo ([Kasutani et Yamada, 2001](#)). Le *descripteur de structure de couleur* (CSD, *Color Structure Descriptor*) ([Messing et al., 2001](#)), capture de l'information spatiale en comptant le nombre des éléments structurants $h_s(m)$ chevauchants qui peuvent être positionnés dans l'image, tel qu'ils contiennent au moins un pixel de couleur m , où s est la taille de l'élément structurant. Le CSD exploite l'espace de couleur *hue-min-max-difference* (HMMD), dans lequel les couleurs sont quantifiées non uniformément. Cet espace de couleur, a montré une grande performance dans *MPEG-7 core experiments* ([Manjunath et al., 2001](#)). D'une manière générale, il a été prouvé dans plusieurs études comparatives que le CSD est le descripteur le plus performant parmi les descripteurs du MPEG-7 ([Penatti et al., 2012](#)).

Via la combinaison du DCD et CSD, [Ka-Man et al. \(2007\)](#) ont créé un nouveau descripteur appelé *descripteur de structure de couleur dominante* (DCSD, *Dominant Color Structure Descriptor*). Par conséquent, le DCSD a hérité la compacité du DCD et la haute performance du CSD. Dans ([Stehling et al., 2003](#)), une méthode qui consiste à établir un *histogramme de cellule* (Cell Histogram) pour chaque couleur réellement présente dans l'image a été proposée. Cette méthode baptisée Histogramme de Cellule/Couleur (CCH, *Cell/Color Histogram*). CCH est reprise par [Salmi et Boucheham \(2014\)](#) et combinée avec le CCV pour créer leur nouvelle méthode appelée *Cell-CCV*. Cette dernière a montré de bonnes performances par rapport aux CCH et CCV.

Toutes les méthodes d'extraction des caractéristiques de la couleur discutées dans cette section sont récapitulées dans le tableau ci-dessous (Tableau III-1).

Tableau III-1 Tableau récapitulatif des descripteurs de couleur discutés.

Référence	Nom de la Méthode (ou une courte description)	Type de caractéristiques (Global/Local)	Application	Nom de la Base utilisée (taille)
Swain et Ballard, 1991	Histogramme de couleurs	Global	Indexation	N/A
Stricker et Orengo, 1995	Histogramme cumulatif de couleurs	Global	Recherche	N/A (3000)
	Caractéristiques dominantes de l'histogramme (moments)	Global	Recherche	N/A
Manjunath et al., 2001	<i>Dominant Color Descriptor</i> (DCD)	Global	Recherche	N/A (5000)
Stricker et Dimai, 1996	Division de l'image en blocs dont un est au milieu.	Local	Recherche	N/A (11650)
Appas et al., 1998	Division de l'image en blocs dont un est au milieu.	Local	Recherche	N/A (205)
Pass et Zabih, 1996	<i>Color Coherent Vector</i> (CCV)	Local	Recherche	N/A (14554)
Jing et al., 1997	<i>Color Correlogram</i>	Local	Recherche	N/A (14554)
Kasutani et Yamada, 2001	<i>Color Layout Descriptor</i> (CLD)	Local	Recherche (images/vidéos)	N/A (5466)
Messing et al., 2001	<i>Color Structure Descriptor</i> (CSD)	Local	Recherche	N/A (>5400)
Ka-Man et al., 2007	<i>Dominant Color Structure Descriptor</i> (DCSD)	Local	Recherche	N/A (5466)
Stehling et al., 2003	<i>Cell/Color Histogram</i> (CCH)	Local	Recherche	N/A (20000)
Salmi et Boucheham, 2014	Cell-CCV	Local	Recherche	Corel 1K
Bougueroua et Boucheham, 2015	GLI-Color	Local	Recherche	Corel 1K

III.2 La Texture

La texture joue un rôle primordial dans la vie quotidienne de l'être humain, dû, non seulement à son omniprésence, mais aussi à son pouvoir discriminatif. En effet, la texture est partout, elle est présente dans les matériaux, les surfaces d'objets, sur les peaux d'animaux, elle est également présente dans les paysages urbains (Figure III-4, également Figure II-4 - Figure II-7). En sus de son omniprésence, la texture a des capacités discriminatives. Par conséquent, grâce à la texture, les objets, les animaux, peuvent être distingués de l'environnement dans lequel ils se trouvent (Figure III-5).

Néanmoins, en dépit qu'il soit facile de percevoir la texture (visuellement, par le toucher, et même par la langue), elle reste encore sans définition formelle et précise consensuelle, bien que la recherche dans ce sujet ait commencé il y a plus de cinq décennies.

En effet, depuis les années 1960s l'*analyse de texture* a connu des recherches intensives. La texture a été exploitée pour la classification, la segmentation, la synthèse de la texture... etc. ainsi, un nombre important de méthodes permettant l'extraction de caractéristiques de cet attribut ont été proposées. Ces méthodes peuvent être classifiées en méthodes *statistiques*, *structurelles*, *basées sur les modèles* et *basées sur les transformations* (pour plus de détails cf. sec. IV.2).

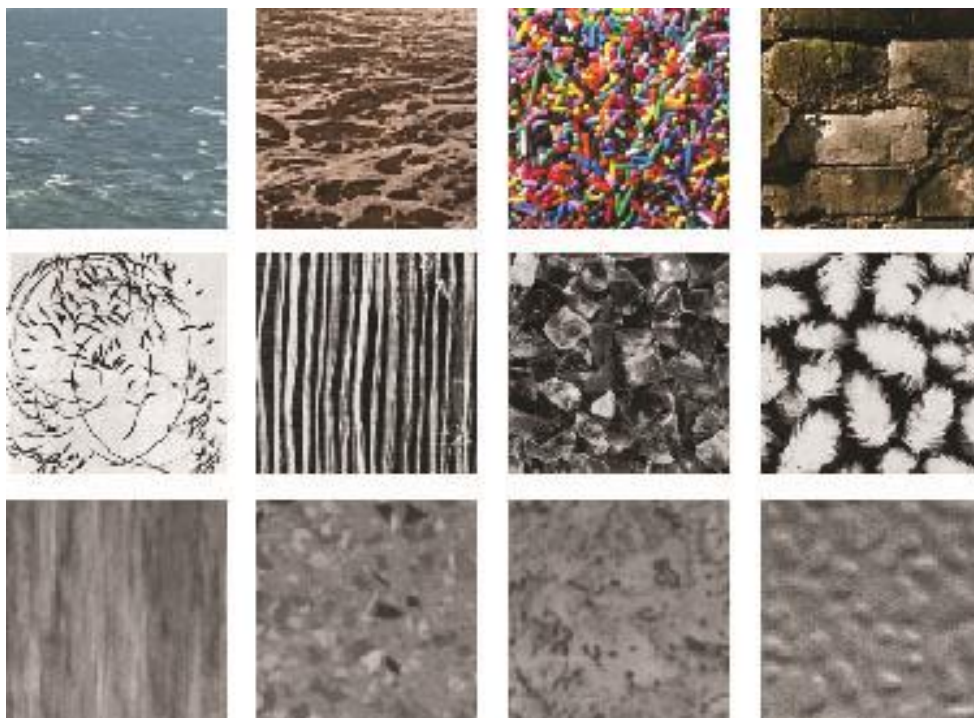


Figure III-4 Quelques textures de différentes bases de données.

Dans le reste de cette section, nous présentons et discutons la texture. Il est à noter que nous nous sommes basés, pour le développement des sous-sections suivantes, sur les références suivantes : ([Davies, 2012](#); [Petrou et García Sevilla, 2006a](#); [Tuceryan et Jain, 2011](#)).



Figure III-5 Rôle de la texture.

La distinction des deux animaux de leurs environnements respectifs semble très difficile voire impossible sans l'utilisation de la texture.

(Source : Wikipédia)

III.2.1 Discussions

III.2.1.1 De quoi résulte la texture ?

Il est reconnu que la texture résulte des variations de l'intensité ou de la couleur. Ces variations dues généralement à deux raisons, qui sont ([Petrou et García Sevilla, 2006a](#)) : (1) les variations de l'*albédo*, c.-à-d. les variations des couleurs de la surface.

L'*albédo* d'une surface est défini par [Petrou et García Sevilla \(2006a\)](#) comme étant *une fonction qui caractérise le matériau à partir duquel la surface est réalisée. Il donne la fraction de lumière incidente que ce matériau réfléchit à chaque longueur d'onde. Quand on dit qu'une surface est multicolore, on dit effectivement qu'elle a un albédo qui varie d'un endroit à l'autre*¹⁴.

Ou (2) les variations de *la forme (la topographie)*. C.-à-d. si une surface est rugueuse, des textures vont se manifester même si la surface est uniformément colorée.

¹⁴ "The albedo of a surface is a function that characterises the material from which the surface is made. It gives the fraction of incident light this material reflects at each wavelength. When we say that a surface is multicoloured, effectively we say that it has an albedo that varies from location to location."

D'un autre côté, les images de textures dépendent d'un ensemble de conditions lors de l'acquisition, telles que la distance, l'orientation, l'éclairage... etc. Ces paramètres ont un impact crucial sur l'apparence des textures (Figure III-6 (UMD HR dataset¹⁵)), et donc sur les performances des tâches subséquentes.



Figure III-6 Impact de la distance sur l'apparence de texture (UMD HR dataset).

III.2.1.2 Texture stationnaire et non-stationnaire

Du point de vue du nombre de textures contenues dans une image, on peut distinguer entre deux types de textures, des textures *stationnaires* et *non stationnaires* :

La texture stationnaire : c'est une image qui contient un seul type de texture ([Petrou et García Sevilla, 2006c](#)). Par exemple, les textures de la Figure III-4, ainsi celles des Figures (Figure II-4 à Figure II-7) sont des textures stationnaires. Dans ce type de textures, les méthodes globales sont suffisantes pour la description, du fait que les informations locales sont les mêmes dans toute l'image.

La texture non stationnaire : c'est une image qui contient plusieurs types de texture ([Petrou et García Sevilla, 2006b](#)). Les images rapportées dans la Figure III-5 illustrent des exemples de textures non stationnaires. Dans ce type de textures, les méthodes locales sont plus appropriées.

III.2.1.3 Quelques termes décrivant les primitives de texture

Lorsqu'on parle d'un pixel d'image, on peut facilement savoir sa couleur. En revanche, on ne peut pas parler de la texture au niveau d'un pixel parce qu'elle n'est définie que dans un voisinage (un ensemble de pixels).

¹⁵ http://www.cfar.umd.edu/~fer/High-resolution-data-base/hr_database.htm

En effet, les *primitives de textures*, qui sont appelées aussi *texels* (1) (*texture elements*) sont extraites à partir de la texture typiquement par les méthodes basées sur l'approche structurale. Une primitive de texture est définie dans (Wechsler, 1980), comme étant un ensemble de pixels, connecté de (ou presque) même niveau de gris. Une surface est texturée si elle contient un grand nombre (indénombrable) de texels, sinon on parle d'un ensemble d'objets (Davies, 2012). Les texels sont des régions de l'image qui peuvent être extraites par des procédures simples, telles que le seuillage (Stockman et Shapiro, 2001).

Les *textons* (2) sont les résultats des travaux pionniers de Julesz (1981). En effet, dans (Julesz, 1981), Julesz a corrigé sa conjoncture originale (Julesz, 1962) stipulant que les textures qui ont le même premier et deuxième ordre statistique (cf. IV.2.1.1) ne peuvent pas être discriminables¹⁶ (Figure III-7(a)). La nouvelle conjoncture de Julesz est comme suit : “*The human pre-attentive textural system cannot globally compute third- or higher- order statistics.*” C.-à-d. que le système visuel humain (SVH) préattentif n'utilise pas nécessairement les statistiques supérieures du deuxième ordre pour la discrimination entre des textures qui ont le même deuxième ordre, mais plutôt des statistiques du deuxième ordre de quelques caractéristiques visibles, appelées *les textons*, telles que les terminaisons, les fermetures... etc. (Figure III-7(b)). Il a également déclaré que le système visuel préattentif humain n'utilise que les statistiques du premier ordre de ces textons.

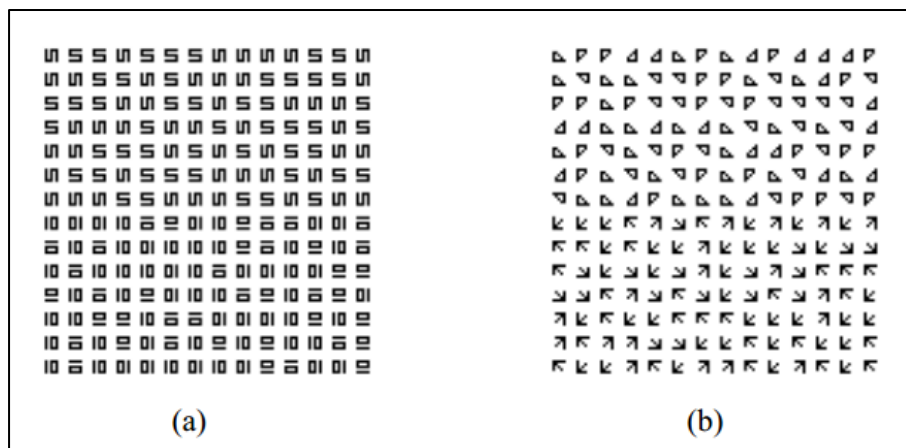


Figure III-7 Deux images de Textures, chacune est constituée de deux régions ont les mêmes statistiques du 2ème ordre.(a) : les deux régions ne sont pas discriminables par le SVH préattentif, (b) : les deux régions sont facilement discriminables. (Tuceryan et Jain, 2011)

Dans un autre contexte, Wang et He (1990) ont introduit une nouvelle méthode appelée *Texture spectrum* qui est basée sur le calcul de la distribution des occurrences de petits éléments appelés *Texture units* (3). Ces unités de texture sont établies par le seuillage des pixels voisins de chaque bloc 3×3 dans l'image par rapport au pixel central. Trois valeurs (0, 1, 2) sont possibles pour chacun des huit voisins ($3^8=6561$ *texture units* au total). La méthode (LBP) *Local Binary Patterns* (Ojala et al., 1996) (cf. sec. IV.1.3) est une version à deux niveaux de *texture units*. Par conséquent, ces unités de texture sont appelées *local binary patterns* (4).

¹⁶ “*iso-second-order textures are indistinguishable*”

[Xie et Mirmehdi \(2007\)](#) ont supposé lors du développement de leur approche que chaque image de texture est générée par la superposition de patches d'image, de différentes tailles. Les auteurs ont référé à ces patches et leurs correspondantes variances par *texems* (5) (*textural exemplars*).

III.2.1.4 Quelques définitions de la texture

Comme déjà mentionné, il n'existe pas un consensus sur la définition du concept de la texture. Dans ce cadre, nous rapportons quelques définitions proposées par différentes sources émanant de différentes expériences humaines. Pour cela, nous avançons, en premier, les définitions de la texture suivantes, telles qu'elles sont rapportées par quelques dictionnaires connus. Il est à noter que plusieurs définitions de la texture peuvent être trouvées dans le même dictionnaire :

- ❖ Selon Larousse¹⁷ : « *Mode d'entrecroisement des fils de tissage.* » ; « *État d'une étoffe ou d'un matériau qui est tissé.* » ; « *Constitution générale d'un matériau solide.* »...etc.
- ❖ Selon Le Grand Robert : « *Arrangement, disposition des éléments d'une matière* » ; « *Arrangement de la matière et des pores à l'intérieur d'un solide divisé (en grains) ou poreux* »...etc.
- ❖ Selon Oxford¹⁸ : « *The feel, appearance, or consistency of a surface or a substance.* »...etc.

Plusieurs chercheurs qui se sont intéressés à cet attribut, ont également proposé des définitions de la texture. Dans ce qui suit, nous avançons quelques-unes :

- ❖ [Haralick \(1979\)](#) : “*An image texture is described by the number and types of its primitives and the spatial organization or layout of its primitives.*”
- ❖ [Tamura et al. \(1978\)](#) : “*A strict definition for visual texture is difficult. However, from an analytical point of view, we may regard texture as what constitutes a macroscopic region. Its structure is simply attributed to the repetitive patterns in which elements or primitives are arranged according to a placement rule.*”
- ❖ [Mäenpää \(2003\)](#) : “*Texture can be broadly defined as the visual or tactile surface characteristics and appearance of something. Textures can consist of very small elements like sand, or of huge elements like stars in the Milky Way.*”
- ❖ [Petrou et García Sevilla \(2006a\)](#) : “*The texture is the variation of data at scales smaller than the scales of interest.*”
- ❖ [Paget \(2012\)](#) : “*... However it is generally accepted that texture is a pattern that can be characterised by its local spatial behaviour, and is statistically stationary.*”

III.2.2 Applications de la texture

Vu l'omniprésence de la texture, cette dernière est exploitée dans différents domaines. Entre autres, nous recensons :

¹⁷ <http://www.larousse.fr/dictionnaires/francais-monolingue>

¹⁸ <https://en.oxforddictionaries.com/>

- **La télédétection (remote sensing)** : la télédétection est parmi les premières applications de la texture. Dans ce type d'images, les régions homogènes doivent être identifiées. Ces régions peuvent représenter des parkings, des régions urbaines, des lacs, zones agricoles... etc (Figure III-8, de la base USC-SIPI¹⁹). Parmi les travaux qui se sont intéressés à ce type d'images : ([Akbarizadeh, 2012](#); [Haralick, 1971](#); [Sidorova, 2015](#); [Weszka et al., 1976](#)).



Figure III-8 Exemples d'images aériennes de la base USC-SIPI.

- **Le traitement de documents (document processing)** : dans ce domaine d'applications, la texture est utilisée pour la segmentation de documents (les pages des documents) en différentes régions telles que : *textes*, *figures* et *fond* (l'arrière-plan). *La segmentation* est une étape capitale pour les éventuelles étapes subséquentes telle que *la compréhension de documents* (document understanding). Entre autres, nous citons les travaux suivants : ([Jain et Zhong, 1996](#); [Payne et al., 1994](#); [Vil'kin et al., 2013](#)).
- **L'inspection** : Dans ce type d'applications, on tente de détecter et de localiser des régions inhomogènes dans les images texturées. Ces régions représentent des défauts de fabrication. En effet, beaucoup d'applications se sont intéressées à la détection des défauts de tissu. Parmi ces travaux : ([Ngan et al., 2005](#); [Tong et al., 2016](#)). D'autres se sont intéressés aux défauts sur la surface de bois tels que ([YongHua et Jin-Cong, 2015](#)).
- **L'analyse des images médicales** : les techniques d'analyse de l'image en général et les techniques d'analyse de texture en particulier jouent un grand rôle dans l'analyse des images issues de l'imagerie médicale. Par exemple, les caractéristiques de texture sont exploitées pour classifier les tissus normaux/anormaux de différents organes, tels que le cœur, les poumons, le foie... etc., issus des différentes techniques d'imagerie médicale telles que les rayons X, l'imagerie par résonance magnétique (IRM), l'échographie... etc. Par exemple, [Wong et al. \(2013\)](#) ont appliqué leur méthode sur la classification des images de la radiographie pulmonaire. [Zhang et al.](#)

¹⁹ <http://sipi.usc.edu/database/database.php>

(2015), ont utilisé les caractéristiques de la texture pour la classification du degré de la fibrose hépatique. Ainsi, pour la détection de masse mammographique (Oliver *et al.*, 2007)... etc.

- **La biométrie** : quelques travaux de la littérature ont également utilisé la texture dans le domaine de la biométrie. En effet, des performances intéressantes ont été obtenues en exploitant la texture de différentes caractéristiques physiologiques telles que l'iris (Huang *et al.*, 2010; Li, Zhou, *et al.*, 2014; Li *et al.*, 2003), le visage (Ahonen *et al.*, 2004; Hadid *et al.*, 2004), la paume (Malathi Latha et Prasad, 2015; Ribaric et Lopar, 2012)... etc. Ainsi pour la détection de lunettes (Fernández *et al.*, 2015), la classification des personnes par leur sexe depuis des images de visages occlus ou non occlus (Rai et Khanna, 2014).

III.3 La forme

Les êtres humains sont connus par leur capacité de reconnaître les objets grâce au fait que plusieurs objets possèdent des formes distinguées (Figure III-9, (MPEG400²⁰)). Ceci a fait de cet attribut visuel un point de concentration pour plusieurs chercheurs tout au long de plusieurs années. Par conséquent, un nombre important de descripteurs de formes ont été proposés pour la recherche d'images, la classification, la reconnaissance... etc., dans différents domaines d'application (Frighetto-Pereira *et al.*, 2016; Iivarinen et Visa, 1996; Jain et Vailaya, 1998; Lande *et al.*, 2014; Li et Guan, 2006; Mohd Anuar *et al.*, 2013; Sokic et Konjicija, 2016; Wu et Yan, 2016; Zhang et Lu, 2002).

Cependant, la description d'un objet en s'appuyant sur les caractéristiques de la forme, certainement, exige que l'objet soit segmenté de son entourage, ce qui n'est pas garanti en pratique. En général, dans les travaux existants, les expériences ont été conduites sur des bases d'images contenant des objets sur un arrière-plan simple ou tout simplement des images binaires telles celles illustrées dans les figures ci-dessous (Figure III-9, Figure III-10). Ainsi, l'une des intéressantes applications de la forme est la recherche des logos (Figure III-10) et les marques déposées (Trademark) afin de prévenir les contrefaçons (Jain et Vailaya, 1998; Mohd Anuar *et al.*, 2013).



Figure III-9 Quelques formes représentant différents objets (images de la base MPEG400).

²⁰ <http://congyang.de/downloads.html>



Figure III-10 Exemples d'images de Logos (Logo dataset²¹).

[Yang et al. \(2008\)](#) énumèrent quelques propriétés d'un bon descripteur, qui sont : l'identifiabilité, l'invariance à la translation, rotation et l'échelle, l'invariance affine, la résistance au bruit, l'invariance à l'occultation, l'indépendance statistique, et la fiabilité.

Malgré le nombre important des descripteurs, ces derniers peuvent être classifiés en quelques classes seulement. [Pavlidis \(1978\)](#) a proposé différentes classifications de descripteurs suivant différents critères. Dans la première classification les descripteurs sont classifiés en *basé-Contour (external)* ou *basé-Région (internal)*, selon qu'un descripteur exploite les points du contour ou ceux internes. Dans cette classification, [Zhang et Lu \(2004\)](#) ont divisé davantage les méthodes de chaque classe en deux approches : *structural* et *global* (Figure III-11). La deuxième classification de [Pavlidis](#), est basée sur le type du résultat. En conséquence, deux classes sont distinguées : *space-domaine* et *scalar transform*. Les méthodes de la première classe transforment une image à une autre, tandis que les méthodes de la deuxième classe créent des scalaires (ou vecteurs). La troisième classification distingue entre les méthodes préservant l'information ou non (*information-preserving* et *non-preserving*). C.-à-d. s'il est possible de reconstruire la forme originale à partir du descripteur ou non.

Dans les sections qui vont suivre (sec.III.3.1-III.3.2), nous présenterons quelques descripteurs de forme en suivant la classification Contour/Région. Pour les lecteurs intéressés par cet attribut, ils peuvent recourir à plusieurs études de l'état de l'art, comme ([Loncaric, 1998](#); [MEHTRE et al., 1997](#); [Yang et al., 2008](#); [Zhang et Lu, 2004](#)).

²¹ University of Maryland, Laboratory for Language and Media Processing (LAMP) , Logo Dataset, <http://lamp.cfar.umd.edu> (2016)

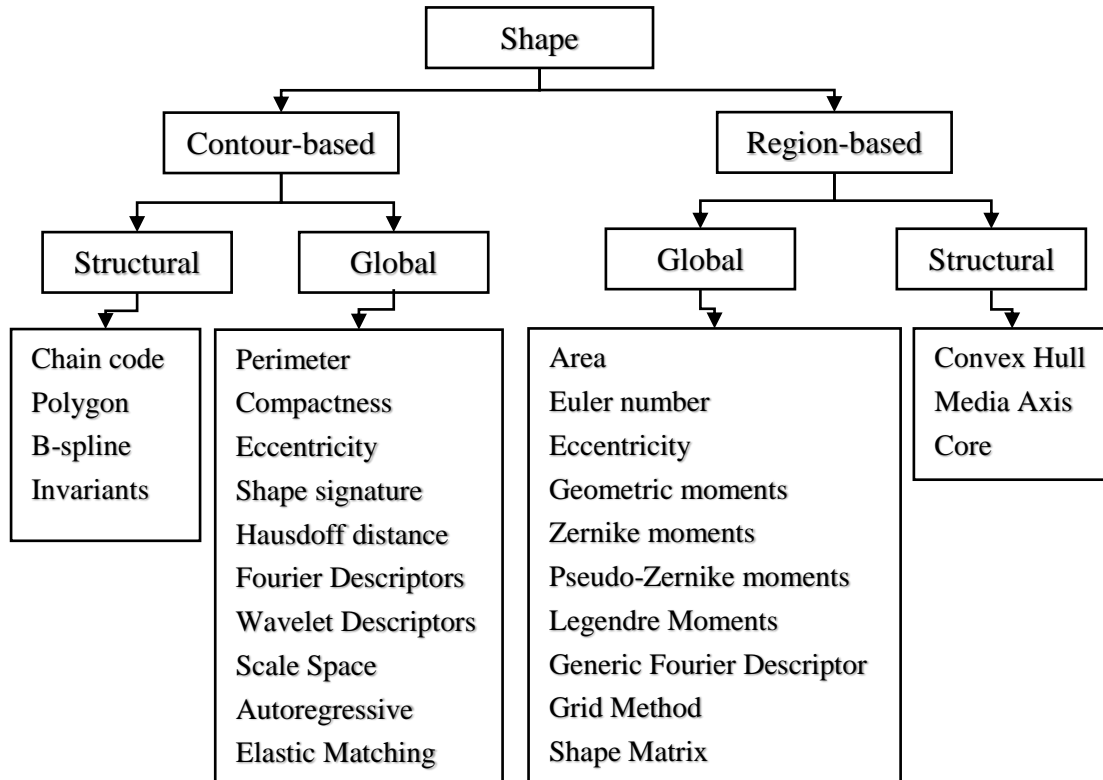


Figure III-11 Classification des représentations de la forme et les techniques de description (Zhang et Lu, 2004).

III.3.1 Descripteurs basés-Contour

III.3.1.1 Code de Freeman (Chain coded string)

Ce code connu aussi par le *code de Freeman* a été introduit par [Freeman et Davis \(1977\)](#). Il s'est basé sur la représentation du contour de la forme par une suite de codes, en parcourant les pixels du contour dans un sens donné (dans le sens antihoraire, par exemple). Le code du prochain pixel est déterminé selon sa position par rapport au pixel actuel. Huit directions sont considérées par les auteurs. Alors que dans ([MEHTRE et al., 1997](#)), le code a été réduit à quatre directions seulement.

Clairement, le code de Freeman n'est pas invariant aux rotations ni aux changements d'échelle. Pour cela, [MEHTRE et al. \(1997\)](#) ont considéré aussi une version modifiée appelée *Reduced Difference Chain code (RDC)*, qui consiste à prendre la différence des codes adjacents. Autres travaux proposent l'histogramme du code de Freeman ([Iivarinen et Visa, 1996](#)) et aussi le calcul des mesures statistiques à partir l'histogramme réarrangé ([Park et al., 2011](#)). Quant aux récents travaux exploitant ce code, nous citons celui de [Bo et al. \(2010\)](#), dans lequel les auteurs proposent une méthode basée sur le code Freeman pour la correspondance entre les courbes. Plus récemment, [Bicego et Lovato \(2016\)](#) proposent trois manières pour encoder la chaîne (Freeman code) d'une forme en une séquence d'acides aminés. Selon les auteurs, ce nouvel encodage permettra de l'exploitation des outils d'alignement des séquences biologiques.

III.3.1.2 Descripteur de Fourier

En plus de sa simplicité, le descripteur de Fourier est très robuste au bruit. Il est obtenu en appliquant la transformée de Fourier sur une *signature de forme* (*Shape signature*). La signature de forme représente la forme par une fonction unidimensionnelle déduite des points du contour (Zhang et Lu, 2003, 2004). Il existe plusieurs signatures telles que *centroidal profile*, *complex coordinates*, *centroid distance*, *tangent angle*, *cumulative angle*, *curvature*, *area* et *chord-length* (Zhang et Lu, 2004).

Soit une forme composée de N points de coordonnées $(x(t), y(t))$, où $t = 0, 1, \dots, N-1$. En considérant la signature *centroid distance*, la distance entre les points du contour et le centre de la forme (x_c, y_c) est obtenue comme suit :

$$d(t) = \sqrt{(x(t) - x_c)^2 + (y(t) - y_c)^2} \quad (\text{III.16})$$

Où

$$x_c = \frac{1}{N} \sum_{t=0}^{N-1} x(t), \quad y_c = \frac{1}{N} \sum_{t=0}^{N-1} y(t) \quad (\text{III.17})$$

Ainsi, la transformée de Fourier discrète de d est donnée par l'équation suivante :

$$a_n = \frac{1}{N} \sum_{t=0}^{N-1} d(t) \exp\left(-\frac{j2\pi nt}{N}\right), n = 0, 1, \dots, N-1 \quad (\text{III.18})$$

a_n sont les coefficients de Fourier. La transformée de Fourier est discutée davantage dans la section IV.2.1.4.1.

Les performances des descripteurs de Fourier dérivés de différentes signatures sont très variées (Yang et al., 2008). Par exemple, les coefficients de Fourier obtenus précédemment sont invariants aux translations grâce à l'invariance de la signature à cette opération. Afin d'obtenir l'invariance aux autres opérations, les coefficients de Fourier doivent être normalisés (Yang et al., 2008; Zhang et Lu, 2003). L'invariance aux rotations et changements de point de départ est obtenue par la considération de la magnitude des coefficients de Fourier, sous prétexte que ces opérations n'ont pas d'effet que sur la phase. Sokic et Konjicija (2016) ont vu que le sacrifice par les informations de phase en suivant une telle approche simple peut entraîner la perte de précieuses informations. Par conséquent, ils proposent d'améliorer les performances du descripteur en conservant les informations de phase, d'où le nom de leur descripteur *Phase-Including Fourier Descriptor* (PIFD). Afin de garder l'invariance par rapport aux rotations, ils introduisent ce qu'ils ont appelé les points *pseudomirror*, pour déterminer l'orientation nominale de la forme.

III.3.2 Descripteurs basés-Région

III.3.2.1 Les moments invariants

Les moments invariants appelés aussi les moments géométriques invariants, ou encore les moments de Hue. Ils sont très répandus dans les différentes tâches de la reconnaissance. Ces moments ont été introduits par ([Ming-Kuei, 1962](#)). Le moment d'ordre $(p+q)^{\text{ème}}$ de la fonction bidimensionnelle $f(x, y)$ est défini comme suit :

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y), \quad p, q = 0, 1, 2, \dots \quad (\text{III.19})$$

Cependant, les moments obtenus par l'équation précédente ne sont pas invariants aux translations, rotations, changements d'échelle ([Yang et al., 2008](#)).

[Ming-Kuei](#), a introduit aussi les moments géométriques centraux qui sont invariants aux translations :

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y), \quad p, q = 0, 1, 2, \dots \quad (\text{III.20})$$

$$\text{Où } \bar{x} = m_{10}/m_{00}, \quad \bar{y} = m_{01}/m_{00}$$

Ci-dessous, l'ensemble de 7 moments invariants tels qu'ils ont été rapportés dans ([Yang et al., 2008](#)) :

$$\begin{aligned} \phi_1 &= \eta_{20} + \eta_{02} \\ \phi_2 &= (\eta_{20} + \eta_{02})^2 + 4\eta_{11}^2 \\ \phi_3 &= (\eta_{30} + 3\eta_{12})^2 + (3\eta_{21} + \eta_{03})^2 \\ \phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\ \phi_5 &= (\eta_{30} + 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ &+ (3\eta_{21} + \eta_{03})(\eta_{21} + \eta_{03})[(3\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ \phi_6 &= (\eta_{20} + \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}^2 \\ &(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ \phi_7 &= (3\eta_{21} + \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ &+ (3\eta_{21} + \eta_{03})(\eta_{21} + \eta_{03})[(3\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (\text{III.21})$$

$$\text{Où : } \eta_{pq} = \mu_{pq} / \mu_{00}^\gamma \text{ et } \gamma = 1 + (p+q)/2, \text{ pour } p+q = 2, 3, \dots$$

III.3.2.2 Les moments de Zernike

[Teague \(1980\)](#) introduit les moments de Zernike en exploitant l'idée de la possibilité de la généralisation des moments de Hue. La généralisation est obtenue en remplaçant le noyau $x^p y^q$ dans l'équation (III.19) par une forme plus générale $P_p(x)P_q(y)$. En effet, les moments de

Zernike sont des moments orthogonaux obtenus par le remplacement du composant $x^p y^q$ par les polynômes de Zernike ([Zhang et Lu, 2003](#)) :

$$V_{nm}(x, y) = V_{nm}(\rho \cos \theta, \rho \sin \theta) = R_{nm}(\rho) \exp(jm\theta) \quad (\text{III.22})$$

Où :

$$R_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \frac{(n-s)!}{s! \left(\frac{n+|m|}{2} - s\right)! \left(\frac{n-|m|}{2} - s\right)!} \rho^{n-2s} \quad (\text{III.23})$$

Où ρ est la longueur de (x, y) par rapport au centre de la forme, θ est l'angle entre ρ et l'axe des x , n et m sont des entiers tel que : $n - |m| = \text{pair}$, et $|m| \leq n$.

Donc, les moments de Zernike complexe d'ordre n et de répétition m sont définis comme suit ([Zhang et Lu, 2003](#)) :

$$A_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x, y) V_{nm}^*(x, y), \quad x^2 + y^2 \leq 1 \quad (\text{III.24})$$

III.4 Les points d'intérêt

L'exploitation de quelques *points clés* extraits de l'image est l'une des approches prometteuses dans le domaine de la vision par ordinateur. En effet, un point clé qui est connu plus par *point d'intérêt* est un pixel de l'image qui est saillant ou unique par rapport à son voisinage ([Torres-Huitzil, 2016](#)). Bien que les caractéristiques extraites d'un point d'intérêt soient suffisantes dans quelques applications, son étendue (le voisinage) est généralement très nécessaire, à cause de l'insuffisance de la localisation seulement du point d'intérêt. Dans ce cas, le terme *région* est utilisé au lieu de point d'intérêt ([Mikolajczyk et Tuytelaars, 2007](#)). Par conséquent, les auteurs ont privilégié dans leur Survey le terme *local feature* (caractéristique locale), pour référer à des points, régions, et même des segments de bord.

Selon [Mikolajczyk et Tuytelaars \(2007\)](#), les caractéristiques locales idéales doivent correspondre à des parties sémantiques d'objets. Cependant, cette correspondance est infaisable en pratique parce que la détection est basée sur des propriétés de bas niveau de l'image, telles que la couleur et la texture. Selon les mêmes auteurs, une bonne caractéristique doit avoir les propriétés suivantes :

- **Répétabilité** : la majorité des caractéristiques doivent être détectées à partir de deux images de même scène ou objet prises sous différentes conditions d'observation.
- **Distinctiveness/informativeness** : les intensités des patterns des caractéristiques détectées doivent montrer de grandes variations afin qu'elles puissent être distinguées et comparées.

- **Localité** : les caractéristiques doivent être locales en vue de réduire la probabilité d'occlusion et pour permettre des approximations en modèle simple des déformations géométriques et photométriques entre deux images prises sous différentes conditions d'observation.
- **Quantité** : le nombre des caractéristiques détectées doit être suffisamment grand.
- **Précision** : la localisation des caractéristiques détectées doit être précise.
- **Efficacité** : la détection doit se faire le plus rapidement possible, afin de supporter les applications temps réel.

Pratiquement, ce type de caractéristiques ont été exploitées dans différents axes de recherche de la vision par ordinateur, comme la recherche d'image ([Wang et al., 2014](#); [Yang et al., 2014](#); [Zhong et al., 2009](#)), la classification d'images ([Azhar et al., 2015](#); [Lu et al., 2016](#)), la reconnaissance d'iris ([Alvarez-Betancourt et Garcia-Silvente, 2016](#)) la reconnaissance de visages ([A et al., 2015](#); [Li, Huang, et al., 2014](#)).

Quant aux méthodes qui permettent l'extraction de ce type de caractéristiques, elles opèrent en deux étapes. La première étape est l'étape de détection dans laquelle les points d'intérêt vont être détectés par un *détecteur*. Puis dans la deuxième étape, un *descripteur* est établi pour chaque point d'intérêt. En effet, plusieurs méthodes de détection de points d'intérêt existent. [Hassaballah et al. \(2016\)](#) les ont classifiées en trois larges catégories : *single-scale detectors*, *multi-scale detectors*, et *affine invariant detectors*. Entre autres, nous citons Harris detector ([Harris et Stephens, 1988](#)), FAST detector (*Features from Accelerated Segment Test*) ([Rosten et Drummond, 2006](#)), Hessian detector ([Beaudet, 1978](#)), Laplace of Gaussian (LoG) ([Lindeberg, 1998](#)), Difference of Gaussian (DoG) ([Lowe, 2004b](#)), Harris-Laplace ([Mikolajczyk et Schmid, 2004](#)). Quant aux descripteurs, l'état de l'art en dispose également d'une large gamme, entre autres : SIFT (*Scale-Invariant feature Transform*) ([Lowe, 2004b](#)), SURF (*Speeded-Up Robust Features descriptor*) ([Bay et al., 2008](#); [Bay et al., 2006](#)), BRIEF (*Binary Robust Independent Elementray Features*) ([Calonder et al., 2010](#))... etc.

Dans ce qui suit, nous présenterons brièvement quelques détecteurs et descripteurs parmi les plus populaires. Pour plus de détecteurs ou de descripteurs, ou tout simplement pour plus ample d'informations, le lecteur peut se référer aux travaux suivants ([Hassaballah et al., 2016](#); [Mikolajczyk et Tuytelaars, 2007](#)), ou encore aux études comparatives ([Mikolajczyk et Schmid, 2005](#); [Mikolajczyk et al., 2005](#); [Mukherjee et al., 2015](#)).

III.4.1 Détecteurs

III.4.1.1 Détecteur de Harris

[Harris et Stephens \(1988\)](#) ont introduit un détecteur de corners et de bords basé sur la matrice d'autocorrélation. La matrice 2×2 symétrique d'autocorrélation utilisée pour la détection des caractéristiques locales de l'image est définie comme suit :

$$M(x, y) = \sum_{u,v} w(u, v) * \begin{bmatrix} I_x^2(x, y) & I_x I_y(x, y) \\ I_x I_y(x, y) & I_y^2(x, y) \end{bmatrix} \quad (III.25)$$

Où I_x et I_y sont les dérivatives locales de l'image dans les deux directions x et y , respectivement. $w(u, v)$ est une fenêtre de pondération binaire, elle est généralement remplacée par une fenêtre circulaire, par exemple un noyau *Gaussien*, dont la réponse est isotropique. Donc, pour trouver les points d'intérêt, les valeurs propres de la matrice M sont calculées pour chaque pixel. Les coins sont localisés dans les positions dont les valeurs propres sont grandes dans les deux directions. Afin de remédier au problème du coût de calcul des valeurs propres, [Harris et Stephens](#), proposent une mesure de *Cornerness* (C). Cette dernière utilise le déterminant (\det) et la trace (tr) de la matrice M :

$$C(x, y) = \det(M) - \xi(tr(M))^2 \quad (\text{III.26})$$

Où, $\det(M) = \lambda_1 \cdot \lambda_2$, et $tr(M) = \lambda_1 + \lambda_2$

λ_1 et λ_2 sont les valeurs propres, ξ est un paramètre d'ajustement.

III.4.1.2 Détecteur Hessien

Le détecteur *Hessien* ([Beaudet, 1978](#)) est un détecteur de blob, basé sur la matrice des dérivatives de second ordre des intensités d'images. Cette matrice est appelée *la matrice Hessienne* :

$$H(x, y, \sigma) = \begin{bmatrix} I_{xx}(x, y, \sigma) & I_{xy}(x, y, \sigma) \\ I_{xy}(x, y, \sigma) & I_{yy}(x, y, \sigma) \end{bmatrix} \quad (\text{III.27})$$

Où I_{xx}, \dots etc. sont les dérivatives de second ordre, calculées par une fonction Gaussienne d'écart type σ . Les caractéristiques locales se trouvent là où les réponses des dérivatives sont supérieures dans les deux directions. À cet effet, le déterminant de la matrice est calculé :

$$\det(H) = I_{xx}I_{yy} - I_{xy}^2 \quad (\text{III.28})$$

III.4.1.3 Laplacien de Gaussien

Laplacian-of-Gaussian (LoG) a été utilisé par [Lindeberg \(1993\)](#); [Lindeberg, 1998](#) pour la détection de blobs. Pour trouver ces derniers, les maxima sont recherchés dans la représentation 3D, *espace-échelle* (*space-scale*) d'une image. La représentation espace-échelle peut être obtenue par le lissage (blur) en utilisant des dérivatives des noyaux Gaussiens de différentes tailles.

Soit une image I , sa représentation espace-échelle L est obtenue comme suit :

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (\text{III.29})$$

Où

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (\text{III.30})$$

Puis, l'opérateur Laplacien est calculé comme suit :

$$\nabla^2 L(x, y, \sigma) = L_{xx}(x, y, \sigma) + L_{yy}(x, y, \sigma) \quad (\text{III.31})$$

Le problème de cet opérateur est que sa réponse est fortement dépendante de la relation entre la taille de blobs et celle du noyau Gaussien ([Hassaballah et al., 2016](#)).

III.4.1.4 Différence de Gaussien

Difference-of-Gaussian (DoG) est un détecteur de blobs, invariant aux changements d'échelle (scale-invariant), a été introduit par [Lowe \(2004a\)](#); [Lowe, 1999](#). Ce détecteur performe, en premier, quelques étapes de lissage de l'image par convolution avec un filtre Gaussien, ce qui permet de générer des images de différentes échelles. Ensuite, les différences entre les images lissées adjacentes sont établies. Donc, les caractéristiques sont localisées par rapport à l'espace et l'échelle en même temps (3D). Les points clés peuvent être détectés efficacement en soustrayant deux Gaussiens séparés par un constant multiplicatif k , puis le résultat est convoluté avec l'image ([Lowe, 2004a](#)) :

$$\begin{aligned} G(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad (\text{III.32})$$

III.4.2 Descripteurs

III.4.2.1 SIFT

[Lowe \(1999\)](#) a proposé une méthode basée sur l'extraction d'un ensemble de vecteurs des caractéristiques locales (*local feature vector*), baptisée *Scale Invariant Feature Transform* (SIFT). La méthode est invariante à un ensemble de transformations comme la translation, le changement d'échelle et la rotation. SIFT utilise DoG (présenté ci-dessus) pour la détection des points d'intérêt.

Afin de calculer le descripteur SIFT, le gradient à chaque pixel dans une région (16×16) centrée au point d'intérêt détecté est calculé. Puis, un histogramme d'orientations de 8-bins pour chaque région 4×4 est établi. Ce qui donne un vecteur de dimension 128 (8×4×4) pour chaque point d'intérêt.

III.4.2.2 SURF

Speeded Up Robust Features (SURF), est un descripteur invariant aux rotations et changements d'échelle, a été proposé par [Bay et al. \(2008\)](#); [Bay et al., 2006](#). Le descripteur est minutieusement construit afin d'obtenir une version plus rapide que le SIFT tout en gardant la performance comparable avec ce dernier. À cet effet, SURF est basé sur l'approximation de la matrice Hessienne pour la détection des points d'intérêt. Ceci se prête à l'utilisation des *images intégrales* (integral images) qui permettent le calcul des convolutions plus rapidement avec des '*box filters*'.

