

**L'Analyse des Images Astronomiques**  
Master OMEGA 2 - Option Astrophysique

**Albert Bijaoui**

Observatoire de la Côte d'Azur  
B.P. 229 F-06304 Nice Cedex 4

Novembre 2005



# Table des matières

<b>1</b>	<b>Les Bases de l'imagerie numérique astronomique</b>	<b>9</b>
1.1	L'Acquisition de l'Information en Astronomie . . . . .	9
1.1.1	Observation et Théorie . . . . .	9
1.1.2	Schéma Général. . . . .	9
1.1.3	Les collecteurs. . . . .	10
1.1.4	Les Analyseurs au foyer. . . . .	11
1.1.5	Les récepteurs d'images optiques. . . . .	11
1.1.6	La réduction photométrique. . . . .	11
1.1.7	L'analyse des flux. . . . .	12
1.2	L'image numérique. . . . .	12
1.2.1	Le théorème de l'échantillonnage. . . . .	13
1.2.2	La formation de l'image. . . . .	13
1.2.3	Relation objet-image. . . . .	13
1.2.4	Systèmes convolutifs. . . . .	14
1.2.5	Profil instrumental. . . . .	14
1.2.6	La fonction de transfert. . . . .	15
1.2.7	Echantillonnage d'une image. . . . .	15
1.2.8	La discrétisation. . . . .	15
1.2.9	Echantillonnage en présence de bruit. . . . .	16
1.3	Image numérique et Image réelle. . . . .	16
1.3.1	Structure de l'image numérique. . . . .	16
1.3.2	Valeur en un point. . . . .	17
<b>2</b>	<b>Généralités sur l'analyse des images</b>	<b>21</b>
2.1	Objectifs de l'Analyse des images . . . . .	21
2.1.1	Analyse des images et information . . . . .	21
2.1.2	Analyse des images en Astronomie. . . . .	21
2.1.3	Possibilités offertes par l'analyse des images. . . . .	22
2.2	Les différents types d'analyse. . . . .	22
<b>3</b>	<b>L'analyse interactive</b>	<b>25</b>
3.1	Objectifs. . . . .	25
3.2	Matériel et logiciels associés. . . . .	25
3.2.1	Le matériel. . . . .	25
3.2.2	Les logiciels généraux. . . . .	26

3.2.3	Les logiciels complets. . . . .	26
3.2.4	Les logiciels spécialisés. . . . .	27
3.3	L'interface et son contenu. . . . .	27
3.3.1	Commandes ou actions graphiques. . . . .	27
3.4	Commandes principales. . . . .	27
3.5	Débruitage. . . . .	28
3.5.1	Filtre de Shannon . . . . .	28
3.5.2	Filtre de Wiener-Kolmogorov. . . . .	28
3.5.3	Lissages. . . . .	28
3.5.4	Méthodes non linéaires. . . . .	29
3.6	Amélioration de la résolution. . . . .	29
3.7	Amélioration du Contraste. . . . .	31
<b>4</b>	<b>Analyse globale de l'image</b>	<b>33</b>
4.1	Principes de l'analyse globale. . . . .	33
4.2	Utilisation de l'histogramme. . . . .	33
4.2.1	Détermination de l'histogramme. . . . .	34
4.3	Paramètres fondamentaux. . . . .	34
4.3.1	Paramètres déduits des moments. . . . .	34
4.3.2	Estimation avec rejet. . . . .	35
4.3.3	Statistiques d'ordre. . . . .	36
4.4	Utilisation des modes. . . . .	36
4.4.1	Les modes de l'histogramme. . . . .	36
4.4.2	Recherche des modes. . . . .	36
4.4.3	Utilisation du lissage. . . . .	36
4.4.4	Somme de gaussiennes. . . . .	37
4.5	Seuils et informations extraites. . . . .	37
4.5.1	Cas d'un histogramme monomode. . . . .	37
4.5.2	Cas d'un histogramme à deux modes. . . . .	38
4.5.3	Cas général. . . . .	38
4.6	Histogrammes de fonctions dérivées. . . . .	38
4.6.1	Utilisation des dérivées premières. . . . .	38
4.6.2	Histogrammes de fonctions de second ordre. . . . .	38
4.7	Histogrammes doubles, Cooccurrence. . . . .	39
4.7.1	Histogramme de fonctions locales. . . . .	39
4.7.2	Matrice de Cooccurrence. . . . .	39
4.7.3	Choix du vecteur. . . . .	39
4.7.4	Paramètres extraits de la matrice de cooccurrence. . . . .	39
4.8	Utilisation de la fonction d'autocorrélation. . . . .	40
4.9	Utilisation de la Densité spectrale de l'énergie. . . . .	41
4.10	Champ de Markov. . . . .	42
4.11	Analyse de texture. . . . .	42