L'étape de la description du SURF est basée sur la description des réponses de l'ondelette de Haar. Donc, une région carrée centrée au point d'intérêt et orientée vers l'orientation présélectionnée (afin d'être invariant aux rotations, une étape de sélection de l'orientation

dominante est effectuée), de taille $20s$ est prise en compte, où s représente l'échelle. Cette région est divisée en 4×4 sous-régions. Ensuite, l'ondelette de Haar est appliquée sur chaque sous-région dans les deux directions ; horizontale et verticale. Soit dx et dy les réponses de l'ondelette de Haar suivant les deux directions, respectivement. Pour des raisons de robustesse par rapport aux déformations géométriques et aux erreurs de localisation, les réponses dx et dy sont pondérées par une Gaussienne ($\sigma = 3.3s$) centrée au point d'intérêt. Finalement, le vecteur de caractéristiques au niveau de chaque sous-région est composé de quatre différentes mesures $v = (\sum dx, \sum dy, \sum |dx|, \sum |dy|)$. Donc, la longueur du vecteur de caractéristiques d'un point d'intérêt est de longueur 64 ($4 \times 4 \times 4$).

III.4.2.3 GLOH

Le descripteur *Gradient Location-Orientation Histogram* (GLOH) a été introduit par [Mikolajczyk et Schmid \(2005\)](#) comme une extension de SIFT pour améliorer sa robustesse et son caractère distinctif. GLOH est également basé sur le calcul des orientations du gradient. Cependant, différemment de SIFT qui considère des régions de forme rectangulaire autour des points d'intérêt, GLOH considère une structure Log-polaire (Log-polar grid) (Figure III-12). Plus précisément, les auteurs dans [\(Mikolajczyk et Schmid, 2005\)](#) ont quantifié la direction radiale en trois bins et la direction angulaire à huit, par conséquent, le nombre de bins est 17 (le bin central n'est pas divisé). Quant aux orientations de gradient sont quantifiées à 16 bins. La taille totale d'un histogramme est donc 272 bins. Ce dernier est encore réduit par les auteurs en utilisant la méthode PCA (Principal Components Analysis).

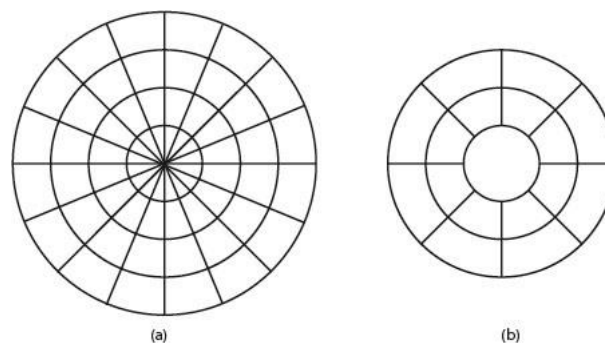


Figure III-12 Grille polaire, (a): 4 bins sur la direction radiale et 16 sur la direction angulaire; (b) : 3 bins sur la direction radiale et 8 sur la direction angulaire (sans division du bin central)

III.4.2.4 LPDF

[Song et Li \(2013\)](#) proposent un descripteur basé sur l'extraction des caractéristiques de la DCT du patch de l'image locale (c.-à-d. région autour du point d'intérêt) dans une structure polaire (Figure III-12). Ce descripteur est baptisé *Local Polar DCT Features* (LPDF). LPDF performe comme suit, après la détection des points d'intérêt en utilisant n'importe quel détecteur de l'état de l'art et la normalisation et le prétraitement des régions, une grille polaire de 8 bins dans la direction radiale et 16 dans la direction angulaire est centrée au point d'intérêt, pour échantillonner les pixels du patch local. Chaque bin est représenté par la moyenne des intensités de ses pixels. Ces échantillons sont ensuite transformés en une matrice 2D dont les entrées sont (θ, ρ) . Où θ et ρ dénotent les différentes positions d'échantillonnage angulaire

et radiale, respectivement, dans la structure géométrique polaire. Finalement, la DCT est appliquée sur cette matrice pour l'extraction des caractéristiques.

III.5 Combinaison des descripteurs de texture et de couleur

Nous avons présenté et discuté à travers les précédentes sections de ce chapitre les différentes propriétés d'image et leurs descripteurs respectifs. Cependant, malgré la diversité des méthodes et les améliorations qui sont y apportées, elles restent insuffisantes surtout lors de la présence de différentes propriétés dans l'image.

Afin de capturer les caractéristiques de différentes propriétés et donc booster les performances des systèmes basés sur le contenu visuel de l'image, la combinaison des descripteurs est sans doute l'une des alternatives les plus triviales. En effet, la combinaison est considérée dans beaucoup de travaux ([Frighetto-Pereira et al., 2016](#); [Jacob et al., 2014](#); [Jiayin et Wenjuan, 2012](#); [Lande et al., 2014](#); [Masoudi, 2016](#); [Murala et al., 2009](#); [Palm, 2004](#); [Wang et al., 2011](#); [Wang et al., 2012](#); [Yue et al., 2015](#); [Yue et al., 2011](#)).

En revanche, très peu sont les travaux qui se sont intéressés aux différentes approches de combinaison des descripteurs (typiquement, ceux de couleur et de texture). Parmi ces travaux, on trouve celui de [Mäenpää et Pietikäinen \(2004\)](#). Dans leur article, ils ont catégorisé puis évalué empiriquement les méthodes combinant la couleur et la texture en deux catégories : celles utilisant les informations de texture et de couleur séparément, et celles qui les utilisent conjointement. La séparation consiste à diviser le signal de couleur en chrominance et luminance, puis à utiliser un descripteur spécifique pour la couleur (par exemple l'histogramme) et un autre propre à l'extraction de l'information texturale. Quant à la deuxième catégorie, les deux propriétés sont gérées en même temps, par exemple, l'utilisation d'un descripteur de texture sur les différents axes de l'espace de couleur.

Après la conduction d'une série d'expériences, les auteurs ont conclu que les mesures de couleur sont suffisantes dans les conditions d'éclairage statique. Cependant dans un éclairage variable, la texture en niveaux de gris fonctionne bien mieux que ce soit la couleur, la texture couleur canal par canal (*channelwise*) ou couleurs opposées (*opponent color*). Donc, [Mäenpää et Pietikäinen \(2004\)](#) n'ont pas recommandé la description jointive de la couleur et la texture vu que dans leurs expériences *Tous les descripteurs joints de texture et de couleur, et toutes les méthodes de combinaison de la couleur et de la texture à un niveau supérieur sont surpassés par la couleur ou la texture en niveaux de gris seule*²².

De son côté, [Palm \(2004\)](#) a distingué entre trois types de méthodes combinant la couleur et la texture, qui sont les méthodes *parallèles*, *séquentielles*, et *intégratives*. Ces trois catégories seront présentées ci-dessous, selon ([Palm, 2004](#)).

²² "All joint color texture descriptors and all methods of combining color and texture on a higher level are outperformed by either color or gray-scale texture alone."

III.5.1 Les méthodes parallèles

Les méthodes de cette catégorie extraient les caractéristiques de couleur, en ignorant le voisinage local du pixel, et celles de texture, en prenant compte des relations entre les pixels voisins (Figure III-13). Donc, les caractéristiques des deux propriétés sont extraites séparément. Parmi les travaux suivant cette approche, nous citons ([Murala et al., 2009](#); [Wang et al., 2012](#); [Yue et al., 2015](#); [Yue et al., 2011](#)).

Selon [Palm](#), l'avantage des méthodes de cette catégorie est que les méthodes d'analyse de texture en niveaux de gris et celles d'analyse de couleur peuvent être appliquées directement. Cependant, cette approche très simplifiée peut ignorer les primitives des textures couleur avec une constante intensité.

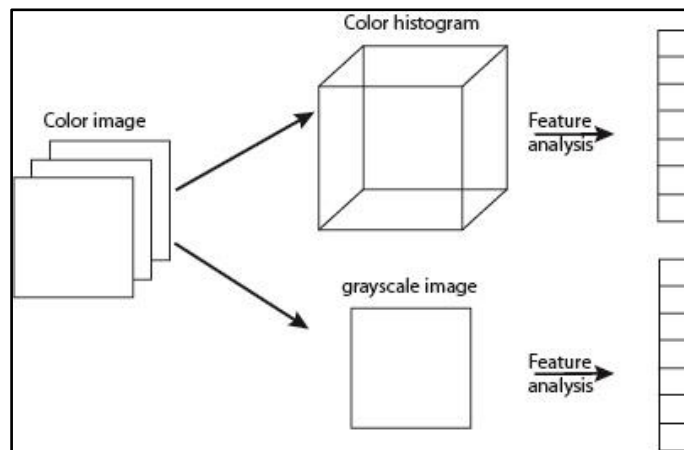


Figure III-13 Approche parallèle.

III.5.2 Les méthodes séquentielles

Les méthodes séquentielles analysent en premier l'information de couleur. Puis, l'histogramme de couleur est partitionné (clustering), et donc le partitionnement de l'image est obtenu. Ensuite, les différents segments sont analysés comme des textures de niveaux de gris (Figure III-14).

Ce type de méthodes se montre utile, néanmoins, elles sont basées sur la segmentation qui dépend de plusieurs paramètres.

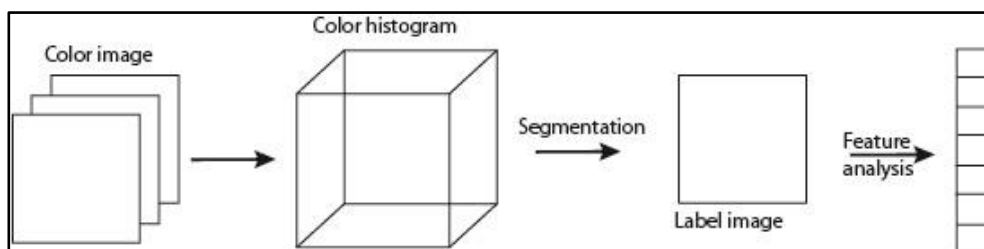


Figure III-14 Approche séquentielle.

III.5.3 Les méthodes intégratives

Les méthodes de cette catégorie appelées intégratives puisqu'elles prennent en compte la dépendance entre les informations des deux propriétés. Deux différentes stratégies des méthodes de cette catégorie, sont distinguées par [Palm](#), :

III.5.3.1 Monocanal (single-channel)

Les méthodes monocanal analysent la texture par rapport à chaque axe séparément (Figure III-15(a)). Donc, les caractéristiques de couleur sont considérées indirectement en se restreignant à un canal donné de couleur. Comme exemples, nous citons l'application des *filtres de Gabor* sur les différents axes ([Drimbarean et Whelan, 2001](#); [Paschos, 2001](#)), le calcul des *matrices de cooccurrence* ([Drimbarean et Whelan, 2001](#)), et les *matrices de cooccurrence de motifs* ([Jhanwar et al., 2004](#)).

L'intérêt de cette stratégie est la facilité d'adaptation des méthodes de texture opérant en niveaux de gris.

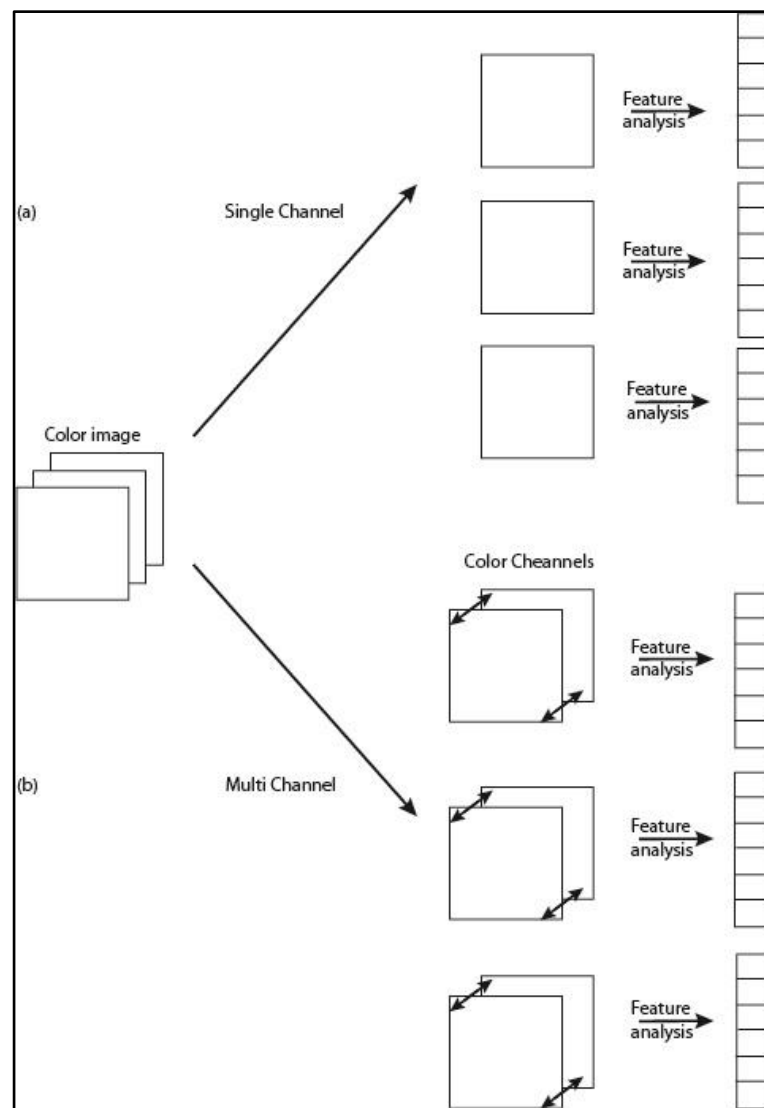


Figure III-15 Approche intégrative. (a) : single channel, (b) : multi channel.

III.5.3.2 Multicanal (multi-channel)

Les méthodes suivant cette stratégie gèrent deux canaux de couleur (ou plus) en même temps (Figure III-15(b)). Par conséquent, la dépendance entre les textures dans les différents canaux est prise en compte. Parmi les travaux suivant cette stratégie, nous avançons ([Jacob *et al.*, 2014](#); [Mäenpää et Pietikäinen, 2004](#); [Palm, 2004](#); [Sandid et Douik, 2016](#); [Subrahmanyam *et al.*, 2013](#)).

III.6 Conclusion

Dans ce chapitre, nous avons survolé plusieurs méthodes de caractérisation des principaux attributs visuels de l'image. En premier, nous avons présenté l'attribut couleur ainsi que quelques espaces étant donné l'impact de ces derniers sur les performances des descripteurs. Nous avons également discuté plusieurs de ses descripteurs parmi les plus connus selon le type d'information capturée. Dans ce contexte, nous avons remarqué que les descripteurs locaux sont les plus adoptés par les auteurs. Ceci peut être interprété par la diversité des stratégies possibles afin de capturer ce type d'information, mais surtout aux bonnes performances des descripteurs locaux par rapport aux descripteurs globaux. En ce qui concerne la texture, nous l'avons présentée et discutée puis nous avons mis l'accent sur plusieurs domaines d'application de cet attribut. Nous avons vu que l'analyse de la texture est un ancien axe de recherche dont ses débuts remontent aux années 60s. La texture a été ensuite utilisée pour la classification, la segmentation de différents types d'images. Dans ce chapitre, nous avons également abordé quelques descripteurs de la forme, puis ceux des points d'intérêt. Enfin, nous avons présenté les différentes approches de combinaison de la texture et la couleur.

Le prochain chapitre sera consacré aux différentes méthodes de description de la texture, l'attribut auquel nous nous sommes intéressés. Pour cela, nous discuterons les méthodes pionnières ainsi que les améliorations y apportées par les chercheurs. Nous discuterons ensuite les différentes classifications de méthodes.

Chapitre IV. Descripteurs de texture

« La théorie, c'est quand on sait tout et que rien ne fonctionne. La pratique, c'est quand tout fonctionne et que personne ne sait pourquoi. Ici, nous avons réuni théorie et pratique : Rien ne fonctionne... et personne ne sait pourquoi ! »

A. Einstein

Sommaire

IV.1	LES DESCRIPTEURS	48
IV.1.1	Les matrices de cooccurrence de Haralick	48
IV.1.2	Les descripteurs de Tamura	52
IV.1.2.1	La grossièreté.....	52
IV.1.2.2	Le contraste	53
IV.1.2.3	La directivité	53
IV.1.3	Les motifs binaires locaux.....	54
IV.1.3.1	Variantes du LBP	57
IV.1.3.1.1	Les variantes populaires	57
IV.1.3.1.1.1	LBP invariante à la rotation et LBP uniforme (Ojala et al., 2002)	57
IV.1.3.1.1.2	Improved LBP (Jin et al., 2004)	58
IV.1.3.1.1.3	Median Binary Patterns (Hafiane et al., 2007)	58
IV.1.3.1.1.4	Local Ternary Patterns (Tan et Triggs, 2007)	58
IV.1.3.1.1.5	Dominant LBP (Liao et al., 2009)	59
IV.1.3.1.1.6	Local Derivative Patterns (Zhang et al., 2010).....	59
IV.1.3.1.1.7	Completed LBP (Guo et al., 2010)	60
IV.1.3.1.1.8	Opponent Color LBP (Mäenpää et Pietikäinen, 2004).....	61
IV.1.3.1.1.9	Block-based LBP (Takala et al., 2005)	62
IV.1.3.1.2	Les variantes récentes	63
IV.1.3.1.2.1	Dominant Rotated Local Binary Patterns (Mehta et Egiazarian, 2016)	63
IV.1.3.1.2.2	Quaternionic Local Ranking Binary Pattern (Lan et al., 2016)	63
IV.1.3.1.2.3	Modified LBP (Naresh et Nagendraswamy, 2016).....	64
IV.1.3.1.2.4	Local Tri-Directionnal Patterns (Verma et Raman, 2016)	65
IV.1.3.1.2.5	LBP by neighborhood and Directionnal LBP (Kaya et al., 2015)	66
IV.1.3.1.2.6	Power LBP (Smolka et Nurzynska, 2015)	67
IV.1.3.1.2.7	SS-3D-LTP (Murala et Wu, 2015)	68
IV.1.3.1.2.8	Enhanced LBP (Wong et al., 2013)	69
IV.1.3.2	Applications	69
IV.1.4	Les ondelettes.....	70
IV.1.4.1	Présentation des ondelettes.....	70
IV.1.4.2	Les ondelettes et les images	72
IV.1.4.3	Applications et descripteurs des ondelettes pour la caractérisation des images	73
IV.1.5	Les filtres de Gabor	74
IV.1.5.1	Présentation	74
IV.1.5.2	Applications	77
IV.2	CLASSIFICATION DES DESCRIPTEURS.....	81
IV.2.1.1	Approches statistiques.....	81
IV.2.1.1.1	Les matrices de cooccurrence de motifs	82
IV.2.1.1.2	Sum and Difference Histograms	82
IV.2.1.2	Approches structurelles.....	83

IV.2.1.3	Approches basées sur les modèles	84
IV.2.1.3.1	Les fractales	84
IV.2.1.3.2	Les modèles stochastiques	84
IV.2.1.4	Approches basées sur les transformations	86
IV.2.1.4.1	La transformée de Fourier	86
IV.2.1.4.2	La transformée en Cosinus discrète.....	87
IV.3	CONCLUSION.....	88

L'omniprésence de la texture est sans doute l'une des raisons qui lui permettent de prendre l'attention de la communauté des chercheurs dans le domaine de la vision par ordinateur, depuis le jeune âge de l'informatique. Cependant, et contrairement aux autres attributs, la texture manque d'une définition formelle et précise. Ce manque a rendu cet attribut parmi les plus difficiles à caractériser. Dans une autre optique, la multiplicité des tentatives conduites pour capturer ces caractéristiques ainsi la diversité des approches suivies, font de l'état de l'art de cet attribut vaste et riche.

Le but de ce chapitre est de dresser un état de l'art sur les méthodes d'extraction des caractéristiques de la texture. Pour cela, en premier, nous présenterons et discuterons les méthodes pionnières ainsi que les améliorations apportées sur ces méthodes. Puis dans la deuxième partie, nous discuterons la classification des descripteurs en général. Ce qui permet de mettre en évidence les méthodes pionnières.

IV.1 Les descripteurs

IV.1.1 Les matrices de cooccurrence de Haralick

[Haralick](#), a suggéré en 1971, de caractériser l'information texturale par des matrices de cooccurrences de niveaux de gris (MCNG). Ces matrices décrivent les dépendances spatiales entre les niveaux de gris 'Gray-Tone Dependence Matrices', pour chaque distance fixée et/ou relation spatiale angulaire. D'après le survey de [Haralick \(1979\)](#), cette manière de décrire l'information texturale existait déjà à cette époque-là, et leur existence remonte aux travaux de [Julesz \(1962\)](#). En 1973, [Haralick et al.](#), ont proposé de résumer l'information texturale contenue dans les MCNG par des mesures extraites de ces matrices.

Cette méthode est parmi les plus célèbres et les plus utilisées grâce aux performances atteintes, vu sa capacité de caractériser la relation spatiale entre les niveaux de gris. En plus, son invariance par rapport aux variations monotonic du gris. Néanmoins, elle n'est pas capable de capturer la forme des primitives.

Une matrice de cooccurrence contient les fréquences spatiales relatives des niveaux de gris suivant une direction et une distance données. Formellement parlant, étant donné une image I de taille $N_r \times N_c$, et de N_g niveaux de gris. La matrice de cooccurrence $P_{d,\theta}$ associée à cette image, pour une direction θ (la Figure IV-1, illustrant les différentes directions) et une distance d , est une matrice carrée de taille $N_g \times N_g$, établie comme suit ([Haralick, 1979](#)) :

$$P_{d,0^\circ} = \# \left\{ \left((k,l), (m,n) \right) \in (L_r \times L_c) \times (L_r \times L_c) \mid k - m = 0, \right. \\ \left. |l - n| = d, I(k,l) = i, I(m,n) = j \right\} \quad (IV.1)$$

$$P_{d,45^\circ} = \# \left\{ \left((k,l), (m,n) \right) \in (L_r \times L_c) \times (L_r \times L_c) \mid (k - m = d, l - n = -d) \right. \\ \left. \text{or} (k - m = -d, l - n = d), I(k,l) = i, I(m,n) = j \right\} \quad (IV.2)$$

$$P_{d,90^\circ} = \# \left\{ \left((k,l), (m,n) \right) \in (L_r \times L_c) \times (L_r \times L_c) \mid |k - m| = d, \right. \\ \left. l - n = 0, I(k,l) = i, I(m,n) = j \right\} \quad (IV.3)$$

$$P_{d,135^\circ} = \# \left\{ \left((k,l), (m,n) \right) \in (L_r \times L_c) \times (L_r \times L_c) \mid (k - m = d, l - n = d) \right. \\ \left. \text{or} (k - m = -d, l - n = -d), I(k,l) = i, I(m,n) = j \right\} \quad (IV.4)$$

Où $L_r = \{1,2,\dots, N_r\}$ et $L_c = \{1,2,\dots, N_c\}$.

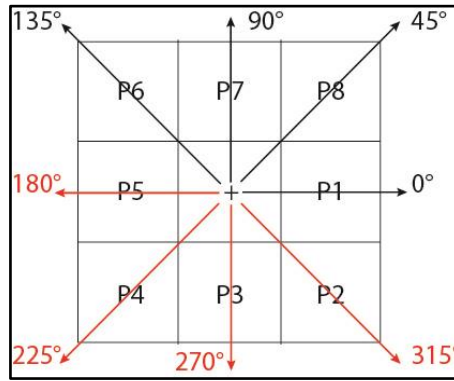


Figure IV-1 Illustration de différentes directions des MCNG.

La Figure IV-2, montre un exemple illustratif sur le calcul des MCNG d'une matrice 4x4, dont les niveaux de gris varient entre 0 et 3.

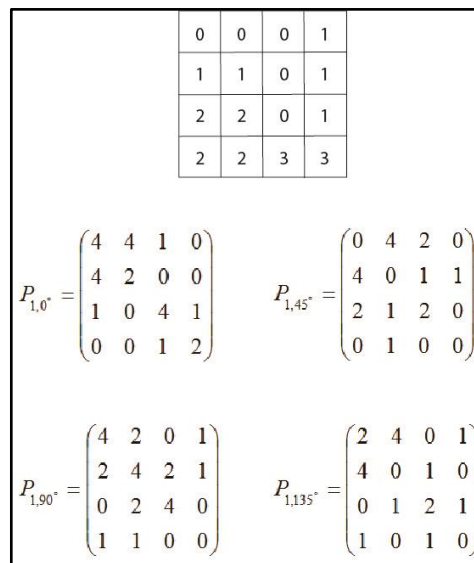


Figure IV-2 Exemple de calcul des MCNG.

En se basant sur son hypothèse consistante que toutes les informations texturales sont contenues dans les MCNG, [Haralick et al. \(1973\)](#) ont proposé 14 mesures pour extraire ces informations. Les équations qui permettent de définir ces mesures sont comme suit :

Angular second moment (ASM)

$$f_1 = \sum_i \sum_j p_{d,\theta}(i, j)^2 \quad (IV.5)$$

Contrast

$$f_2 = \sum_{n=0}^{N_g-1} n^2 \left(\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p_{d,\theta}(i, j) \right), |i - j| = n \quad (IV.6)$$

Corrélation

$$f_3 = \frac{\sum_i \sum_j (i, j) p_{d,\theta}(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (IV.7)$$

Sum of squares : variance

$$f_4 = \sum_i \sum_i (i - \mu)^2 p_{d,\theta}(i, j) \quad (IV.8)$$

Inverse difference Moment

$$f_5 = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p_{d,\theta}(i, j) \quad (IV.9)$$

Sum average

$$f_6 = \sum_{i=2}^{2N_g} i p_{x+y}(i) \quad (IV.10)$$

Sum variance

$$f_7 = \sum_{i=2}^{2N_g} (i - f_6)^2 p_{x+y}(i) \quad (IV.11)$$

Sum entropy :

$$f_8 = - \sum_{i=2}^{2N_g} p_{x+y}(i) \log(p_{x+y}(i)) \quad (IV.12)$$

Entropy

$$(IV.13)$$

$$f_9 = -\sum_i \sum_j p(i, j) \log(p(i, j))$$

Difference variance

$$f_{10} = \text{var } p_{x-y} \quad (\text{IV.14})$$

Difference entropy

$$f_{11} = -\sum_{i=0}^{N_g-1} p_{x-y}(i) \log(p_{x-y}(i)) \quad (\text{IV.15})$$

Information measures of correlation 1

$$f_{12} = \frac{HXY - HXY1}{\max\{HX, HY\}} \quad (\text{IV.16})$$

Information measures of correlation 2

$$f_{13} = \left(1 - \exp(-2.0(HXY2 - HXY))\right)^{\frac{1}{2}} \quad (\text{IV.17})$$

Maximal correlation coefficient

$$f_{14} = (\text{second plus grande valeur propre de } Q)^{\frac{1}{2}} \quad (\text{IV.18})$$

La méthode de [Haralick et al. \(1973\)](#) a montré des performances, en général, meilleures vis-à-vis des méthodes de l'état de l'art de son époque selon plusieurs études comparatives ([Ohanian et Dubes, 1992](#); [Weszka et al., 1976](#)). En revanche, elle souffre d'une complexité temporelle relativement grande.

Du point de vue d'amélioration de la performance et/ou l'efficacité des MCNG, quelques tentatives ont été faites dans ce sens. Dans ce contexte, et étant donné que les éléments des matrices de cooccurrences peuvent avoir des contributions différentes aux caractéristiques globales, [Walker et al. \(1995\)](#) génèrent une matrice de discrimination qui permet de décrire la force de discrimination de chaque élément de la matrice de cooccurrence. Ainsi, l'influence des éléments de la matrice de cooccurrence qui ont une faible puissance de discrimination sera supprimée et celle des éléments avec une grande puissance de discrimination sera améliorée. Pour une recherche plus rapide d'images basée sur le contenu, [Aksoy et Haralick \(1998\)](#) proposent d'exploiter une seule caractéristique parmi les quatorze décrites par [Haralick et al.](#), calculée pour cinq distances (de 1 à 5) et quatre orientations. Les auteurs prennent en compte la variance, sous prétexte que cette caractéristique prend de grandes valeurs pour les images avec grandes variations locales, et vice-versa. [Arivazhagan et Ganesan \(2003\)](#) combinent les caractéristiques statistiques des ondelettes avec celles des matrices de cooccurrences établies sur les différentes sous-bandes d'ondelettes. [Palm \(2004\)](#) s'intéresse à la classification d'autre type de texture qui est la texture couleur. L'auteur a étudié les approches existantes utilisant les caractéristiques de cooccurrence. Puis, il a proposé une nouvelle extension de MCNG qui exploite l'information couleur suivant deux approches intégratives (*single and multi channel*)

(cf. sec. III.5.3). [Gelzinis et al. \(2007\)](#) ont vu que les coefficients de Haralick sont généralement calculés à partir de la matrice de cooccurrence par rapport à une seule distance. Pour remédier au problème de sélection de distances, ils ont proposé une nouvelle approche pour l'utilisation efficace d'informations disponibles de différentes matrices de cooccurrence calculées selon différentes distances. Pour cela, les auteurs ajustent un polynôme de degré n pour chacun des quatorze coefficients de Haralick. Les paramètres du polynôme sont les nouvelles caractéristiques. Elles caractérisent la tendance de variation des coefficients de Haralick selon la variation de la distance. Les caractéristiques locales sont exploitées par [Ribaric et Lopar \(2012\)](#) pour la reconnaissance de l'empreinte palmaire. Pour cela, l'image de la région d'intérêt est divisée en plusieurs sous-images, puis une MCNG est établie pour chacune de ces sous-images. L'énergie, le contraste et la corrélation sont pris en compte par les auteurs pour décrire les MCNG.

Par ailleurs, les MCNG ont été largement exploitées dans plusieurs domaines d'application. Par exemple, les auteurs dans ([Xuejun et al.,](#)) exploitent la texture pour classifier le degré de la fibrose hépatique. Pour cela, ils calculent 15 caractéristiques sur des MCNG extraites des régions d'intérêts dans le foie. Dans ([YongHua et Jin-Cong, 2015](#)), les auteurs combinent les caractéristiques de Tamura et des MCNG pour l'identification de défauts sur la surface du bois. [Tang et al. \(2015\)](#) adaptent et combinent la méthode LBP et les MCNG pour la classification en temps réel des feuilles de thé. [Partio et al. \(2002\)](#) ont investigué les MCNG pour la recherche d'images de rochers, les résultats des expériences conduites montrent des résultats supérieurs à ceux obtenus par les filtres de Gabor.

IV.1.2 Les descripteurs de Tamura

[Tamura et al. \(1978\)](#) ont proposé six caractéristiques qui sont : la grossièreté, le contraste, la directivité, la ressemblance de ligne, la régularité, et la rugosité. Leur travail est distingué par le fait que ces caractéristiques sont des approximations computationnelles de la perception visuelle humaine. Parmi ces six caractéristiques, seulement les trois premières sont les plus exploitées à cause de leur grande force de discrimination ([Tamura et al., 1978](#)).

IV.1.2.1 La grossièreté

La première étape pour calculer la grossièreté est le calcul de la moyenne des valeurs voisines pour chaque point $f(i, j)$ comme suit :

$$A_k(x, y) = \sum_{i=x-2^{k-1}}^{x+2^{k-1}-1} \sum_{j=y-2^{k-1}}^{y+2^{k-1}-1} f(i, j) / 2^{2k} \quad (\text{IV.19})$$

Où $k = 1 \dots \max_k$

Puis, le calcul des différences entre les paires de moyennes qui s'oppose (horizontal et vertical) pour chaque point. Par exemple, pour la direction horizontale la différence est calculée comme suit :

$$E_{k,h}(x, y) = |A_k(x + 2^{k-1}, y) - A_k(x - 2^{k-1}, y)| \quad (\text{IV.20})$$

Troisièmement, la valeur de k qui maximise E dans l'équation (Eq.IV.22), est prise pour l'équation suivante (Eq.IV.21) :

$$S_{best}(x, y) = 2^k \quad (IV.21)$$

$$E_k = E_{max} = \max(E_1, E_2, \dots, E_L) \quad (IV.22)$$

Finalement, la mesure F_{crs} de la grossièreté est la moyenne des S_{best} par rapport à l'image :

$$F_{crs} = \frac{1}{m \times n} \sum_i^m \sum_j^n S_{best}(i, j) \quad (IV.23)$$

IV.1.2.2 Le contraste

Le calcul du contraste est obtenu via l'équation suivante :

$$F_{con} = \sigma / (\alpha_4)^n \quad (IV.24)$$

Où n est un nombre positive (la valeur 0.25 est trouvée qu'elle est la meilleure), et :

$$\alpha_4 = \frac{\mu_4}{\sigma^4} \quad (IV.25)$$

Où μ_4 est le 4^{ème} moment autour de la moyenne, et σ^2 est la variance.

IV.1.2.3 La directivité

Pour trouver la directivité, [Tamura et al.](#), ont utilisé la magnitude et la direction du vecteur de gradient. Ces deux mesures sont approximées par les deux équations suivantes :

$$|\Delta G| = (|\Delta_H| + |\Delta_V|) / 2 \quad (IV.26)$$

$$\theta = \tan^{-1}(\Delta_V / \Delta_H) + \pi / 2 \quad (IV.27)$$

Ainsi, la directivité est :

$$F_{dir} = 1 - r \cdot n_p \cdot \sum_p^{n_p} \sum_{\Phi \in w_p} (\phi - \phi_p)^2 \cdot H_D(\phi) \quad (IV.28)$$

Où n est le nombre des pics, ϕ_p est la position du pic de H_D , w_p est la portée du p^{ème} pic entre vallées et H_D est l'histogramme obtenu par θ en comptant seulement les points dont la magnitude est supérieure à t .

Tandis que les caractéristiques de Tamura sont basées sur l'approximation de la perception visuelle de l'être humain, elles semblent insuffisantes et souffrent, en plus, de quelques inconvénients. En effet, cette méthode a montré de mauvaises performances par rapport à d'autres méthodes ([Doshi et Schaefer, 2012](#); [Wei-Ying et Hongjiang, 1998](#)).

[Wei-Ying et Hongjiang \(1998\)](#) comparent un ensemble de descripteurs de couleur et de texture, parmi les descripteurs de texture était une version modifiée des caractéristiques de Tamura. Les modifications consistent dans la représentation de la grossièreté et la directivité par un histogramme pour chacune au lieu d'une seule valeur. [Islam et al. \(2008\)](#) découvrent un inconvénient sérieux dans l'équation de la directivité, où les bins lointains du pic contribuent plus à la directivité (le terme $(\phi - \phi_p)^2$ dans l'équation IV.28). Alors ils proposent de calculer la directivité par la prise en compte de quelques propriétés géométriques telles le calcul des pentes des segments composants la colline (*hill*) dans l'histogramme. Ces pentes vont être pondérées et totalisées afin de déterminer son acuité.

[Li et al. \(2013\)](#); [Liu et al. \(2013\)](#) proposent une amélioration pour le calcul de contraste en utilisant les moments statistiques, l'uniformité et l'entropie de l'histogramme d'intensité. Quant à [Chih-Yi et al. \(2003\)](#) ont profité du fait que les caractéristiques de Tamura sont perceptibles, en reliant ces caractéristiques de bas niveau avec des concepts humains de haut niveau. Cette liaison est assurée par une fonction d'appartenance floue.

Afin qu'ils puissent exploiter les caractéristiques de Tamura dans les images de fluorescence, [Majtner et Svoboda \(2012\)](#) ont étendu ces caractéristiques à l'espace 3D, en adaptant les formules de la grossièreté et la directivité. Lors de leur comparaison d'un ensemble de métriques d'histogramme, [Zhang et Canosa \(2014\)](#) gèrent les valeurs de Tamura comme une jointure spatiale de la distribution Grossièreté-Contraste-Directivité (GCD/CND).

Récemment, des auteurs ont fait recours à la combinaison des caractéristiques de Tamura avec des descripteurs de même attribut (Texture) comme celle dans ([YongHua et Jin-Cong, 2015](#)), et aussi, avec des descripteurs de couleur comme celle dans ([Yue et al., 2015](#)).

IV.1.3 Les motifs binaires locaux

Les *motifs binaires locaux* (LBP, *Local Binary Patterns*) est l'une des rares méthodes qui ont réuni la simplicité, l'efficacité et les hautes performances. Grâce à ces caractéristiques rarement réunies dans une méthode, la LBP s'est prodigieusement répandue dans la communauté des chercheurs en vision par ordinateur. D'où un nombre important de variantes qui ne cesse de s'accroître.

La méthode LBP est introduite sous sa version de base par [Ojala et al. \(1996\)](#) dans une étude comparative en tant qu'opérateur agissant sur des blocs 3×3 pixels. Leur travail s'est basé sur l'hypothèse stipulant que la texture possède localement deux aspects complémentaires qui sont le motif et sa puissance (*the pattern and its strenght*). La LBP procède par un seuillage du voisinage en prenant le pixel central comme seuil, les pixels dont les valeurs sont supérieures ou égales à celle du pixel central prendront la valeur 1, et la valeur 0 est prise par les autres pixels (inférieures) (Figure IV-3(b)). Ces valeurs sont ensuite multipliées par des puissances de deux (Figure IV-3(d)) et totalisées pour avoir la nouvelle valeur du pixel central. Ainsi, la LBP peut être considérée comme la concaténation des directions de gradients binaires (binary gradient directions) est appelée *micropattern* ([Zhang et al., 2010](#)). La Figure IV-4 illustre les images LBP (LBP maps) des deux images à droite de la figure.

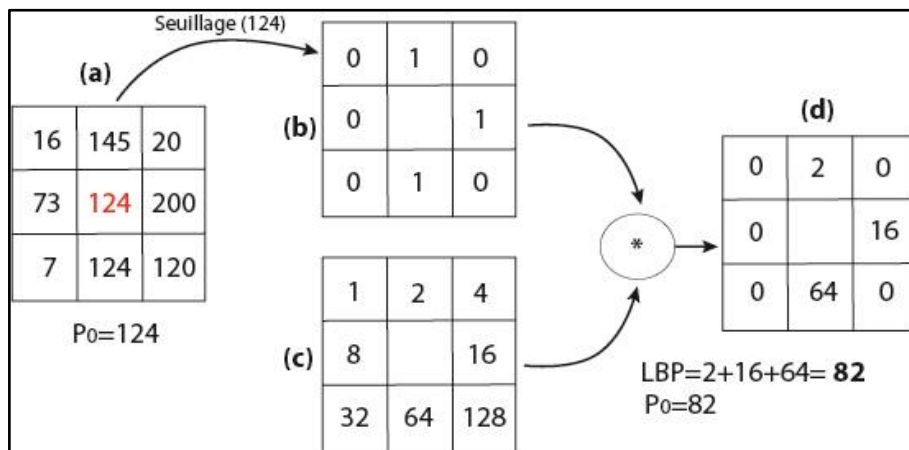


Figure IV-3 LBP de base. Exemple de calcul.

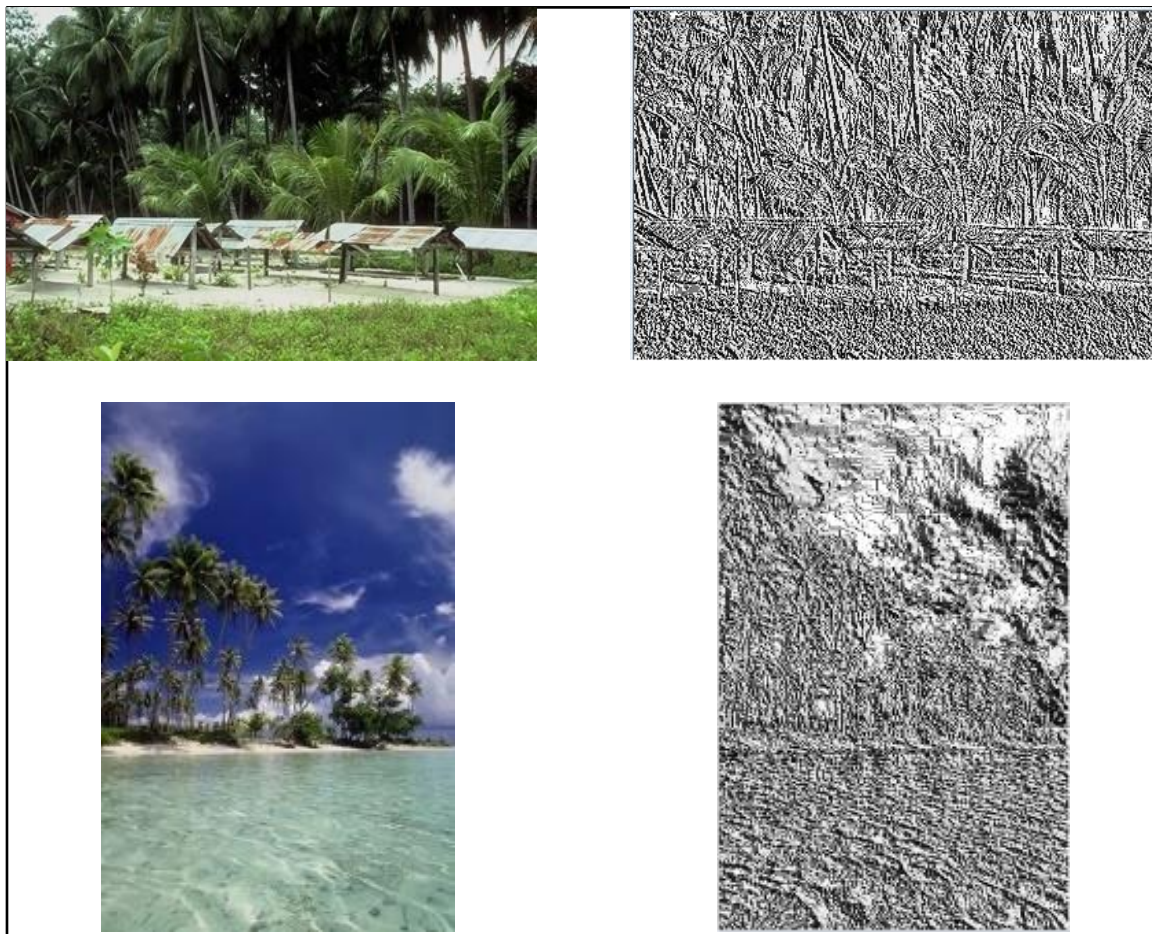


Figure IV-4 Exemples d'image LBP, (à gauche) : deux images en couleur, (à droite) : les images LBP correspondantes

Il est digne de mentionner ici que LBP est une version à deux niveaux d'*unité de texture* (TU, *Texture Unit*) introduite dans (Wang et He, 1990). Le voisinage de TU est tri-value (0, 1, 2), ce qui donne $3^8 = 6561$ valeurs possibles.

Les restrictions sur le rayon et le voisinage posées par LBP de base ont été ensuite levées dans sa version révisée (Ojala et al., 2002). Par conséquent, un nombre quelconque de pixels voisins peuvent être à n'importe quelle distance par rapport au centre (Figure IV-5).

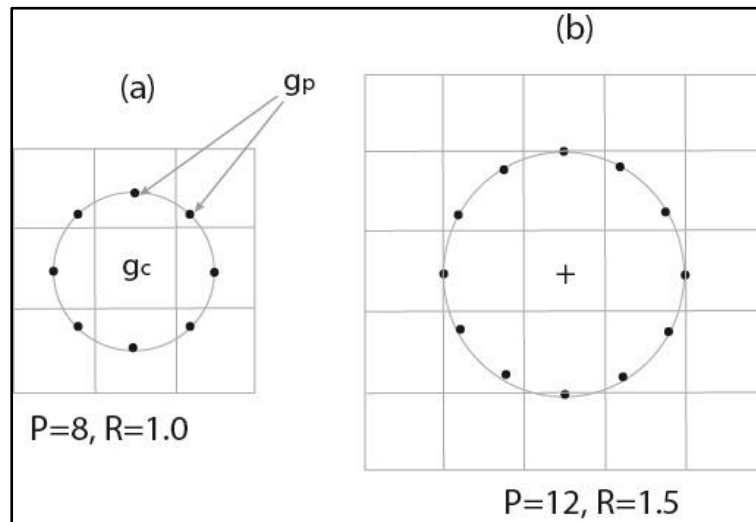


Figure IV-5 LBP générale.

Supposant que P voisins se trouvent à une distance R autour un point $g_c = (x_c, y_c)$. Les positions des voisins $g_p = (x_p, y_p)$ sont déterminées par les équations ci-dessous :

$$x_p = x_c + R \cos(2\pi p / P) \quad (\text{IV.29})$$

$$y_p = y_c - R \sin(2\pi p / P) \quad (\text{IV.30})$$

Où $p = 0, \dots, P-1$

Le nombre unique qui caractérise la structure spatiale de l'image locale de texture $LBP_{P,R}$ est alors (Ojala et al., 2002) :

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad (\text{IV.31})$$

Où

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (\text{IV.32})$$

Comme il a été rapporté au début de cette section, grâce à l'efficacité et les bonnes performances de la LBP, elle s'est popularisée rapidement dans la communauté des chercheurs.

Par conséquent, elle est ensuite devenue l'objet de plusieurs améliorations et extensions. D'où un nombre énorme de variantes qui ne cesse de s'accroître. Généralement, les variantes de la LBP sont développées en modifiant le voisinage, le seuillage, le nombre de seuils, la dimensionnalité, la prise en compte d'un type spécifique de motifs, ou tout simplement la combinaison avec d'autres méthodes. Dans ce contexte, [Huang et al. \(2011\)](#) distinguent entre cinq catégories des améliorations sur la LBP : 1) l'amélioration des capacités discriminatives ; 2) renforcement de la robustesse ; 3) sélection du voisinage ; 4) extension aux données 3D ; et 5) combinaison avec d'autres approches.

IV.1.3.1 Variantes du LBP

Comme il a été avancé ci-dessus, un nombre important de variantes LBP ont été vues le jour depuis l'apparition de cette dernière. Donc, vu l'impossibilité de prendre en compte toutes les variantes, nous avons décidé de présenter quelques méthodes que nous avons vues, qu'il est digne de les mentionner. Dans la seconde sous-section (sec. IV.1.3.1.2), nous concentrerons sur la nouveauté des variantes. À cet effet, seulement les variantes les plus récentes vont être présentées. Enfin, nous pouvons citer ici quelques références qui s'intéressent à la méthode LBP et ses variantes, comme le livre de [Pietikäinen et al. \(2011a\)](#) et aussi [Local Binary Patterns: New Variants and Applications, 2014](#).

IV.1.3.1.1 Les variantes populaires

IV.1.3.1.1.1 LBP invariante à la rotation et LBP uniforme ([Ojala et al., 2002](#))

En plus de la version révisée d'LBP de base, [Ojala et al. \(2002\)](#) proposent dans le même papier deux versions supplémentaires d' $LBP_{P,R}$; $LBP_{P,R}^{ri}$ et $LBP_{P,R}^U$.

$LBP_{P,R}^{ri}$ est une extension invariante à la rotation. Elle est obtenue en décalant circulairement le code LBP vers sa représentation minimale par l'équation (Eq.IV.33). Où $ROR(x, i)$ signifie le décalage circulaire de la séquence de bits x , i fois vers la droite.

Par exemple, les codes LBP de 8-bits suivants : '01000011b' (67d), '10000110b' (134d), et '10100001b' (161d) sont tous transformés à la séquence suivante '00001101b' (13d). Donc, 36 motifs binaires locaux uniques sont distingués pour $P = 8$.

$$LBP_{P,R}^{ri} = \min\{ROR(LBP_{P,R}, i) \mid i = 0, 1, \dots, P-1\} \quad (IV.33)$$

[Ojala et al.](#), ont constaté que certains motifs binaires locaux se manifestent avec grande quantité, parfois plus de 90% de tous les motifs 3×3 présents dans les images de texture utilisées. Par conséquent, les auteurs proposent une deuxième extension qui prend en compte cette observation. Pour cela, la mesure d'uniformité $U(\text{'motif'})$ est introduite. Cette mesure compte le nombre de transitions de 0 à 1 ou vice-versa. Ainsi un motif binaire local est uniforme si la mesure d'uniformité est au plus 2. Par exemple, $U(\text{'00011111'}) = U(\text{'11100011'}) = 2$, $U(\text{'00011011'}) = 4$, $U(\text{'10101011'}) = 6$. En effet, $LBP_{P,R}^{u2}$ produit $P(P-1) + 3$ sorties différentes. Par exemple, si $P = 8$ alors le nombre de sorties est 59 (58 étiquètes pour les différents motifs uniformes et une pour ceux non uniformes). En outre de la dominance des motifs uniformes dans les images, la prise en compte des motifs uniformes améliore les résultats de

reconnaissance dans plusieurs applications, due à la résistance au bruit et le nombre réduit des codes LBP possibles ([Pietikäinen et al., 2011b, p. 17](#)). Le nombre des codes LBP uniformes peut être réduit encore plus, en considérant les motifs binaires locaux uniformes (la valeur de U est au plus 2) invariants à la rotation $LBP_{P,R}^{riu2}$ ([Ojala et al., 2002](#)). Cette dernière possède $P + 2$ sorties différentes seulement.

IV.1.3.1.1.2 Improved LBP ([Jin et al., 2004](#))

Improved LBP (ILBP), est une autre variante d’LBP, proposée par [Jin et al.](#), pour la détection de visages. Son principe est basé sur la considération du pixel central en lui donnant le grand poids (8) sous prétexte que le pixel central fournit plus d’informations que ses voisins, dans la plupart des cas. En conséquence, le nombre total de motifs est 511 au lieu des 256 motifs d’LBP. Quant au seuillage, la moyenne du bloc 3×3 est prise comme seuil.

IV.1.3.1.1.3 Median Binary Patterns ([Hafiane et al., 2007](#))

Median Binary Patterns (MBP) ([Hafiane et al., 2007](#)) est une autre variante qui s’est basée aussi sur la modification du seuil. En effet, les auteurs ont pris la valeur médiane dans un bloc 3×3 comme seuil pour le seuillage des valeurs du bloc, y compris la valeur du pixel central, ce qui aboutit à 512 motifs différents.

IV.1.3.1.1.4 Local Ternary Patterns ([Tan et Triggs, 2007](#))

[Tan et Triggs.](#) se sont intéressés à la reconnaissance de visages dans les situations incontrôlées (difficiles). Ils ont vu que l’utilisation de la valeur du pixel central seule comme seuil rend le LBP sensible au bruit, surtout pour les surfaces presque uniformes. Pour cela, les auteurs proposent *Local Ternary Patterns* (LTP). Cette extension quantifie les valeurs dans la zone $\pm t$ autour g_c à 0, tandis que celles supérieures ou égales, à 1, et celles inférieures ou égales, à -1 (Eq.IV.34).

Par conséquent, un code de motif issu de cette opération de seuillage est trivalué. Pour des raisons de simplicité, les auteurs proposent de diviser le code en deux parties : une partie positive et une autre négative. C.-à-d. les (1s) sont gardés comme tels tandis que les (-1s) sont convertis à des zéros pour la partie positive (*Upper Code*). Quant à la partie négative (*Lower Code*), les (-1s) sont convertis à des (1s), alors que les (1s) sont convertis en zéros. En gardant les bits de valeur 0 comme tels dans les deux cas (Figure IV-6).

$$s'(x, g_c, t) = \begin{cases} 1, & x \geq g_c + t \\ 0, & |x - g_c| < t \\ -1, & x \leq g_c - t \end{cases} \quad (\text{IV.34})$$

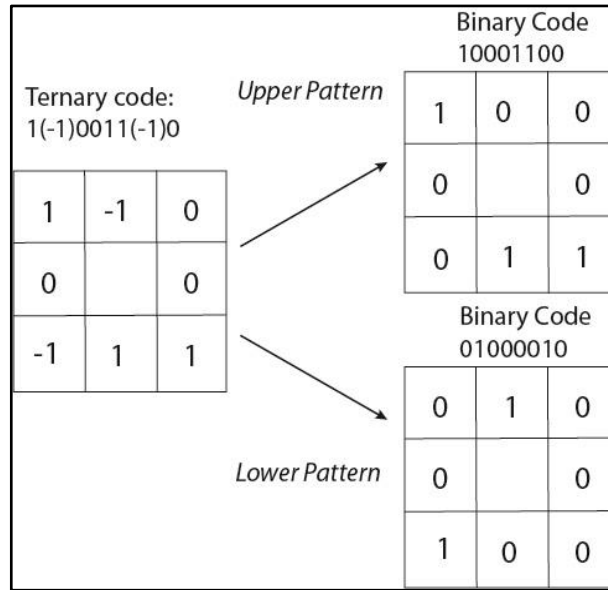


Figure IV-6 LTP, Upper et Lower Patterns.

IV.1.3.1.1.5 Dominant LBP (Liao et al., 2009)

En se basant sur l’observation que les motifs uniformes ne représentent pas nécessairement la grande partie de l’image due que les images en pratique peuvent contenir de fortes courbures, des bordures croisées, et/ou des corners alors que l’LBP uniforme capture les informations des courbures faibles et des bordures droites. Également, vu que sous certains bruits un motif peut changer sa nature ((non)uniforme), Liao et al. (2009) propose *Dominant LBP* (DLBP). Cette méthode se base sur l’apprentissage du nombre de motifs qui représentent 80% du nombre total des motifs présents dans les images. Ce nombre est alors le nombre des motifs considérés lors de la phase de test. Selon les résultats rapportés par les auteurs, DLBP est une méthode performante. En effet, elle a montré des résultats meilleurs que ceux obtenus par quelques méthodes connues de l’état de l’art, tels les filtres de Gabor, les ondelettes... etc.

IV.1.3.1.1.6 Local Derivative Patterns (Zhang et al., 2010)

Zhang et al. (2010) ont tenté de partir au-delà de l’information du premier ordre que LBP peut capturer. De là, ils ont proposé *Local derivative Patterns* (LDP), qui est capable de capturer de l’information d’ordre supérieur, qui contient plus de caractéristiques discriminantes.

Formellement, étant donné les dérivatives du premier-ordre le long des directions 0°, 45°, 90° et 135° dans le point Z_0 , tel que ses voisins sont notés Z_i où $i = 1, \dots, 8$ (Figure IV-7) :

$$I'_{0^\circ}(Z_0) = I(Z_0) - I(Z_4) \tag{IV.35}$$

$$I'_{45^\circ}(Z_0) = I(Z_0) - I(Z_3) \tag{IV.36}$$

$$I'_{90^\circ}(Z_0) = I(Z_0) - I(Z_2) \tag{IV.37}$$

$$I'_{135^\circ}(Z_0) = I(Z_0) - I(Z_1) \tag{IV.38}$$

Z ₁	Z ₂	Z ₃
Z ₈	Z ₀	Z ₄
Z ₇	Z ₆	Z ₅

Figure IV-7 Voisinage du LDP.

LDP directionnel du second ordre dans la direction α est défini comme suit :

$$LDP_{\alpha}^2(Z_0) = \left\{ \begin{array}{l} f(I'_{\alpha}(Z_0), I'_{\alpha}(Z_1)), f(I'_{\alpha}(Z_0), I'_{\alpha}(Z_2)), \\ \dots, f(I'_{\alpha}(Z_0), I'_{\alpha}(Z_8)) \end{array} \right\} \quad (IV.39)$$

Où :

$$f(I'_{\alpha}(Z_0), I'_{\alpha}(Z_i)) = \begin{cases} 0, & \text{if } I'_{\alpha}(Z_i) \cdot I'_{\alpha}(Z_0) > 0 \\ 1, & \text{if } I'_{\alpha}(Z_i) \cdot I'_{alpha}(Z_0) \leq 0 \end{cases} \quad (IV.40)$$

Ainsi, *Local Derivative Pattern* du second ordre $LDP^2(Z)$ est la concaténation des quatre 8-bits directionnels :

$$LDP^2(Z) = \{LDP_{\alpha}^2(Z) | \alpha = 0^{\circ}, 45^{\circ}, 90^{\circ}, 135^{\circ}\} \quad (IV.41)$$

Cette variante est également utilisée avec succès pour la recherche d'images par le contenu ([Reddy et Prasad, 2011](#)).

IV.1.3.1.1.7 Completed LBP ([Guo et al., 2010](#))

Dans leur article, [Guo et al.](#), propose *Completed LBP* (CLBP). Une variante (framework) qui explore trois types de caractéristiques : les deux composants complémentaires des différences, c.-à-d. le signe et la magnitude, et aussi, les niveaux de gris locaux de l'image originale. Par conséquent, trois opérateurs ont été proposés : *CLBP_Sign* (CLBP_S), *CLBP_Magnitude* (CLBP_M), et *CLBP_Center* (CLBP_C).

Sachant que d_p est la différence entre la valeur d'un pixel voisin et celle du centre du bloc auquel il appartient :

$$d_p = g_p - g_c \quad (IV.42)$$

Cette différence peut être décomposée en deux parties, signe et magnitude :

$$d_p = s_p \cdot m_p \quad (IV.43)$$

Où :

$$\begin{cases} s_p = \text{sign}(d_p) \\ m_p = |d_p| \end{cases} \quad (\text{IV.44})$$

Et

$$\text{sign}(x) = \begin{cases} 1, x \geq 0 \\ -1, x < 0 \end{cases} \quad (\text{IV.45})$$

Pour le premier opérateur, CLBP_S qui s'intéresse aux signes des différences est similaire à l'LBP originale (Equation IV.31). Quant au deuxième opérateur, CLBP_M, qui s'intéresse aux magnitudes des différences et afin de garder le codage consistant avec celui du CLBP_S (et donc LBP), les auteurs proposent d'utiliser l'équation suivante :

$$CLBP_M_{P,R} = \sum_{p=0}^{P-1} t(m_p, c) 2^p \quad (\text{IV.46})$$

Où c est un seuil qui sera déterminé d'une manière adaptative. À cet effet, les auteurs adoptent la moyenne des magnitudes m_p sur toute l'image comme seuil. t est une fonction de seuillage définie comme suit :

$$t(x, c) = \begin{cases} 1, x \geq c \\ 0, x < c \end{cases} \quad (\text{IV.47})$$

Quant au CLBP_C qui exprime le niveau de gris local dans l'image et afin de rester consistant avec les codages précédents, le codage suivant a été proposé :

$$CLBP_C_{P,R} = t(g_c, c_l) \quad (\text{IV.48})$$

Où la fonction de seuillage t est définie ci-dessus (Eq.IV.47) et c_l est la moyenne des niveaux de gris sur toute l'image.

IV.1.3.1.1.8 Opponent Color LBP ([Mäenpää et Pietikäinen, 2004](#))

La méthode *Opponent Color LBP* (OCLBP) ([Mäenpää et Pietikäinen, 2004](#)) a été proposée afin d'extraire de l'information texturale à partir des images couleurs. Dans cette variante, la méthode LBP est appliquée sur les trois canaux de l'espace de couleur (r, g, b, pour l'espace de couleur RGB), puis sur chaque paire de canaux en prenant le centre à partir d'un canal et le voisinage à partir d'un autre. Par conséquent, six caractéristiques sont extraites de : r, g, b, r-g, r-b, et g-b (Figure IV-8), vu que les paires opposantes (ex : r-b et b-r) sont redondantes. En faisant cela, l'information texturale et celle de couleur sont gérées conjointement (cf. sec.III.5).

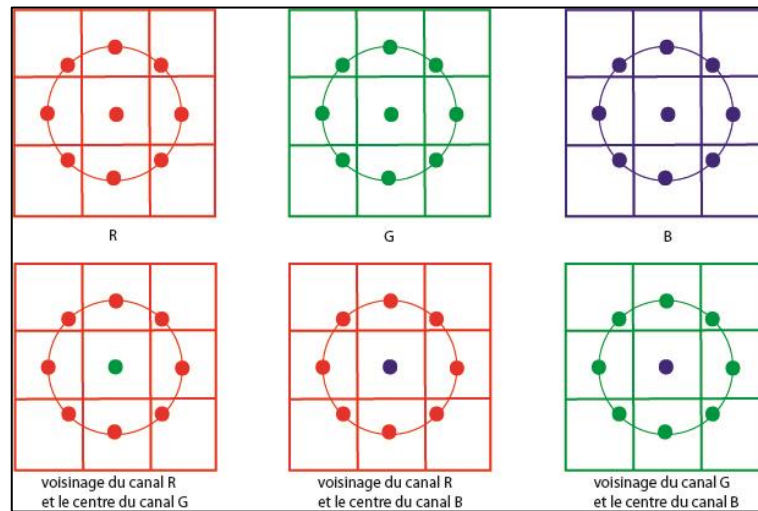


Figure IV-8 Opponent Color LBP.

IV.1.3.1.1.9 Block-based LBP (Takala et al., 2005)

En revanche des autres variantes LBP, Takala et al. (2005) ont proposé d’extraire de l’information texturale à partir des sous-images (blocs), d’où le nom de la méthode. En effet, les auteurs ont proposé deux manières d’utilisation de leur approche. La première consiste à diviser les images de la base ainsi que la requête en blocs. Lors de la comparaison, chaque bloc de la requête est comparé avec tous les blocs d’une image de la base (Figure IV-9) en utilisant l’équation suivante :

$$D = \sum_{i=0}^{N-1} \min_j (D_{i,j}) \tag{IV.49}$$

La deuxième manière d’utilisation de blocs consiste à diviser les images de la base en blocs plus petits et de constante taille (des blocs primitifs). Ces blocs primitifs sont combinés pour correspondre à la taille de la requête (Figure IV-10). Ainsi la somme des histogrammes de ces blocs combinés, représente le vecteur de caractéristiques d’une région dans l’image. Ce vecteur va être comparé avec celui de la requête. En plus de la précision raisonnable lors de la correspondance à la taille de la requête, la méthode est également plus rapide que les méthodes à fenêtres coulissantes basées sur les pixels (c.-à-d. déplacement de la fenêtre pixel par pixel).

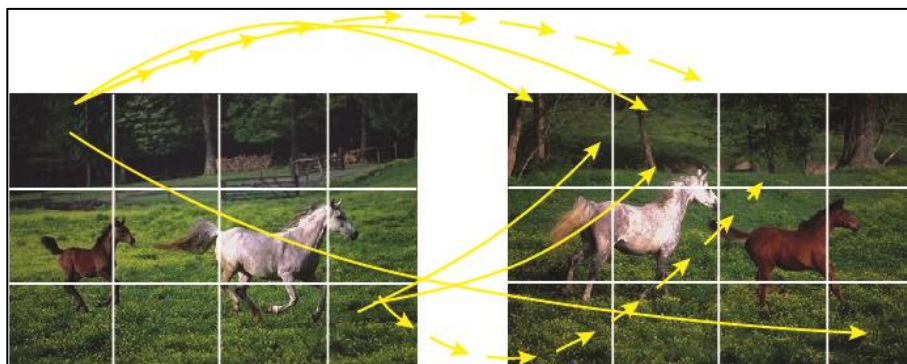


Figure IV-9 La méthode basée sur les blocs.

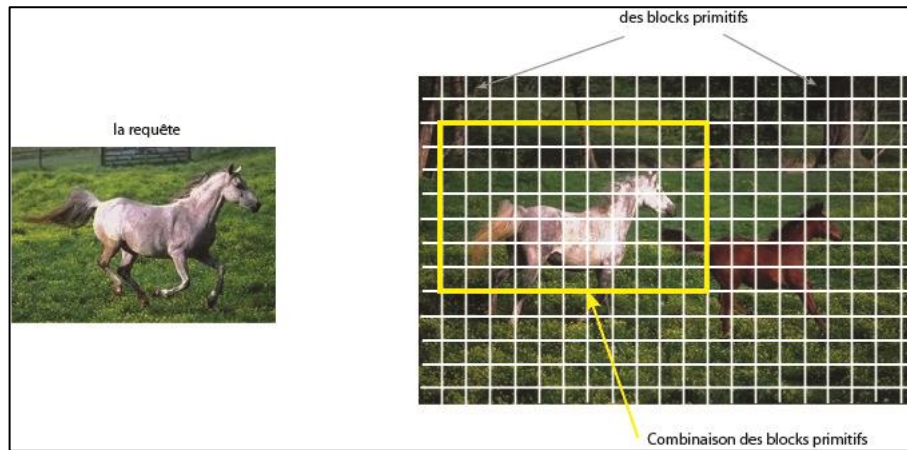


Figure IV-10 La méthode basée sur les blocs primitifs.

À notre connaissance, cette méthode est la seule dans l'état de l'art qui prend en compte l'information locale dans des blocs. Dans ce contexte, nous avons proposé une méthode qui délimite et manipule des régions de l'image différemment. À cet effet, nous avons proposé un schéma original de ces régions, en utilisant des régions elliptiques centrées dans l'image au lieu de simples blocs. Notre méthode sera vue en détail dans la première partie du Chapitre V.

IV.1.3.1.2 Les variantes récentes

IV.1.3.1.2.1 Dominant Rotated Local Binary Patterns ([Mehta et Egiazarian, 2016](#))

[Mehta et Egiazarian](#), ont développé une variante d'LBP qui est invariante à la rotation. Par contre au riLBP qui est basée sur le décalage de la séquence binaire vers sa représentation minimale, [Mehta et Egiazarian](#), proposent un arrangement adaptatif des poids, en les pivotant vers une référence locale au bloc (Eq.IV.51). Cette référence est l'index D du pixel voisin avec lequel la différence est maximale du centre (Eq.IV.50), appelée 'Dominant Direction'. D'où le nom de la méthode *Dominant Rotated Local Binary Patterns* (DRLBP). Vu la considération de la magnitude de la différence, RLBP incorpore plus d'informations que riLBP.

$$D = \arg \max_{p=0,1,\dots,P-1} |g_p - g_c| \quad (\text{IV.50})$$

$$RLBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^{\text{mod}(p-D,P)} \quad (\text{IV.51})$$

IV.1.3.1.2.2 Quaternionic Local Ranking Binary Pattern ([Lan et al., 2016](#))

[Lan et al.](#), se sont intéressés à l'extraction des caractéristiques des images en couleurs. Vu qu'un pixel est représenté par plusieurs composants (par exemple : r, g, b pour l'espace RGB), les auteurs introduisent leur méthode appelée *Quaternionic Local Ranking Binary Pattern* (QLRBP), qui exploite la *représentation quaternionique* (QR : *Quaternionic Representation*). Cette dernière permet de coder les pixels en utilisant des nombres hypercomplexes appelés : *quaternions*²³. Quelques avantages de cette représentation ont été compilés par [Lan et al. \(2016\)](#)

²³ Quaternion est un nombre complexe de quatre dimensions, introduit par Hamilton en 1843. Un quaternion est composé d'une partie réelle et trois parties imaginaires :

comme suit : (1) la combinaison holistique des canaux de couleurs d'une image ; (2) elle a aussi une complexité computationnelle relativement inférieure à celles des autres approches vectorielles ; (3) possède une implémentation de la corrélation croisée de vecteurs ; (4) effectue des transformations dans l'espace 3D ou l'espace 4D commodément ; (5) a fructueux quaternion algèbre théories qui peuvent être utilisés dans le traitement de l'image couleur.

Donc, pour obtenir le descripteur QLRBP, l'image est transformée, en premier, à sa QR. Les auteurs proposent d'appliquer *Clifford Translation of Quaternion* (CTQ) comme fonction de classement quaternionique. Le CTQ à droite de \dot{q} par \dot{p} est montré par l'équation IV.52. Où le quaternion référence \dot{p} est un quaternion unitaire pur.

$$\begin{aligned} CTQ_r(\dot{q}, \dot{p}) &= (ir + jg + kb)(ir' + jg' + kb') \\ &= -(rr' + gg' + bb') + i(gb' - bg') + j(br' - rb') + k(rg' - gr') \end{aligned} \quad (IV.52)$$

Ensuite, *phase image* est calculée à partir de l'image CTQ obtenue en utilisant un quaternion de référence spécifique \dot{p}_0 . En effet, les auteurs proposent d'utiliser L_1 pondéré afin de gérer la variation d'importances de différents canaux de couleurs, calculée par l'équation suivante :

$$\delta = \tan^{-1} \frac{\alpha_1 |gb' - bg'| + \alpha_2 |br' - rb'| + \alpha_3 |rg' - gr'|}{-(rr' + gg' + bb')} \quad (IV.53)$$

Finalement, le codage QLRBP au niveau du pixel central \dot{q}_m d'un bloc 3×3 (comme celui d'LBP), en considérant un quaternion de référence \dot{p}_0 , est réalisé par l'équation suivante :

$$QLRBP_{\dot{q}_m} = \sum_{n=0}^{|S_m|-1} h(R_{QLRBP}(\dot{q}_m, \dot{q}_n)) 2^n \quad (IV.54)$$

Où R_{QLRBP} est la fonction de classement (*ranking function*), utilisée pour le classement du \dot{q}_n et \dot{q}_m :

$$R_{QLRBP}(\dot{q}_m, \dot{q}_n) = \delta_{CTQ}(\dot{q}_n - \dot{p}_0) - \delta_{CTQ}(\dot{q}_m - \dot{p}_0) \quad (IV.55)$$

Pour plus de détails, le lecteur peut se référer à ([Lan et al., 2016](#)).

IV.1.3.1.2.3 Modified LBP ([Naresh et Nagendraswamy, 2016](#))

[Naresh et Nagendraswamy](#), proposent une version modifiée d'LBP appelée *Modified LBP* (MLBP). Cette modification tente de pallier les problèmes dus au seuillage dur que l'LBP utilise (Eq.IV.32). À cet effet, ils proposent d'utiliser la moyenne μ du bloc 3×3 en tant que seuil (Eq.IV.56) et son écart type σ (Eq.IV.57) en tant que marge.

$$\dot{q} = a + ib + jc + kd$$

$$\mu = \frac{\sum_{p=0}^{P-1} g_p + g_c}{p+1} \tag{IV.56}$$

$$\sigma = \sqrt{\frac{\sum_{p=0}^{P-1} (g_p - \mu)^2 + (g_c - \mu)^2}{P+1}} \tag{IV.57}$$

Par conséquent, la fonction de seuillage devient :

$$s_{MLBP}(x) = \begin{cases} 1, & (\mu + \sigma) \geq x \geq (\mu - \sigma) \\ 0, & \text{sinon} \end{cases} \tag{IV.58}$$

IV.1.3.1.2.4 Local Tri-Directionnal Patterns (Verma et Raman, 2016)

Local Tri-Directionnal Pattern (LTriDP) est une variante LBP, qui tente d'extraire plus ample d'informations locales à partir du voisinage. Pour cela, Verma et Raman, calculent la différence de chaque pixel du voisinage avec ses trois voisins dans trois directions (Figure IV-11) d'où le nom de la variante. A cet effet, l'équation suivante est utilisée :

$$\begin{cases} D_1 = g_p - g_{p-1}, D_2 = g_p - g_{p+1}, D_3 = g_p - g_c, \forall p = 2, \dots, 7 \\ D_1 = g_p - g_8, D_2 = g_p - g_{p+1}, D_3 = g_p - g_c, \text{ for } p = 1 \\ D_1 = g_p - g_{p-1}, D_2 = g_p - g_1, D_3 = g_p - g_c, \text{ for } p = 8 \end{cases} \tag{IV.59}$$

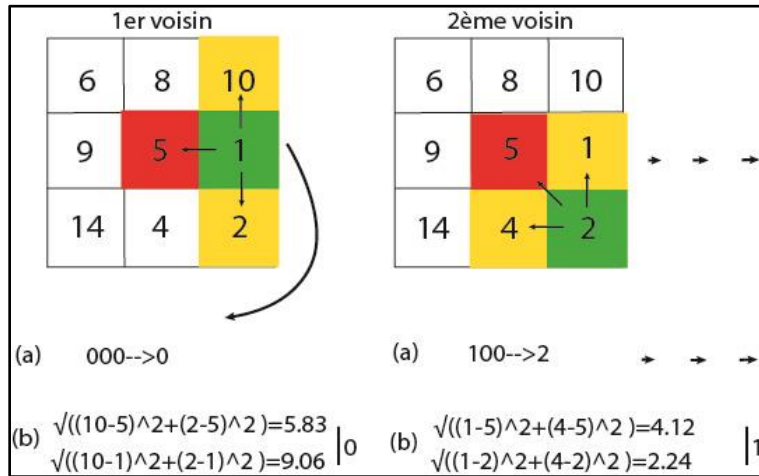


Figure IV-11 LTriDP, exemple.

Le nombre des différences inférieures de zéro est dénombré pour chaque pixel du voisinage (Eq. IV.60) (Figure IV-11(a)). Cette opération aboutit à un motif tri-valué, qui est converti à son tour en deux motifs binaires (Eq.IV.61-IV.62). En conséquence, le pixel central est représenté par deux motifs (Eq.IV.63).

$$f(D_1, D_2, D_3) = \{\#(D_k < 0)\} \text{mod} 3, \quad \forall k = 1, 2, 3 \tag{IV.60}$$

$$LTriDP_1(g_c) = \{S(f_1), S(f_2), \dots, S(f_8)\}$$

$$S(x) = \begin{cases} 1, & \text{si } x = 1 \\ 0, & \text{sinon} \end{cases} \quad (IV.61)$$

$$LTriDP_2(g_c) = \{S'(f_1), S'(f_2), \dots, S'(f_8)\}$$

$$S'(x) = \begin{cases} 1, & \text{si } x = 2 \\ 0, & \text{sinon} \end{cases} \quad (IV.62)$$

$$LTriDP(g_c) \Big|_{i=1,2} = \sum_{l=0}^7 2^l * LTriDP_l(g_c)(l+1) \quad (IV.63)$$

En outre des informations extraites jusqu'à présent, qui sont, en effet, basées sur le signe des différences, les auteurs proposent d'intégrer plus d'informations complémentaires en calculant les magnitudes des différences des deux pixels voisins de chacun des huit pixels avec ces derniers (Eq.IV.65) et avec le pixel central (Eq.IV.64) (Figure IV-11(b)).

$$M_1 = \begin{cases} \sqrt{(I_{i-1} - I_c)^2 + (I_{i+1} - I_c)^2}, & \forall i = 2, 3, \dots, 7 \\ \sqrt{(I_8 - I_c)^2 + (I_{i+1} - I_c)^2}, & \text{for } i = 1 \\ \sqrt{(I_{i-1} - I_c)^2 + (I_1 - I_c)^2}, & \text{for } i = 8 \end{cases} \quad (IV.64)$$

$$M_2 = \begin{cases} \sqrt{(I_{i-1} - I_i)^2 + (I_{i+1} - I_i)^2}, & \forall i = 2, 3, \dots, 7 \\ \sqrt{(I_8 - I_i)^2 + (I_{i+1} - I_i)^2}, & \text{for } i = 1 \\ \sqrt{(I_{i-1} - I_i)^2 + (I_1 - I_i)^2}, & \text{for } i = 8 \end{cases} \quad (IV.65)$$

Les deux magnitudes de chaque pixel voisin sont ensuite comparées (Eq.IV.66), puis le motif associé au pixel central est extrait (Eq.IV.67-IV.68). Finalement, la concaténation des trois histogrammes donne le descripteur final.

$$Mag_i(M_1, M_2) = \begin{cases} 1, & \text{si } M_1 \geq M_2 \\ 0, & \text{sinon} \end{cases} \quad (IV.66)$$

$$LTriDP_{mag}(g_c) = \{Mag_1, Mag_2, \dots, Mag_8\} \quad (IV.67)$$

$$LTriDP(g_c) \Big|_{mag} = \sum_{l=0}^7 2^l \cdot LTriDP_{mag}(g_c) \quad (IV.68)$$

IV.1.3.1.2.5 LBP by neighborhood and Directionnal LBP ([Kaya et al., 2015](#))

[Kaya et al.](#), proposent deux variantes d'LBP, en prenant différentes combinaisons de pixels voisins en relation. La première variante appelée *Local Binary Patterns by neighborhoods* (nLBP_d). Elle dépend de la relation entre les pixels voisins autour du pixel central qui se trouvent à une distance d entre eux (Figure IV-12(a)) :

$$P_C = \left\{ \begin{array}{l} s(p_0 > p_1), s(p_1 > p_2), s(p_2 > p_3), s(p_3 > p_4), s(p_4 > p_5), \\ s(p_5 > p_6), s(p_6 > p_7), s(p_7 > p_0) \end{array} \right\} \quad (IV.69)$$

Où :

$$s(p_i > p_j) = \begin{cases} 1, & \text{si } p_i > p_j \\ 0, & \text{sinon} \end{cases} \quad (IV.70)$$

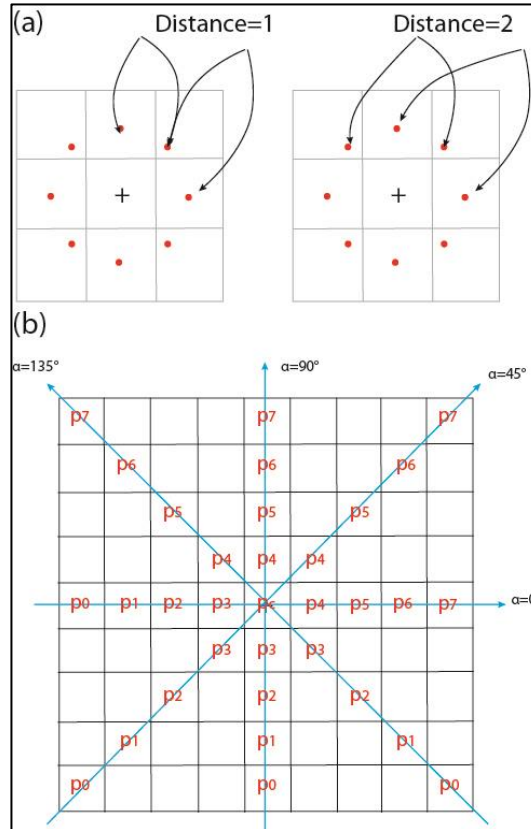


Figure IV-12 (a) : nLBP, (b) : dLBP.

La deuxième variante appelée *Directional Local Binary Patterns* ($dLBP_\alpha$), elle prend en compte la relation des pixels qui sont sur le même support passant par le centre (c.-à-d. qui ont la même direction). A cet effet, quatre directions distinctes sont prises en compte : $\alpha = 0^\circ, 45^\circ, 90^\circ, 135^\circ$ (Figure IV-12(b)).

IV.1.3.1.2.6 Power LBP (Smolka et Nurzynska, 2015)

Une autre variante d’LBP basée sur un nouveau schéma a été proposée par Smolka et Nurzynska. Dans cette méthode, l’ordre des pixels est basé sur les distances cumulées entre la valeur de chaque pixel voisin et celles de tous les autres pixels (Eq.IV.71), au lieu de l’intensité.

$$D_{x_i} = \sum_{k=0}^P |g_i - g_k|^\gamma \quad (IV.71)$$

Où $i = 0, 1, \dots, P$ et la puissance $\gamma > 0$, d’où le nom de la méthode *Power LBP*.

Par conséquent, le code *power-LBP* d'un pixel x_0 est obtenu comme suit :

$$powerLBP_{p,R}(x_0) = \sum_{p=1}^P s(D_p - D_0) 2^{p-1} \tag{IV.72}$$

IV.1.3.1.2.7 SS-3D-LTP ([Murala et Wu, 2015](#))

[Murala et Wu](#), proposent une nouvelle variante d'LBP appelée *Spherical Symmetric three Dimensional Local Ternary Patterns* (SS-3D-LTP). Cette dernière et contrairement aux autres variantes qui établissent les codes de motifs à partir du plan de l'image (2D), le SS-3D-LTP établit les codes des motifs à partir d'un plan trois dimensionnels généré de l'image en utilisant une banque de filtres gaussiens multi-résolutions. Les codes de motifs qui encodent la relation du pixel central avec ses voisins (dans le plan 3D) sont établis suivant cinq directions (Figure IV-13). Clairement, le codage est obtenu en utilisant la variante LTP (cf. sec. IV.1.3.1.1.4). Pour plus de détails le lecteur peut se référer à [Murala et Wu, 2015](#).

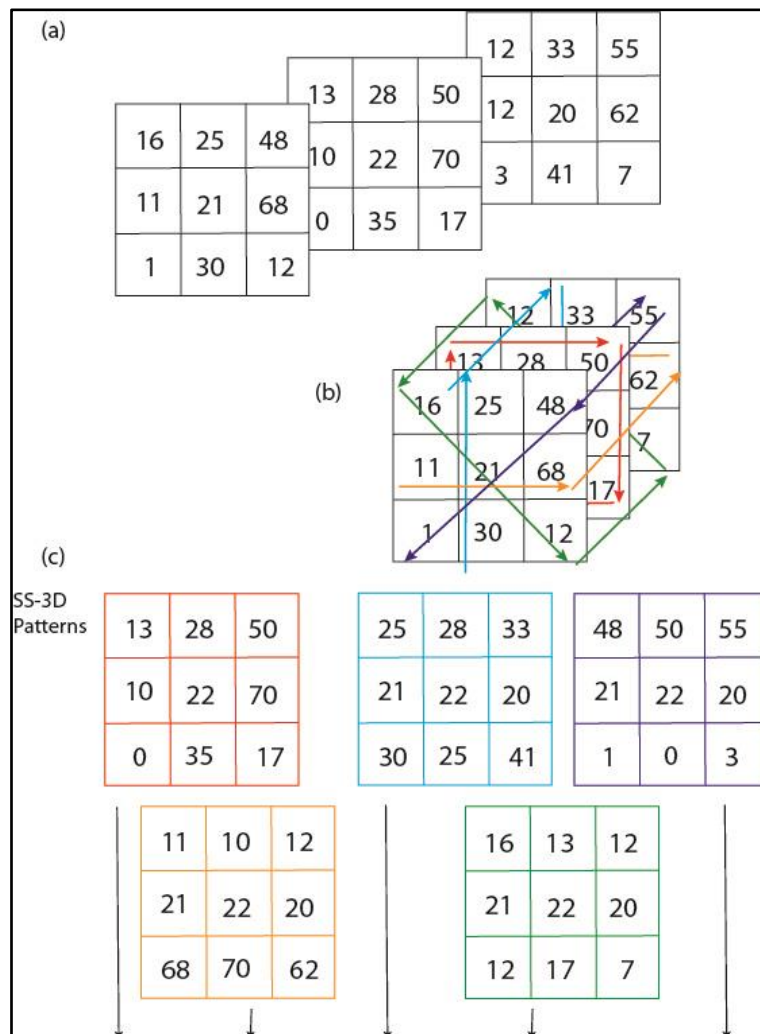


Figure IV-13 SS-3D-LTP.

IV.1.3.1.2.8 Enhanced LBP (Wong et al., 2013)

Wong et al., s'intéressent à la classification des images de la radiographie pulmonaire. Pour cela, ils proposent de prendre en compte, en même temps, les deux pixels qui s'opposent par rapport au pixel central dans un voisinage donné (Figure IV-14), afin de capturer les propriétés des contours (bords) saillants dans la région de la cage thoracique et qui peuvent être utiles pour la caractérisation de telles images (Wong et al., 2013). À cet effet, les auteurs proposent leur variante appelée *Enhanced LBP* (ELBP). Elle adopte l'équation suivante pour le calcul de codes de motifs :

$$ELBP_{P,R}(x, y) = \sum_{i=0}^{P/2-1} (s((g_r - g_c), (g_l - g_c)))2^i \quad (IV.73)$$

Où :

$$s(v_1, v_2) = \begin{cases} 1, & \text{if } v_1 > 0 \text{ and } v_2 \leq 0 \text{ or } v_1 = 0 \text{ and } v_2 < 0 \\ 0, & \text{if } v_1 < 0 \text{ and } v_2 \geq 0 \text{ or } v_1 = 0 \text{ and } v_2 > 0 \\ -1, & \text{otherwise} \end{cases} \quad (IV.74)$$

g_r et g_l représentent les pixels voisins droits et gauches, respectivement. Clairement, d'après l'équation ci-dessus (Eq.IV.73), il y aura $P/2$ paires de pixels voisins (Figure IV-14).

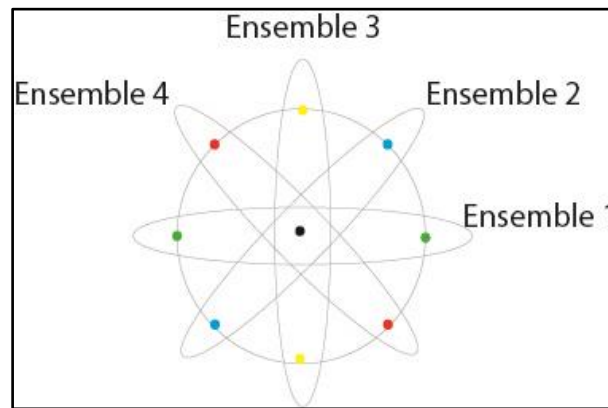


Figure IV-14 Enhanced LBP.

IV.1.3.2 Applications

Vu la simplicité et les intéressantes performances de la méthode LBP, les chercheurs en ont profité pour résoudre les différents problèmes rencontrés. Par exemple, la méthode LBP (ou variante) a été utilisée pour la détection de visage (Jin et al., 2004) ; la reconnaissance de visage (Ahonen et al., 2006; Liu et al., 2016; Zhang et al., 2010) ; la classification démographique (Yang et Ai, 2007) ; la classification de genre de la musique (Costa et al., 2012; Nanni et al., 2016) ; la classification de genre (Hadid et al., 2015; Patel et al., 2016) ; la reconnaissance visuelle d'incendie (Jing et al., 2010) ; la reconnaissance des expressions faciales (Chao et al., 2015) ; la détection des lunettes sur des images de visages (Fernández et al., 2015) ; la

reconnaissance d'iris ([Li, Zhou, et al., 2014](#)); la reconnaissance d'émotions ([Smolka et Nurzynska, 2015](#)); la classification des feuilles de thé ([Tang et al., 2015](#)); la classification des plantes médicinales ([Naresh et Nagendraswamy, 2016](#)); la classification des espèces de bois dur ([Yadav et al., 2015](#)); l'estimation de la composition botanique dans les champs mixtes ([McRoberts et al., 2016](#)). Également, elle a été utilisée dans le domaine médical pour la classification de la radiographie pulmonaire ([Wong et al., 2013](#)); l'analyse des images médicales ([Nanni et al., 2010](#)); et la détection de masse mammographique ([Oliver et al., 2007](#)).

IV.1.4 Les ondelettes

IV.1.4.1 Présentation des ondelettes

Les ondelettes sont des outils d'analyse de signal. Elles se distinguent par leur capacité de localisation en fréquence et en espace (temps). Mathématiquement, les ondelettes sont définies en tant qu'une famille de fonctions qui sont générées par dilatation (a) et translation (b) d'une seule fonction h ([Daubechies, 1988](#)) :

$$h_{a,b}(x) = |a|^{-1/2} h\left(\frac{x-b}{a}\right), \quad a, b \in \mathbb{R}, a \neq 0 \quad (\text{IV.75})$$

Pour certaines applications, telle l'analyse du signal, ces coefficients devraient prendre des valeurs discrètes. Par conséquent, l'équation précédente devient selon [Daubechies](#), comme suit :

$$h_{m,n}(x) = a_0^{-m/2} h(a_0^{-m}x - nb_0), \quad m, n \in \mathbb{Z} \quad (\text{IV.76})$$

Dans ([Daubechies, 1988](#)), a_0 et b_0 sont fixés à 2 et 1, respectivement. Généralement, la dilatation et la translation sont basées sur la puissance de 2, notamment pour le traitement du signal.