<b>5</b>	<b>Analyse paramétrique de l'image</b>	<b>43</b>
5.1	Modélisation de l'image. . . . .	43
5.2	Méthodes d'estimation. . . . .	43
5.2.1	Estimateur des Moindres Carrés. . . . .	43
5.2.2	Principe du maximum de vraisemblance. . . . .	44
5.2.3	Autres méthodes. . . . .	44
5.2.4	Intervalle de confiance. . . . .	44
5.3	Tests des modèles. . . . .	44
5.3.1	Cas gaussien. . . . .	44
5.3.2	Test du $X^2$ . . . . .	45
5.4	L'utilisation du rejet. . . . .	45
5.5	Choix des pixels. . . . .	46
5.6	Modification de la pondération. . . . .	46
5.7	Correction de l'environnement. . . . .	47
5.8	La modélisation restreinte. . . . .	47
5.9	La cartographie du fond. . . . .	48
5.9.1	Concept général. . . . .	48
5.9.2	Cas d'un fond constant. . . . .	48
5.9.3	Cartographie à partir de l'histogramme local. . . . .	49
5.9.4	Recherche des points de fond. . . . .	49
5.9.5	Recherche des points du fond et cicatrisation. . . . .	49
5.9.6	Ajustement fonctionnel. . . . .	49
5.10	Méthodes de détermination de la position. . . . .	50
5.10.1	Ajustement. . . . .	50
5.10.2	Maximum de corrélation. . . . .	52
5.10.3	Position par le maximum. . . . .	52
5.10.4	Barycentre. . . . .	52
5.10.5	Centre de symétrie. . . . .	52
5.11	Calcul des flux. . . . .	53
5.12	Estimation du paramètre d'échelle. . . . .	53
5.12.1	Ajustement. . . . .	53
5.12.2	Estimation à partir des moments. . . . .	54
5.13	Autres paramètres de forme. . . . .	54
<b>6</b>	<b>La vision par ordinateur</b>	<b>55</b>
6.1	La vision par une machine. . . . .	55
6.1.1	Les éléments du problème. . . . .	55
6.1.2	Les objets. . . . .	55
6.1.3	Domaines et Objets. . . . .	56
6.2	Appartenance d'un pixel à un domaine. . . . .	56
6.2.1	Aspect multiple de la vision. . . . .	56
6.2.2	Critère de couleur. . . . .	56
6.2.3	Critère de contraste. . . . .	57
6.2.4	Critère de luminosité. . . . .	57
6.2.5	Critère de texture. . . . .	57
6.2.6	Critères du second ordre. . . . .	58