Dans cette optique, les ondelettes sont capables d'analyser un signal (une fonction) dans différentes résolutions (multi-résolutions). En effet, l'approximation d'une fonction f à la résolution 2^{-j} est définie comme la projection orthogonale dans l'espace V_j . Cette approximation est caractérisée par la fonction d'échelles ϕ (*Scaling Function*) qui génère une base orthogonale pour chaque espace V_j ([Stéphane, 2009](#)). Elle est définie comme suit (cf. [théorème 7.1 dans Stéphane, 2009](#)) :

$$\phi_{j,n}(t) = \frac{1}{\sqrt{2^j}} \phi\left(\frac{t-n}{2^j}\right) \quad (\text{IV.77})$$

Où la famille $\{\phi_{j,n}\}_{n \in \mathbb{Z}}$ est une base orthonormée de V_j pour tout $j \in \mathbb{Z}$.

Maintenant, étant donné V_j et V_{j-1} deux espaces d'approximations, sachant que $V_j \subset V_{j-1}$ (voir Figure IV-15), l'approximation de l'espace V_j est plus grossière que celle de l'espace

V_{j-1} . Donc, le passage de V_j à V_{j-1} nécessite quelques informations (détails). En effet, ces détails se trouvent dans W_j le complément orthogonal de V_j dans V_{j-1} :

$$V_{j-1} = V_j \oplus W_j \quad (\text{IV.78})$$

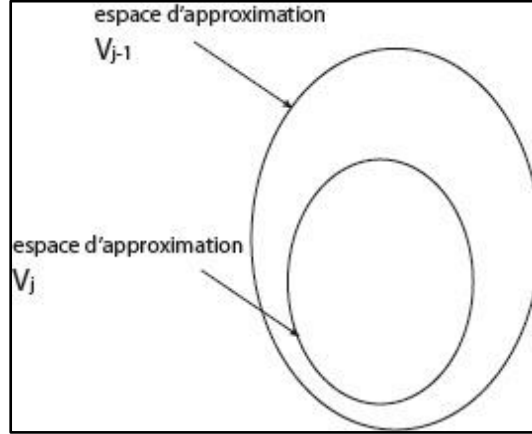


Figure IV-15 Les espaces d'approximation.

La base orthonormée de W_j , est obtenue par mise à échelle et translation de l'ondelette ψ (cf. [théorème 7.3 dans Stéphane, 2009](#)) :

$$\psi_{j,n}(x) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{t - 2^j n}{2^j}\right) \quad (\text{IV.79})$$

Où la famille $\{\psi_{j,n}\}_{n \in \mathbb{Z}}$ est une base orthonormée de W_j pour tout $j \in \mathbb{Z}$.

En pratique, la transformée en ondelettes, plus précisément la transformée en ondelettes rapide (*Fast Discrete Wavelet Transform*, (FDWT)) est obtenue en cascasant des filtres à miroir conjugué (CMF : *Conjugate Mirror Filters*) ([Mallat, 1989](#)). Donc, la fonction (le signal) est convolutive par deux types de filtres : *un filtre passe-bas* (h) et *un filtre passe-haut* (g) comme il est montré par les équations (Eq.IV.80-IV.81) (cf. [le théorème 7.10 dans Stéphane, 2009](#)). La convolution est suivie par une opération de sous-échantillonnage.

$$a_{j+1}[p] = \sum_{n=-\infty}^{+\infty} h[n-2p] \cdot a_j[n] = a_j * \bar{h}[2p] \quad (\text{IV.80})$$

$$d_{j+1}[p] = \sum_{n=-\infty}^{+\infty} g[n-2p] \cdot a_j[n] = a_j * \bar{g}[2p] \quad (\text{IV.81})$$

Où $\bar{x}[n] = x[-n]$

Lors de la cascade des filtres, les coefficients qui représentent l'approximation à l'échelle 2^j (c.-à-d. a_j) sont utilisés pour le calcul de l'approximation et les détails à l'échelle suivante (c.-à-d. 2^{j+1}).

IV.1.4.2 Les ondelettes et les images

Vu que les images sont des données 2-dimensionnelles, les ondelettes doivent être adaptées à ce type de données de dimension supérieure. Pour cela, l'approximation de l'image $f(x_1, x_2)$ à la résolution 2^{-j} est définie comme la projection orthogonale de f dans V_j^2 , sachant que $V_j^2 = V_j \otimes V_j$. Ainsi, la base orthonormée de V_j^2 pour $x = (x_1, x_2)$ et $n = (n_1, n_2)$, est définie comme suit ([Stéphane, 2009](#)) :

$$\left\{ \begin{array}{l} \phi_{j,n}^2(x) = \phi_{j,n_1}(x_1)\phi_{j,n_2}(x_2) \\ = \frac{1}{2^j} \phi\left(\frac{x_1 - 2^j n_1}{2^j}\right) \phi\left(\frac{x_2 - 2^j n_2}{2^j}\right) \end{array} \right\}_{n \in \mathbb{Z}^2} \quad (\text{IV.82})$$

La même chose pour les signaux d'une seule dimension, les détails se trouvent dans l'espace W_j^2 le complément orthogonal de V_j^2 dans V_{j-1}^2 :

$$V_{j-1}^2 = V_j^2 \oplus W_j^2 \quad (\text{IV.83})$$

Selon le théorème 7.25 dans ([Stéphane, 2009](#)), trois ondelettes sont définies :

$$\begin{aligned} \psi^1(x) &= \phi(x_1)\psi(x_2), \\ \psi^2(x) &= \psi(x_1)\phi(x_2), \\ \psi^3(x) &= \psi(x_1)\psi(x_2). \end{aligned} \quad (\text{IV.84})$$

Tel que :

$$\psi_{j,n}^k(x) = \frac{1}{2^j} \psi^k\left(\frac{x_1 - 2^j n_1}{2^j}, \frac{x_2 - 2^j n_2}{2^j}\right) \quad (\text{IV.85})$$

Pour $1 \leq k \leq 3$.

La famille d'ondelettes $\{\psi_{j,n}^1, \psi_{j,n}^2, \psi_{j,n}^3\}_{n \in \mathbb{Z}^2}$ est une base orthonormée de W_j^2 .

En pratique, les coefficients d'ondelette à l'échelle 2^{j+1} sont calculés de l'approximation a_j par des convolutions 2-dimensionnelles séparables et sous-échantillonnage :

$$a_{j+1}[n] = a_j * \bar{h}\bar{h}[2n] \quad (\text{IV.86})$$

$$d_{j+1}^1[n] = a_j * \bar{h}\bar{g}[2n] \quad (\text{IV.87})$$

$$d_{j+1}^2[n] = a_j * \bar{g}\bar{h}[2n] \quad (\text{IV.88})$$

$$d_{j+1}^3[n] = a_j * \bar{g}\bar{g}[2n] \quad (\text{IV.89})$$

Où $\bar{x}[n] = x[-n]$. Sachant que h et g sont les filtres à miroir conjugué associés à l'ondelette ψ , tel que h et g agissent comme des filtres passe-bas et passe-haut, respectivement. a_j, d_j^1, d_j^2 , et d_j^3 correspondent respectivement à l'approximation, les hautes fréquences verticales (les détails horizontaux (bords)), les hautes fréquences horizontales (les détails verticaux), et les hautes fréquences dans les deux directions (corners) à l'échelle 2^j . Les coefficients des filtres passe-haut associés à quelques types d'ondelettes sont rapportés dans le Tableau IV-1.

En bref, pour analyser (décomposer) une image, ces deux filtres sont appliqués sur les lignes puis sur les colonnes de l'image (Figure IV-16). La sous-image d'approximation a_{j+1} est utilisée pour le calcul des coefficients de l'approximation et des détails de l'échelle 2^{j+2} .

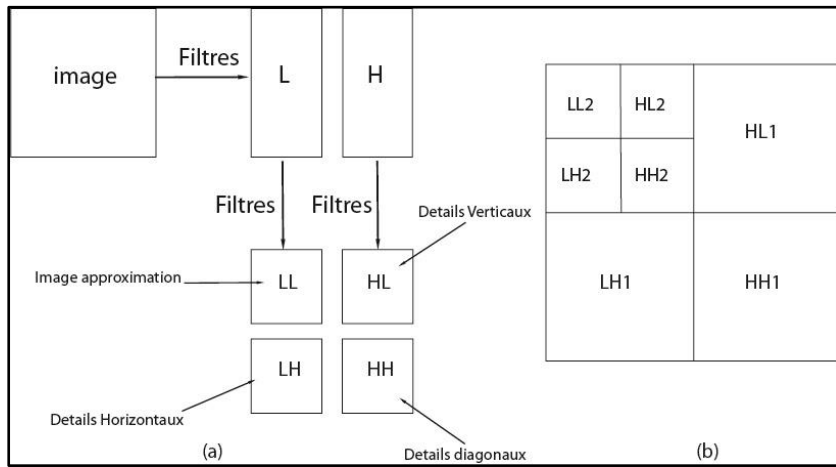


Figure IV-16 Les ondelettes. La décomposition pyramidale.

Tableau IV-1 Quelques filtres passe-haut de décomposition pour quelques types d'ondelettes

Type d'ondelette	Coefficients des filtres de décomposition (high-pass filters)							
Haar	-0.7071 0.7071							
Db2	-0.4830 0.8365 -0.2241 -0.1294							
Db3	-0.3327 0.8069 -0.4599 -0.1350 0.0854 0.0352							
Db4	-0.2304 0.7148 -0.6309 -0.0280 0.1870 0.0308 -0.0329 -0.0106							
Bior2.4	0 0 0 0.3536 -0.7071 0.3536 0 0 0 0							
Bior3.3	0 0 -0.1768 0.5303 -0.5303 0.1768 0 0							

IV.1.4.3 Applications et descripteurs des ondelettes pour la caractérisation des images

Une fois, la décomposition d'image est obtenue, les caractéristiques qui se trouvent dans les différentes sous-bandes (sous-images) sont inexploitablement directement à cause de leur dimensionnalité et leur sensibilité au bruit. Pour cela, une étape de description en utilisant une mesure qui permet de résumer et de tirer profit de ces informations est primordiale.

Dans ce contexte, plusieurs auteurs ont investigué différentes mesures afin de déterminer la mesure la plus discriminante pour les différents problèmes auxquels ils se sont intéressés. Effectivement, [Laine et Fan \(1993\)](#) ont comparé l'énergie et l'entropie pour le problème de la

classification de textures et ils sont arrivés que la signature basée sur l'énergie est légèrement meilleure que celle basée sur l'entropie calculée à partir des mêmes paquets d'ondelettes (*wavelet packets*). La moyenne de la norme l_1 d'énergie (Eq.IV.90) est proposée par [Chang et Kuo \(1993\)](#) en tant que critère pour décider si un paquet d'ondelettes est décomposable ou non.

La moyenne de la norme l_1 d'énergie de la sous-bande c est calculée comme suit :

$$E^{(c)} = \frac{\sum_i^N |e_i^{(c)}|}{N^{(c)}} \quad (\text{IV.90})$$

Où $e_i^{(c)}$ représente le $i^{\text{ième}}$ coefficient dans la sous-bande (c).

[Pichler et al. \(1996\)](#) ont utilisé la variance (Eq.IV.91) comme critère pour la segmentation de textures. Cependant cette mesure a montré une mauvaise performance. L'énergie et grâce à ses bonnes performances a été considérée dans plusieurs travaux subséquents tels que ([Jiang et al., 2006](#); [Manthalkar et al., 2003](#); [Porter et Canagarajah, 1997](#)). Dans l'intéressant travail de [Yuan et Zhang \(2010\)](#), Gaussian Mixture Model (GMM) et aussi, Generalized GMM (GGMM) ont été utilisés pour décrire les caractéristiques statistiques des coefficients d'ondelettes. Vu la nature de la distribution des coefficients d'ondelettes, la fonction de densité de probabilité est modélisée par *multistate Gaussian Mixture*. Plus précisément, par deux états. Une pour décrire la distribution des coefficients de grandes valeurs et l'autre pour la distribution des coefficients de petites valeurs.

Également, la combinaison des mesures est l'une des solutions qui ont prouvé leurs efficacités pour booster les performances ([Arivazhagan et Ganesan, 2003](#); [Xin-Wu et Yu-Xi, 2010](#)).

$$\text{var}^{(c)} = \frac{\sum_i (e_i^{(c)} - \mu^{(c)})^2}{N^{(c)}} \quad (\text{IV.91})$$

Où $\mu^{(c)}$ est la moyenne de la sous-bande (c).

$$\mu^{(c)} = \frac{\sum_i e_i^{(c)}}{N^{(c)}} \quad (\text{IV.92})$$

IV.1.5 Les filtres de Gabor

IV.1.5.1 Présentation

Les filtres de Gabor ont pris une attention particulière, notamment depuis les années 90s, où un nombre important de travaux qui ont commencé à y s'intéresser : ([Bovik et al., 1990](#); [Jain et Farrokhnia, 1990](#); [Manjunath et Ma, 1996](#)). Cette attention particulière revient du fait que les filtres de Gabor suivent une approche de *filtrage multicanal*. Cette approche de filtrage multicanal est inspirée par la théorie de filtrage multicanal pour le traitement de l'information

visuelle dans les premières étapes du système visuel humain ([Jain et Farrokhnia, 1990](#)). En effet, cette théorie remonte aux travaux de [Campbell et Robson \(1968\)](#), où ils ont découvert que le système visuel ne se comporte pas comme un seul mécanisme de détection précédé d'un seul filtre spatial à large bande, mais plutôt comme un certain nombre de mécanismes détecteurs indépendants, chacun précédé d'un filtre à bande relativement étroite « accordé » à une fréquence différente ([Campbell et Robson, 1968](#)). En outre, l'utilisation des filtres de Gabor est motivée par le fait qu'ils ont une orientation, des largeurs de bande de fréquences radiales, et des fréquences centrales tous accordables. Encore, ils obtiennent de manière optimale la résolution conjointe dans l'espace et en fréquence spatiale ([Bovik et al., 1990](#)).

En réalité, la fonction 2-Dimensionnelle de Gabor est constituée d'une onde plane sinusoidale d'une certaine fréquence et orientation, modulée par une enveloppe gaussienne bidimensionnelle ([Jain et Farrokhnia, 1990](#)). La Figure IV-17 illustre une fonction gaussienne et la modulation d'une fonction sinusoidale par cette fonction. Une fonction 2D est illustrée par la Figure IV-18.

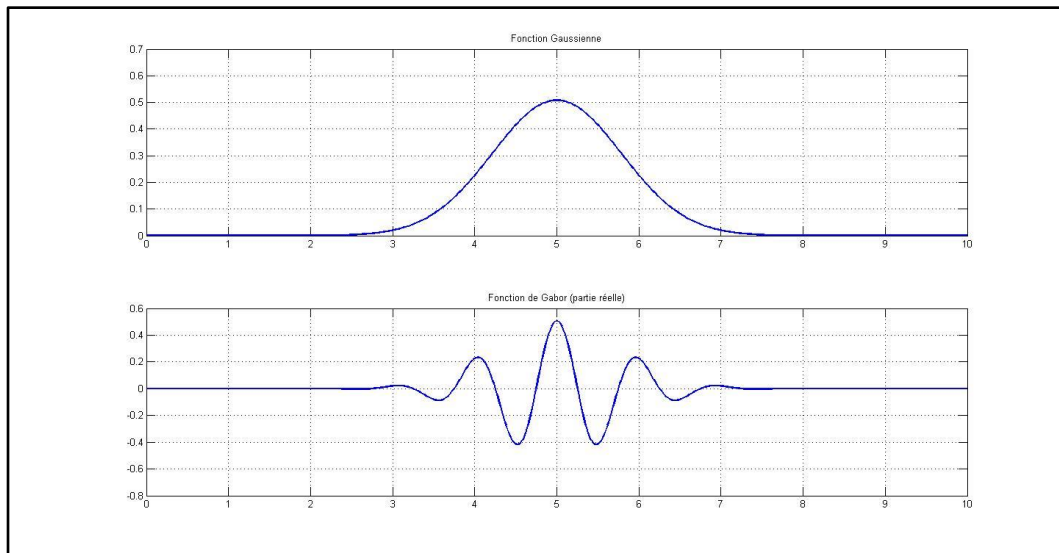


Figure IV-17 La courbe d'une fonction Gaussienne, et celle de la fonction de Gabor (partie réelle).

La version complexe de la fonction de Gabor est définie dans ([Bovik et al., 1990](#); [Porter et Canagarajah, 1997](#)) comme 'a complex sinusoidal grating modulated by a 2-D Gaussian function'²⁴.

La fonction gaussienne elliptique est décrite par l'équation IV.93. Dans le cas où les écarts types σ_x et σ_y prennent la même valeur, la fonction gaussienne devient circulaire.

$$f_g(x, y) = \left(\frac{1}{2\pi\sigma_x\sigma_y} \right) \exp \left[-\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) \right] \quad (\text{IV.93})$$

²⁴ 'Un réseau sinusoidal complexe modulé par une fonction Gaussienne bidimensionnelle'

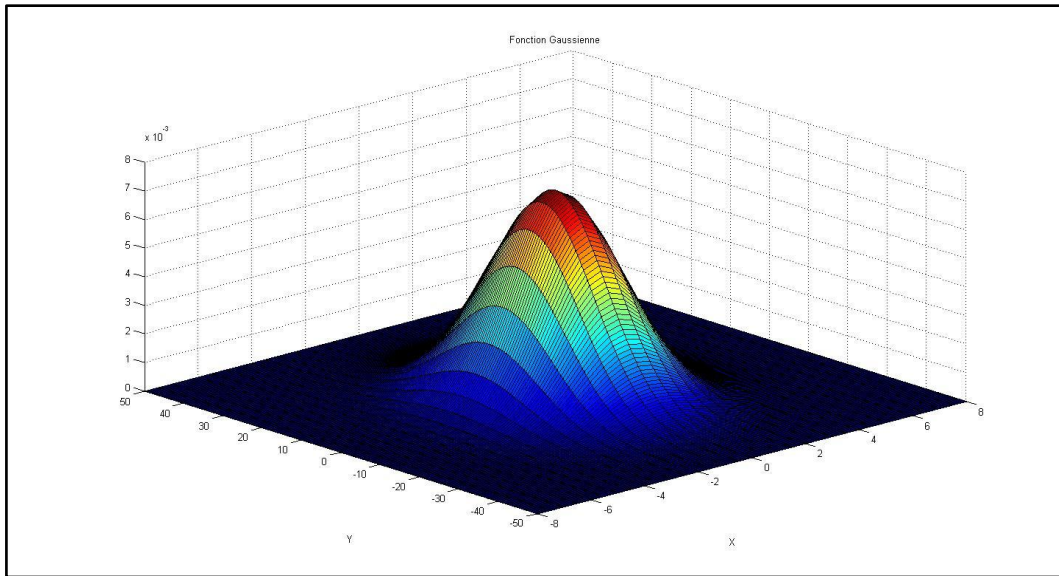


Figure IV-18 Une fonction Gaussienne (2D).

Formellement, la fonction 2D complexe de Gabor peut être s'écrite comme suit ([Ayres et al., 2010, p. 8](#); [Manjunath et Ma, 1996](#)) :

$$g(x, y) = \left(\frac{1}{2\pi\sigma_x\sigma_y} \right) \exp \left[-\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) \right] \exp(2\pi jFx) \quad (IV.94)$$

La partie réelle et la partie imaginaire de la fonction de Gabor sont illustrées par la Figure IV-19 et la Figure IV-20, respectivement.

La fonction $g(x, y)$ peut être vue comme une fonction de Gabor mère (*mother Gabor wavelet*). Donc, un ensemble de fonctions auto-similaires (Gabor wavelets) peuvent être générées en utilisant des opérations de translations et de rotations ([Manjunath et Ma, 1996](#)) :

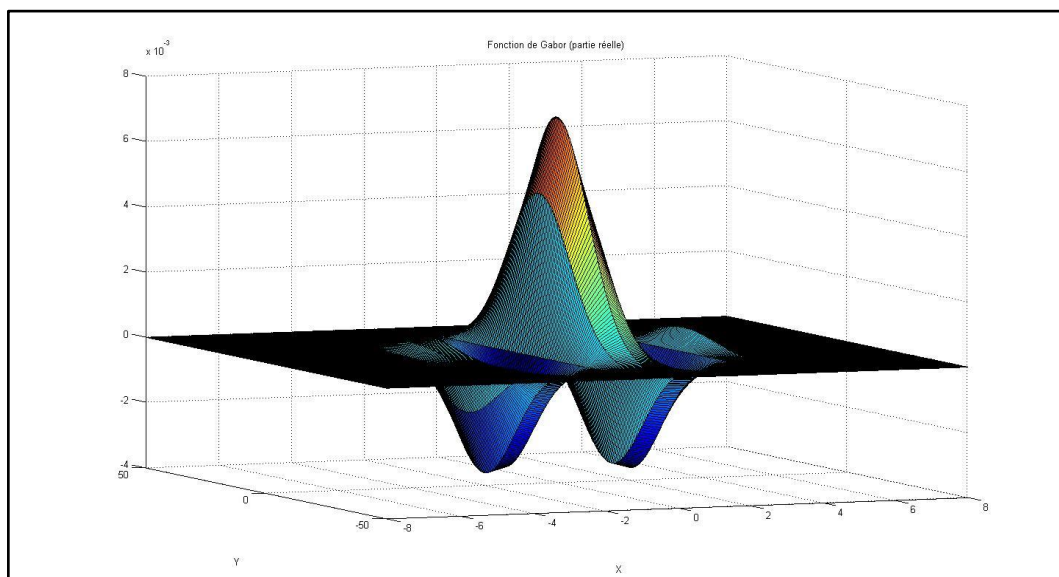


Figure IV-19 Plot de la fonction de Gabor (partie réelle).

$$g_{m,n}(x, y) = a^{-m} g(x', y'), \quad a > 1 \quad (\text{IV.95})$$

$$x' = a^{-m} (x \cos \theta + y \sin \theta) \quad (\text{IV.96})$$

$$y' = a^{-m} (-x \sin \theta + y \cos \theta)$$

Où $\theta = n\pi/K$, m et n sont des nombres entiers, et K est le nombre total d'orientations. L'ensemble des fonctions auto-similaires constituent une banque de filtres.

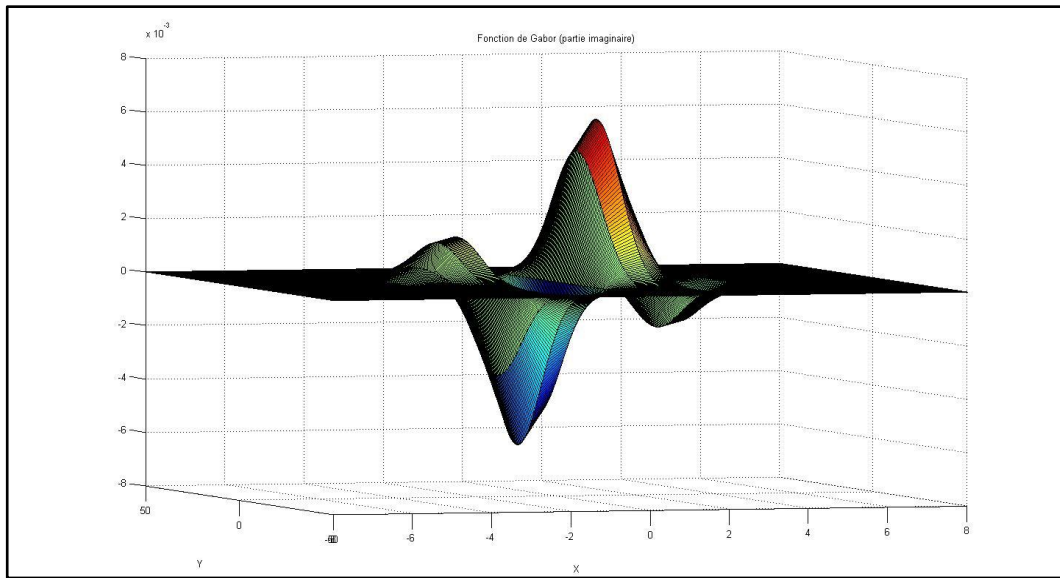


Figure IV-20 Plot de la fonction de Gabor (partie imaginaire).

IV.1.5.2 Applications

Dans les domaines du traitement d'images et la vision par ordinateur, les filtres de Gabor ont été utilisés dans différentes tâches. Ils ont été également exploités dans différents domaines d'application telle l'analyse de la texture ([Bovik et al., 1990](#)) ; la segmentation de texture ([Jain et Farrokhnia, 1990](#)) ; la détection de bords (contours) et lignes ([Manjunath et Chellappa, 1993](#)) ; la recherche d'images basée sur le contenu (texture) ([Han et Ma, 2007](#); [Ma et Manjunath, 1996](#); [Manjunath et Ma, 1996](#); [Sastry et al., 2007](#); [Zhang et al., 2000](#)) ; la classification de textures ([Arivazhagan et al., 2006](#); [Hadizadeh, 2015](#); [Idrissa et Acheroy, 2002](#)) ; la reconnaissance de visage ([Lu et Linghua, 2016](#)) ; la classification des personnes par leur sexe depuis des images de visages occlus ou non occlus ([Rai et Khanna, 2014](#)) ; La détection des vaisseaux dans les angiographies ([Cruz-Aceves et al., 2016](#)) ; la détection du niveau de bacs à déchets solides ([Hannan et al., 2016](#)).

Les différentes méthodes discutées dans cette première partie de ce chapitre sont données par les tableaux récapitulatifs ci-dessous (Tableau IV-2-Tableau IV-5).

Tableau IV-2 Tableau récapitulatif des variantes LBP discutées.

Référence	Nom de la méthode (ou une courte description)	Problème	Nom de la base utilisée (taille)
Ojala et al., 1996	Local Binary Patterns (LBP) (de base)	Classification	Brodatz (subset)
Ojala et al., 2002	LBP (général)	Classification	Brodatz (subset) Outex
Ojala et al., 2002	$LBP_{P,R}^{ri}$ et $LBP_{P,R}^U$	Classification	Brodatz (subset) Outex
Jin et al., 2004	ILBP	Détection de visage	FERET Yale B MIT-CMU
Hafiane et al., 2007	MBP	Classification	Outex
Tan et Triggs, 2007	LTP	Reconnaissance faciale	FRGC-104 Extended Yale-B CMU PIE
Liao et al., 2009	DLBP	Classification	Brodatz Outex Meastex CUREt
Zhang et al., 2010	LDP	Reconnaissance de visage	FERET CAS-PEAL CMU-PIE Extended Yale B FRGC
Guo et al., 2010	CLBP	Classification	Outex CUREt
Mäenpää et Pietikäinen, 2004	OCLBP	Classification	Outex VisTex
Takala et al., 2005	Block-based LBP	Recherche d'images	Corel(1350)
Mehta et Egiazarian, 2016	DRLBP	Classification	Outex KTH-TIPS
Lan et al., 2016	QLRBP	Identification de personnes Reconnaissance faciale	ETHZ i-LIDS MCTS RWFR AR SCface
Naresh et Nagendraswamy, 2016	MLBP	Classification des feuilles de plantes	UoM medicinal plant Flavia Foliage Swedish plant leaf Outex
Verma et Raman, 2016	LTriDP	Recherche d'images	Brodatz VisTex AT&T (face)
Kaya et al., 2015	$nLBP_d$, $dLBP_\alpha$	Classification	Brodatz-1 Brodatz-2 Butterfly Kylber
Smolka et Nurzynska, 2015	Power LBP	Classification de visages (souriant/neutre)	Cohn-Kanade Utrecht All
Murala et Wu, 2015	SS-3D-LTP	Indexation	Corel-10K

		Recherche	Brodatz OASIS
Wong et al., 2013	ELBP	Classification des images de la radiographie pulmonaire	ImageCLEFmed 2009

Tableau IV-3 Tableau récapitulatif des descripteurs de texture discutés-(MCNG).

Référence	Courte description des modifications	Problème (/domaine ou type d'images)	Nom de la base utilisée (taille)
Haralick et al., 1973	MCNG	Classification (photomicrographie, aérien, satellite) images	N/A
Walker et al., 1995	Génération d'une matrice de discrimination qui permet de décrire la force de discrimination de chaque élément de la matrice de cooccurrence	Classification (noyaux de cellules du col utérin)	N/A
Aksoy et Haralick, 1998	La variance a été calculée pour cinq distances (de 1 à 5) et quatre orientations.	Recherche d'images	N/A
Arivazhagan et Ganesan, 2003	Combinaison avec les caractéristiques des ondelettes	Classification	VisTex (20)
Palm, 2004	Extension des MCNG afin de prendre en compte l'information de la couleur.	Classification	VisTex (30) BarkTex (408)
Gelzinis et al., 2007	Ajustement d'un polynôme pour chacun des coefficients de Haralick.	Classification des images de textures générales et de larynx	N/A(785) Outex(180)
Ribaric et Lopar, 2012	Une MCNG est établie pour chaque sous-image.	Reconnaissance de l'empreinte palmaire	N/A(550) N/A(1324)
Zhang et al., 2015	Le calcul de 15 caractéristiques sur des MCNG extraites des régions d'intérêts dans le foie	Classification du degré de la fibrose hépatique	
YongHua et Jin-Cong, 2015	Combinaison de Tamura et des MCNG	Identification de défauts sur la surface du bois	N/A(300)
Tang et al., 2015	Combinaison du LBP et MCNG	Classification de feuilles de thé	N/A Brodatz
Partio et al., 2002		Recherche d'images de rochers	N/A(168) N/A(1512)

Tableau IV-4 Tableau récapitulatif des descripteurs de texture discutés-(Tamura).

Référence	Courte description des modifications	Problème (/domaine ou type d'images)	Nom de la base utilisée (taille)
Tamura et al., 1978		/	Brodatz
Wei-Ying et Hongjiang, 1998	La représentation de la grossièreté et la directivité par un histogramme pour chacune au lieu d'une seule valeur	Recherche	Corel (19800)
Islam et al., 2008	Le calcul de la directivité par la prise en compte de quelques propriétés géométriques	Recherche d'images	Brodatz(112)
Li et al., 2013; Liu et al., 2013	le calcul du contraste en utilisant les moments statistiques, l'uniformité, et l'entropie de l'histogramme d'intensité	Recherche d'images	N/A (200)
Chih-Yi et al., 2003	Le liage des caractéristiques de Tamura avec des concepts humains de haut niveau	Recherche	Corel (1436) VisTex(405)
Majtner et Svoboda, 2012	Adaptation des formules de la grossièreté et la directivité à l'espace 3D.	/	N/A(167) N/A(167)
Zhang et Canosa, 2014	Gestion des valeurs de Tamura comme une jointure spatiale de la distribution Grossièreté-Contraste-Directivité (GCD/CND).	Recherche	Wang Corel (sous ensemble) CalTech
Yue et al., 2015	Combinaison avec les descripteurs de couleur	Recherche	N/A(100)

Tableau IV-5 Tableau récapitulatif des travaux exploitants les ondelettes.

Référence	Courte description des modifications	Problème (/domaine ou type d'images)	Nom de la base utilisée (taille)
Laine et Fan, 1993	Comparaison de signatures basées sur l'entropie vs l'énergie en utilisant les wavelet packets	Classification	Brodatz(23)+2
Chang et Kuo, 1993	Montre de l'utilité des ondelettes tree-structurées pour l'analyse de la texture.	Analyse Classification	Brodatz(30)
Pichler et al., 1996	Comparaison de différentes méthodes d'extraction des caractéristiques de texture	Segmentation de texture	N/A
Porter et Canagarajah, 1997	Combinaison des sous-bandes opposées afin d'aboutir à l'invariance par rapport à la rotation	Classification	Brodatz(16)
Jiang et al., 2006	Création d'une nouvelle sous-bande en intégrant la sous-bande horizontale et verticale	Recherche d'images	Brodatz(117)
Manthalkar et al., 2003	Proposition de calculer la moyenne des caractéristiques des sous-bandes (horizontale et verticale). L'énergie et	Classification	Brodatz(15)

	l'écart type autour de l'énergie sont calculés.		
Xin-Wu et Yu-Xi, 2010	Combinaison de différentes mesures	Recherche d'images	Brodatz(109)

IV.2 Classification des descripteurs

Bien qu'un nombre important de descripteurs (méthodes) ont été introduits depuis les années 1960s, ils peuvent être catégorisés en quelques grandes classes seulement suivant l'approche adoptée. En général, ces classifications se différencient aussi, mais légèrement dans le nombre de classes et leurs nomenclatures. [Tuceryan et Jain \(2011\)](#) ont proposé quatre catégories de méthodes : les méthodes *statistiques*, *géométriques*, *basées sur les modèles* et de *traitement de signal*. [Materka et Strzelecki \(1998\)](#) ont considéré aussi les quatre approches suivantes : *structurelle*, *statistique*, *basée sur les modèles*, et *basée sur les transformations*. De son côté, [Manjunath et Ma \(2002\)](#) ont distingué entre trois catégories de méthodes : *domaine spatial*, *les modèles de champs aléatoires* et *domaine de transformation*.

Au cours de nos recherches, nous avons constaté que les méthodes basées sur les approches statistiques et celles basées sur les transformations sont les plus répandues pour l'extraction des caractéristiques, mais ça n'ignore pas l'existence de méthodes d'autres classes. À cet effet, nous reprendrons, ci-dessous, la classification de [Materka et Strzelecki \(1998\)](#) pour discuter les différentes approches et méthodes.

IV.2.1.1 Approches statistiques

Les approches statistiques se sont basées sur la représentation de la texture en utilisant les propriétés régissant la distribution et la relation entre les niveaux de gris dans l'image ([Castellano et al., 2004](#)). Dans ce contexte, deux types de statistiques sont distinguées ; les statistiques du premier ordre et du second ordre.

Statistiques du premier ordre

Mesure la probabilité d'observer un niveau de gris à un endroit aléatoirement choisi dans l'image ([Tuceryan et Jain, 2011](#)). Le fameux histogramme de niveaux de gris d'une image est un exemple de mesures de cet ordre, car il représente la distribution des niveaux de gris sans la considération des interactions entre eux. D'autres mesures calculées à partir de l'histogramme comme : la moyenne, la variance... etc. sont aussi du premier ordre.

Statistiques du second ordre

Mesure la probabilité d'observer une paire de valeurs de gris apparaissant au niveau des points d'extrémité d'un dipôle (ou aiguille) de longueur aléatoire placé dans l'image à un emplacement et une orientation aléatoire ([Tuceryan et Jain, 2011](#)). Les matrices de cooccurrences représentent un exemple typique de cet ordre.

Donc, en plus des descripteurs statistiques déjà décrits en détail dans la première partie de ce chapitre, à savoir : *les matrices de cooccurrence*, *les descripteurs de Tamura*, et *les motifs*

binaires locaux (LBP), d'autres descripteurs (méthode) existent aussi tels que *les matrices de cooccurrence de motifs* (Jhanwar et al., 2004), *sum and difference histograms* (Unser, 1986)... etc. dans ce qui suit, nous présenterons brièvement ces deux méthodes.

IV.2.1.1.1 Les matrices de cooccurrence de motifs

Les matrices de cooccurrence de motifs *MCM* (*Motif Cooccurrence Matrix*) ont été proposées par Jhanwar et al. (2004) pour la recherche d'images par le contenu. Différemment des *MCNG* et du *corrélogramme de couleurs* dont les cooccurrences concernent les niveaux de gris et les couleurs, respectivement. Les *MCM* établissent les cooccurrences de primitives appelées *motifs*. Ces motifs sont obtenus par la décomposition de l'image en grilles de 2×2 pixels. Puis, chaque grille est remplacée par un motif particulier de *Peano scan*, en traversant la grille dans un sens optimal. L'optimalité considérée par les auteurs est selon la différence incrémentale des intensités sur la ligne de scan minimisant la variation des intensités dans le voisinage local (Jhanwar et al., 2004). Parmi les 24 motifs possibles dans une grille de 2×2, seulement les motifs qui commencent de la partie supérieure gauche sont considérés par les auteurs. Ce qui a permis de réduire le nombre de motifs possibles à 6 motifs seulement (Figure IV-21). Les auteurs ont intégré la couleur, en calculant la *MCM* de chaque plan de couleur.

Pour mesurer la similarité entre deux images, la distance entre leurs *MCM* est calculée. Plus précisément, les auteurs ont calculé la somme des différences pondérées entre les entrées des deux matrices. Pour plus de détails le lecteur peut se référer à (Jhanwar et al., 2004).

Pour plus de performance, Subrahmanyam et al. (2013) ont proposé une version modifiée de *MCM* baptisée *MCMCM* (*Modified Color MCM*). Cette nouvelle version intègre la couleur, par la considération de l'inter-corrélation entre les plans de l'espace de couleur RGB. Par conséquent, neuf images de motifs transformées sont créées.

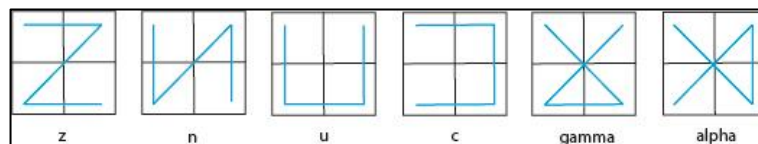


Figure IV-21 Six motifs sur une grille 2×2.

IV.2.1.1.2 Sum and Difference Histograms

Unser (1986) a proposé une méthode simplifiant les matrices de cooccurrence, appelée *Sum and Difference Histograms* (SDH). Dans cette méthode, les matrices de cooccurrence sont remplacées par les histogrammes associés des sommes et des différences. L'auteur a relaté que la méthode est presque aussi puissante que les matrices de cooccurrence, mais elle est moins couteuse en termes de temps de calcul et mémoire de stockage.

La somme et la différence, associées au déplacement (d_1, d_2) sont définies comme suit :

$$S_{k,l} = y_{k,l} + y_{k+d_1, l+d_2} \tag{IV.97}$$

$$D_{k,l} = y_{k,l} - y_{k+d_1, l+d_2} \tag{IV.98}$$

D'une manière similaire aux MCNG, les histogrammes de somme et de différence sont définis comme suit :

$$h_s(i; d_1, d_2) = h_s(i) = \#\{(k, l) \in D, S_{k,l} = i\} \quad (\text{IV.99})$$

$$h_d(j; d_1, d_2) = h_d(j) = \#\{(k, l) \in D, D_{k,l} = j\} \quad (\text{IV.100})$$

Où $i = 2, \dots, 2N_g$ et $j = -N_g + 1, \dots, N_g - 1$.

À la fin, les histogrammes sont normalisés en divisant par la somme des bins.

3D-ASDH (*3-Dimensionnal Adaptive Sum and Difference Histograms*) est une variante récente de la SDH, a été introduite par [Sandid et Douik \(2016\)](#). En effet, en plus de l'intégration de la couleur, elle est également robuste aux rotations. Cette robustesse est obtenue en prenant la somme et la différence maximale entre le pixel central et ses voisins. Par conséquent, les deux équations (IV.97-IV.98) peuvent être réécrites comme suit :

$$S_{k,l} = y_{k,l} + y_v \quad (\text{IV.101})$$

$$D_{k,l} = y_{k,l} - y_v \quad (\text{IV.102})$$

Où v représente les coordonnées des pixels voisins du pixel central (k, l) . Pour plus de robustesse, les auteurs ont proposé également de prendre la moyenne des pixels dans leurs voisinages locaux respectifs au lieu de la valeur exacte.

IV.2.1.2 Approches structurelles

Les méthodes de cette approche supposent que la texture est composée de primitives qui se manifestent d'une manière répétitive presque régulière ([Haralick, 1979](#)). Généralement, l'analyse de textures suivant cette approche consiste en deux étapes majeures qui sont l'extraction des éléments de texture (primitives, cf. sec. III.2.1.3) puis l'inférence des règles de placement. Cependant, des méthodes qui se limitent à la description des primitives extraites existent. Habituellement, l'extraction de ces primitives est faite par le seuillage en utilisant un ou plusieurs seuils. Quant à la caractérisation de ces primitives est faite à travers le calcul de quelques propriétés telles que la taille, le contraste, l'orientation, la surface, la courbure... etc. par exemple, [Voorhees et Poggio \(1988\)](#) filtrent l'image par un filtre *centre-surround*, puis ils effectuent un seuillage de l'image résultante. Les blobs (primitives) obtenus sont ensuite caractérisés en calculant cinq propriétés pour chacun. [Chen et al. \(1995\)](#) calculent une séquence d'images binaires en faisant le seuillage de l'image originale par différents seuils. Plusieurs statistiques sont ensuite calculées pour chaque région connectée. Quant aux travaux utilisant l'inférence des règles de placement, nous citons celui de [Zucker \(1976\)](#), dont l'auteur suppose que les textures réelles sont des versions déformées de textures idéales (imaginaires). Ces textures idéales sont hautement structurées, créées par la combinaison de primitives d'une manière parfaite. Le passage de textures idéales aux textures réelles est effectué en utilisant des règles de transformations.

L'avantage de cette approche est qu'elle fournit une bonne description symbolique de l'image. Cependant, elle est plus appropriée pour la synthèse de la texture que l'analyse ([Materka et Strzelecki, 1998](#)). En général, les méthodes structurales sont limitées à moins que les textures soient très régulières.

IV.2.1.3 Approches basées sur les modèles

Les méthodes de cette approche tentent de représenter la texture par des modèles mathématiques tels que les modèles fractals et stochastiques. Les paramètres estimés d'un modèle sont utilisés pour l'analyse. Cependant, les méthodes suivant cette approche souffrent de la complexité computationnelle pour l'estimation des paramètres.

IV.2.1.3.1 Les fractales

Les *objets fractals* ou tout simplement *les fractales* ont été originalement introduits en [1983](#) par [Mandelbrot](#). Comme il a été relaté dans ([Mandelbrot, 1983](#)), le terme «*fractal*» est en relation avec le verbe latin : *frangere*, qui signifie : *casser*, et qui peut évoquer l'action de «*casser les pierres*», d'où l'adjectif *fractus* est dérivé. Cet adjectif décrit les deux propriétés évidentes dans les pierres cassées : l'irrégularité et la fragmentation.

Différemment de la dimension topologique qui est de type entier, les fractales se distinguent par leurs dimensions fractionnaires. Ainsi, l'une des principales propriétés des fractals est l'*autosimilarité* (self-similarity). Étant donné un ensemble borné A dans un espace Euclidien n -dimensionnel, l'ensemble A est dit auto-similaire quand A est l'union de N copies non chevauchantes d'elle-même. Chacune est réduite par un ratio de r ([Kpalma, 1994](#); [Tuceryan et Jain, 2011](#)). La dimension fractale D est liée au nombre N et le ratio r comme suit ([Keller et al., 1989](#); [Mandelbrot, 1983](#)) :

$$D = \frac{\log(N)}{\log(1/r)} \quad (\text{IV.103})$$

En fait, cette dimension fractale est une mesure de grossièreté ([Tuceryan et Jain, 2011](#)). Plus cette dimension est grande, plus la texture est grossière.

En revanche de la simplicité du calcul de la dimension fractale (en utilisant l'équation ci-dessus), son estimation directement de l'image est difficile. À cet effet, plusieurs méthodes ont été proposées telle que *box dimension* ([Keller et al., 1989](#)). D'un autre côté, la dimension seule ne suffit pas pour la description efficace de la texture ([Keller et al., 1989](#)).

IV.2.1.3.2 Les modèles stochastiques

Dans cette approche, on suppose que les niveaux de gris des pixels soient de nature aléatoire (c.-à-d. des nombres aléatoires). Cependant, les pixels ne sont pas indépendants de leurs voisinages. Parmi les modèles basés sur l'aspect stochastique, on trouve les champs aléatoires de Markov (Markov Random Fields : MRF). Ces derniers sont très utilisés dans le domaine du traitement d'image, que ce soit pour l'analyse de la texture ([Chellappa et Chatterjee, 1985](#); [Chen et Huang, 1993](#); [Hassner et Sklansky, 1980](#); [Solberg et al., 1996](#); [Wang et Liu, 1999](#)), ou pour sa synthèse ([Cross et Jain, 1983](#); [Paget et Longstaff, 1995](#)) en profitant de leurs capacités génératives.

Soit $X(i)$ dénote la valeur de niveau de gris du pixel dont la position est i dans un réseau L (*Lattice*) de taille $M \times N$, sachant que $i \in S = \{1, 2, \dots, MN\}$. L'ensemble A est l'ensemble des niveaux de gris possibles, tel que $A = \{0, 1, \dots, G-1\}$, et $G = 256$ pour une image. Soit aussi, Ω , l'ensemble de tous les motifs de texture possibles.

Comme le stipule la *définition 1* dans ([Cross et Jain, 1983](#)), la coloration de L par G niveaux de gris est une fonction \mathbf{X} de l'ensemble des points de L vers A .

Donc, le MRF est un champ aléatoire dont la fonction de masse de probabilité satisfait les trois conditions suivantes ([Chen et Huang, 1993](#); [Cross et Jain, 1983](#)) :

1. *Positivité* :

$$P(X = x) > 0 \text{ pour tout } x \in \Omega \tag{IV.104}$$

2. *Markoviénité* : pour tout $t \in S$,

$$\begin{aligned} P(X_t = x_t | X_r = x_r, r \neq t) \\ = P(X_t = x_t | X_r = x_r, r \in R_t) \end{aligned} \tag{IV.105}$$

Où R_t est l'ensemble ordonné des voisins du point t (Figure IV-22).

3. *Homogénéité* : $P(x_t | R_t)$ ne dépend pas de la position t , elle dépend seulement de la configuration du voisinage. C.-à-d. elle est invariante par rapport à la translation.

Le MRF peut être de différents ordres suivant l'ordre du voisinage pris en compte. La Figure IV-22 illustre les voisinages de différents ordres du point t .

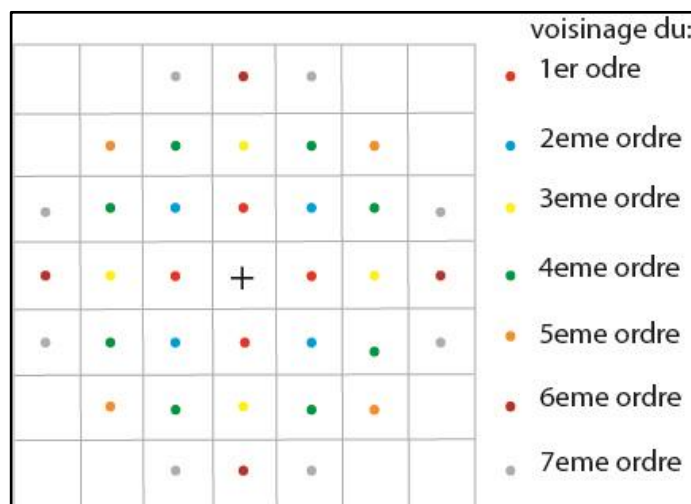


Figure IV-22 Les voisinages de différents ordres-MRF.

Tandis que la notion du voisinage est cruciale dans le MRF, le discrète Gibbs Random Field (GRF) fournis un modèle global d'une image en spécifiant la fonction de masse de probabilité comme suit ([Dubes et Jain, 1989, p. 137](#)) :

$$P(X = x) = \frac{e^{-U(x)}}{Z} \quad (\text{IV.106})$$

Où la fonction $U(x)$ est la fonction d'énergie, dont x est un vecteur de couleurs (niveaux de gris). Z est un constant de normalisation, appelé *fonction de partition*. Pour plus de détails le lecteur peut se référer à ([Dubes et Jain, 1989](#)).

IV.2.1.4 Approches basées sur les transformations

L'exploitation d'autres domaines que le domaine spatial est une manière intéressante pour tirer profit de nouvelles caractéristiques. En effet, l'analyse fréquentielle est utilisée depuis longtemps pour l'analyse de textures. Cette utilisation est motivée par les travaux de [Campbell et Robson \(1968\)](#) qui stipulent que le système visuel humain effectue une analyse fréquentielle de l'image.

L'état de l'art dispose d'une multitude de transformées, parmi d'autres, la transformée de Fourier (FT : Fourier Transform), la transformée en cosinus discret (DCT : Discrete Cosine Transform), et la transformée en ondelettes (WT : Wavelet Transform). Nous mentionnons également dans cette catégorie de méthodes les filtres de Gabor, ainsi que quelques filtres qui opèrent directement sur le domaine spatial.

En pratique, les méthodes basées sur la transformée de Fourier manifestent de faibles résultats à cause de l'absence de l'information spatiale ([Materka et Strzelecki, 1998](#)). Ci-dessous, la FT et DCT sont présentées brièvement. Quant aux WT et Gabor sont discutées en détail dans la première partie de ce chapitre (cf. sec. IV.1.4 et sec. IV.1.5).

IV.2.1.4.1 La transformée de Fourier

La représentation fréquentielle d'un signal 1D discret $f(x)$ de taille X en utilisant la transformée de Fourier discrète (DFT : Discret FT) est obtenue par l'équation suivante ([Maani et al., 2015](#)) :

$$F(u) = \sum_{x=0}^{X-1} f(x) e^{-2\pi i x u / X} \quad (\text{IV.107})$$

De la même manière, la DFT d'un signal 2-Dimensionnel $f(x, y)$ (le cas d'une image), dont le nombre de colonnes est X et celui de lignes est Y , est obtenue par l'équation suivante :

$$F(u, v) = \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} f(x, y) e^{-2\pi i \left(\frac{xu}{X} + \frac{yv}{Y} \right)} \quad (\text{IV.108})$$

Parmi les travaux qui ont exploité la FT pour la description de la texture, nous citons le travail de [Zhou et al. \(2001\)](#), dans lequel les auteurs ont proposé un nouveau descripteur basé sur la transformée de Fourier locale (*Local Fourier Transform*). Pour obtenir cette transformée locale, [Zhou et al.](#), appliquent la DFT sur le voisinage de chaque pixel de l'image (Figure IV-23.a). Ensuite, un histogramme est calculé à partir de l'ensemble de *maps* des coefficients de Fourier Locale. En plus du vecteur de caractéristiques de la méthode précédente, [Ursani et al. \(2009\)](#) ont proposé de prendre en compte la valeur du pixel central, donc ils ont considéré quatre séries de quatre pixels comme il est montré dans la Figure IV-23(b) les résultats obtenus par les auteurs montrent la supériorité de la méthode par rapport à la version originale et aussi par rapport aux filtres de Gabor.

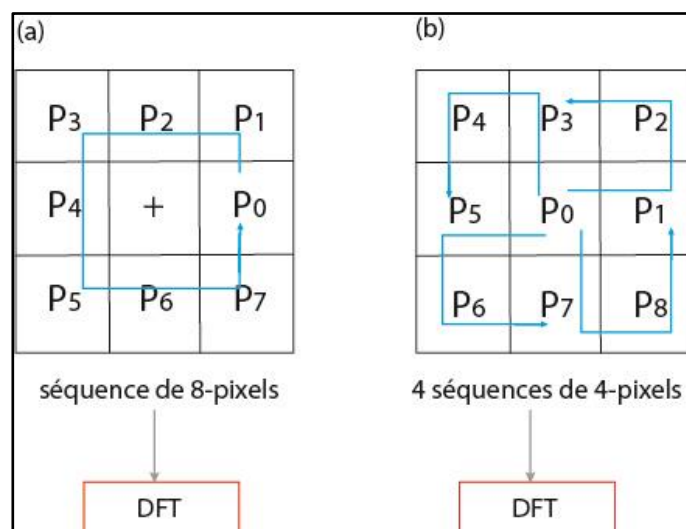


Figure IV-23 Fourier Local. (a) : une séquence de 8 pixels, (b) : 4 séquences de 4 pixels.

IV.2.1.4.2 La transformée en Cosinus discrète

Différemment de la transformée de Fourier dont le noyau de projection est complexe, celui de la transformée en cosinus discrète est un cosinus. Ainsi, la transformation d'un signal 2-Dimensionnel $f(x, y)$ (une image) au domaine fréquentiel est obtenue par l'équation suivante ([Song et Li, 2013](#)) :

$$C(u, v) = \frac{2}{\sqrt{MN}} \alpha(u) \alpha(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \cos\left[\frac{\pi(2x+1)u}{2M}\right] \cos\left[\frac{\pi(2y+1)v}{2N}\right] \quad (\text{IV.109})$$

Où

$$\alpha(u) = \begin{cases} \frac{1}{\sqrt{2}}, & u = 0 \\ 1, & u = 1, 2, \dots, M-1 \end{cases} \quad (\text{IV.110})$$

$$\alpha(v) = \begin{cases} \frac{1}{\sqrt{2}}, & v = 0 \\ 1, & v = 1, 2, \dots, N - 1 \end{cases} \quad (\text{IV.111})$$

La taille du signal dans le domaine fréquentiel est la même que celle du signal original. Or les coefficients significatifs se situent dans les fréquences basses ([Song et Li, 2013](#)). De là, plusieurs travaux considèrent seulement le composant DC (*Direct Current*, c.-à-d. $C(0,0)$) et les premiers composants AC (*Alternating Current*) de la DCT. Qui sont utilisés pour construire les vecteurs des caractéristiques. En outre, la DCT est généralement calculée pour chaque bloc de pixels (par exemple. 8×8 px), ce qui permet d'avoir quelques informations spatiales.

Parmi les premiers travaux qui ont pris en considération le domaine du DCT pour la description de l'image, nous citons celui de [Hee-Jung et Sung-Hwan \(1997\)](#). Dans lequel les auteurs ont calculé la DCT sur des blocs de pixels, puis ils ont constitué 9 vecteurs de caractéristiques à partir des 64 composants de la DCT. [Tsai et al. \(2006\)](#) ont divisé l'image en quatre blocs, ensuite la DCT est appliquée sur le composant (Y) de l'espace de couleurs (YUV). Le vecteur de caractéristiques est constitué de 16 coefficients de chaque bloc (64 au total) : un composant DC, cinq horizontaux, verticaux, et diagonaux. [Song et Li \(2013\)](#) proposent un intéressant descripteur dans lequel les caractéristiques de la DCT sont extraites à partir des régions d'intérêt. Les patches locaux sont quantifiés par rapport à une grille polaire afin de capturer d'informations spatiales et meilleure localisation ([Song et Li, 2013](#)) (cf. sec.III.4.2.4).

IV.3 Conclusion

Dans ce chapitre, nous avons dressé un état de l'art sur les différentes méthodes de caractérisation de la texture. Nous avons mis également l'accent sur les méthodes pionnières dans le domaine de caractérisation de l'image par la texture. En particulier la méthode LBP, qui est considérée comme l'une des rares méthodes qui réunissent la simplicité et la haute performance, a été présentée et discutée. Cependant, vu le nombre important de ses variantes, il n'a pas été possible de les présenter toutes. Pour cette raison, nous avons établi deux listes de variantes, une pour les variantes qui semblent populaires et l'autre pour les plus récentes. À travers les listes de variantes présentées, nous avons vu que la plupart des variantes sont créées par des modifications qui restent toujours au niveau du voisinage. Dans la deuxième partie du présent chapitre, nous étions plus généralistes dans la discussion qui a été faite par rapport aux différentes classes de descripteurs.

Par le prochain chapitre, la deuxième partie de ce manuscrit commence, et dans laquelle nos contributions seront présentées. Donc, le prochain chapitre est consacré à la présentation de l'ensemble de nos contributions pour l'extraction des caractéristiques de texture.

Chapitre V. *Vers des descripteurs visuels plus performants pour la recherche de l'image*

« L'imagination est plus importante que le savoir. »
A. Einstein

« Les détails font la perfection, et la perfection n'est pas un détail. »
Léonard De Vinci

Sommaire

V.1	INTEGRATION GRADUELLE DE L'INFORMATION LOCALE	90
V.1.1	<i>Motivation</i>	90
V.1.2	<i>Gradual Locality Integration of Binary Patterns (GLIBP)</i>	91
V.1.2.1	La rotation d'ellipses	92
V.1.2.2	Mise en échelle et l'enchaînement d'ellipses	92
V.1.2.3	La comparaison	93
V.1.2.3.1	Exemple	93
V.1.2.4	Expérimentation	94
V.1.2.5	Résultats et discussion	94
V.1.3	<i>GLI pour la couleur</i>	98
V.1.3.1	GLI-Color et ses deux variantes GLI-Color ^(HSV) et M-GLI-Color	98
V.1.3.2	Résultats et discussion	99
V.1.3.2.1	L'efficacité de M-GLI-Color	101
V.2	EXPLOITATION DE L'ÉNERGIE DES SOUS-BANDES D'ONDELETTES	102
V.2.1	<i>Motivation</i>	102
V.2.2	<i>L'énergie et la moyenne de sous-bande d'ondelettes</i>	102
V.2.2.1	Remarques et discussion	104
V.2.3	<i>L'impact de l'intégration de l'énergie dans les moments d'ondelettes</i>	104
V.2.4	<i>Expérimentation</i>	105
V.2.5	<i>Résultats et discussion</i>	105
V.3	CONCLUSION	108

V.1 Intégration graduelle de l'information locale

V.1.1 Motivation

En survolant l'état de l'art dressé dans le Chapitre IV. , on peut remarquer la simplicité de la méthode LBP, qui utilise des opérations simples pour le calcul des motifs binaires. En outre, elle utilise un seul histogramme qui reflète la distribution statistique globale des motifs binaires. D'un autre côté, on peut également remarquer sa popularité à travers le grand nombre de variantes proposées. Cependant, la plupart des variantes LBP tentent d'améliorer sa performance via des modifications qui restent toujours au niveau local (au niveau du voisinage). L'exception était avec la méthode '*Block-based LBP*' (cf. sec.IV.1.3.1.1.9) de [Takala et al. \(2005\)](#). Cette méthode comme nous l'avons vue, divise l'image en régions (blocs) puis établit un histogramme pour chaque bloc. Lors de la comparaison, chaque bloc dans l'image requête est comparé avec tous les blocs de chaque image de la base. De là, nous croyions que l'information locale peut être capturée d'une manière plus sophistiquée en utilisant un schéma soigneusement conçu. Cependant, le type de ce schéma sera sûrement non trivial.

Notre méthode proposée dans ce contexte est basée sur un nouveau schéma de délimitation de régions, où ces dernières prennent des formes elliptiques, organisées d'une manière circulaire et concentrique par rapport au centre de l'image et de taille croissante du centre à l'extérieur (Figure V-1). La centralisation du schéma dans l'image est basée sur notre supposition que les régions de grand intérêt, voire les objets, se trouvent dans le centre (ou proches du centre). En effet, cette supposition est adoptée dans plusieurs travaux dans la littérature, tels que ([Kim et al., 2003](#); [Rodhetbhai, 2009](#)). D'un autre côté, la structure du schéma proposé permettra une meilleure localisation par rapport à l'approche en blocs. Dans cette optique, nous recensons plusieurs travaux qui ont adopté la grille polaire (Figure III-12) (une structure circulaire visuellement plus proche à notre schéma que la grille en blocs), en particulier, pour améliorer les performances des descripteurs des points d'intérêt (cf. sec. III.4.2).

Quant à la mise en correspondance entre les régions est de nature sélective, c.-à-d. une région (elliptique) dans l'image requête sera comparée avec les régions de même taille seulement de l'image cible. Donc, l'information est intégrée graduellement, d'où le nom de notre méthode '*Gradual Locality Integration of Binary Patterns*' (GLIBP) ([Bougueroua et Boucheham, 2016a](#)).

En effet, la méthode GLIBP est le résultat de la révision et la maturation de l'idée contenue dans ([Bougueroua et Boucheham, 2014](#)), intitulée '*Ellipse based local binary pattern*' (EILBP). La génération du schéma est faite à travers une procédure basée sur des équations simples. La procédure de génération sera vue en détail dans ce qui suit.

V.1.2 Gradual Locality Integration of Binary Patterns (GLIBP)

Comme il a été mentionné ci-dessus, notre méthode GLIBP est basée sur des régions de forme elliptique, tel que l'ensemble de toutes les régions superposées sur une image prennent une forme similaire à celle illustrée dans la Figure V-1.

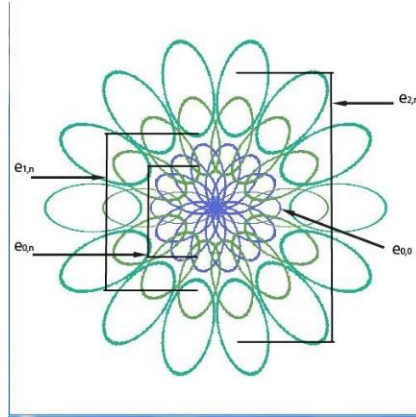


Figure V-1 Positionnement d'ellipses ($S=3, N=7$).

Donc, soit $I(l, m)$ une image LBP. Un pixel $p(i, j)$ tel que $i = 1..l, j = 1..m$, appartient à une région elliptique délimitée par une ellipse e , s'il satisfait l'inéquation suivante :

$$\left(\frac{i^2}{a^2} + \frac{j^2}{b^2} \right) \leq 1 \quad (\text{V.1})$$

Où a et b représentent l'axe semi-majeur et semi-mineur, respectivement. Soit également, $e_{s,n}$ une ellipse d'échelle s , où $s = 0, \dots, S-1$, et d'orientation n , où $n = 0, \dots, 2N-1$.

Les ellipses internes $e_{0,n}$ sont générées par rotation de $e_{0,0}$ par rapport au centre de l'image, puis par rapport au centre de l'ellipse. Les ellipses externes $e_{s,n}$ ($s > 0$) sont créées par dilatation d'ellipses $e_{0,n}$. Pour chaque région, un histogramme normalisé est calculé (Figure V-2). Lors de la comparaison, l'historgramme normalisé d'une région délimitée par l'ellipse $e'_{s,n}$ est comparé avec tous les histogrammes des régions délimitées par des ellipses de même échelle s dans $e''_{s,j}$ de l'image cible (l'image de la base), où $j = 0, \dots, 2N-1$. À la fin, seulement la distance minimale est gardée. Donc, la distance finale entre une requête Q et une image B de la base est la somme de ces distances minimales. La méthode sera expliquée davantage à travers les sous-sections qui vont suivre.

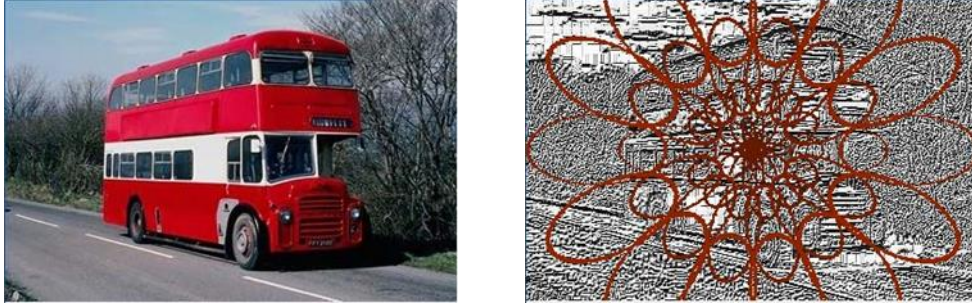


Figure V-2 Illustration des régions elliptiques sur l'image LBP ($S=3, N=7$). À gauche : l'image originale, à droite : les régions elliptiques sur l'image LBP.

V.1.2.1 La rotation d'ellipses

Afin de calculer les ellipses internes $e_{0,n}$ de paramètres a_0 et b_0 , on procède par une simple rotation du centre $c_0(cx_0, cy_0)$ de $e_{0,0}$ en utilisant les deux équations ci-dessous (V.2-V.3), par rapport au centre de l'image. Puis on détermine les pixels appartenant à cette ellipse en utilisant l'équation (V.1). Finalement, les pixels de l'ellipse sont pivotés par rapport son centre.

$$cx'_0 = (cx_0 - i_0) \cdot \cos(\Theta) + (cy_0 - j_0) \cdot \sin(\Theta) \quad (\text{V.2})$$

$$cy'_0 = -(cx_0 - i_0) \cdot \sin(\Theta) + (cy_0 - j_0) \cdot \cos(\Theta) \quad (\text{V.3})$$

$$\Theta = n \cdot \pi / N \quad (\text{V.4})$$

Où Θ représente l'orientation d'ellipse.

V.1.2.2 Mise en échelle et l'enchaînement d'ellipses

Les régions délimitées par les ellipses d'échelle supérieure ($s > 0$), clairement, elles vont prendre des tailles supérieures. Donc, en premier, on doit calculer leurs paramètres a_s et b_s . Pour cela, on utilise les deux équations suivantes :

$$a_s = A^s \cdot a_0 \quad (\text{V.5})$$

$$b_s = A^s \cdot b_0 \quad (\text{V.6})$$

Où

$$A = U^{1/s-1} \quad (\text{V.7})$$

Et U est une valeur empirique.

Pour l'enchaînement des ellipses, le centre $c_{s+1,0}$ de l'ellipse $e_{s+1,0}$ est calculé comme suit :

$$c_{s+1,0} = c_{s,0} + a_{s+1,0} \quad (\text{V.8})$$

La Figure V-3 présente une illustration de la manière d'enchaînement.

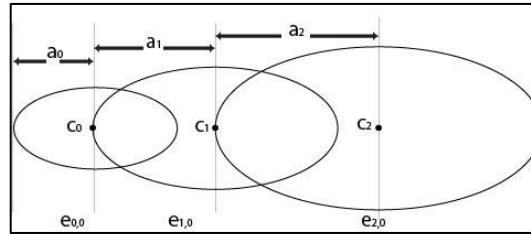


Figure V-3 Enchaînement des ellipses.

V.1.2.3 La comparaison

La distance D entre deux images, la requête Q et une image cible B est calculée par l'équation suivante :

$$D(Q, B) = \sum_{s=0}^{S-1} \sum_{n=0}^{2N-1} \min_k d(Hq_{s,n}, Hb_{s,k}) \quad (\text{V.9})$$

Où $Hq_{s,n}$ (respectivement, $Hb_{s,k}$) représente l'histogramme normalisé des régions délimitées par les ellipses $e_{s,n}$ ($e_{s,k}$) dans l'image Q (respectivement, B), et d est une métrique de distance. Dans ce travail, nous avons investigué *Manhattan* et D_1 .

V.1.2.3.1 Exemple

La Figure V-4 illustre la comparaison entre deux images prises de la base Corel Dataset²⁵. La région délimitée par l'ellipse rouge dans l'image requête sera comparée avec toutes les régions de sa classe dans l'image cible (c.-à-d. avec les régions délimitées par les ellipses verts). À la fin, seulement la distance minimale qui est gardée. Dans notre illustration, cette distance est en rouge qui correspond à l'ellipse en rouge également. La procédure est répétée pour chacune des régions en vert dans la requête, et également, pour les régions délimitées par les ellipses d'échelle supérieure. Bien sûr, ces dernières seront comparées avec celles de même échelle dans l'image cible (les régions délimitées par les ellipses marron).

²⁵ Sous-ensemble de la base : http://vision.stanford.edu/resources_links.html

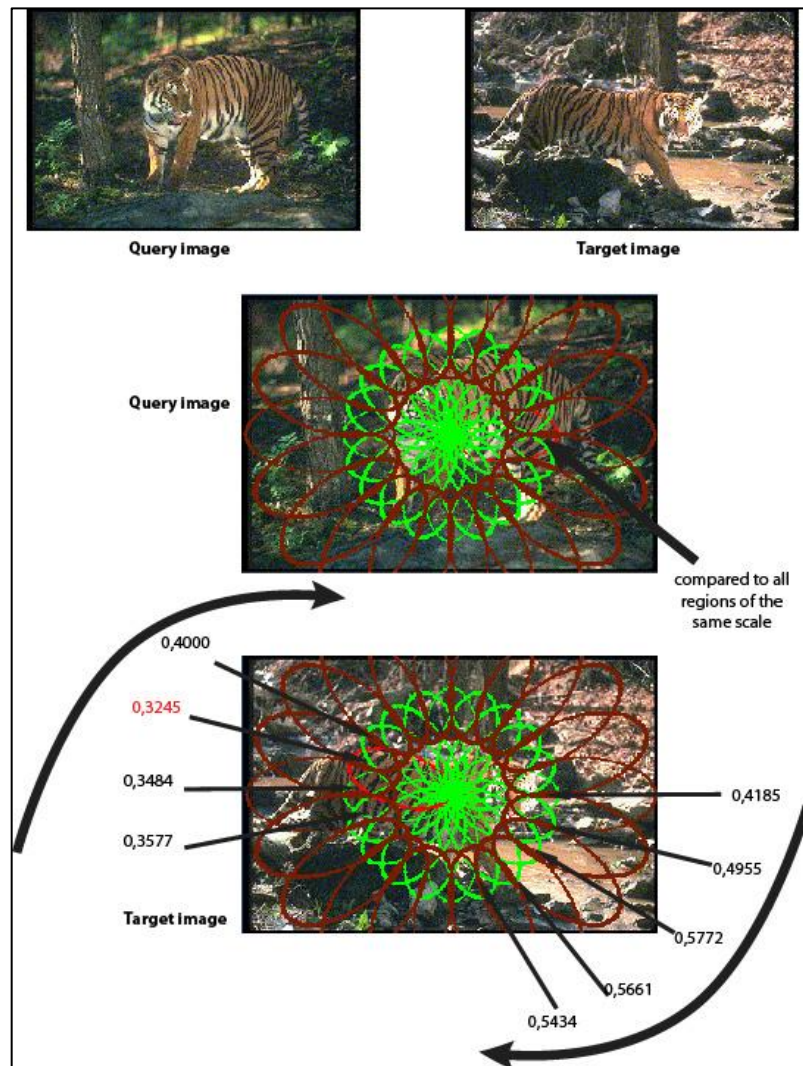


Figure V-4 Illustration de la comparaison de la méthode GLIBP, $(a,b,U,N,S)=(50,25,1.4,10,2)$

V.1.2.4 Expérimentation

Afin d'évaluer les performances de notre méthode, nous avons exploité la base Corel 1k (cf. Tableau II-1) l'une des bases les plus utilisées en CBIR. À savoir, cette base est composée de 1000 images, classifiées en 10 classes : *Africans*, *beaches*, *buildings*, *buses*, *dinosaurs*, *elephants*, *flowers*, *horses*, *mountains* et *food*. Toutes les images ont la taille 256×384 ou 384×256.

La précision moyenne AP (Average Precision) et la précision moyenne pondérée AWP ('Average Weighted Precision') sont calculées pour déterminer les performances de notre méthode.

V.1.2.5 Résultats et discussion

Les deux premiers tableaux ci-dessous (Tableau V-1-Tableau V-2) rapportent les résultats obtenus par la méthode GLIBP en utilisant différents ensembles de paramètres. Également, deux mesures de similarité ont été utilisées. En effet, ces résultats montrent l'impact des paramètres sur les performances de la méthode.

Par exemple, à partir du Tableau V-1, on peut noter des améliorations de (+0.13%, +0.22% et +0.22%) en termes d'AWP lorsque N prend les valeurs (9, 10 et 11 respectivement) par rapport la précision obtenue par N=8. Quant à l'ensemble de paramètres (A, B, U, S) = (30, 15, 2, 3) dont les résultats sont rapportés dans le Tableau V-2, on peut noter des améliorations de (+0.15% et +0.32%) lorsque N est égale à (8 et 9 respectivement) par rapport au résultat obtenu par N=7. D'un autre côté, l'analyse des résultats classe par classe, montre que les classes "Dinosaurs", "Buses", "Horses" et "Flowers" sont celles qui ont enregistré une AWP supérieure de 90%. En contraste, les classes "Africans", "Mountains" et "Foods" sont celles qui ont enregistré une AWP inférieure de 65%. La plus faible précision a été enregistrée par la classe "Mountains".

Tableau V-1 les résultats de la méthode GLIBP en termes de précision moyenne pondérée, avec les paramètres (a,b,U,S)=(50,25,1.4,2) et différentes valeurs pour N, en utilisant la métrique de distance Manhattan et D1 (w=20)

	N=8		N=9		N=10		N=11	
	Manhattan	D1	Manhattan	D1	Manhattan	D1	Manhattan	D1
Africans	62.66	63.18	63.25	63.62	63.50	63.93	63.71	64.17
Beaches	68.69	69.15	68.58	69.16	69.02	69.67	68.63	69.25
Buildings	70.51	71.44	70.60	71.71	70.45	71.49	70.63	71.64
Buses	97.12	97.35	97.14	97.33	97.30	97.46	97.07	97.35
Dinosaurs	99.46	99.51	99.48	99.54	99.50	99.48	99.44	99.48
Elephants	73.43	73.89	73.18	73.49	73.40	74.00	74.10	74.52
Flowers	90.25	90.47	90.44	90.79	90.40	90.68	90.50	90.70
Horses	90.63	90.72	91.02	91.13	91.00	91.06	90.36	90.50
Mountains	47.46	47.71	47.50	47.97	47.12	47.49	47.08	47.81
Food	61.21	61.63	61.53	62.02	61.92	62.38	62.12	62.59
Moyenne	76.14	76.50	76.27	76.67	76.36	76.76	76.36	76.80

Tableau V-2 Les résultats de la méthode GLIBP en termes de précision moyenne pondérée, avec les paramètres (a,b,U,S)=(30,15,2,3) et différentes valeurs pour N, en utilisant la métrique de distance Manhattan et D1 (w=20).

	a=30, b=15, u=2, N=7, S=3		a=30, b=15, u=2, N=8, S=3		a=30, b=15, u=2, N=9, S=3	
	Manhattan	D1	Manhattan	D1	Manhattan	D1
Africans	59.77	60.15	59.90	60.34	60.63	60.99
Beaches	66.49	67.30	67.58	68.64	67.26	68.09
Buildings	69.69	70.66	70.32	71.43	70.57	71.47
Buses	96.72	96.98	96.57	96.86	96.96	97.19
Dinosaurs	99.54	99.59	99.61	99.64	99.67	99.70
Elephants	70.38	70.94	69.63	70.21	70.31	70.66
Flowers	88.54	88.91	88.95	89.44	89.05	89.38
Horses	90.81	90.88	90.38	90.53	90.78	90.78
Mountains	48.04	48.29	48.34	48.50	48.06	47.99

Food	58.03	58.35	57.78	57.89	58.66	58.93
Moyenne	74.80	75.20	74.90	75.35	75.19	75.52

La comparaison de la méthode GLIBP avec d'autres méthodes de l'état de l'art montre également les performances intéressantes de la méthode GLIBP. Par exemple, la comparaison faite dans le Tableau V-3 en termes de l'AWP contre les méthodes EILBP ([Bougueroua et Boucheham, 2014](#)), LBP ([Ojala et al., 1996](#)), et ILBP ([Jin et al., 2004](#)), montre clairement la supériorité de la méthode GLIBP qui atteint une AWP de 82.13% dans une fenêtre de taille 10, contre 80.45%, 76.99%, et 79.15% seulement pour les trois méthodes citées ci-dessus, respectivement.

La comparaison dans le Tableau V-4 en termes de précision moyenne AP, dans une fenêtre de taille 9, contre les méthodes EILBP, et la méthode de ([Malik et Baharudin, 2013b](#)) utilisant les différents filtres : DCT (Median Filter), DCT (Median with Edge Extraction), et DCT (Laplacian Filter), montre encore une deuxième fois la bonne performance de la méthode GLIBP par rapport aux méthodes considérées.

Le dernier Tableau V-5, rapporte les précisions moyennes obtenues par plusieurs méthodes de l'état de l'art dans des fenêtres de taille 10 et 20. Donc, en plus des méthodes EILBP ([Bougueroua et Boucheham, 2014](#)) (A), LBP ([Ojala et al., 1996](#)) (D), et ILBP ([Jin et al., 2004](#)) (E), nous avons considéré également les méthodes suivantes : LDP_{16,2} ([Reddy et Prasad, 2011](#)) (F), 'DLEP' de [Murala et al. \(2012a\)](#) (B), et également la méthode 'block-based LBP' ([Takala et al., 2005](#)) (C) qui est basée sur les histogrammes locaux et dont les résultats sur la base Wang (Corel 1K) sont rapportés dans ([Murala et al., 2012a](#)).

Tableau V-3 Comparaison de la méthode proposée avec quelques méthodes existantes en termes de précision moyenne pondérée en utilisant la métrique de distance Manhattan, (a,b,U,N,S)=(50,25,1.4,10,2).

Classes	(w=10)				(w=20)			
	GLIBP (Bougueroua et Boucheham, 2016a)	EILBP (Bougueroua et Boucheham, 2014)	LBP (Ojala et al., 1996)	ILBP (Jin et al., 2004)	GLIBP (Bougueroua et Boucheham, 2016a)	EILBP (Bougueroua et Boucheham, 2014)	LBP (Ojala et al., 1996)	ILBP (Jin et al., 2004)
Africans	71.40	66.61	75.19	73.44	63.50	59.24	67.17	66.43
Beaches	76.67	74.40	67.01	68.63	69.02	66.62	58.75	61.46
Buildings	77.69	76.47	68.57	70.58	70.45	69.17	59.23	61.33
Buses	97.89	97.37	96.85	97.77	97.30	96.05	95.19	96.95
Dinosaurs	99.56	99.88	98.36	98.16	99.50	99.64	97.71	97.53
Elephants	82.33	78.28	63.16	74.01	73.40	68.72	51.51	63.00
Flowers	94.06	92.88	92.15	93.94	90.40	88.89	89.12	90.62
Horses	94.78	94.46	87.24	90.29	91.00	90.34	79.25	84.31
Mountains	56.66	51.62	52.61	53.66	47.12	43.26	43.86	44.70
Food	70.24	72.49	68.79	71.05	61.92	64.45	60.04	62.44
Moyenne	82.13	80.45	76.99	79.15	76.36	74.64	70.18	72.88

Tableau V-4 Comparaison de la méthode proposée avec quelques méthodes existantes en termes de précision moyenne, $(a,b,U,N,S) = (50,25,1.4,10,2)$ ($w=9$)

Classes	GLIBP (Bougueroua et Boucheham, 2016a)	EILBP (Bougueroua et Boucheham, 2014)	DCT (Median Filter) (Malik et Baharudin, 2013b)	DCT (Median with Edge Extraction) (Malik et Baharudin, 2013b)	DCT (Laplacian Filter) (Malik et Baharudin, 2013b)
Africans	62.1	56.6	100	100	100
Beaches	68.3	65	67	78	80
Buildings	70.4	67.3	56	56	57
Buses	97.7	96	79	89	82
Dinosaurs	99.4	99.6	100	100	100
Elephants	74	67.7	51	57	57
Flowers	90.4	89.3	61	62	78
Horses	91.6	90.3	80	89	92
Mountains	43.9	39.1	39	34	42
Food	60.9	62.3	31	30	51
Moyenne	75.87	73.32	66.4	69.5	73.9

Tableau V-5 Comparaison de la méthode proposée avec quelques méthodes existantes en termes de précision moyenne, $(a,b,U,N,S) = (50,25,1.4,10,2)$.

Méthodes Classes	(w=10)							(w=20)			
	GLIBP	EILBP (A)	DLEP (B)	BLOCK based LBP (C)	LBP (D)	ILBP (E)	LDP _{1,6,2} (F)	GLIBP	EILBP (A)	LBP (D)	ILBP (E)
Africans	60.9	56.1	69.3	53.7	63.7	63.1	66	52.6	49.5	56.2	57
Beaches	66.6	64.2	60.5	52.8	55.2	58.2	58	59.5	55.7	47.3	51.3
Buildings	69.7	66.2	72.0	64.3	55.9	58.4	66	61.1	58	45.4	48.6
Buses	97.3	95.8	97.9	89.8	95	96.9	95.4	96.5	93.5	92.1	95.5
Dinosaurs	99.4	99.4	98.5	99.6	97.8	97.1	96.6	99.5	99.4	96.4	96.6
Elephants	72.4	65.5	55.9	58.4	46.4	59.9	60	60.6	54.9	36.3	47.2
Flowers	89.6	88.1	91.9	91.9	87.9	89.9	89.1	84.5	82.9	84.2	85.9
Horses	91.1	89.6	76.9	83.8	78.9	82.9	75.8	84.2	83.3	66.5	73.8
Mountains	42.4	38.2	42.7	44.5	38.8	39	44.3	35.2	32.3	32.6	33.4
Food	59.6	60.5	82	62.2	56.3	59.5	76.9	51.5	54	47.8	50.5
Moyenne	74.9	72.36	74.8	70.1	67.59	70.49	72.81	68.52	66.35	60.48	63.98

V.1.3 GLI pour la couleur

Notre approche d'intégration de l'information locale est également appliquée avec succès pour l'extraction des caractéristiques de la couleur ([Bougueroua et Boucheham, 2015](#)). Pour cela, notre schéma présenté dans les sections précédentes (section.V.1.2) est appliqué directement sur les différents canaux de l'espace de couleur RGB (Figure V-5), par conséquent, la méthode a été baptisée 'GLI-Color'. Néanmoins, la méthode est facilement généralisable sur les autres espaces de couleur, ce qui permet de profiter de différentes propriétés et capacités de différents espaces de couleur. Effectivement, la méthode a été appliquée sur l'espace de couleur HSV à travers une nouvelle variante plus performante, appelée 'GLI-Color^(HSV)', et les résultats étaient très satisfaisants. En outre, nous avons proposé également l'utilisation des moments d'histogramme ce qui a permis de créer une autre variante plus efficace, baptisée 'M-GLI-Color'.

Vu que la manière de génération du schéma a été déjà présentée, ainsi que la généralisation est clairement très simple et intuitive, nous nous limitons dans ce qui suit à une courte description, pour passer immédiatement aux résultats.

V.1.3.1 GLI-Color et ses deux variantes GLI-Color^(HSV) et M-GLI-Color

La généralisation de l'approche utilisée dans la méthode GLIBP pour l'extraction de caractéristiques de texture à la couleur est très simple et intuitive. En effet, notre première variante GLI-Color délimite les régions elliptiques sur chacune des trois images associées aux différents canaux de l'espace RGB (Figure V-5), puis extrait les caractéristiques via les histogrammes. En conséquence, trois histogrammes normalisés sont établis pour chaque région, chaque histogramme est de taille 256 bins.

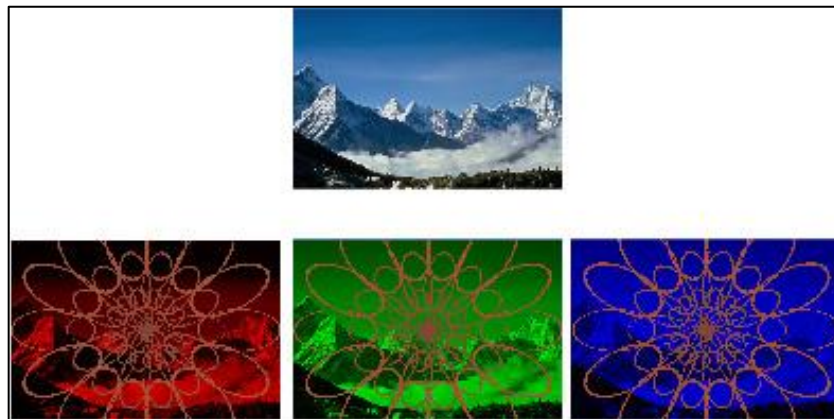


Figure V-5 La méthode GLI-Color, illustration des régions elliptiques sur les trois canaux de l'espace RGB.

- **GLI-Color^(HSV)** : La première variante que nous avons proposée exploite l'espace de couleur HSV. Donc, les régions sont extraites des canaux (Hue, Saturation, et Valeur) comme le montre la Figure V-6, ci-dessous. Vu que les valeurs des pixels selon les différents composants appartiennent à l'intervalle $[0.0...1.0]$, nous les avons quantifiées en 101 valeurs, ce qui représente la taille de chaque histogramme.

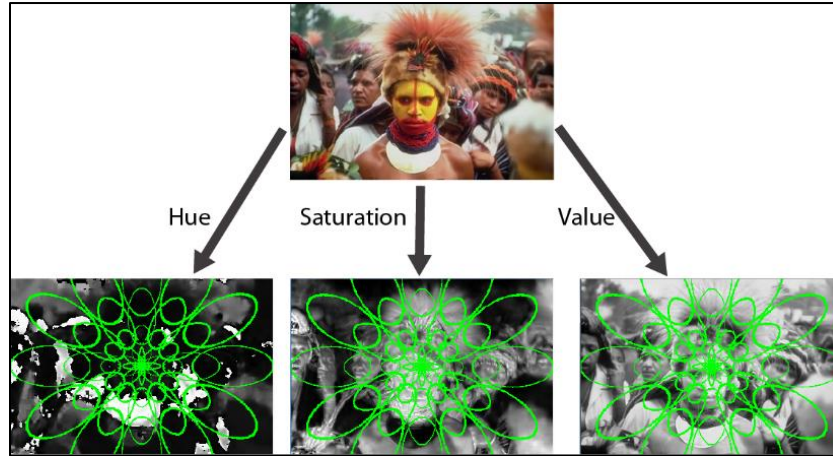


Figure V-6 Illustration des régions elliptiques sur les trois canaux : Hue, Saturation et Value de l'espace de couleur HSV.

- **M-GLI-Color** : dans la deuxième variante, nous nous sommes plus intéressés à l'efficacité. Pour cela, nous avons remplacé les histogrammes du GLI-Color par leurs moments. Autrement dit, un histogramme normalisé $H_{s,n}^c$ du canal de couleur c , et de la région délimitée par l'ellipse $e_{s,n}$ est remplacé par les trois premiers moments calculés comme suit :

$$m_1 = \left[\sum_{i=0}^{255} i \times H_{s,n}^c[i] \right] \quad (\text{V.10})$$

$$m_2 = \sqrt{\sum_{i=0}^{255} (i \times H_{s,n}^c[i] - m_1)^2} \quad (\text{V.11})$$

$$m_3 = \sqrt[3]{\sum_{i=0}^{255} (i \times H_{s,n}^c[i] - m_1)^3} \quad (\text{V.12})$$

En faisant cela, le vecteur de caractéristiques (histogramme) est réduit d'une manière significative de 256 à 3 composants seulement.

V.1.3.2 Résultats et discussion

Les résultats obtenus des expériences conduites sur la base Corel 1K sont rapportés dans le Tableau V-6. Ce dernier présente les résultats en termes de la précision moyenne (AP) et d'AWP, en utilisant deux métriques de distance.

Donc, d'après les résultats de la méthode GLI-Color rapportés dans le Tableau V-6, nous remarquons que les schémas constitués de deux niveaux de régions elliptiques (deux échelles, $S=2$) sont plus performants que le premier qui utilise une seule couche ($S=1$). De point de vue des métriques de distance, les deux métriques considérées atteignent des précisions très proches.

Le Tableau V-7 rapportant une comparaison entre GLI-Color et GLI-Color^(HSV), montre clairement la supériorité de la deuxième variante qui exploite l'espace de couleur HSV.