6.2.7	Autres critères. . . . .	58
6.3	Bords et fonctions de bord. . . . .	58
6.3.1	Bords d'un tracé à une dimension. . . . .	58
6.3.2	Bords par les maxima des gradients. . . . .	59
6.3.3	Gradients et fonctions de bord. . . . .	59
6.3.4	Points de bord déterminés de manière dyadique. . . . .	61
6.4	Appartenance à un domaine. . . . .	62
6.4.1	Critère basé sur la détection par fonction. . . . .	62
6.4.2	Critère basé sur la délimitation des domaines. . . . .	62
6.4.3	Appartenance floue. . . . .	62
6.4.4	Relaxation dans la fonction d'appartenance. . . . .	63
6.5	Segmentation de l'image. . . . .	63
6.5.1	Étiquetage des pixels d'une image. . . . .	63
6.5.2	Algorithme associé à une lecture vidéo de l'image. . . . .	63
6.5.3	Variante vectorielle. . . . .	64
6.5.4	Procédures par dilatations successives. . . . .	65
6.6	Caractères structuraux associés aux domaines. . . . .	65
6.6.1	Caractères accessibles pendant l'étiquetage. . . . .	65
6.6.2	Moments et moments réduits. . . . .	65
6.6.3	Valeurs extrêmes. . . . .	68
6.6.4	Quantités déduites ultérieurement à l'étiquetage. . . . .	69
6.7	Morphologie et Paramétrisation. . . . .	69
6.8	Analyse avec le Contour. . . . .	70
6.8.1	Extraction du contour du domaine. . . . .	70
6.8.2	Analyse du contour. . . . .	70
6.8.3	Courbure et équation intrinsèque. . . . .	72
6.8.4	Convexité. . . . .	72
6.8.5	Paramètres associés au contour. . . . .	73
6.9	Reconnaissance par corrélation. . . . .	74
6.9.1	Principe. . . . .	74
6.9.2	Procédure par lissage. . . . .	74
6.9.3	Utilisation de la segmentation. . . . .	75
6.9.4	Corrélation et contour. . . . .	75
<b>7</b>	<b>Etudes morphologiques.</b>	<b>77</b>
7.1	Morphologie mathématique. . . . .	77
7.1.1	Erosion et Dilation. . . . .	77
7.1.2	Composition des opérations. . . . .	77
7.1.3	Distance au contour. . . . .	77
7.1.4	Squelette d'un domaine. . . . .	78
7.2	Filtre morphologique. . . . .	78
7.3	Morphologie en niveaux de gris. . . . .	79
7.3.1	Germes de la ligne. . . . .	79
7.3.2	Volume image et filtre morphologique. . . . .	79
7.3.3	Germes d'une image. . . . .	79
7.4	Structures de la surface image. . . . .	79

7.4.1	La surface image. . . . .	79
7.4.2	Les extrema. . . . .	79
7.4.3	Les lignes associées aux dérivées partielles premières. . . . .	80
7.4.4	Graphe de liaison entre extrema. . . . .	80
7.4.5	Lignes de vallée. . . . .	80
7.4.6	Les lignes de points d'inflexion. . . . .	81
7.4.7	Autres lignes. . . . .	82
7.5	Profils associés aux extrema. . . . .	82
7.5.1	Profils radiaux. . . . .	82
7.5.2	Cas d'un objet de forme elliptique. . . . .	82
7.5.3	Décomposition à partir des profils. . . . .	82
7.5.4	Profil angulaire. . . . .	83
<b>8</b>	<b>La Vision Multiéchelle.</b>	<b>85</b>
8.1	Les transformations Multi-échelles . . . . .	85
8.2	L'Algorithme à trous . . . . .	86
8.3	Visualisation . . . . .	87
8.4	Coefficients Significatifs . . . . .	87
8.5	Le filtrage adaptatif multiéchelle . . . . .	88
8.6	La restauration Multiéchelle. . . . .	88
8.7	La vision multiéchelle . . . . .	89
8.8	Conclusions. . . . .	89
<b>9</b>	<b>Analyse d'un ensemble d'Images</b>	<b>91</b>
9.1	Introduction. . . . .	91
9.2	Le Rééchantillonnage. . . . .	91
9.3	Réduction Photométrique. . . . .	92
9.4	Comparaison des Images 2 à 2. . . . .	92
9.5	La comparaison de 3 images. . . . .	93
9.6	Analyse d'une collection importante d'images. . . . .	93
9.7	L'analyse en composantes principales. . . . .	93
9.8	Détermination de la Base. . . . .	94
9.9	Diagramme des Images. . . . .	95
9.10	Diagramme des pixels. . . . .	95
9.11	Analyse des Nuées. . . . .	95
9.12	Présentation des résultats. . . . .	97
9.13	Réseaux de neurones artificiels. . . . .	97
9.14	Analyse des Mouvements. . . . .	98