Tableau V-6 La précision moyenne pondérée et la précision moyenne de GLI-Color pour $w=10$ et $w=20$, avec différents ensembles de paramètres (a-b-U-N-S), en utilisant les métriques de distances Manhattan et D1.

(a-b-U-N-S)	95-50-1-8-1		50-25-1.4-8-2		50-25-1.4-9-2		50-25-1.4-10-2	
Moyenne	Manhattan	D1	Manhattan	D1	Manhattan	D1	Manhattan	D1
AWP($w=10$)	79.99	80.05	82.15	82.18	82.27	82.31	82.26	82.32
AWP($w=20$)	73.21	73.29	75.56	75.62	75.65	75.73	75.72	75.80
AP($w=10$)	70.84	70.98	73.69	73.71	73.68	73.68	73.98	74
AP($w=20$)	63.36	63.41	65.64	65.78	65.75	65.87	65.99	66.1

Tableau V-7 Comparaison de GLI-Color(HSV) avec GLI-Color en termes de la précision moyenne pondérée et la précision moyenne pour $w=10$ et $w=20$, en utilisant différents ensembles de paramètres (a-b-U-N-S), et la métrique de distance Manhattan.

(a-b-U-N-S)	95-50-1-8-1		50-25-1.4-8-2		50-25-1.4-9-2		50-25-1.4-10-2	
Moyenne	GLI-Color ^(HSV)	GLI-Color	GLI-Color ^(HSV)	GLI-Color	GLI-Color ^(HSV)	GLI-Color	GLI-Color ^(HSV)	GLI-Color
AWP($w=10$)	82.52	79.99	83.43	82.15	83.42	82.27	83.48	82.26
AWP($w=20$)	76.73	73.21	77.67	75.56	77.73	75.65	77.82	75.72
AP($w=10$)	75.14	70.84	75.89	73.69	76.06	73.68	76.08	73.98
AP($w=20$)	68.09	63.36	69.35	65.64	69.29	65.75	69.49	65.99

Le Tableau V-8 ci-dessous, dresse une comparaison entre les performances de la méthode GLI-Color et ses deux variantes M-GLI-Color et GLI-Color^(HSV) contre un ensemble de méthodes de l'état de l'art.

D'après les résultats, notre méthode (GLI-Color^(HSV)) qui exploite l'espace de couleur HSV est la meilleure avec une précision de 76.08%, suivie par GLI-Color avec une précision 74%. Ces deux méthodes ont clairement surmonté le problème des histogrammes globaux, car cette dernière approche n'a atteint que 65.41% (GCH_{rgb}). Nos deux méthodes sont également meilleures que CH+GT et la méthode introduite dans (Jiayin et Wenjuan, 2012). Cependant, notre deuxième variante (M-GLI-Color) n'a pas pu atteindre la barre de 61%. En plus, même la méthode GCH_{rgb} (sans quantification de couleur) qui utilise les histogrammes globaux est supérieure de cette variante. Ce qui est très logique vu que cette dernière (M-GLI-Color) est fixée sur l'efficacité que les performances (cf. sec.V.1.3.2.1 ci-dessous). Cependant, M-GLI-Color reste meilleure que plusieurs méthodes de l'état de l'art : l'histogramme global GCH (avec quantification), et aussi CCV (Pass et Zabih, 1996), CCH (Stehling et al., 2003), et Cell-CCV (Salmi et Boucheham, 2014).

Tableau V-8 Comparaison de la méthode proposée GLI-Color avec quelques méthodes existantes en termes de précision moyenne ($w=10$).

Classes	GCH _r gb	GLI-Color ^(H SV)	GLI-Color	M-GLI-Color	CH+ GT	Méthode (Jiayin et Wenjuan, 2012)	GCH	CCV	CCH	Cell-CCV
Africans	84.1	82.7	86.1	53.5	75.90	69.07	45	45.8	42.70	47
Beaches	43.5	43	48.3	37.2	50.00	55.32	30.80	28.10	34.20	31.20
Buildings	47.9	72.7	66.6	46	46.80	56.45	33.50	34.50	33.40	34
Buses	50.3	82.7	66.7	56.7	51.70	89.36	42.24	44.40	39.30	45
Dinosaurs	98.9	99.7	99.1	98.2	100	93.27	95.75	95.90	97.30	97.90
Elephants	66.5	64.8	72.3	60.8	61.70	70.84	50.10	52.80	53.90	56.40
Flowers	75.4	78.8	77.3	63.3	79.50	88.47	70.10	77.70	73.50	83.60
Horses	88.3	97.7	90.5	76.3	90.80	81.37	84.70	82.70	83.90	83.70
Mountains	29.3	52.1	51.6	55.1	28.40	64.58	47.90	41.20	45.30	44.60
Food	69.9	86.6	81.5	62.3	62.80	69.83	49.60	48.60	49.80	51.40
Moyenne	65.41	76.08	74	60.94	64.76	73.86	54.96	55.17	55.54	57.48

V.1.3.2.1 L'efficacité de M-GLI-Color

Afin de montrer l'efficacité de la variante M-GLI-Color vis-à-vis de GLI-Color et GLI-Color^(HSV), nous avons mesuré le temps pris pour mesurer la similarité entre la requête et toutes les images de la base Corel 1K, puis retourner les images les plus pertinentes. Pour cela, dix requêtes différentes ont été utilisées. Selon la Figure V-7 montrant les temps moyens des dix exécutions, la deuxième variante M-GLI-Color est la plus efficace avec un temps moyen de 0.23 seconde, suivie par GLI-Color^(HSV) avec 4.7 secondes, et enfin GLI-Color qui est la plus lente avec 12.69 secondes. Les détails des dix exécutions sont rapportés dans le Tableau V-9.

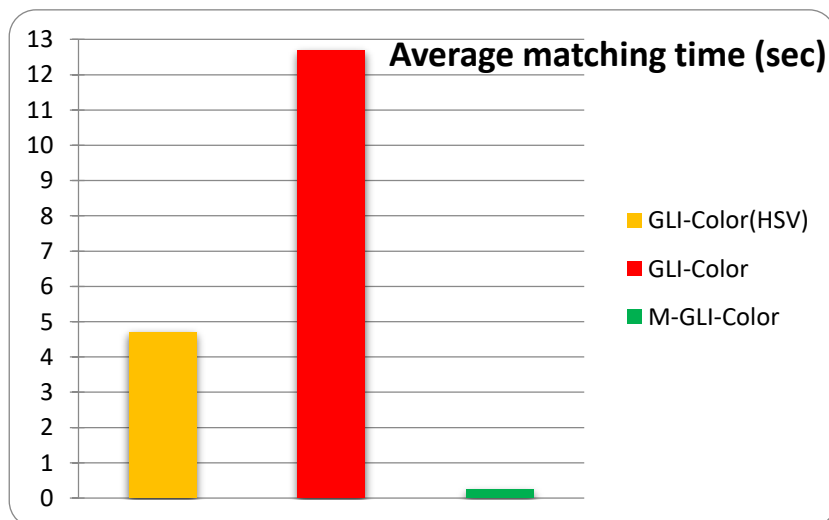


Figure V-7 Comparaison des trois variantes en termes de temps de comparaison moyenne (Seconds).

Tableau V-9 Comparaison en termes de temps de comparaison (seconds).

Numéro d'exécution	GLI-Color(HSV)	GLI-Color	M-GLI-Color
#1	4.18	14.42	0.266
#2	4.27	14.75	0.204
#3	4.26	10.92	0.257
#4	4.2	10.85	0.197
#5	5.81	14.48	0.213
#6	5.76	10.67	0.245
#7	4.21	14.47	0.253
#8	5.71	14.75	0.185
#9	4.33	10.93	0.261
#10	4.22	10.64	0.253
Moyenne	4.7	12.69	0.23
max	5.81	14.75	0.266
min	4.18	10.64	0.185

V.2 Exploitation de l'énergie des sous-bandes d'ondelettes

V.2.1 Motivation

Notre deuxième contribution ([Bougueroua et Boucheham](#); [Bougueroua et al., 2015](#)) est basée sur la constatation qu'il est rapporté dans la littérature que la moyenne de sous-bande d'ondelettes est proche de zéro. En se basant sur cette remarque et sachant que les caractéristiques sont généralement extraites de ces sous-bandes via les moments statistiques, nous avons posé l'hypothèse suivante : *le fait de calculer les moments en utilisant leurs formules standards (le calcul est fait autour de la moyenne) pour décrire les caractéristiques de sous-bandes va inhiber leurs capacités descriptives*. Donc, notre contribution consiste, en premier, à montrer expérimentalement la différence entre la distribution de l'énergie et celle de la moyenne de différentes sous-bandes. Puis, à mettre en évidence l'impact de l'intégration de l'énergie au lieu de la moyenne dans les formules des moments.

V.2.2 L'énergie et la moyenne de sous-bande d'ondelettes

Comme il a été mentionné ci-dessus, notre idée s'est basée sur la remarque qu'il est rapporté dans la littérature, en particulier dans ([Livens et al., 1997](#); [Lonnestad, 1992](#)), que *'the average of wavelet coefficients is close to zero'*. Ainsi, aucun des travaux publiés n'a accordé de l'attention à cette propriété à l'exception le travail de [Manthalkar et al. \(2003\)](#). Dans ce dernier, les auteurs ont calculé l'écart type autour de l'énergie de sous-bande au lieu de la moyenne. Cependant, aucune argumentation n'a été relatée.

Donc, afin de montrer expérimentalement la différence entre la distribution des valeurs de l'énergie et celle des valeurs de la moyenne, nous avons utilisé la base Outex (Outex_tc_00000 test suite). Nous avons utilisé également les ondelettes de Haar comme exemple de différents types d'ondelettes. Puis, nous avons calculé la moyenne des moyennes et la moyenne des

énergies pour chacune des sous-bandes de détails, obtenues de la décomposition en 2 niveaux (c.-à-d. HL1, LH1, HH1, HL2, LH2, et HH2) pour chacune des 24 classes de la base Outex_tc_00000.

Les résultats obtenus sont illustrés dans la Figure V-8 (Figure V-9) en termes de la moyenne et d'écart type de la moyenne (de l'énergie, respectivement) des canaux par rapport aux images de chaque classe (intra-class).

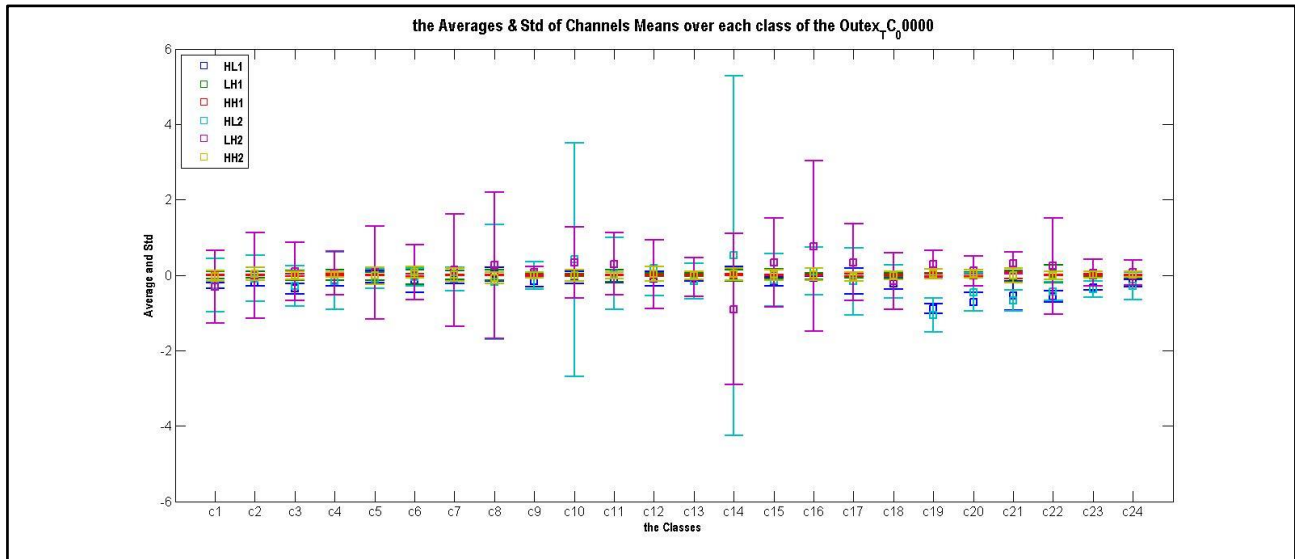


Figure V-8 Les moyennes et les écarts type des moyennes des canaux pour chacune des classes de la base Outex_TC_00000.

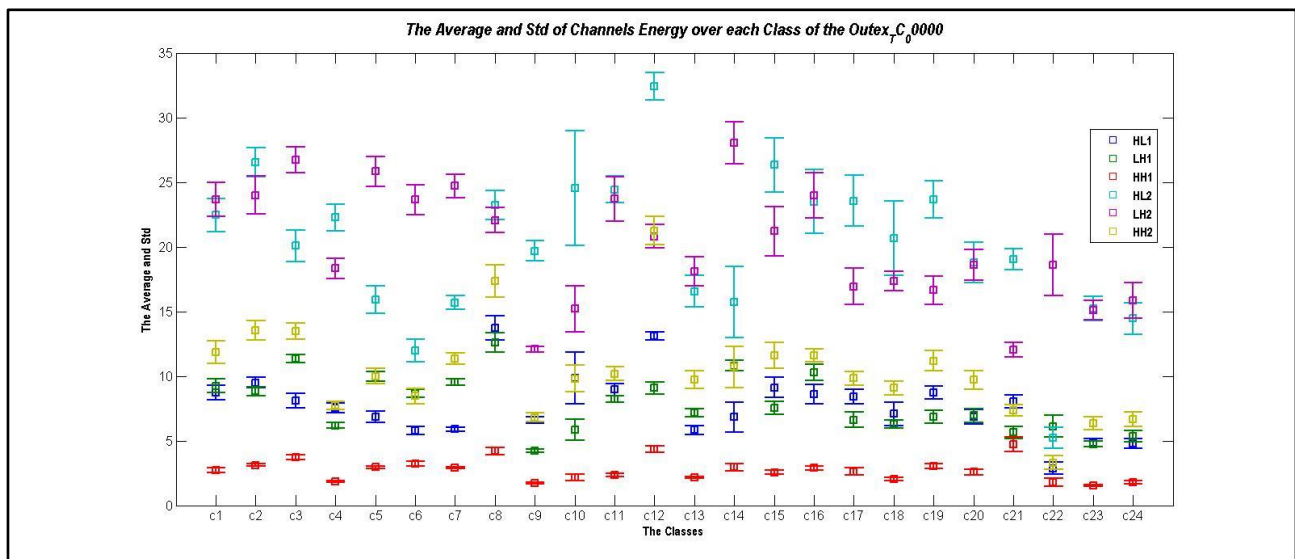


Figure V-9 Les moyennes et les écarts type des énergies des canaux pour chacune des classes de la base Outex_TC_00000.

V.2.2.1 Remarques et discussion

Une différence très claire peut être remarquée dès la première vue des deux plots. D'après la Figure V-8, nous remarquons que les moyennes des canaux sont très proches de zéro comme il a été prévu. En outre, il semble que les valeurs de la moyenne des moyennes pour chaque canal ont une basse variance par rapport à toute la base. La première colonne du Tableau V-10 rapportant les écarts type de la moyenne des moyennes sur toute la base (inter-class) de différents canaux prouve cette observation. Effectivement, toutes les valeurs d'écarts types sont proches de zéro. En revanche, la Figure V-9, montre que les valeurs d'énergie de canaux sont plus dispersées. En outre, il semble qu'elles sont plus variantes que les valeurs de la moyenne. Cette fois-ci, la deuxième colonne du Tableau V-10 prouve cette observation. Effectivement, les valeurs d'écarts types de la moyenne d'énergies sur toute la base (inter-class) des différents canaux sont clairement plus grandes.

Tableau V-10 L'écart type des moyennes de moyennes et des moyennes d'énergies sur toute la base (inter-class).

	l'écart type de la moyenne des moyennes	l'écart type des moyennes d'énergies
HL1	0.2295	2.42
LH1	0.0218	2.2068
HH1	0.0045	0.8492
HL2	0.3222	5.6818
LH2	0.2997	4.5301
HH2	0.0306	3.6676

V.2.3 L'impact de l'intégration de l'énergie dans les moments d'ondelettes

En s'appuyant sur les résultats précédents, nous avons mis en comparaison les quatre combinaisons de caractéristiques :

- **F1** : Energy-Std(mean).
- **F2** : Energy-Std(energy) ([Manthalkar et al., 2003](#)).
- **F3** : Energy-Std(energy)-Skew(mean).
- **F4** : Energy-Std(energy)-Skew(energy).

Pour cela, l'énergie (*Energy*) est calculée par l'équation Eq.IV.90. L'écart type autour de l'énergie (*Std(energy)*) est calculé par l'équation V.13, ci-dessous.

$$std^{(c)} = \frac{\sqrt{\sum_i (|e_i^{(c)}| - E^{(c)})^2}}{N^{(c)}} \quad (V.13)$$

Le skewness autour de l'énergie (*Skew(energy)*) et le skewness autour de la moyenne (*Skew(mean)*) sont calculés par les équations (V.14-V.15) ci-dessous, respectivement.

$$s^{(c)} = \sqrt[3]{\frac{\sum_i (|e_i^{(c)}| - E^{(c)})^3}{N^{(c)}}} \quad (\text{V.14})$$

$$skewness^{(c)} = \sqrt[3]{\frac{\sum_i (e_i^{(c)} - \mu^{(c)})^3}{N^{(c)}}} \quad (\text{V.15})$$

Où $\mu^{(c)}$ est la moyenne de la sous-bande (c), calculée comme suit :

$$\mu^{(c)} = \frac{\sum_i e_i^{(c)}}{N^{(c)}} \quad (\text{V.16})$$

V.2.4 Expérimentation

Les expérimentations ont été conduites sur deux bases d'images de texture. La première est Outex, plus précisément la suite de test Outex_TC_00000 (cf. Tableau II-1). Cette base est composée de 480 images de texture, distribuées sur 24 classes. La deuxième base est la base d'images Brodatz. Cette dernière est composée de 112 grandes images de texture (640*640 px). Donc, nous avons divisé chaque image en 16 sous-images non chevauchantes. Par conséquent, le nombre total d'images devient 1792 images. Nous avons également investigué plusieurs types d'ondelettes : Haar, Db2, Db3, Db4, Bior2.4, Bior3.3. Les filtres associés à ces ondelettes ont été rapportés dans le Tableau IV-1.

V.2.5 Résultats et discussion

En examinant les tableaux ci-dessous (Tableau V-11-Tableau V-12), rapportant les résultats obtenus sur la base Outex_TC_00000 en performant 2 et 3 décompositions, respectivement. On peut observer que les combinaisons qui intègrent le composant d'énergie sont plus performantes. Par exemple, les deux combinaisons F1 et F2 exploitent le même nombre de caractéristiques, cependant, F2 qui calcule l'écart type autour de l'énergie est plus performant que F1 (F2 : 78.74% vs F1 : 75.51%). Similairement, les combinaisons F3 et F4 exploitent trois caractéristiques, cependant la combinaison F4 qui calcule le skewness autour de l'énergie est plus performante. Les mêmes remarques peuvent être extraites à partir du Tableau V-12 mais avec des précisions plus élevées, à cause du nombre de sous-bandes considérées, et donc plus d'informations ont été intégrées.

Les performances des mesures intégrant l'énergie dans les formules des moments, sont également montrées à travers la Figure V-10. Dans cette figure, les courbes représentent la précision moyenne des différentes combinaisons en fonction du nombre d'images retournées.

Tableau V-11 Comparaison de la méthode proposée avec quelques combinaisons de caractéristiques en utilisant 2 niveaux de décompositions, (Outex_TC_00000, et $k=20$).

	F1: energy-std(mean)		F2: energy-std(energy) (Manthalkar et al., 2003)		F3: energy-std(Energy)-skew(mean)		F4: energy-std(Energy)-skew(energy)	
	<i>Manhattan</i>	<i>Canberra</i>	<i>Manhattan</i>	<i>Canberra</i>	<i>Manhattan</i>	<i>Canberra</i>	<i>Manhattan</i>	<i>Canberra</i>
Haar	69.11	72.60	65.85	77.57	72.93	76.81	80.25	80.40
Db2	71.90	76.02	<u>68.11</u>	79.31	74.37	78.50	78.73	80.85
Db3	<u>72.66</u>	77.44	67.89	80.14	<u>75.50</u>	79.73	79.01	81.34
Db4	71.81	<u>77.78</u>	67.33	<u>80.62</u>	74.99	<u>80.06</u>	79.14	82.03
Bior2.4	68.64	75.18	63.76	78.15	70.49	77.17	74.86	79.21
Bior3.3	67.24	74.02	63.51	76.64	69.58	76.08	73.34	77.43
Moyenne	70.23	75.51	66.08	78.74	72.98	78.06	77.56	80.21

Tableau V-12 Comparaison de la méthode proposée avec quelques combinaisons de caractéristiques en utilisant 3 niveaux de décompositions (Outex_TC_00000, et $k=20$).

	F1: energy-std(mean)		F2: energy-std(energy) (Manthalkar et al., 2003)		F3: energy-std(Energy)-skew(mean)		F4: energy-std(Energy)-skew(energy)	
	<i>Manhattan</i>	<i>Canberra</i>	<i>Manhattan</i>	<i>Canberra</i>	<i>Manhattan</i>	<i>Canberra</i>	<i>Manhattan</i>	<i>Canberra</i>
Haar	70.25	76.95	68.91	80.71	71.56	80.07	78.2	82.57
Db2	72.08	79.81	<u>70.70</u>	82.73	<u>74.57</u>	82.27	76.98	83.99
Db3	72.15	81.01	70.40	83.52	73.79	83.05	75.95	83.95
Db4	<u>72.53</u>	<u>81.54</u>	70.22	<u>83.88</u>	73.86	<u>83.46</u>	75.83	84.38
Bior2.4	67.45	79.58	65.71	82.83	68.68	82.02	73.98	83.23
Bior3.3	70.18	81.23	68.23	82.92	70.67	82.64	74.93	83.50
Moyenne	70.77	80.02	69.03	82.77	72.19	82.25	75.98	83.60

Quant aux résultats obtenus sur la base Brodatz, rapportés dans les tableaux ci-dessous (Tableau V-13 et Tableau V-14), ils montrent également un comportement similaire à celui obtenu sur la base Outex, mais avec des précisions inférieures. Cela peut être interprété par le fait que la taille de la base est plus grande que celle d'Outex_TC_00000. En outre, quelques classes de la base Brodatz sont visuellement très proches et se montrent difficiles à distinguer.

À partir du Tableau V-13 (respectivement Tableau V-14) la précision moyenne atteinte par F2 est 61.33% (65.75%) contre 57.05% (61.67%) seulement pour F1. Également, une précision moyenne de 63.29% (67.50%) est atteinte par F4 contre 60.71% (65.11%) seulement pour F3. Ces résultats prouvent que l'intégration de l'énergie a un impact positif sur les performances des caractéristiques.

La Figure V-11 montre également la supériorité des combinaisons intégrant l'énergie en termes de la précision moyenne et en fonction du nombre d'images retournées.

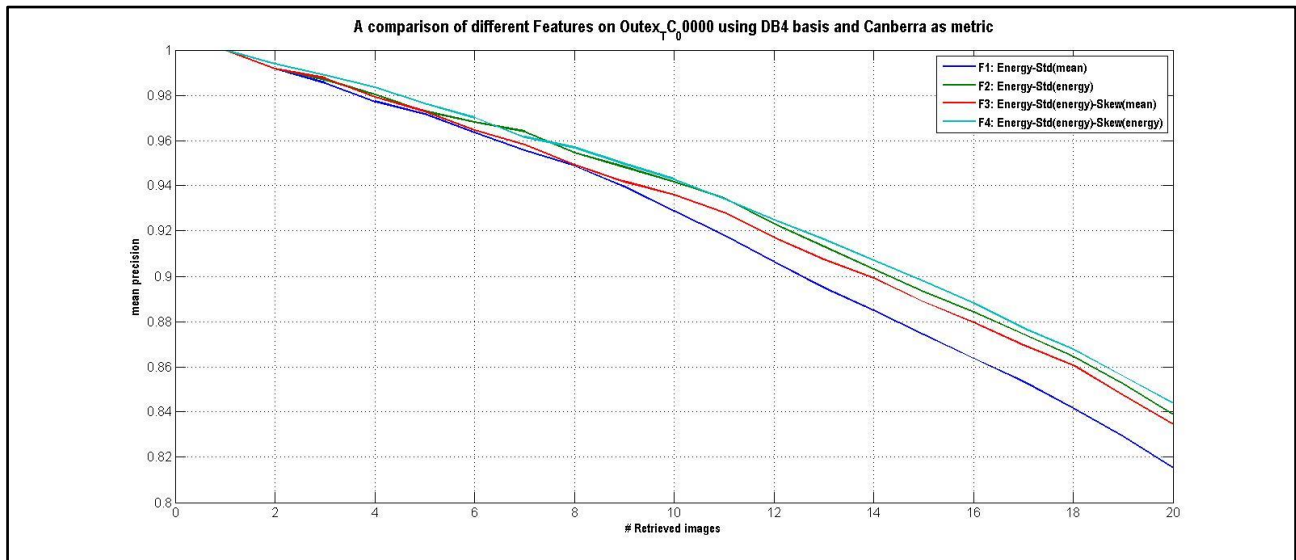


Figure V-10 Comparaison de performances des différentes méthodes en termes de la précision moyenne, (Outex_TC_00000).

Tableau V-13 Comparaison de la méthode proposée avec quelques combinaisons de caractéristiques en utilisant 2 niveaux de décompositions, (Brodatz, and $k=16$).

	F1: energy-std(mean)		F2: energy-std(energy) (Manthalkar et al., 2003)		F3: energy-std(Energy)-skew(mean)		F4: energy-std(Energy)-skew(energy)	
	Manhattan	Canberra	Manhattan	Canberra	Manhattan	Canberra	Manhattan	Canberra
Haar	55.18	57.08	50.09	61.59	57.99	61.07	62.60	63.80
Db2	55.34	57.81	51.23	62.27	58.57	<u>61.64</u>	63.40	64.26
Db3	<u>55.87</u>	58.22	52.17	62.33	<u>58.74</u>	61.56	63.33	64.08
Db4	55.58	<u>58.34</u>	<u>52.35</u>	<u>62.40</u>	58.38	61.63	63.15	64.04
Bior2.4	53.14	56.27	48.64	60.74	56.22	60.19	61.75	62.99
Bior3.3	50.29	54.56	45.87	58.63	52.67	58.16	57.71	60.55
Moyenne	54.23	57.05	50.06	61.33	57.10	60.71	61.99	63.29

Tableau V-14 Comparaison de la méthode proposée avec quelques combinaisons de caractéristiques en utilisant 3 niveaux de décompositions, (Brodatz, and $k=16$).

	F1: energy-std(mean)		F2: energy-std(energy) (Manthalkar et al., 2003)		F3: energy-std(Energy)-skew(mean)		F4: energy-std(Energy)-skew(energy)	
	Manhattan	Canberra	Manhattan	Canberra	Manhattan	Canberra	Manhattan	Canberra
Haar	57.28	61.68	54.49	65.86	59.82	65.27	64.94	67.49
Db2	<u>58.12</u>	62.57	55.39	66.84	<u>60.63</u>	66.25	65.58	68.71
Db3	58	62.87	<u>55.88</u>	66.84	59.99	66.21	64.44	68.33
Db4	57.23	<u>62.96</u>	55.24	<u>66.94</u>	59.55	<u>66.31</u>	64.27	68.60
Bior2.4	54.02	60.68	51.12	64.92	56.12	64.30	61.02	66.95
Bior3.3	51.95	59.23	49.77	63.11	53.89	62.31	57.91	64.89
Moyenne	56.1	61.67	53.65	65.75	58.33	65.11	63.03	67.50

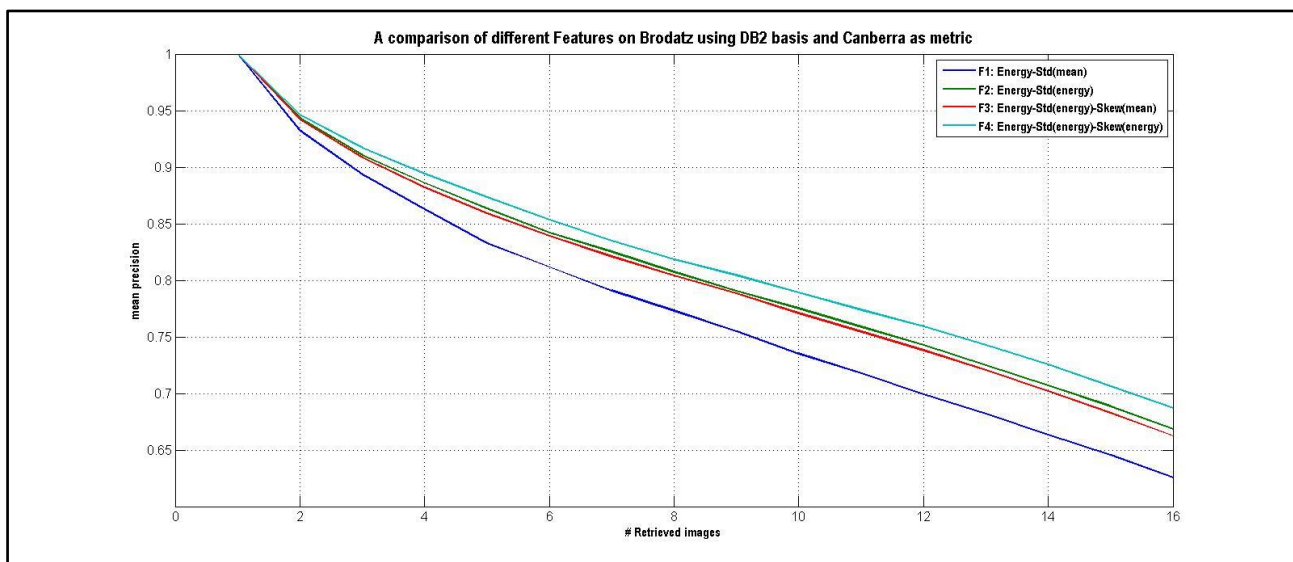


Figure V-11 Comparaison de performances des différentes méthodes en termes de la précision moyenne, (Brodatz).

V.3 Conclusion

Dans ce chapitre, nous avons présenté nos contributions qui ont été faites dans le cadre de cette thèse, où l'attribut visuel 'texture' était le point de focalisation. Principalement, deux grandes idées ont été présentées. La première consiste dans l'intégration graduelle de l'information locale. En particulier, son instance qui s'intéresse à l'extraction des caractéristiques de texture 'Gradual Locality Integration of Binary Patterns' (GLIBP) est le résultat de nos travaux et tentatives pour l'amélioration de la fameuse méthode LBP. En effet, GLIBP utilise des histogrammes locaux de régions prises d'une façon originale et révisée par rapport à notre première contribution (EILBP). Nous avons vu également que cette approche

d'intégration graduelle de l'information locale est généralisable pour extraire des caractéristiques de couleur. Dans ce contexte, nous avons développé plusieurs variantes : (GLI-Color, GLI-Color^(HSV), M-GLI-Color).

Notre deuxième axe de contribution était les ondelettes, où nous avons remarqué que quelques auteurs dans littérature avancent que les coefficients des sous-bandes d'ondelettes sont proches de zéro. En partant de ces propos et vu que les mesures calculées généralement des sous-bandes d'ondelettes sont l'énergie et les moments, nous avons étudié l'impact du remplacement de la moyenne par l'énergie dans les formules des moments. Un impact positif a été remarqué des résultats obtenus.

Vu les bons résultats obtenus en se basant sur les caractéristiques de la texture, nous pouvons confirmer ses capacités discriminatives intéressantes.

Dans le prochain chapitre, nous allons présenter notre dernière contribution, dans laquelle, nous proposons une nouvelle approche pour la recherche d'images par le contenu basée sur une stratégie gloutonne.

Chapitre VI. *Une nouvelle méthode Gloutonne pour la recherche d'images basée MSR (Mean Squared Residue)*

« Un problème créé ne peut être résolu en réfléchissant de la même manière qu'il a été créé. »

A. Einstein

Sommaire

VI.1	MOTIVATION	111
VI.2	RAPPEL	112
VI.2.1	<i>Biclustering</i>	112
VI.2.2	<i>Mean Squared Residue (MSR)</i>	112
VI.3	LE MSR POUR LA RECHERCHE D'IMAGES	113
VI.3.1	<i>Exemple</i>	114
VI.3.2	<i>Expérimentation</i>	115
VI.3.3	<i>Résultats et discussion</i>	115
VI.4	CONCLUSION	117

VI.1 Motivation

Notre contribution présentée dans ce chapitre est distinguée des précédentes par le fait que celle-ci agit au niveau de l'étape de comparaison. Notre méthode se distingue également des méthodes classiques qui sont basées sur la comparaison de requêtes avec les images cibles (*requête-images cibles*) par le fait qu'elle utilise une nouvelle stratégie qui consiste à exploiter, dans une itération donnée, toutes les images déjà récupérées, pour récupérer l'image suivante (c.-à-d. *l'ensemble des images retournées-images cibles*). A notre connaissance cette stratégie est totalement nouvelle vu que le seul travail existant qui a proposé une stratégie relativement similaire à la nôtre est celui de [Chen et al. \(2016\)](#). En effet, c'est dans ledit travail le concept d'*image-class* similarité a été introduit. Cette similarité est définie comme étant la somme entre une image de la base et chaque image de la classe de la

requête, où la classe de requête est composée de la requête elle-même et les images récupérées actuelles ([Chen et al., 2016](#)).

En outre, notre méthode adapte l'une des techniques du biclustering (biclustering des expressions génétiques) pour la recherche d'images par le contenu. À cet effet, nous avons posé l'hypothèse suivante : *les vecteurs de caractéristiques des images qui sont visuellement similaires ont la même variation (Co régulation) par rapport aux caractéristiques.* ([Bougueroua et Boucheham, 2016b](#))

VI.2 Rappel

VI.2.1 Biclustering

Différemment des techniques de clustering qui classifient les instances selon toutes leurs caractéristiques, les techniques du biclustering tentent de faire le clustering simultané d'instances et des caractéristiques en même temps (les lignes et les colonnes). En analyse des données d'expression, les lignes correspondent aux gènes et les colonnes aux conditions. Donc, le biclustering des données d'expression de gènes vise l'extraction du plus grand sous ensemble de gènes qui ont des activités très corrélées sous un ensemble de conditions.

Formellement parlant, soit A une matrice qui représente des données d'expression, composée de l'ensemble X de gènes (les lignes) et l'ensemble de conditions Y (les colonnes).

Donc, les techniques du biclustering visent l'extraction de la sous-matrice (le bicluster) A_{IJ} , où $I \subseteq X$ et $J \subseteq Y$ de sorte que les gènes en I soient très Co régulés par rapport à l'ensemble de conditions J .

Pour plus de détails, le lecteur intéressé est invité à se reporter à l'intéressant survey de [Madeira et Oliveira \(2004\)](#).

VI.2.2 Mean Squared Residue (MSR)

[Cheng et Church \(2000\)](#) ont introduit un algorithme efficace de suppression de nœuds pour le biclustering des données d'expression. Leur algorithme tente de trouver le bicluster maximum dont le Mean Squared Residue (MSR) est le minimum.

Le MSR H du bicluster composé de l'ensemble des lignes I et l'ensemble de colonnes J est calculé comme suit :

$$H(I, J) = \frac{1}{|I||J|} \sum_{i \in I, j \in J} (a_{ij} - a_{iJ} - a_{Ij} + a_{IJ})^2 \quad (\text{VI.1})$$

Où :

$$a_{iJ} = \frac{1}{|J|} \sum_{j \in J} a_{ij} \quad (\text{VI.2})$$

$$a_{Ij} = \frac{1}{|I|} \sum_{i \in I} a_{ij} \quad (\text{VI.3})$$

Et

$$a_{IJ} = \frac{1}{|I||J|} \sum_{i \in I, j \in J} a_{ij} \quad (\text{VI.4})$$

VI.3 Le MSR pour la recherche d'images

Dans ce travail, nous avons adapté l'utilisation du MSR pour la recherche d'images par le contenu. Donc, la matrice qui représente les données d'expression est remplacée par l'index (c.-à-d. l'ensemble des vecteurs de caractéristiques des images de la base de données). Par conséquent, les lignes correspondent aux images et les colonnes aux caractéristiques.

Étant donné que nous nous sommes intéressés à la recherche d'images en se basant sur l'attribut 'texture', nous avons exploité la méthode LBP (cf. sec.IV.1.3). Plus précisément, le LBP uniforme ($LBP_{P,R}^{U2}$), avec ($P = 8$ et $R = 1$) (cf. sec.IV.1.3.1.1.1), qui est plus compact, a été utilisé. Cependant, l'utilisation d'autres descripteurs est prévue, tels que l'histogramme de couleur. Donc, si la base contient M images, dont les caractéristiques sont extraites par $LBP_{8,1}^{U2}$ alors l'index est de taille $M \times 59$ valeurs.

L'algorithme proposé est composé de deux étapes principales, comme suit :

- Étape 1 : *Initialisation*. Soit I , le cluster qui va contenir les vecteurs de caractéristiques des images pertinentes par rapport à la requête Q . Dans cette étape, cet ensemble est initialisé à *vide* (étape 1, dans la Figure VI-1).
- Étape 2 : *Elargissement*. Dans cette étape, I est élargi en ajoutant les images sélectionnées de la base de données, une image par itération, de sorte que l'image à ajouter soit celle qui préserve l'homogénéité globale de l'ensemble I . Pour cela, une recherche gloutonne est utilisée. Donc, l'image à insérer dans le cluster, est celle dont l'inclusion au cluster produit le moins *coût*. En effet, le *coût* n'est que la *différence* entre le MSR du cluster *avant* et le MSR *après* l'ajout du vecteur de caractéristique d'une image i (Eq.VI.5-VI.6).

$$\Delta MSR_i(Q, I) = H(Q \cup I, Y) - H(Q \cup I \cup \{i\}, Y) \quad (\text{VI.5})$$

$$\arg \min_i (\Delta MSR_i(Q, I)) \quad (\text{VI.6})$$

La description de l'algorithme est la suivante :

Algorithm: texture image retrieval using the MSR;

Input: $\text{index:matrix}[m,n]$;

//m the number of images in the database.

//n the number of patterns

Q: vector; // Query Feature vector

sizeCR: integer; // the maximum number of rows (images) in //the result cluster **I**.

Output: the set **I** of images which are visually most //similar to the query **Q**.

Begin

$I = \emptyset$;

While ($\text{size}(I) < \text{sizeCR}$) **do**

begin

 //Search for image **i** which brings the minimum ΔMSR //among the images not included in the current cluster.

$i = \text{argmin}(\Delta\text{MSR}_i(Q, I))$

$I = I \cup \{F_{V_i}\}$;

End;

End.

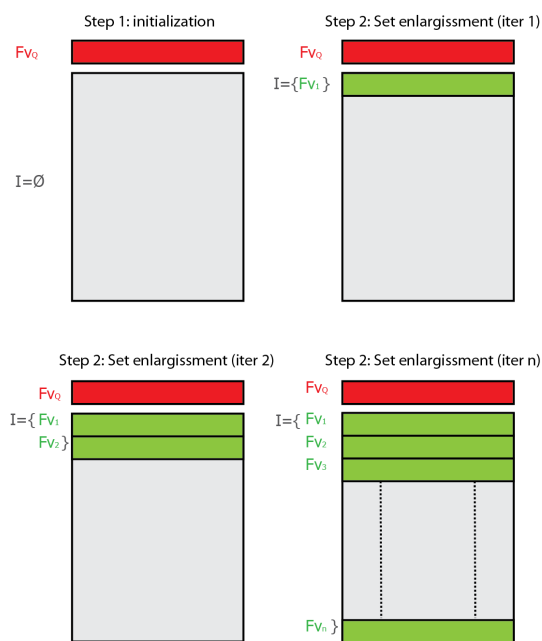


Figure VI-1 Les étapes d'élargissement de l'ensemble I .

VI.3.1 Exemple

Soit les cinq vecteurs V_1, V_2, V_3, V_4, V_5 , définis comme suit :

- $V_1 = [7, 8, 9, 8, 10, 7, 8]$.
- $V_2 = [4, 5, 6, 5, 7, 4, 5]$.
- $V_3 = [1, 3, 4, 2, 5, 1, 2]$.
- $V_4 = [9, 7, 8, 9, 8, 7, 9]$.
- $V_5 = [8, 7, 8, 9, 10, 9, 8]$.

On suppose également d'avoir une requête dont le vecteur est :

$$V_Q = [7, 8, 9, 8, 10, 7, 8].$$

En utilisant une métrique de distance classique telle que la distance Euclidienne, ces vecteurs seront classés selon leurs distances de la requête comme suit V_1 ($d=0$), V_5 ($d=2.83$), V_4 ($d=3.46$), V_2 ($d=7.94$), et V_3 ($d=14.8$). En revanche, en utilisant notre méthode basée sur le MSR, ces vecteurs seront classés comme suit : V_1 ($\Delta MSR_{V1}=0$), V_2 ($\Delta MSR_{V2}=0$), V_3 ($\Delta MSR_{V3}=0.0459$), V_5 ($\Delta MSR_{V5}=0.1957$), et V_4 ($\Delta MSR_{V4}=0.1926$).

VI.3.2 Expérimentation

Les expérimentations ont été conduites sur deux bases de données, à savoir la suite de test Outex_TC_00000 de la base Outex, et aussi la base Brodatz. Dans cette dernière, chacune des 112 images de texture a été divisée en 16 sous-images (4×4), non chevauchantes. Quant à l'évaluation, la précision moyenne (AP) est utilisée.

VI.3.3 Résultats et discussion

La comparaison entre les résultats obtenus par l'approche proposée et ceux obtenus par la distance Euclidienne sur la base Outex (Tableau VI-1) montre que notre approche est plus performante que la distance Euclidienne. En effet, notre méthode basée sur le MSR atteint 87.68% alors que la distance Euclidienne atteint 85.18% seulement. En outre, l'analyse des résultats classe par classe, montre une autre intéressante remarque ; lorsque la distance Euclidienne atteint 90% d'AP, le MSR donne une AP parfaite (exception de la classe C3).

Les résultats obtenus sur la base Brodatz montrent également des performances très compétitives par rapport à celles de la distance Euclidienne (Tableau VI-2). En effet, une précision moyenne de 73.59% est atteinte par notre méthode contre 72.67% seulement pour la distance Euclidienne. En outre, l'analyse des résultats classe par classe montre que notre méthode est plus performante que lorsqu'on se base sur la distance Euclidienne dans 57 classes. Alors qu'elle est moins performante dans 15 classes seulement.

Tableau VI-1 Comparaison de la méthode proposée avec la distance Euclidienne en termes de précision moyenne (la base Outex_tc_00000, window=20).

Texture	Proposée	Euclidean	Texture	Proposée	Euclidean
C0	100	96.5	C12	100	95.5
C1	100	96	C13	50	50.3
C2	100	90.5	C14	40	50.5
C3	99.5	90.5	C15	74	67.5
C4	100	100	C16	56.3	57
C5	100	100	C17	58	55.3
C6	100	100	C18	100	97.3
C7	100	100	C19	78.8	74.5
C8	100	98.3	C20	100	96
C9	77.8	68.3	C21	100	99.8
C10	100	100	C22	95	89
C11	100	99	C23	75	72.5
			Total	87.68	85.18

Tableau VI-2 Comparaison de la méthode proposée avec la distance Euclidienne en termes de précision moyenne (la base Brodatz (16 sub-images×112)).

Texture	Proposée	Euclidean	Texture	Proposée	Euclidean	Texture	Proposée	Euclidean
D1	94.5	89.8	D39	25	32.8	D77	100	100
D2	43.4	44.1	D40	99.2	86.7	D78	100	85.2
D3	63.7	72.7	D41	93.8	85.2	D79	94.1	85.5
D4	74.2	73	D42	58.6	66.4	D80	100	89.1
D5	45.3	49.6	D43	13.3	17.6	D81	100	94.9
D6	100	100	D44	19.1	21.9	D82	100	95.3
D7	43	46.1	D45	14.1	23	D83	100	100
D8	39.8	41.4	D46	94.1	82.4	D84	100	99.2
D9	59.4	55.9	D47	100	96.5	D85	100	99.2
D10	94.5	85.2	D48	71.1	62.9	D86	100	98.8
D11	60.2	63.3	D49	100	100	D87	94.1	94.9
D12	89.5	78.9	D50	94.5	80.9	D88	26.2	40.6
D13	25	36.7	D51	83.6	81.6	D89	39.8	49.6
D14	94.1	94.9	D52	100	99.2	D90	48.8	50
D15	58.2	56.6	D53	100	100	D91	33.2	40.2
D16	100	100	D54	44.5	48.4	D92	98	94.1
D17	100	94.5	D55	100	100	D93	50	66
D18	100	89.5	D56	100	100	D94	61.3	62.9
D19	67.6	69.9	D57	100	100	D95	89.1	85.2
D20	100	100	D58	28.9	33.2	D96	87.9	84
D21	100	100	D59	14.8	23.4	D97	50.4	50.4
D22	53.9	52.7	D60	72.7	66.4	D98	61.7	57
D23	33.2	41.4	D61	69.1	63.3	D99	36.7	41
D24	82.8	72.3	D62	65.6	63.7	D100	36.7	39.5
D25	63.3	66.4	D63	31.3	35.5	D101	100	97.3
D26	87.5	80.9	D64	100	100	D102	100	98.8
D27	50	49.2	D65	100	100	D103	100	97.3
D28	94.1	85.2	D66	94.9	89.5	D104	94.1	87.1
D29	100	92.2	D67	88.7	80.5	D105	76.6	79.7
D30	41	48	D68	97.7	96.1	D106	60.5	70.3
D31	29.3	36.3	D69	48	49.6	D107	47.3	53.1
D32	100	100	D70	100	97.3	D108	68	67.2
D33	100	98.8	D71	71.1	68.8	D109	99.2	91.8
D34	98	86.7	D72	30.1	38.7	D110	100	99.2
D35	28.9	33.2	D73	32.8	41.4	D111	100	79.7
D36	29.7	41	D74	93	77.3	D112	83.2	74.6
D37	82	80.9	D75	92.6	86.7			
D38	36.7	42.2	D76	100	93.4	Total	73.59	72.67

VI.4 Conclusion

Nous avons présenté dans ce chapitre la dernière de nos contributions faites dans le cadre de cette thèse. Cette dernière est distinguée des premières par le fait qu'elle s'intéresse à l'étape de comparaison. Donc, comme nous avons vu notre contribution est pleine d'originalités. Effectivement, notre approche a été développée en s'appuyant sur l'hypothèse que nous avons posée, stipulant que *“les vecteurs de caractéristiques des images qui sont visuellement similaires ont la même variation (Co régulation) par rapport aux caractéristiques.”* D'où notre utilisation de la mesure MSR qui a été originalement proposée pour le biclustering des expressions génétiques. À cet effet, nous avons adapté l'algorithme à notre domaine qui est le CBIR. Par conséquent, toutes les images retournées contribuent à la sélection de la suivante, ce qui diffère de l'approche typique utilisée en CBIR où seulement la requête est considérée lors de la recherche.

Les résultats des expériences conduites sur différentes bases de données montrent des performances intéressantes vis-à-vis de la mesure considérée.

Chapitre VII. Conclusion

générale

« En toute chose, c'est la fin qui est essentiel. »
Aristote

Nous nous sommes intéressés dans ce travail à la réduction du *fossé sémantique*, un problème majeur dans les systèmes de recherche d'images par le contenu visuel. Ce problème provient de l'appui sur les caractéristiques de bas niveau telles que la couleur, la texture, la forme... etc. pour décrire les sémantiques d'images de haut niveau. Nous nous sommes particulièrement intéressés à l'attribut 'texture', où nous avons montré, dans un premier temps, son apport à la caractérisation des images en montrant la popularité de ce descripteur dans la littérature. Ensuite, nous avons proposé des descripteurs de bas niveau, mais plus discriminants, afin de réduire ce fossé, en utilisant cet attribut.

La première partie de ce manuscrit a été organisée comme suit : en premier lieu, nous avons présenté et discuté les deux étapes principales des systèmes CBIR, à savoir l'étape d'extraction des caractéristiques et l'étape de comparaison. Dans le Chapitre III. , nous avons survolé les différents attributs visuels ainsi qu'un ensemble de méthodes d'extraction de caractéristiques pour chacun des attributs visuels. Vu que l'attribut texture est le point central dans notre travail, nous nous sommes limités dans ce chapitre à la présentation et la discussion de la texture ainsi qu'à la mise en évidence de ses différentes applications. Alors que dans le Chapitre IV. , nous avons abordé une liste, relativement longue, de descripteurs de texture. En particulier, la méthode LBP et les ondelettes ont eu une attention particulière. À travers ces travaux présentés ainsi que la diversité des problèmes abordés dans la littérature en utilisant la texture, nous avons mis l'accent sur l'apport de cet attribut à la caractérisation des images en général.

Comme vu d'ensemble de cette première partie, nous pouvons avancer les remarques suivantes : la prise en compte de l'information locale est très avantageuse, vu la considération de l'information spatiale qui permet d'améliorer les performances des méthodes de CBIR. En particulier, nous avons vu que la plupart des descripteurs de couleur suivent une approche locale. En ce qui concerne la texture, nous avons remarqué qu'en général, les approches statistiques et les approches basées sur les transformations (les méthodes spatio-fréquentielles) ont montré de bonnes performances. Nous avons aussi remarqué que, à la différence de l'attribut couleur, la texture est exploitée pour différents problèmes telles la classification, la segmentation... etc. Cette diversité montre clairement le rôle et l'apport de la texture à la caractérisation de l'image. Particulièrement, la méthode LBP est largement exploitée dans ce cadre, où des dizaines de variantes ont été proposées. Cependant, nous avons constaté que les modifications apportées restent généralement au niveau du voisinage et très peu sont les travaux qui ont accordé l'attention à la considération des motifs binaires locaux dans des régions.

Dans la deuxième partie de ce manuscrit, nous avons présenté nos contributions proposées dans le cadre de cette thèse. Dans le premier chapitre de cette partie (Chapitre V.), nous avons présenté nos contributions au niveau de l'étape d'extraction des caractéristiques. En premier lieu, notre méthode d'intégration graduelle de l'information locale des motifs binaires (GLIBP) a été présentée. GLIBP propose un nouveau schéma de délimitation de régions permettant de meilleures localisations des caractéristiques de l'image. Nous avons, ainsi, illustré qu'elle manifeste de bonnes performances vis-à-vis de plusieurs méthodes de l'état de l'art. De même, de bonnes performances ont été enregistrées sur la couleur par notre méthode GLI-Color, ainsi que ses variantes GLI-Color^(HSV) et M-GLI-Color. Dans le même chapitre, nous avons montré aussi l'impact de l'intégration de l'énergie dans les formules des moments statistiques des sous-bandes d'ondelettes au lieu de la moyenne. Les résultats obtenus montrent que cette modification est assez efficace.

Dans le deuxième chapitre de la partie des contributions, nous avons présenté une nouvelle méthode pour la recherche et la récupération des images. Notre méthode repose sur une approche gloutonne basée sur la mesure MSR importée du domaine du biclustering des expressions de gènes. La méthode proposée diffère de l'approche classique de comparaison qui compare la requête avec chaque image de la base par le fait qu'elle utilise dans une itération donnée toutes les images déjà récupérée (donc jugées pertinentes par le système) pour rechercher l'image suivante. Cette approche de recherche a montré d'intéressantes performances vis-à-vis de la distance Euclidienne des histogrammes de la LBP. Par conséquent, nous pouvons conclure que la considération des résultats partiels pour la suite de la recherche en gardant l'homogénéité des résultats peut parvenir à l'amélioration des performances des systèmes CBIR.

VII.1 Perspectives

En ce qui concerne les futurs travaux, nous avançons l'ensemble suivant d'idées et d'axes à emprunter.

Comme première perspective, nous projetons d'appliquer notre méthode GLIBP sur des bases généralistes de taille plus grande (voire des bases réelles) afin de la valoriser davantage.

Vu l'approche distincte que les techniques du biclustering suivent, l'exploitation de ces dernières dans le domaine de CBIR pourrait être un sujet d'étude très intéressant. Dans ce contexte, une extension de notre travail est envisagée, en particulier, par la considération d'autres attributs et descripteurs.

La troisième de nos perspectives est la combinaison des attributs, vu que les images, en particulier celles du grand public, sont complexes et les différents attributs peuvent apporter de distinctes informations ; ce qui permet d'améliorer la qualité des résultats de la recherche des images.

Bibliographie

- A, V., Hebbar, D., Shekhar, V. S., Murthy, K. N. B. et Natarajan, S. (2015). Two Novel Detector-Descriptor Based Approaches for Face Recognition Using SIFT and SURF. *Procedia Computer Science*, 70, 185-197. doi: 10.1016/j.procs.2015.10.070
- Ahonen, T., Hadid, A. et Pietikainen, M. (2006). Face description with local binary patterns: application to face recognition. *IEEE Trans Pattern Anal Mach Intell*, 28(12), 2037-2041. doi: 10.1109/TPAMI.2006.244
- Ahonen, T., Hadid, A. et Pietikainen, M. (2004). Face Recognition with Local Binary Patterns. Dans T. Pajdla & J. Matas (dir.), *Computer Vision - ECCV 2004: 8th European Conference on Computer Vision, Prague, Czech Republic, May 11-14, 2004. Proceedings, Part I* (p. 469-481). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Akbarizadeh, G. (2012). A New Statistical-Based Kurtosis Wavelet Energy Feature for Texture Recognition of SAR Images. *IEEE Transactions on Geoscience and Remote Sensing*, 50(11), 4358-4368. doi: 10.1109/tgrs.2012.2194787
- Aksoy, S. et Haralick, R. M. (1998). Content-based image database retrieval using variances of gray level spatial dependencies. Dans H. H. S. Ip & A. W. M. Smeulders (dir.), *Multimedia Information Analysis and Retrieval: IAPR International Workshop, MINAR'98 Hong Kong, China, August 13-14, 1998 Proceedings* (p. 3-19). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Alvarez-Betancourt, Y. et Garcia-Silvente, M. (2016). A keypoints-based feature extraction method for iris recognition under variable image quality conditions. *Knowledge-Based Systems*, 92, 169-182. doi: <http://dx.doi.org/10.1016/j.knosys.2015.10.024>
- Appas, A. R., Darwish, A. M., El-Desouki, A. I. et Shaheen, S. I. (1998). *Image indexing using composite regional color channel features*. Repéré à <http://dx.doi.org/10.1117/12.333868>
- Arevalillo-Herráez, M., Domingo, J. et Ferri, F. J. (2008). Combining similarity measures in content-based image retrieval. *Pattern Recognition Letters*, 29(16), 2174-2181. doi: 10.1016/j.patrec.2008.08.003
- Arivazhagan, S. et Ganesan, L. (2003). Texture classification using wavelet transform. *Pattern Recognition Letters*, 24(9-10), 1513-1521. doi: 10.1016/s0167-8655(02)00390-2
- Arivazhagan, S., Ganesan, L. et Priyal, S. P. (2006). Texture classification using Gabor wavelets based rotation invariant features. *Pattern Recognition Letters*, 27(16), 1976-1982. doi: 10.1016/j.patrec.2006.05.008
- Ayres, F. J., Rangayyan, R. M. et Desautels, J. E. L. (2010). Analysis of Oriented Texture with Applications to the Detection of Architectural Distortion in Mammograms. *Synthesis Lectures on Biomedical Engineering*, 5(1), 1-162. doi: 10.2200/s00301ed1v01y201010bme038
- Azhar, R., Tuwohingide, D., Kamudi, D., Sarimuddin et Suciati, N. (2015). Batik Image Classification Using SIFT Feature Extraction, Bag of Features and Support Vector Machine. *Procedia Computer Science*, 72, 24-30. doi: 10.1016/j.procs.2015.12.101
- Barrena, M., Caro, A., Durán, M. L., Rodríguez, P. G., Arias-Nicolás, J. P. et Alonso, T. (2015). Qatris iManager: a general purpose CBIR system. *Machine Vision and Applications*, 26(4), 423-442. doi: 10.1007/s00138-015-0672-3
- Bay, H., Ess, A., Tuytelaars, T. et Van Gool, L. (2008). Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*, 110(3), 346-359. doi: 10.1016/j.cviu.2007.09.014
- Bay, H., Tuytelaars, T. et Van Gool, L. (2006). SURF: Speeded Up Robust Features. Dans A. Leonardis, H. Bischof & A. Pinz (dir.), *Computer Vision – ECCV 2006: 9th European*

- Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I* (p. 404-417). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Beudet, P. R. (1978). *Rotationally invariant image operators*. Communication présentée International Joint Conference on Pattern Recognition.
- Bianconi, F. et Fernández, A. (2014). An appendix to “Texture databases – A comprehensive survey”. *Pattern Recognition Letters*, 45, 33-38. doi: 10.1016/j.patrec.2014.02.017
- Bicego, M. et Lovato, P. (2016). A bioinformatics approach to 2D shape classification. *Computer Vision and Image Understanding*, 145, 59-69. doi: 10.1016/j.cviu.2015.11.011
- Bo, Y., Lei, G., Tianyun, Z. et Xiaoliang, Q. (2010). *A curve matching algorithm based on Freeman Chain Code*. doi: 10.1109/icip.2010.5658385
- Bougueroua, S. et Boucheham, B. Impact of Energy integration in Wavelet sub-bands Moments in image retrieval: an experimental study. (*To be Submitted*).
- Bougueroua, S. et Boucheham, B. (2014, 9-10 Nov. 2014). *Ellipse based local binary pattern for color image retrieval*. Communication présentée ISKO-Maghreb: Concepts and Tools for knowledge Management (ISKO-Maghreb), 2014 4th International Symposium, Algiers, Algeria. doi: 10.1109/ISKO-Maghreb.2014.7033459
- Bougueroua, S. et Boucheham, B. (2015). *GLI-Color: Gradual Locality Integration of Color features for image retrieval*. Communication présentée Proceedings of the International Conference on Intelligent Information Processing, Security and Advanced Communication, Batna, Algeria. doi: 10.1145/2816839.2816920
- Bougueroua, S. et Boucheham, B. (2016a). GLIBP: Gradual Locality Integration of Binary Patterns for Scene Images Retrieval. *Journal of Information Processing Systems (Accepted 2016)*.
- Bougueroua, S. et Boucheham, B. (2016b). Greedy Mean Squared Residue for Texture Images Retrieval. Dans S. Chikhi, A. Amine, A. Chaoui, K. M. Kholliadi & E. D. Saidouni (dir.), *Modelling and Implementation of Complex Systems: Proceedings of the 4th International Symposium, MISC 2016, Constantine, Algeria, May 7-8, 2016, Constantine, Algeria* (p. 123-136). Cham: Springer International Publishing.
- Bougueroua, S., Boucheham, B. et Belaoued, M. (2015, 10-12 Nov 2015). *Enhancement of Image Characterization with Discrete Wavelet Coefficients using the Third Statistical Moment for Texture-Based Image Retrieval*. Communication présentée Conférence Nationale sur les Nouvelles Technologies Informatiques (CN2TI'15), Guelma, Algeria.
- Bovik, A. C., Clark, M. et Geisler, W. S. (1990). Multichannel texture analysis using localized spatial filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(1), 55-73. doi: 10.1109/34.41384
- Calonder, M., Lepetit, V., Strecha, C. et Fua, P. (2010). BRIEF: Binary Robust Independent Elementary Features. Dans K. Daniilidis, P. Maragos & N. Paragios (dir.), *Computer Vision – ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part IV* (p. 778-792). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Campbell, F. W. et Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *J Physiol*, 197(3), 551-566.
- Carson, C., Thomas, M., Belongie, S., Hellerstein, J. M. et Malik, J. (1999). *Blobworld: A system for region-based image indexing and retrieval*. Communication présentée International Conference on Advances in Visual Information Systems.
- Castellano, G., Bonilha, L., Li, L. M. et Cendes, F. (2004). Texture analysis of medical images. *Clin Radiol*, 59(12), 1061-1069. doi: 10.1016/j.crad.2004.07.008
- Chang, T. et Kuo, C. J. (1993). Texture analysis and classification with tree-structured wavelet transform. *IEEE Trans Image Process*, 2(4), 429-441. doi: 10.1109/83.242353

- Chao, W.-L., Ding, J.-J. et Liu, J.-Z. (2015). Facial expression recognition based on improved local binary pattern and class-regularized locality preserving projection. *Signal Processing*, 117(0), 1-10. doi: 10.1016/j.sigpro.2015.04.007
- Chellappa, R. et Chatterjee, S. (1985). Classification of textures using Gaussian Markov random fields. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 33(4), 959-963. doi: 10.1109/tassp.1985.1164641
- Chen, C.-C. et Huang, C.-L. (1993). Markov random fields for texture classification. *Pattern Recognition Letters*, 14(11), 907-914. doi: 10.1016/0167-8655(93)90155-7
- Chen, J., Wang, Y., Luo, L., Yu, J.-G. et Ma, J. (2016). Image retrieval based on image-to-class similarity. *Pattern Recognition Letters*, 83, 379-387. doi: 10.1016/j.patrec.2016.01.017
- Chen, Y. Q., Nixon, M. S. et Thomas, D. W. (1995). Statistical geometrical features for texture classification. *Pattern Recognition*, 28(4), 537-552. doi: 10.1016/0031-3203(94)00116-4
- Cheng, Y. et Church, G. M. (2000). *Biclustering of Expression Data*. Communication présentée Proceedings of the Eighth International Conference on Intelligent Systems for Molecular Biology.
- Chih-Yi, C., Hsin-Chih, L. et Shi-Nine, Y. (2003, 25-28 May 2003). *A fuzzy logic CBIR system*. Communication présentée Fuzzy Systems, 2003. FUZZ '03. The 12th IEEE International Conference on. doi: 10.1109/fuzz.2003.1206597
- Costa, Y. M. G., Oliveira, L. S., Koerich, A. L., Gouyon, F. et Martins, J. G. (2012). Music genre classification using LBP textural features. *Signal Processing*, 92(11), 2723-2737. doi: 10.1016/j.sigpro.2012.04.023
- Cross, G. R. et Jain, A. K. (1983). Markov random field texture models. *IEEE Trans Pattern Anal Mach Intell*, 5(1), 25-39. doi: 10.1109/tpami.1983.4767341
- Cruz-Aceves, I., Oloumi, F., Rangayyan, R. M., Aviña-Cervantes, J. G. et Hernandez-Aguirre, A. (2016). Automatic segmentation of coronary arteries using Gabor filters and thresholding based on multiobjective optimization. *Biomedical Signal Processing and Control*, 25, 76-85. doi: 10.1016/j.bspc.2015.11.001
- Datta, R., Joshi, D., Li, J. et Wang, J. Z. (2008). Image retrieval. *ACM Computing Surveys*, 40(2), 1-60. doi: 10.1145/1348246.1348248
- Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 41(7), 909-996. doi: 10.1002/cpa.3160410705
- Davies, E. R. (2012). Introduction to Texture Analysis (*Handbook of Texture Analysis* (p. 1-31): PUBLISHED BY IMPERIAL COLLEGE PRESS AND DISTRIBUTED BY WORLD SCIENTIFIC PUBLISHING CO.
- Doshi, N. P. et Schaefer, G. (2012, 11-15 Nov. 2012). *A comprehensive benchmark of local binary pattern algorithms for texture retrieval*. Communication présentée Pattern Recognition (ICPR), 2012 21st International Conference on.
- Drimbarean, A. et Whelan, P. F. (2001). Experiments in colour texture analysis. *Pattern Recognition Letters*, 22(10), 1161-1167. doi: 10.1016/s0167-8655(01)00058-7
- Dubes, R. C. et Jain, A. K. (1989). Random field models in image analysis. *Journal of Applied Statistics*, 16(2), 131-164. doi: 10.1080/02664768900000014
- Eakins, J. P. et Graham, M. E. (1999). Content based image retrieval: A report to the JISC technology applications programme.
- Ebner, M. (2007). *Color Constancy*.
- Fairchild, M. D. (2005). *Color Appearance Models*. (2nd^e éd.). John Wiley & Sons Ltd.
- Fernández, A., García, R., Usamentiaga, R. et Casado, R. (2015). Glasses detection on real images based on robust alignment. *Machine Vision and Applications*, 26(4), 519-531. doi: 10.1007/s00138-015-0674-1
- Ford, A. et Roberts, A. (1998). Colour space conversions.

- Ford, A. et Roberts, A. (1998(b)). Colour Space Conversions.
- Freeman et Davis. (1977). A Corner-Finding Algorithm for Chain-Coded Curves. *IEEE Transactions on Computers*, C-26(3), 297-303. doi: 10.1109/tc.1977.1674825
- Frighetto-Pereira, L., Rangayyan, R. M., Metzner, G. A., de Azevedo-Marques, P. M. et Nogueira-Barbosa, M. H. (2016). Shape, texture and statistical features for classification of benign and malignant vertebral compression fractures in magnetic resonance images. *Comput Biol Med*, 73, 147-156. doi: 10.1016/j.combiomed.2016.04.006
- Gelzinis, A., Verikas, A. et Bacauskiene, M. (2007). Increasing the discrimination power of the co-occurrence matrix-based features. *Pattern Recognition*, 40(9), 2367-2372. doi: 10.1016/j.patcog.2006.12.004
- Guo, Z., Zhang, L. et Zhang, D. (2010). A completed modeling of local binary pattern operator for texture classification. *IEEE Trans Image Process*, 19(6), 1657-1663. doi: 10.1109/TIP.2010.2044957
- Hadid, A., Pietikainen, M. et Ahonen, T. (2004, 27 June-2 July 2004). *A discriminative feature space for detecting and recognizing faces*. Communication présentée Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on. doi: 10.1109/CVPR.2004.1315246
- Hadid, A., Ylioinas, J., Bengherabi, M., Ghahramani, M. et Taleb-Ahmed, A. (2015). Gender and texture classification: A comparative analysis using 13 variants of local binary patterns. *Pattern Recognition Letters*, 68(0), 231-238. doi: 10.1016/j.patrec.2015.04.017
- Hadizadeh, H. (2015). Multi-resolution local Gabor wavelets binary patterns for gray-scale texture description. *Pattern Recognition Letters*, 65, 163-169. doi: 10.1016/j.patrec.2015.07.038
- Hafiane, A., Seetharaman, G. et Zavidovique, B. (2007). Median Binary Pattern for Textures Classification. Dans M. Kamel & A. Campilho (dir.), *Image Analysis and Recognition: 4th International Conference, ICIAR 2007, Montreal, Canada, August 22-24, 2007. Proceedings* (p. 387-398). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Hafner, J., Sawhney, H. S., Equitz, W., Flickner, M. et Niblack, W. (1995). Efficient color histogram indexing for quadratic form distance functions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(7), 729-736. doi: 10.1109/34.391417
- Han, J. et Ma, K.-K. (2007). Rotation-invariant and scale-invariant Gabor features for texture image retrieval. *Image and Vision Computing*, 25, 1474-1481.
- Hannan, M. A., Arebey, M., Begum, R. A., Basri, H. et Al Mamun, M. A. (2016). Content-based image retrieval system for solid waste bin level detection and performance evaluation. *Waste Manag*, 50, 10-19. doi: 10.1016/j.wasman.2016.01.046
- Haralick, R. M. (1971, 15-17 Dec. 1971). *On a texture-context feature extraction algorithm for remotely sensed imagery*. Communication présentée Decision and Control, 1971 IEEE Conference on. doi: 10.1109/cdc.1971.271084
- Haralick, R. M. (1979). Statistical and Structural Approaches to Texture. *proceeding of the IEEE*, 67(5), 786-804.
- Haralick, R. M., Shanmugam, K. et Dinstein, I. h. (1973). Textures features for image classification. *IEEE transactions on systems, man and cybernetics*, SMC-3(6), 610-621.
- Harris, C. et Stephens, M. (1988). *A combined corner and edge detector*. Communication présentée Alvey vision conference.
- Hassaballah, M., Abdelmgeid, A. A. et Alshazly, H. A. (2016). Image Features Detection, Description and Matching. Dans I. A. Awad & M. Hassaballah (dir.), *Image Feature Detectors and Descriptors : Foundations and Applications* (p. 11-45). Cham: Springer International Publishing.

- Hassner, M. et Sklansky, J. (1980). The use of Markov Random Fields as models of texture. *Computer Graphics and Image Processing*, 12(4), 357-370. doi: 10.1016/0146-664x(80)90019-2
- Hee-Jung, B. et Sung-Hwan, J. (1997, 9-12 Sep 1997). *Image retrieval using texture based on DCT*. Communication présentée Information, Communications and Signal Processing, 1997. ICICS., Proceedings of 1997 International Conference on. doi: 10.1109/icics.1997.652144
- Hossain, S. et Serikawa, S. (2013). Texture databases – A comprehensive survey. *Pattern Recognition Letters*, 34(15), 2007-2022. doi: 10.1016/j.patrec.2013.02.009
- Huang, D., Shan, C., Ardabilian, M., Wang, Y. et Chen, L. (2011). Local Binary Patterns and Its Application to Facial Image Analysis: A Survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 41(6), 765-781. doi: 10.1109/tsmcc.2011.2118750
- Huang, J., You, X., Yuan, Y., Yang, F. et Lin, L. (2010). Rotation invariant iris feature extraction using Gaussian Markov random fields with non-separable wavelet. *Neurocomputing*, 73(4-6), 883-894. doi: 10.1016/j.neucom.2009.09.016
- Huang, P. W. et Dai, S. K. (2003). Image retrieval by texture similarity. *Pattern Recognition*, 36(3), 665-679. doi: 10.1016/s0031-3203(02)00083-3
- Idrissa, M. et Acheroy, M. (2002). Texture classification using Gabor filters. *Pattern Recognition Letters*, 23(9), 1095-1102. doi: 10.1016/s0167-8655(02)00056-9
- Iivarinen, J. et Visa, A. J. E. (1996). *Shape recognition of irregular objects*. Repéré à <http://dx.doi.org/10.1117/12.256280>
- Islam, M. M., Dengsheng, Z. et Guojun, L. (2008, June 23 2008-April 26 2008). *A geometric method to compute directionality features for texture images*. Communication présentée Multimedia and Expo, 2008 IEEE International Conference on. doi: 10.1109/icme.2008.4607736
- Jacob, I. J., Srinivasagan, K. G. et Jayapriya, K. (2014). Local Oppugnant Color Texture Pattern for image retrieval system. *Pattern Recognition Letters*, 42, 72-78.
- Jain, A. K. et Farrokhnia, F. (1990, 4-7 Nov 1990). *Unsupervised texture segmentation using Gabor filters*. Communication présentée Systems, Man and Cybernetics, 1990. Conference Proceedings., IEEE International Conference on. doi: 10.1109/icsmc.1990.142050
- Jain, A. K. et Vailaya, A. (1998). Shape-Based Retrieval: A Case Study with Trademark Image Databases. *Pattern Recognition*, 31(9), 1369-1390. doi: 10.1016/s0031-3203(97)00131-3
- Jain, A. K. et Zhong, Y. (1996). Page segmentation using texture analysis. *Pattern Recognition*, 29(5), 743-770. doi: 10.1016/0031-3203(95)00131-x
- Jasmine, K. P. et Kumar, P. R. (2014). Integration of HSV Color Histogram and LMEBP Joint Histogram for Multimedia Image Retrieval. Dans *Intelligent Computing, Networking, and Informatics* (Vol. 243, p. 753-762). doi: 10.1007/978-81-322-1665-0_76.
- Jhanwar, N., Chaudhuri, S., Seetharaman, G. et Zavidovique, B. (2004). Content based image retrieval using motif cooccurrence matrix. *Image and Vision Computing*, 22(14), 1211-1220. doi: 10.1016/j.imavis.2004.03.026
- Jiang, J., Weng, Y., Guo, B. et Feng, Y. (2006). Robust-to-rotation texture descriptor for image retrieval in wavelets domain. *Journal of Electronic Imaging*, 15(1), 013013. doi: 10.1117/1.2179421
- Jiayin, K. et Wenjuan, Z. (2012, 23-25 May 2012). *A framework for image retrieval with hybrid features*. Communication présentée Control and Decision Conference (CCDC), 2012 24th Chinese. doi: 10.1109/ccdc.2012.6244213

- Jin, H., Liu, Q., Lu, H. et Tong, X. (2004). *Face Detection Using Improved LBP Under Bayesian Framework*. Communication présentée Proceedings of the Third International Conference on Image and Graphics (ICIG'04).
- Jing, H., Jianhui, Z., Weiwei, G., Chengjiang, L., Lu, X., Zhiyong, Y. et Shizhong, H. (2010, 16-18 Oct. 2010). *Local binary pattern based texture analysis for visual fire recognition*. Communication présentée Image and Signal Processing (CISP), 2010 3rd International Congress on. doi: 10.1109/cisp.2010.5647609
- Jing, H., Kumar, S. R., Mitra, M., Wei-Jing, Z. et Zabih, R. (1997, 17-19 Jun 1997). *Image indexing using color correlograms*. Communication présentée Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on. doi: 10.1109/cvpr.1997.609412
- Julesz, B. (1962). Visual Pattern Discrimination. *IEEE Transactions on Information Theory*, 8(2), 84-92. doi: 10.1109/tit.1962.1057698
- Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature*, 290(5802), 91-97.
- Ka-Man, W., Lai-Man, P. et Kwok-Wai, C. (2007, Sept. 16 2007-Oct. 19 2007). *Dominant Color Structure Descriptor for Image Retrieval*. Communication présentée Image Processing, 2007. ICIP 2007. IEEE International Conference on. doi: 10.1109/icip.2007.4379597
- Kasutani, E. et Yamada, A. (2001, 2001). *The MPEG-7 color layout descriptor: a compact image feature description for high-speed image/video segment retrieval*. Communication présentée Image Processing, 2001. Proceedings. 2001 International Conference on. doi: 10.1109/icip.2001.959135
- Kaya, Y., Ertuğrul, Ö. F. et Tekin, R. (2015). Two novel local binary pattern descriptors for texture analysis. *Applied Soft Computing*, 34(0), 728-735. doi: 10.1016/j.asoc.2015.06.009
- Keller, J. M., Chen, S. et Crownover, R. M. (1989). Texture description and segmentation through fractal geometry. *Computer Vision, Graphics, and Image Processing*, 45(2), 150-166. doi: 10.1016/0734-189x(89)90130-8
- Kerr, D. A. (2010). Digital Camera Sensor Colorimetry.
- Kim, S., Park, S. et Kim, M. (2003). Central Object Extraction for Object-Based Image Retrieval. Dans E. M. Bakker, M. S. Lew, T. S. Huang, N. Sebe & X. S. Zhou (dir.), *Image and Video Retrieval: Second International Conference, CIVR 2003 Urbana-Champaign, IL, USA, July 24-25, 2003 Proceedings* (p. 39-49). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Kpalma, K. (1994). *Caractérisation de textures par l'anisotropie de la dimension fractale*. Communication présentée Proceedings of the 2nd African Conference on Research in Computer Science (CARI).
- Kuehni, R. G. (2003). *Color space and its divisions : color order from antiquity to the present*. John Wiley & Sons
- Laine, A. et Fan, J. (1993). Texture classification by wavelet packet signatures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11), 1186-1191. doi: 10.1109/34.244679
- Lan, R., Zhou, Y. et Tang, Y. Y. (2016). Quaternionic Local Ranking Binary Pattern: A Local Descriptor of Color Images. *IEEE Trans Image Process*, 25(2), 566-579. doi: 10.1109/TIP.2015.2507404
- Lande, M. V., Bhanodiya, P. et Jain, P. (2014). An Effective Content-Based Image Retrieval Using Color, Texture and Shape Feature. *Intelligent Computing, Networking, and Informatics*, 243, 1163-1170. doi: 10.1007/978-81-322-1665-0_119

- Li, C., Zhou, W. et Yuan, S. (2014). Iris recognition based on a novel variation of local binary pattern. *The Visual Computer*, 31(10), 1419-1429. doi: 10.1007/s00371-014-1023-5
- Li, H., Huang, D., Morvan, J.-M., Wang, Y. et Chen, L. (2014). Towards 3D Face Recognition in the Real: A Registration-Free Approach Using Fine-Grained Matching of 3D Keypoint Descriptors. *International Journal of Computer Vision*, 113(2), 128-142. doi: 10.1007/s11263-014-0785-6
- Li, M., Tieniu, T., Yunhong, W. et Dexin, Z. (2003). Personal identification based on iris texture analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(12), 1519-1533. doi: 10.1109/tpami.2003.1251145
- Li, Y. et Guan, L. (2006, May 2006). *An Effective Shape Descriptor for the Retrieval of Natural Image Collections*. Communication présentée Canadian Conference on Electrical and Computer Engineering, 2006. CCECE '06. , Ottawa, Ont. doi: 10.1109/CCECE.2006.277550
- Li, Z., Liu, Y. et Li, D. (2013, 9-11 June 2013). *A new texture feature extraction method for image retrieval*. Communication présentée Intelligent Control and Information Processing (ICICIP), 2013 Fourth International Conference on. doi: 10.1109/icicip.2013.6568122
- Liao, S., Law, M. W. et Chung, A. C. (2009). Dominant local binary patterns for texture classification. *IEEE Trans Image Process*, 18(5), 1107-1118. doi: 10.1109/TIP.2009.2015682
- Lindeberg, T. (1993). Detecting salient blob-like image structures and their scales with a scale-space primal sketch: A method for focus-of-attention. *International Journal of Computer Vision*, 11(3), 283-318. doi: 10.1007/bf01469346
- Lindeberg, T. (1998). Feature Detection with Automatic Scale Selection. *International Journal of Computer Vision*, 30(2), 79-116. doi: 10.1023/a:1008045108935
- Liu, L., Fieguth, P., Zhao, G., Pietikäinen, M. et Hu, D. (2016). Extended local binary patterns for face recognition. *Information Sciences*, 358-359, 56-72. doi: 10.1016/j.ins.2016.04.021
- Liu, Y., Li, Z. et Gao, Z.-M. (2013). An Improved Texture Feature Extraction Method for Tyre Tread Patterns. Dans C. Sun, F. Fang, Z.-H. Zhou, W. Yang & Z.-Y. Liu (dir.), *Intelligence Science and Big Data Engineering* (Vol. 8261, p. 705-713): Springer Berlin Heidelberg.
- Liu, Y., Zhang, D., Lu, G. et Ma, W.-Y. (2007). A survey of content-based image retrieval with high-level semantics. *Pattern Recognition*, 40(1), 262-282. doi: 10.1016/j.patcog.2006.04.045
- Livens, S., Scheunders, P., Van de Wouwer, G. et Van Dyck, D. (1997, 14-17 Jul 1997). *Wavelets for texture analysis, an overview*. Communication présentée Image Processing and Its Applications, 1997., Sixth International Conference on. doi: 10.1049/cp:19970958
- Local Binary Patterns: New Variants and Applications* (2014). doi: 10.1007/978-3-642-39289-4
- Loncaric, S. (1998). A survey of shape analysis techniques. *Pattern Recognition*, 31(8), 983-1001. doi: 10.1016/s0031-2023(97)00122-2
- Lonnestad, T. (1992, 30 Aug-3 Sep 1992). *A new set of texture features based on the Haar transform*. Communication présentée Pattern Recognition, 1992. Vol.III. Conference C: Image, Speech and Signal Analysis, Proceedings., 11th IAPR International Conference on. doi: 10.1109/icpr.1992.202077
- Lowe, D. (2004a). Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2), 91-110. doi: 10.1023/b:visi.0000029664.99615.94

- Lowe, D. G. (1999, 1999). *Object recognition from local scale-invariant features*. Communication présentée Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on. doi: 10.1109/iccv.1999.790410
- Lowe, D. G. (2004b). Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2), 91-110. doi: 10.1023/b:visi.0000029664.99615.94
- Lu, Z. et Linghua, Z. (2016). Face recognition algorithm based on discriminative dictionary learning and sparse representation. *Neurocomputing*, 174, 749-755. doi: 10.1016/j.neucom.2015.09.091
- Lu, Z., Wang, L. et Wen, J.-R. (2016). Image classification by visual bag-of-words refinement and reduction. *Neurocomputing*, 173, 373-384. doi: 10.1016/j.neucom.2015.01.098
- ma, J.-q. (2009). *Content-Based Image Retrieval with HSV Color Space and Texture Features*. Communication présentée 2009 International Conference on Web Information Systems and Mining. doi: 10.1109/WISM.2009.20
- Ma, W.-Y. et Manjunath, B. S. (1999). Netra: A toolbox for navigating large image databases. *Multimedia systems*, 7(3), 184-198.
- Ma, W. Y. et Manjunath, B. S. (1996, 18-20 Jun 1996). *Texture features and learning similarity*. Communication présentée Computer Vision and Pattern Recognition, 1996. Proceedings CVPR '96, 1996 IEEE Computer Society Conference on. doi: 10.1109/cvpr.1996.517107
- Maani, R., Kalra, S. et Yang, Y.-H. (2015). A REVIEW OF TEXTURE CLASSIFICATION METHODS AND THEIR APPLICATIONS IN MEDICAL IMAGE ANALYSIS OF THE BRAIN (*Handbook of Pattern Recognition and Computer Vision* (5th^e éd., p. 351-369): WORLD SCIENTIFIC.
- Madeira, S. C. et Oliveira, A. L. (2004). Biclustering algorithms for biological data analysis: a survey. *IEEE/ACM Trans Comput Biol Bioinform*, 1(1), 24-45. doi: 10.1109/TCBB.2004.2
- Mäenpää, T. (2003). *The local binary pattern approach to texture analysis: extensions and applications*. (Oulun yliopisto).
- Mäenpää, T. et Pietikäinen, M. (2004). Classification with color and texture: jointly or separately? *Pattern Recognition*, 37(8), 1629-1640. doi: 10.1016/j.patcog.2003.11.011
- Majtner, T. et Svoboda, D. (2012, 13-15 Oct. 2012). *Extension of Tamura Texture Features for 3D Fluorescence Microscopy*. Communication présentée 3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), 2012 Second International Conference on. doi: 10.1109/3dimpvt.2012.61
- Malathi Latha, Y. L. et Prasad, M. V. N. K. (2015). GLCM Based Texture Features for Palmprint Identification System. Dans C. L. Jain, S. H. Behera, K. J. Mandal & P. D. Mohapatra (dir.), *Computational Intelligence in Data Mining - Volume 1: Proceedings of the International Conference on CIDM, 20-21 December 2014* (p. 155-163). New Delhi: Springer India.
- Malik, F. et Baharudin, B. (2013a). Analysis of distance metrics in content-based image retrieval using statistical quantized histogram texture features in the DCT domain. *Journal of King Saud University - Computer and Information Sciences*, 25(2), 207-218. doi: 10.1016/j.jksuci.2012.11.004
- Malik, F. et Baharudin, B. (2013b). The Statistical Quantized Histogram Texture Features Analysis for Image Retrieval Based on Median and Laplacian Filters in the DCT Domain. *The International Arab Journal of Information Technology*, 10(6).
- Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7), 674-693. doi: 10.1109/34.192463
- Mandelbrot, B. B. (1983). *Fractals and the Geometry of Nature*.

- Manjunath, B. S. et Chellappa, R. (1993). A unified approach to boundary perception: edges, textures, and illusory contours. *IEEE Trans Neural Netw*, 4(1), 96-108. doi: 10.1109/72.182699
- Manjunath, B. S. et Ma, W.-Y. (2002). Texture Features for Image Retrieval (*Image Databases* (p. 313-344): John Wiley & Sons, Inc.
- Manjunath, B. S. et Ma, W. Y. (1996). Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(8), 837-842. doi: 10.1109/34.531803
- Manjunath, B. S., Ohm, J. R., Vasudevan, V. V. et Yamada, A. (2001). Color and texture descriptors. *IEEE Transactions on Circuits and Systems for Video Technology*, 11(6), 703-715. doi: 10.1109/76.927424
- Manthalkar, R., Biswas, P. K. et Chatterji, B. N. (2003). Rotation and scale invariant texture features using discrete wavelet packet transform. *Pattern Recognition Letters*, 24(14), 2455-2462. doi: 10.1016/s0167-8655(03)00090-4
- Masoudi, B. (2016). Classification of color texture images based on modified WLD. *International Journal of Multimedia Information Retrieval*, 5(2), 117-124. doi: 10.1007/s13735-016-0097-4
- Materka, A. et Strzelecki, M. (1998). Texture analysis methods—a review (p. 9-11). Technical university of lodz, institute of electronics, COST B11 report, Brussels.
- McRoberts, K. C., Benson, B. M., Mudrak, E. L., Parsons, D. et Cherney, D. J. R. (2016). Application of local binary patterns in digital images to estimate botanical composition in mixed alfalfa–grass fields. *Computers and Electronics in Agriculture*, 123, 95-103. doi: 10.1016/j.compag.2016.02.015
- Mehta, R. et Egiazarian, K. (2016). Dominant Rotated Local Binary Patterns (DRLBP) for texture classification. *Pattern Recognition Letters*, 71, 16-22. doi: 10.1016/j.patrec.2015.11.019
- MEHTRE, B. M., KANKANHALLI, M. S. et LEE, W. F. (1997). Shape Measures For Content Based Image Retrieval: A Comparison. *Information Processing & Management*, 33(3), 319-337.
- Messing, D. S., van Beek, P. et Errico, J. H. (2001, 2001). *The MPEG-7 colour structure descriptor: image description using colour and local spatial information*. Communication présentée Image Processing, 2001. Proceedings. 2001 International Conference on. doi: 10.1109/icip.2001.959134
- Mikolajczyk, K. et Schmid, C. (2004). Scale & Affine Invariant Interest Point Detectors. *International Journal of Computer Vision*, 60(1), 63-86. doi: 10.1023/B:VISI.0000027790.02288.f2
- Mikolajczyk, K. et Schmid, C. (2005). Performance evaluation of local descriptors. *IEEE Trans Pattern Anal Mach Intell*, 27(10), 1615-1630. doi: 10.1109/TPAMI.2005.188
- Mikolajczyk, K. et Tuytelaars, T. (2007). Local Invariant Feature Detectors: A Survey. *Foundations and Trends® in Computer Graphics and Vision*, 3(3), 177-280. doi: 10.1561/06000000017
- Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., . . . Gool, L. V. (2005). A Comparison of Affine Region Detectors. *International Journal of Computer Vision*, 65(1-2), 43-72. doi: 10.1007/s11263-005-3848-x
- Ming-Kuei, H. (1962). Visual pattern recognition by moment invariants. *IEEE Transactions on Information Theory*, 8(2), 179-187. doi: 10.1109/tit.1962.1057692
- Mohd Anuar, F., Setchi, R. et Lai, Y.-k. (2013). Trademark image retrieval using an integrated shape descriptor. *Expert Systems with Applications*, 40(1), 105-121. doi: 10.1016/j.eswa.2012.07.031

- Mosbah, M. et Boucheham, B. (2016). Distance selection based on relevance feedback in the context of CBIR using the SFS meta-heuristic with one round. *Egyptian Informatics Journal*. doi: <http://dx.doi.org/10.1016/j.eij.2016.09.001>
- MPEG_Requirements_Group. (1999). MPEG-7: Context, Objectives and Technical Roadmap. Vancouver.
- Mukherjee, D., Jonathan Wu, Q. M. et Wang, G. (2015). A comparative experimental study of image feature detectors and descriptors. *Machine Vision and Applications*, 26(4), 443-466. doi: 10.1007/s00138-015-0679-9
- Murala, S., Gonde, A. B. et Maheshwari, R. P. (2009, 6-7 March). *Color and Texture Features for Image Indexing and Retrieval*. Communication présentée 2009 IEEE International Advance Computing Conference, Patiala, India.
- Murala, S., Maheshwari, R. P. et Balasubramanian, R. (2012a). Directional local extrema patterns: a new descriptor for content based image retrieval. *International Journal of Multimedia Information Retrieval*, 1(3), 191-203. doi: 10.1007/s13735-012-0008-2
- Murala, S., Maheshwari, R. P. et Balasubramanian, R. (2012b). Local tetra patterns: a new feature descriptor for content-based image retrieval. *IEEE Trans Image Process*, 21(5), 2874-2886. doi: 10.1109/TIP.2012.2188809
- Murala, S. et Wu, Q. M. J. (2015). Spherical symmetric 3D local ternary patterns for natural, texture and biomedical image indexing and retrieval. *Neurocomputing*, 149, 1502-1514.
- Nanni, L., Costa, Y. M. G., Lumini, A., Kim, M. Y. et Baek, S. R. (2016). Combining visual and acoustic features for music genre classification. *Expert Systems with Applications*, 45, 108-117. doi: 10.1016/j.eswa.2015.09.018
- Nanni, L., Lumini, A. et Brahnam, S. (2010). Local binary patterns variants as texture descriptors for medical image analysis. *Artif Intell Med*, 49(2), 117-125. doi: 10.1016/j.artmed.2010.02.006
- Naresh, Y. G. et Nagendraswamy, H. S. (2016). Classification of medicinal plants: An approach using modified LBP with symbolic representation. *Neurocomputing*, 173, 1789-1797. doi: 10.1016/j.neucom.2015.08.090
- Ngan, H. Y. T., Pang, G. K. H., Yung, S. P. et Ng, M. K. (2005). Wavelet based methods on patterned fabric defect detection. *Pattern Recognition*, 38(4), 559-576. doi: 10.1016/j.patcog.2004.07.009
- Niblack, C. W., Barber, R., Equitz, W., Flickner, M. D., Glasman, E. H., Petkovic, D., . . . Taubin, G. (1993). *QBIC project: querying images by content, using color, texture, and shape*. Repéré à <http://dx.doi.org/10.1117/12.143648>
- Ohanian, P. P. et Dubes, R. C. (1992). Performance evaluation for four classes of textural features. *Pattern Recognition*, 25(8), 819-833. doi: 10.1016/0031-3203(92)90036-i
- Ojala, T., Pietikäinen, M. et Harwood, D. (1996). A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition*, 29(1), 51-59.
- Ojala, T., Pietikäinen, M. et Mäenpää, T. (2002). Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7), 971-987.
- Oliver, A., Lladó, X., Freixenet, J. et Martí, J. (2007). False Positive Reduction in Mammographic Mass Detection Using Local Binary Patterns. Dans N. Ayache, S. Ourselin & A. Maeder (dir.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2007: 10th International Conference, Brisbane, Australia, October 29 - November 2, 2007, Proceedings, Part I* (p. 286-293). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Paget, R. (2012). Texture Modelling and Synthesis (*Handbook of Texture Analysis* (p. 33-60): PUBLISHED BY IMPERIAL COLLEGE PRESS AND DISTRIBUTED BY WORLD SCIENTIFIC PUBLISHING CO.