# Chapitre 1

## Les Bases de l'imagerie numérique astronomique

### 1.1 L'Acquisition de l'Information en Astronomie

#### 1.1.1 Observation et Théorie

Le but des recherches astronomiques consiste dans la détermination de la structure de l'univers à travers ses différentes composantes et d'en donner une interprétation compatible avec les connaissances de la physique. L'observation joue donc un rôle central. L'astronome va tout d'abord rechercher des *faits*. Pour cela il doit disposer d'une instrumentation adaptée à leur nature. Après leur mise en évidence, il doit les interpréter grâce à des modèles mathématiques, résultant d'hypothèses physiques sur les astres étudiés (théorie). Il y a en fait un va et vient constant entre l'observation et la théorie : on ne découvre que très rarement des faits fortuits. L'observateur est toujours guidé par un modèle de l'objet étudié, de même qu'un modèle s'appuie nécessairement sur un ensemble de faits, en un aussi grand nombre que possible.

L'observation, en astronomie est basée en grande partie sur les ondes électromagnétiques. C'est seulement pour le système solaire que des études in-situ peuvent être effectuées. Les rayons cosmiques fournissent aussi quelques informations difficiles à déchiffrer sur notre galaxie. L'analyse des gerbes Cherenkov à partir de la coïncidence entre les flux mesurés avec plusieurs télescopes permet aujourd'hui d'observer les émissions de très hautes énergies provenant de très lointains quasars. L'astronomie des neutrinos se développe, par exemple avec la caméra ANTARES. Enfin la détection des ondes gravitationnelles, avec VIRGO, LIGO ou l'interféromètre spatial LISA devrait ouvrir une nouvelle fenêtre d'observation non photonique.

#### 1.1.2 Schéma Général.

L'observation astronomique est basée sur les éléments suivants :

**Le collecteur** : C'est avec lui qu'on récolte les photons qui vont nous servir à établir des faits.

**L'analyseur** : les photons passent toujours à travers un dispositif plus ou moins compliqué dont le but est de pouvoir spécifier la nature du fait observé. Il s'agit, par exemple, d'un filtre coloré ou d'un spectrographe.

**Le récepteur de photons** : cela peut être, par exemple, un simple photomultiplicateur (monocanal) ou une caméra à transfert de charges (CCD, multicanaux).

**La réduction photométrique :** les mesures brutes sont réduites de tous les facteurs perturbateurs (non-linéarités, variations de transmission de l'optique, variations de sensibilité, distorsion géométrique,...).

**L'analyse des flux :** dans le cas d'un récepteur monocanal il s'agit d'une simple analyse de flux, le plus souvent temporelle (variations), dans le cas d'un récepteur multicanaux on effectue une analyse de ou des images.

**L'analyse des données :** les résultats obtenus à l'étape précédente sont analysés, soit de manière statistique pour mettre en évidence des *lois*, soit par comparaison à des prévisions théoriques ou empiriques.

C'est donc seulement au dernier point que s'effectue l'analyse des faits observés, but de tout travail de recherche en astronomie.

### 1.1.3 Les collecteurs.

Les collecteurs d'onde électromagnétique sont extrêmement divers. Ils dépendent du domaine de longueur d'onde, la panoplie étant la plus riche pour le domaine visible.

**Les très grandes longueurs d'onde :** Pour  $\lambda$  supérieure à 40m environ l'ionosphère réfléchit les ondes électromagnétiques, l'observation s'effectue par satellite. Des mesures interférométriques permettent de reconstituer des images.

**Pour  $\lambda$  entre 1mm et 40m** on est dans le domaine de la radioastronomie au sol. Les radiotélescopes sont des antennes géantes, mobiles ou non, fonctionnant individuellement ou en montage interférométrique. En raison de la taille gigantesque de ces instruments, et de leur coût élevé, il n'en existe que très peu dans le monde, fonctionnant parfois dans un cadre international (IRAM), ou avec un large accès (VLA).