- Paget, R. et Longstaff, D. (1995). Texture synthesis via a non-parametric Markov random Field. *Proceedings of DICTA-95, Digital Image Computing: Techniques and Applications, 1*, 547-552.
- Palm, C. (2004). Color texture classification by integrative Co-occurrence matrices. *Pattern Recognition*, 37(5), 965-976. doi: 10.1016/j.patcog.2003.09.010
- Park, J., Chisty, K. M. M., Lee, J., An, Y. et Choi, Y. (2011, 12-14 Dec. 2011). *Image Retrieval Technique Using Rearranged Freeman Chain Code*. Communication présentée Informatics and Computational Intelligence (ICI), 2011 First International Conference on. doi: 10.1109/ici.2011.54
- Partio, M., Cramariuc, B., Gabbouj, M. et Visa, A. (2002). *Rock texture retrieval using gray level co-occurrence matrix*. Communication présentée Proc. of 5th Nordic Signal Processing Symposium.
- Paschos, G. (2001). Perceptually uniform color spaces for color texture analysis: an empirical evaluation. *IEEE Transactions on Image Processing*, 10(6), 932-937. doi: 10.1109/83.923289
- Pass, G. et Zabih, R. (1996, 2-4 December). *Histogram Refinement for Content-Based Image Retrieval*. Communication présentée Applications of Computer Vision, 1996. WACV '96., Proceedings 3rd IEEE Workshop on Sarasota, FL. doi: 10.1109/ACV.1996.572008
- Patel, B., Maheshwari, R. P. et Balasubramanian, R. (2016). Multi-quantized local binary patterns for facial gender classification. *Computers & Electrical Engineering*, 54, 271-284. doi: 10.1016/j.compeleceng.2015.11.004
- Patil, S. et Talbar, S. (2012). Content Based Image Retrieval Using Various Distance Metrics. Dans R. Kannan & F. Andres (dir.), *Data Engineering and Management* (Vol. 6411, p. 154-161): Springer Berlin Heidelberg.
- Pavlidis, T. (1978). A review of algorithms for shape analysis. *Computer Graphics and Image Processing*, 7(2), 243-258. doi: 10.1016/0146-664x(78)90115-6
- Payne, J. S., Stonham, T. J. et Patel, D. (1994, 9-13 Oct 1994). *Document segmentation using texture analysis*. Communication présentée Pattern Recognition, 1994. Vol. 2 - Conference B: Computer Vision & Image Processing., Proceedings of the 12th IAPR International. Conference on. doi: 10.1109/ICPR.1994.576947
- Penatti, O. A. B., Valle, E. et Torres, R. d. S. (2012). Comparative study of global color and texture descriptors for web image retrieval. *Journal of Visual Communication and Image Representation*, 23(2), 359-380. doi: 10.1016/j.jvcir.2011.11.002
- Pentland, A., Picard, R. W. et Sclaroff, S. (1996). Photobook: Content-based manipulation of image databases. *International journal of computer vision*, 18(3), 233-254.
- Petrou, M. et García Sevilla, P. (2006a). Introduction (*Image Processing* (p. 1-10): John Wiley & Sons, Ltd.
- Petrou, M. et García Sevilla, P. (2006b). Non-Stationary Grey Texture Images (*Image Processing* (p. 297-606): John Wiley & Sons, Ltd.
- Petrou, M. et García Sevilla, P. (2006c). Stationary Grey Texture Images (*Image Processing* (p. 81-295): John Wiley & Sons, Ltd.
- Pichler, O., Teuner, A. et Hosticka, B. J. (1996). A comparison of texture feature extraction using adaptive gabor filtering, pyramidal and tree structured wavelet transforms. *Pattern Recognition*, 29(5), 733-742. doi: 10.1016/0031-3203(95)00127-1
- Pietikäinen, M., Hadid, A., Zhao, G. et Ahonen, T. (2011a). *Computer Vision Using Local Binary Patterns*. doi: 10.1007/978-0-85729-748-8
- Pietikäinen, M., Hadid, A., Zhao, G. et Ahonen, T. (2011b). Local Binary Patterns for Still Images (*Computer Vision Using Local Binary Patterns* (p. 13-47). London: Springer London.

- Porter, R. et Canagarajah, N. (1997). Robust rotation-invariant texture classification: wavelet, Gabor filter and GMRF based schemes. *IEE Proceedings - Vision, Image, and Signal Processing*, 144(3), 180. doi: 10.1049/ip-vis:19971182
- Rai, P. et Khanna, P. (2014). A gender classification system robust to occlusion using Gabor features based (2D)2PCA. *Journal of Visual Communication and Image Representation*, 25(5), 1118-1129. doi: 10.1016/j.jvcir.2014.03.009
- Reddy, P. V. N. et Prasad, K. S. (2011). Content Based Image Retrieval Using Local Derivative Patterns. *Journal of Theoretical and Applied Information Technology*, 28(2), 95-103.
- Ribaric, S. et Lopar, M. (2012, 25-28 March 2012). *Palmprint recognition based on local Haralick features*. Communication présentée Electrotechnical Conference (MELECON), 2012 16th IEEE Mediterranean. doi: 10.1109/melcon.2012.6196517
- Rodhetbhai, W. (2009). *Preprocessing for Content-Based Image Retrieval*. (University of Southampton).
- Rosten, E. et Drummond, T. (2006). Machine Learning for High-Speed Corner Detection. Dans A. Leonardis, H. Bischof & A. Pinz (dir.), *Computer Vision – ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I* (p. 430-443). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Rubner, Y. (2000). The Earth Mover's Distance as a Metric for Image Retrieval. *International Journal of Computer Vision*, 40(2), 99-121. doi: 10.1023/a:1026543900054
- Rubner, Y., Puzicha, J., Tomasi, C. et Buhmann, J. M. (2001). Empirical Evaluation of Dissimilarity Measures for Color and Texture. *Computer Vision and Image Understanding*, 84(1), 25-43. doi: 10.1006/cviu.2001.0934
- Salmi, M. et Boucheham, B. (2014, 9-10 Nov. 2014). *Content based image retrieval based on Cell Color Coherence Vector (Cell-CCV)*. Communication présentée ISKO-Maghreb: Concepts and Tools for knowledge Management (ISKO-Maghreb), 2014 4th International Symposium. doi: 10.1109/ISKO-Maghreb.2014.7033451
- Sandid, F. et Douik, A. (2016). Robust color texture descriptor for material recognition. *Pattern Recognition Letters*, 80, 15-23. doi: 10.1016/j.patrec.2016.05.010
- Sastry, C. S., Ravindranath, M., Pujari, A. K. et Deekshatulu, B. L. (2007). A modified Gabor function for content based image retrieval. *Pattern Recognition Letters*, 28, 293-300.
- Sethi, I. K., Coman, I. L. et Stan, D. (2001). *Mining association rules between low-level image features and high-level concepts*. Repéré à <http://dx.doi.org/10.1117/12.421083>
- Shrivastava, N. et Tyagi, V. (2014). Content based image retrieval based on relative locations of multiple regions of interest using selective regions matching. *Information Sciences*, 259(0), 212-224. doi: <http://dx.doi.org/10.1016/j.ins.2013.08.043>
- Sidorova, V. S. (2015). Global segmentation of textural images on the basis of hierarchical clusterization by predetermined cluster separability. *Pattern Recognition and Image Analysis*, 25(3), 541-546. doi: 10.1134/s1054661815030232
- Smeulders, W. M., Worring, M., Santini, S., Gupta, A. et Jain, R. (2000). Content-Based Image Retrieval at the End of the Early Years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12), 1-32.
- Smith, J. R. (2002). Color for Image Retrieval (*Image Databases* (p. 285-311): John Wiley & Sons, Inc.
- Smolka, B. et Nurzynska, K. (2015). Power LBP: A Novel Texture Operator for Smiling and Neutral Facial Display Classification. *Procedia Computer Science*, 51(0), 1555-1564. doi: 10.1016/j.procs.2015.05.350
- Sokic, E. et Konjicija, S. (2016). Phase preserving Fourier descriptor for shape-based image retrieval. *Signal Processing: Image Communication*, 40, 82-96. doi: 10.1016/j.image.2015.11.002

- Solberg, A. H. S., Taxt, T. et Jain, A. K. (1996). A Markov random field model for classification of multisource satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 34(1), 100-113. doi: 10.1109/36.481897
- Song, T. et Li, H. (2013). Local Polar DCT Features for Image Description. *IEEE Signal Processing Letters*, 20(1), 59-62. doi: 10.1109/lsp.2012.2229273
- Stehling, R. O., Nascimento, M. A. et Falcão, A. X. (2003). Cell Histograms Versus Color Histograms for Image Representation and Retrieval. *Knowledge and Information Systems*, 5(3), 315-336. doi: 10.1007/s10115-003-0084-y
- Stéphane, M. (2009). CHAPTER 7 - Wavelet Bases (A Wavelet Tour of Signal Processing (Third Edition) (p. 263-376). Boston: Academic Press.
- Stockman, G. et Shapiro, L. G. (2001). *Computer Vision*. Prentice Hall PTR.
- Stricker, M. et Orengo, M. (1995). Similarity of color images. *Proceedings of the SPIE Storage and Retrieval for Image and Video Databases*.
- Stricker, M. A. et Dimai, A. (1996). *Color indexing with weak spatial constraints*. Repéré à <http://dx.doi.org/10.1117/12.234802>
- Subrahmanyam, M., Jonathan Wu, Q. M., Maheshwari, R. P. et Balasubramanian, R. (2013). Modified color motif co-occurrence matrix for image indexing and retrieval. *Computers & Electrical Engineering*, 39(3), 762-774. doi: 10.1016/j.compeleceng.2012.11.023
- Sural, S., Qian, G. et Pramanik, S. (2002). *SEGMENTATION AND HISTOGRAM GENERATION USING THE HSV COLOR SPACE FOR IMAGE RETRIEVAL*. Communication présentée Image Processing. 2002. Proceedings. 2002 International Conference on. doi: 10.1109/ICIP.2002.1040019
- Swain, M. J. et Ballard, D. H. (1991). Color indexing. *International Journal of Computer Vision*, 7(1), 11-32.
- Takala, V., Ahonen, T. et Pietikäinen, M. (2005). Block-Based Methods for Image Retrieval Using Local Binary Patterns. Dans H. Kalviainen, J. Parkkinen & A. Kaarna (dir.), *Image Analysis* (Vol. 3540, p. 882-891): Springer Berlin Heidelberg.
- Tamura, H., Mori, S. et Yamawaki, T. (1978). Textural Features Corresponding to Visual Perception. *IEEE Transactions on Systems, Man, and Cybernetics*, 8(6), 460-473. doi: 10.1109/tsmc.1978.4309999
- Tan, X. et Triggs, B. (2007). Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions. *Analysis and Modeling of Faces and Gestures*, 4778, 168-182. doi: 10.1007/978-3-540-75690-3_13
- Tang, Z., Su, Y., Er, M. J., Qi, F., Zhang, L. et Zhou, J. (2015). A local binary pattern based texture descriptors for classification of tea leaves. *Neurocomputing*, 168(0), 1011-1023. doi: 10.1016/j.neucom.2015.05.024
- Teague, M. R. (1980). Image analysis via the general theory of moments*. *Journal of the Optical Society of America*, 70(8), 920. doi: 10.1364/josa.70.000920
- Tong, L., Wong, W. K. et Kwong, C. K. (2016). Differential evolution-based optimal Gabor filter model for fabric inspection. *Neurocomputing*, 173, 1386-1401. doi: 10.1016/j.neucom.2015.09.011
- Torres-Huitzil, C. (2016). A Review of Image Interest Point Detectors: From Algorithms to FPGA Hardware Implementations. Dans I. A. Awad & M. Hassaballah (dir.), *Image Feature Detectors and Descriptors : Foundations and Applications* (p. 47-74). Cham: Springer International Publishing.
- Tsai, T., Huang, Y. P. et Chiang, T. W. (2006, 8-11 Oct. 2006). *Dominant Feature Extraction in Block-DCT Domain*. Communication présentée 2006 IEEE International Conference on Systems, Man and Cybernetics. doi: 10.1109/icsmc.2006.384692
- Tuceryan, M. et Jain, A. K. (2011). TEXTURE ANALYSIS (*Handbook of Pattern Recognition and Computer Vision* (2nd^e éd., p. 207-248): WORLD SCIENTIFIC.

- Unser, M. (1986). Sum and difference histograms for texture classification. *IEEE Trans Pattern Anal Mach Intell*, 8(1), 118-125. doi: 10.1109/TPAMI.1986.4767760
- Ursani, A., Kpalma, K., Ronsin, J., Memon, A. et Chowdhry, B. (2009). Improved Texture Description with Features Based on Fourier Transform. Dans D. M. A. Hussain, A. Rajput, B. Chowdhry & Q. Gee (dir.), *Wireless Networks, Information Processing and Systems* (Vol. 20, p. 19-28): Springer Berlin Heidelberg.
- Veltkamp, R. C. et Tanase, M. (2002). A Survey of Content-Based Image Retrieval Systems. Dans *Content-Based Image and Video Retrieval* (Vol. 21, p. 47-101). doi: 10.1007/978-1-4615-0987-5_5.
- Verma, M. et Raman, B. (2016). Local tri-directional patterns: A new texture feature descriptor for image retrieval. *Digital Signal Processing*, 51, 62-72. doi: 10.1016/j.dsp.2016.02.002
- Vil'kin, A. M., Safonov, I. V. et Egorova, M. A. (2013). Algorithm for segmentation of documents based on texture features. *Pattern Recognition and Image Analysis*, 23(1), 153-159. doi: 10.1134/s1054661813010136
- Voorhees, H. et Poggio, T. (1988). Computing texture boundaries from images. *Nature*, 333(6171), 364-367. doi: 10.1038/333364a0
- Walker, R. F., Jackway, P. et Longstaff, I. (1995). *Improving co-occurrence matrix feature discrimination*. Communication présentée DICTA'95, 3rd Conference on Digital Image Computing: Techniques and Application.
- Wang, J. Z., Li, J. et Wiederhold, G. (2001). SIMPLiCity: Semantics-Sensitive Integrated Matching for Picture LIBraries. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(9), 947-963.
- Wang, L. et He, D.-C. (1990). Texture classification using texture spectrum. *Pattern Recognition*, 23(8), 905-910. doi: 10.1016/0031-3203(90)90135-8
- Wang, L. et Liu, J. (1999). Texture classification using multiresolution Markov random field models. *Pattern Recognition Letters*, 20(2), 171-182. doi: 10.1016/s0167-8655(98)00129-9
- Wang, X.-Y., Yu, Y.-J. et Yang, H.-Y. (2011). An effective image retrieval scheme using color, texture and shape features. *Computer Standards & Interfaces*, 33(1), 59-68. doi: 10.1016/j.csi.2010.03.004
- Wang, X.-Y., Zhang, B.-B. et Yang, H.-Y. (2012). Content-based image retrieval by integrating color and texture features. *Multimedia Tools and Applications*, 68(3), 545-569. doi: 10.1007/s11042-012-1055-7
- Wang, Y., Li, H. et Wang, L. (2014). Similar Image Retrieval Using Color Histogram in HSV Space and SIFT Descriptor with FLANN. Dans *Foundations of Intelligent Systems* (Vol. 277, p. 1085-1093). doi: 10.1007/978-3-642-54924-3_103.
- Wechsler, H. (1980). Texture analysis — a survey. *Signal Processing*, 2(3), 271-282. doi: 10.1016/0165-1684(80)90024-9
- Wei-Ying, M. et Hongjiang, Z. (1998, 1-4 Nov. 1998). *Benchmarking of image features for content-based retrieval*. Communication présentée Signals, Systems & Computers, 1998. Conference Record of the Thirty-Second Asilomar Conference on. doi: 10.1109/acssc.1998.750865
- Westland, S. et Ripamonti, C. (2004). *Computational Colour Science using MATLAB*. John Wiley & Sons Ltd.
- Weszka, J. S., Dyer, C. R. et Rosenfeld, A. (1976). A Comparative Study of Texture Measures for Terrain Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-6(4), 269-285. doi: 10.1109/tsmc.1976.5408777

- Wong, W., Abu-Shareha, A. A., Pasha, M. F. et Mandava, R. (2013). *Enhanced Local Binary Pattern for Chest X-ray Classification*. Communication présentée 2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013).
- Wu, H. et Yan, S. (2016). Computing invariants of Tchebichef moments for shape based image retrieval. *Neurocomputing*, 215, 110-117. doi: 10.1016/j.neucom.2015.05.147
- Xie, X. et Mirmehdi, M. (2007). TEXEMS: texture exemplars for defect detection on random textured surfaces. *IEEE Trans Pattern Anal Mach Intell*, 29(8), 1454-1464. doi: 10.1109/TPAMI.2007.1038
- Xin-Wu, C. et Yu-Xi, L. (2010, 25-26 Dec. 2010). *Contourlet Retrieval System Using Absolute Mean and Kurtosis Features*. Communication présentée Information Engineering and Computer Science (ICIECS), 2010 2nd International Conference on. doi: 10.1109/iciecs.2010.5677689
- Xuejun, Z., Gao, X., Liu, B. J., Ma, K., Yan, W., Liling, L., . . . Fujita, H. Effective staging of fibrosis by the selected texture features of liver: Which one is better, CT or MR imaging? *Computerized Medical Imaging and Graphics*. doi: <http://dx.doi.org/10.1016/j.compmedimag.2015.09.003>
- Yadav, A. R., Anand, R. S., Dewal, M. L. et Gupta, S. (2015). Multiresolution local binary pattern variants based texture feature extraction techniques for efficient classification of microscopic images of hardwood species. *Applied Soft Computing*, 32(0), 101-112. doi: 10.1016/j.asoc.2015.03.039
- Yang, H.-Y., Li, Y.-W., Li, W.-Y., Wang, X.-Y. et Yang, F.-Y. (2014). Content-based image retrieval using local visual attention feature. *Journal of Visual Communication and Image Representation*, 25(6), 1308-1323. doi: 10.1016/j.jvcir.2014.05.003
- Yang, M., Kpalma, K. et Ronsin, J. (2008). A Survey of Shape Feature Extraction Techniques. Dans Y. Peng-Yeng (dir.), *Pattern Recognition* (p. 43-90): IN-TECH.
- Yang, Z. et Ai, H. (2007). Demographic Classification with Local Binary Patterns. Dans S.-W. Lee & S. Li (dir.), *Advances in Biometrics* (Vol. 4642, p. 464-473): Springer Berlin Heidelberg.
- YongHua, X. et Jin-Cong, W. (2015). Study on the identification of the wood surface defects based on texture features. *Optik - International Journal for Light and Electron Optics*, 126(19), 2231-2235. doi: 10.1016/j.ijleo.2015.05.101
- Youngeun, A., Riaz, M. et Jongan, P. (2010, 24-26 March 2010). *CBIR Based on Adaptive Segmentation of HSV Color Space*. Communication présentée 12th International Conference on Computer Modelling and Simulation (UKSim), 2010 Cambridge. doi: 10.1109/UKSIM.2010.53
- Yuan, H. et Zhang, X. P. (2010). Statistical Modeling in the Wavelet Domain for Compact Feature Extraction and Similarity Measure of Images. *IEEE Transactions on Circuits and Systems for Video Technology*, 20(3), 439-445. doi: 10.1109/TCSVT.2009.2031396
- Yue, J., Li, C. et Li, Z. (2015). An Improved Method for Image Retrieval Based on Color and Texture Features. Dans D. Li & Y. Chen (dir.), *Computer and Computing Technologies in Agriculture VIII* (Vol. 452, p. 739-752): Springer International Publishing.
- Yue, J., Li, Z., Liu, L. et Fu, Z. (2011). Content-based image retrieval using color and texture fused features. *Mathematical and Computer Modelling*, 54(3-4), 1121-1127. doi: <http://dx.doi.org/10.1016/j.mcm.2010.11.044>
- Zhang, B., Gao, Y., Zhao, S. et Liu, J. (2010). Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor. *IEEE Trans Image Process*, 19(2), 533-544. doi: 10.1109/TIP.2009.2035882

- Zhang, D. et Lu, G. (2002). Shape-based image retrieval using generic Fourier descriptor. *Signal Processing: Image Communication*, 17(10), 825-848. doi: 10.1016/s0923-5965(02)00084-x
- Zhang, D. et Lu, G. (2003). Evaluation of MPEG-7 shape descriptors against other shape descriptors. *Multimedia Systems*, 9(1), 15-30. doi: 10.1007/s00530-002-0075-y
- Zhang, D. et Lu, G. (2004). Review of shape representation and description techniques. *Pattern Recognition*, 37(1), 1-19. doi: 10.1016/j.patcog.2003.07.008
- Zhang, D., Wong, A., Indrawan, M. et Lu, G. (2000, December 13-15). *Content-based Image Retrieval Using Gabor Texture Features*. Communication présentée First IEEE Pacific-Rim Conference on Multimedia (PCM'00), Sydney, Australia.
- Zhang, Q. et Canosa, R. L. (2014). *A comparison of histogram distance metrics for content-based image retrieval*. doi: 10.1117/12.2042359 Repéré à <http://dx.doi.org/10.1117/12.2042359>
- Zhang, X., Gao, X., Liu, B. J., Ma, K., Yan, W., Liling, L., . . . Fujita, H. (2015). Effective staging of fibrosis by the selected texture features of liver: Which one is better, CT or MR imaging? *Comput Med Imaging Graph*, 46 Pt 2, 227-236. doi: 10.1016/j.compmedimag.2015.09.003
- Zhong, W., Qifa, K., Isard, M. et Jian, S. (2009, 20-25 June 2009). *Bundling features for large scale partial-duplicate web image search*. Communication présentée Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. doi: 10.1109/cvpr.2009.5206566
- Zhou, F., Feng, J.-f. et Shi, Q.-y. (2001, 7-10 October). *Texture feature based on local Fourier transform*. Communication présentée Image Processing, 2001. Proceedings. 2001 International Conference on, Thessaloniki. doi: 10.1109/ICIP.2001.958567
- Zucker, S. W. (1976). Toward a model of texture. *Computer Graphics and Image Processing*, 5(2), 190-202. doi: 10.1016/0146-664x(76)90027-7

Annexe (Bibliographie du candidat)

Travaux publiés :

[1] BOUGUEROUA, S. and BOUCHEHAM, B., 2016. *GLIBP: Gradual Locality Integration of Binary Patterns for Scene Images Retrieval*. Journal of Information Processing Systems (JIPS), (The official international journal of the Korea Information Processing Society, ISSN: 1976-913X (Print), ISSN: 2092-805X (Online)) ([Accepted 2016](#)).

[2] BOUGUEROUA, S. and BOUCHEHAM, B., 2014. *Ellipse based local binary pattern for color image retrieval*. In ISKO-Maghreb: Concepts and Tools for knowledge Management (ISKO-Maghreb), 2014 4th International Symposium IEEE, Algiers, Algeria, 1-8. DOI=<http://dx.doi.org/10.1109/ISKO-Maghreb.2014.7033459>.

[3] BOUGUEROUA, S. and BOUCHEHAM, B., 2015. *GLI-Color: Gradual Locality Integration of Color features for image retrieval*. In Proceedings of the Proceedings of the International Conference on Intelligent Information Processing, Security and Advanced Communication (Batna, Algeria2015), ACM, 2816920, 1-5. DOI=<http://dx.doi.org/10.1145/2816839.2816920>.

[4] BOUGUEROUA, S. and BOUCHEHAM, B., 2016. *Greedy Mean Squared Residue for Texture Images Retrieval*. In Modelling and Implementation of Complex Systems: Proceedings of the 4th International Symposium, MISC 2016, Constantine, Algeria, May 7-8, 2016, Constantine, Algeria, S. CHIKHI, A. AMINE, A. CHAOUI, K.M. KHOLLADI and E.D. SAIDOUNI Eds. Springer International Publishing, Cham, 123-136. DOI=http://dx.doi.org/10.1007/978-3-319-33410-3_9.

[5] BOUGUEROUA, S., BOUCHEHAM, B., and BELAOUED, M., 2015. *Enhancement of Image Characterization with Discrete Wavelet Coefficients using the Third Statistical Moment for Texture-Based Image Retrieval*. In Proceedings of the Conférence Nationale sur les Nouvelles Technologies Informatiques (CN2TI'15) (Guelma, Algeria, 10-12 Nov 2015 2015).

Pages de garde des travaux publiés :

GLIBP: Gradual Locality Integration of Binary Patterns for Scene images retrieval

Salah Bougueroua*, and Bachir Boucheham**

Abstract

We propose an enhanced version of the Local Binary Pattern (LBP) operator for texture extraction in images for the purpose of image retrieval. The novelty of our proposal is the observation that the LBP exploits only the lowest kind of local information through the global histogram, which reflects only the statistical distribution of the various LBP codes in the image. The Block Based LBP of Takala *et al*[1], that uses local histograms of LBP, was one of few tentative to catch higher level textural information. However, we believe that important local information in between the two levels is just ignored by the two schemas. The newly developed method: *Gradual Locality Integration of Binary Patterns* (GLIBP) is a novel attempt to catch as much local information as possible, in a gradual fashion. Indeed, GLIBP aggregates the texture features present in grayscale images extracted by LBP through a complex structure comprised of a multitude of ellipse-shaped regions that are arranged in circular-concentric forms of increasing size. The framework of ellipses is in fact derived through a simple parameterized generator. In addition, the elliptic forms allow targeting texture directionality, which is a very useful property in texture characterization. Also, the general framework of ellipses allows for taking into account the spatial information (specifically rotation). The effectiveness of GLIBP was investigated on the Corel-1K (Wang) dataset [2]. It was also compared to published works including the very effective DLEP of Murala *et al* [3]. Results show significant higher or comparable performance of GLIBP with regard to the other methods, which qualifies it as a good tool for scene images retrieval.

Keywords

CBIR, Elliptic-region, LBP, Local information, Global Information, Texture

1. Introduction

The accumulation of multimedia information on web servers and specifically the images due to increasing uploading from different sources and for different aims in large amounts, made the development of efficient algorithms for exploring and searching of images an important and mandatory task. The first generation of such techniques associated to each image a set of descriptive words. Accordingly, the searching for images similar to a query is guided by this textual information (keywords). This type of search is known by TBIR (Text Based Image Retrieval). Although this technique is widely used by existing search engines, it suffers from a set of disadvantages. Firstly, the annotation, which is mostly performed manually on large databases, is a hard and time consuming task. Secondly, it is largely admitted that manual annotation is subjective and relative to the human

* This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Manuscript received: Month 06, 20XX; accepted Month 06, 20XX.

Corresponding Author: Salah Bougueroua (s_bougueroua@yahoo.fr)

* Department of computer science, 20 août 1955 university of Skikda, B.P.26 route d'El-hadaiek, Skikda, 21000, Algeria (s_bougueroua@yahoo.fr)

** Department of computer science, 20 août 1955 university of Skikda, B.P.26 route d'El-hadaiek, Skikda, 21000, Algeria (boucheham_bachir@yahoo.fr)

Ellipse based local binary pattern for color image retrieval

Salah Bougueroua

s_bougueroua@yahoo.fr

Department of Computer science
20 Août 1955 University of Skikda
Skikda, Algeria

Bachir Boucheham

boucheham_bachir@yahoo.fr

Abstract— A novel region-based method for color scene images characterization is proposed. The method exploits the texture features present in grayscale images extracted by LBP. It superposes a set of ellipses which are generated by rotation and contraction and organized for covering the image, where each ellipse delimits a region of the LBP map. One LBP-histogram is calculated for each region. Finally, texture features present in the images are represented by m LBP-histograms, where m is the number of ellipses (regions) used. Each elliptic region in the query image is then matched with the most similar region of its class in the target image. A global distance between the query and the target images is afterward derived. The method has been compared to published works, including LBP and its improved version ILBP, using the Wang dataset. Results show significant higher or comparable performance of our method with regard to the other methods.

Keywords— CBIR, Ellipse-based, Elliptic-region, LBP, Retrieval.

I. INTRODUCTION

The falling prices of image acquisition devices and scanning through the development of technology and the improved performance of the Internet allowed more people to search for images for professional, touristic or just amusement reasons. Due to the accumulation of such images in large amounts, the task of effective and efficient image searching has emerged as an important problem in the fields of information mining, retrieval and classification. One paradigm for such a task is image annotation which allows image search based on keywords (TBIR: Text Based Image Retrieval). The inconvenience of this solution is the existence of the semantic gap between the images content and their respective descriptions. Another paradigm for image search/retrieval reposes on the visual content of images (CBIR: Content Based Image Retrieval). CBIR searches for relevant images based on properties of the images, such as colors, textures and shapes.

The CBIR field has been under extensive research for the last two decades [1] and many methods were proposed, based on color, texture and/or shape. For instance, intersection of color histograms was proposed by Swain and Ballard [2] for indexing large color image databases. The refinement histogram [3] is another effective color approach. It consists in

subdivision of color histogram under some conditions for creating a new color vector called “Color Coherence Vector” (CCV). From another perspective, texture is a ubiquitous property in objects and images. Indeed, many objects can be distinguished by their texture. However, texture features are hardly captured. There seems to be even a difficulty to come up with a formal and precise definition of texture [4]. A distinguished texture method was introduced by Ojala *et al* [9] is the *Local Binary Pattern* (LBP) on the supposition that the texture has two complementary local aspects: the pattern and its strength. Later, the method was extended [10] to yield rotation invariant LBP and other derivatives. Gabor wavelets are another approach for texture that proved to give interesting results in the context of CBIR [6-8]. For boosting of the effectiveness of retrieval systems, some researchers used combination of color, texture, shape. For instance, Murala *et al* [11] combined the color and texture features extracted by a histogram and Gabor wavelet, respectively. In [12] the authors combined these two properties (color, texture) but they used the Haar wavelet for extracting the texture features, whereas Milind *et al* [14] combined the three properties.

In this paper, we propose a novel texture based method for color images retrieval. The method is an enhancement of the LBP technique. Whereas the LBP uses a global histogram of extracted local texture features, we propose to exploit these same features by an elliptic-shaped regions matching. The method was implemented and compared to existing published works, including LBP and its enhanced version ILBP [15]. Results show comparable or better performance of the proposed method (ELBP) with respect to the other methods.

The rest of this paper is organized as follows. In Section 2, the original LBP is briefly presented. In the 3rd Section, our method is presented along with illustrative cases. The experiments and results are presented and discussed in the 4th Section. Finally, this work is terminated with a conclusion in the 5th section.

II. RELATED WORKS

A. Texture based approaches

Many texture based methods for image retrieval and classification were proposed during the years. One of the

GLI-Color: Gradual Locality Integration of Color features for image retrieval

Salah Bougueroua
Department of computer science,
20 août 1955 university of Skikda
B.P.26 route d'El-hadaiek, Skikda,
21000, Algeria
s_bougueroua@yahoo.fr

Bachir Boucheham
Department of computer science,
20 août 1955 university of Skikda
B.P.26 route d'El-hadaiek, Skikda,
21000, Algeria
boucheham_bachir@yahoo.fr

ABSTRACT

In this paper, we propose a novel method for color features based image retrieval. The proposal exploits the advantage of local histograms using a distinguished design to delimiting the concerned regions through an elliptic form, organized in circular-concentric manner of increasing size. For each elliptic region, normalized color histograms in the RGB color space are established. Through this sophisticated design, our method will gradually integrate the locality information, thus it is called: Gradual Locality Integration of Color features (GLI-Color). The experiments are conducted on Corel-1k database, using Manhattan and d_1 distance metrics. The comparisons of obtained results with global histogram and some published works have shown a significant higher performance.

Categories and Subject Descriptors

H.3.3 [INFORMATION STORAGE AND RETRIEVAL]: Information Search and Retrieval –

General Terms

Algorithms, Performance, Experimentation.

Keywords

Color, CBIR, Local histograms, Local information, elliptic-region, Image similarity.

1. INTRODUCTION

The generalization of the internet around the globe and the enhancement of the download bandwidth, and also the improvement of computers processing and storage capacities, all of that has motivated the users to dealing with the multimedia information in large amounts, such as images, videos and music. The retrieval and search of desired images with specific properties from huge amounts of such data is a quite hard task. The well known scenario to accomplishing the search process is usually as follow: The user enters one (or more) textual keyword(s), and then the system, based on these keywords, searches and returns to the user the relevant images. This manner of retrieving the images is known by text based image retrieval (TBIR). Although TBIR is a well established style of

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page.

Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

IPAC '15, November 23-25, 2015, Batna, Algeria

© 2015 ACM. ISBN 978-1-4503-3458-7/15/11...\$15.00

DOI: <http://dx.doi.org/10.1145/2816839.2816920>

image retrieval, it is characterized by several disadvantages, amongst which the semantic gap is the most important. Indeed, one user can tag a specific image with different keywords in different situations. Furthermore, the task of tagging a large number of images is clearly hard. To overcome these disadvantages, *Content Based Image Retrieval* (CBIR) has been proposed in the early 90's of the past century. Since then, CBIR has known an extensive research and interest [12]. Typically, a CBIR session is performed as follows: (1) A user enters an image as a query; (2) the system extracts the features of one (or more) visual attribute(s); (3) a similarity measure is established between the query and the images in the database, using the extracted features; (4) the system then ranks the images in the database according to their similarity to the query, in descending order; (5) Finally, the top k-images are returned to the user.

Many types of visual attributes were investigated in the literature, including color, texture and shape. For instance, the color attribute has been investigated since the early years of CBIR by Swain and Ballard [16] through the color histograms for indexing large image databases. Later, several methods have been proposed to describe the features of this attribute, such as: Color moments [14], Color Coherent Vector (CCV) [10], Cell/Color Histograms (CCH) [13], and the recently published work: Cell-CCV [11]. The second visual attribute is the texture, which is an omnipresent attribute in objects. Thus, it is a crucial attribute which has been exploited well before CBIR for image classification. For instance, Haralick *et al* [4] have proposed a method for image classification, called "the co-occurrence matrices" in 1973. Then, many other methods have been proposed for dealing with this important attribute, such as: Local Binary Pattern (LBP) [9], Local Derivative Pattern (LDP) [18].

The attributes combination is another issue used by researchers for boosting the effectiveness of their systems. For instance, both the color and texture features are used by Murala *et al* [8], also the color is combined successfully with the shape as in [3]. And also the combination of the three attributes as in [7].

In this paper, we address the problem engendered by the global color histograms using local histograms within elliptic regions organized in circular concentric fashion and with increasing size, to gradually integrate colors location. Hence de name, Gradual Locality Integration of Color features (GLI-Color). The method is experimented on the RGB color space. The above color features are extracted from each channel separately, then all features are exploited to establish a similarity measure between two images (query vs. target image in the database).

Greedy Mean Squared Residue for Texture Images Retrieval

Salah Bougueroua and Bachir Boucheham

Abstract In this paper, we propose a new algorithm for texture retrieval, using clustering strategy. Indeed, it is largely noticed that in existing CBIR systems and methods, the collection of the images similar to the query is realized on the basis of comparison of the database images to the query solely. Hence, the results might not be globally homogeneous. In this paper, the collection of the images most similar to the query is realized considering the global homogeneity of the whole cluster (result). Knowing that this is of an exponential order problem, we use a greedy solution consisting in growing the cluster corresponding to a query, one image at a time, based on the *Mean Squared residue* measure of Cheng and Church (Proceedings of the Eighth International Conference on Intelligent Systems for Molecular Biology, 2000) [1], originally proposed for the biclustering of gene expression data. At each stage, the new added image to the cluster will be that that preserves most the homogeneity of the current cluster. The texture descriptor used in this work is the uniform-LBP. Experimentations were conducted on two texture image databases, Outext and Brodatz. The proposed algorithm shows an interesting performance compared to the uniform-LBP combined to Euclidean metric.

Keywords CBIR • Image retrieval • Biclustering • Mean squared residue • Texture • Similarity measure • Greedy search • Optimization

S. Bougueroua (✉) • B. Boucheham
Department of Computer Science, 20 Août 1955 University of Skikda,
B.P. 26 Route d'El-Hadaiek, 21000 Skikda, Algeria
e-mail: s_bougueroua@yahoo.fr; s.bougueroua@univ-skikda.dz

B. Boucheham
e-mail: boucheham_bachir@yahoo.fr; b.boucheham@univ-skikda.dz

© Springer International Publishing Switzerland 2016
S. Chikhi et al. (eds.), *Modelling and Implementation
of Complex Systems*, Lecture Notes in Networks and Systems 1,
DOI 10.1007/978-3-319-33410-3_9

123

Enhancement of Image Characterization with Discrete Wavelet Coefficients using the Third Statistical Moment for Texture-Based Image Retrieval

Salah Bougueroua

s_bougueroua@yahoo.fr
s.bougueroua@univ-skikda.dz

Bachir Boucheham

Department of computer science
20 août 1955 university of Skikda
Skikda, Algeria
boucheham_bachir@yahoo.fr
b.boucheham@univ-skikda.dz

Mohamed Belaoued

belaoued.mohamed@gmail.com
m.belaoued@univ-skikda.dz

Abstract—Texture is a very important visual attribute in images. Moreover, classification and retrieval based on texture are very active topics among the research community. In this paper, we propose modified third statistical moments (Skewness) as a feature for wavelet sub-bands coefficients, for that purpose. Mainly, the proposed modification consists in calculating the skewness around the energy of each sub-band instead of the mean. We argue that this modification has important positive impact on retrieval accuracy. The experimentations are conducted with Haar basis wavelet, with different levels and with three distance metrics namely, Manhattan, Canberra and d1, using the Outex_TC_00000 database. Results show clearly that the proposed modified feature improves the discrimination power of the discrete wavelet in the context of image search/retrieval.

Keywords—CBIR; image retrieval; discrete wavelet; Haar basis; statistical moment; skewness

I. INTRODUCTION

The retrieval of images based on their contents (CBIR: Content Based Image Retrieval) is a challenging task due to high diversity of images and especially ones uploaded on the network by different users from different localities. CBIR appeared in the early of the 90s as an alternative for the already existing approach based on the keywords (TBIR: Text Based Image Retrieval), which suffers mainly from two serious drawbacks: the hardness of the annotation, especially for large databases, and the subjectivity of the attributed keywords.

Mainly, Color, Texture and Shape are the visual attributes the most studied and exploited ones. Wherein texture taken a large part thanks to its existence in the most objects and images, which qualify it as a ubiquitous property. However, texture is confronted with the difficulty to come up with a formal and precise definition. Consequently texture became a big challenge for the researchers. Generally, the existing method dealing with this attribute can be categorized into

one of the three following categories: spatial, frequency and spatial-frequency domain. For instance, the renowned method “Co-occurrence Matrices” of Haralick *et al* [1], and also the interesting “Local Binary Patterns” (LBP) of Ojala *et al* [2] are both acting in the spatial domain. Later, the LBP method has known many improvements, such as ILBP[3], LDP[4] and EILBP[5]. The published method of Zhou *et al* [6] which uses the Fourier Transform is an example on the frequency methods. In the spatial-frequency category, we find a set of fascinating and sophisticated methods, which are based on the transformation of signal to a set of wavelets using the scaling and shifting of a specific wavelet called the mother wavelet. Indeed, many types of wavelet are proposed, and exploited in computer vision realm, specifically in texture-based image retrieval and classification. For instance, Lonnestad [7] used Haar basis based wavelet for texture classification, and the Daubechies basis are also used by Moghaddam *et al*[8] for image retrieval, and also the Gabor filter [9].

In this paper, we are interested by texture feature extraction using discrete wavelet based on the Haar basis. Thus, we proposed a modified third statistical moment (skewness), which will be used with two other existing measures for each wavelet sub-band. The conducted experimentations, showed improvements on the results compared to some existing methods based on average precision. This proves that the modified skewness brought new discriminate information.

The remains of the paper are organized as follow: in the next section, the discrete wavelet transform (DWT) and its pyramidal decomposition form [10] are presented in the 1st sub section, then the related works in the feature extraction and image classification/retrieval topics are addressed in the 2nd sub section. In the section 3, the proposed method is presented. The experimentation are shown in the 4th section and discussed in the 5th. And finally we conclude this paper in the 6th section.