**Pour  $\lambda$  compris entre  $1\mu\text{m}$  et 1mm** on est dans le domaine Infra-Rouge et submillimétrique. On utilise des grands télescopes optiques ou des grands collecteurs de médiocre qualité optique. Pour quelques bandes passantes, l'atmosphère joue le rôle d'écran, on doit observer avec un satellite. Il s'agit alors souvent de missions internationales (IRAS, ISO).

**Dans le domaine visible 0.3 à  $1\mu\text{m}$**  l'observation s'effectue avec des lunettes (réfracteurs) ou des télescopes (réflecteurs). Des techniques interférométriques ont été mises en oeuvre avec succès avec une échelle de plusieurs dizaines de mètres (GI2T, VLTI). En raison des dégradations de l'information par l'atmosphère, on choisit des sites privilégiées. Le HST (Hubble Space Telescope) permet de faire des observations spatiales d'une résolution et d'une profondeur sans aucune commune mesure avec les observations au sol. La plupart des grands instruments sont gérés dans des consortiums nationaux ou internationaux. Avec l'optique adaptative il est devenu possible d'obtenir des images de résolution similaire au sol avec les plus grands collecteurs (CFHT, VLT, Keck).

**De l'UV aux X** la technique des télescopes embarqués est la seule possible. Il s'agit aussi d'instruments coûteux, développés en coopération par plusieurs pays (IUE, Einstein, Exosat, Chandra, XMM, ...).

**L'astronomie  $\gamma$**  ne s'effectue pas avec des télescopes mais avec des dispositifs permettant de détecter chaque photon avec sa direction (chambre à étincelles, scintillateurs,...). On reconstitue les images à partir des ces informations. Les instruments sont embarqués sur un satellite. Une grande partie de l'information vient de l'aspect temporel (sursauts  $\gamma$ ). Les instruments sont souvent internationaux (COS-B, BATSE, HETE2, INTEGRAL).

Quel que soit le domaine de longueur d'onde on retrouve des éléments similaires :

- Le télescope avec ses caractéristiques optiques ;
- L'interférométrie et ses potentialités ;
- Les problèmes de l'atmosphère
- Le coût des instruments et la coopération internationale.

#### 1.1.4 Les Analyseurs au foyer.

Au foyer d'un télescope il est très rare qu'on n'interpose pas un dispositif avant le récepteur. Il peut s'agir d'éléments très divers :

**Filtres colorés** : ce qui permet d'avoir une définition de la bande passante reçue par le récepteur. Cette notion de filtre s'étend aisément au domaine radio. Dans le domaine optique, on distingue les bandes larges (1000 Å), moyennes (300 Å) et étroites (inférieures à 100 Å). Pour ces dernières on utilise des filtres interférentiels.

**Des polariseurs** : pour étudier le taux de polarisation des sources étudiées.

**Des spectrographes** permettent de décomposer la lumière et d'obtenir une information sur les variations d'intensité avec la longueur d'onde.

Des *bonnettes* permettent d'intégrer l'ensemble des éléments nécessaires à l'acquisition des informations. Sur la bonnette on trouve l'optique de repérage du champ, avec une caméra TV pour une vision déportée, celle de guidage avec aussi sa caméra TV, les éléments permettant d'effectuer une mise au foyer, ceux qui permettent d'illuminer avec une source de référence ou avec une lampe à émission riche en raies fines et tout un ensemble mécanique permettant de déplacer l'ensemble dans les 3 directions et en rotation autour de l'axe optique. Les bonnettes sont les compléments indispensables du télescope.

#### 1.1.5 Les récepteurs d'images optiques.

Les récepteurs d'images sont essentiellement :

**L'émulsion photographique** pour une longueur d'onde inférieure au micron.

**Les récepteurs à photoémission** amplificateur de brillance, tube-image, galette de microcanaux, électronique, etc. Ils ne sont pratiquement plus utilisés pour l'observation astronomique.

**Les détecteurs** du type CCD, matrices de photodiodes au silicium, lues par transfert de charges. L'utilisation de ces détecteurs s'est généralisée au cours des années 80.

**Les caméras à comptage de photons** basées sur une amplification primaire importante permettant de voir l'effet d'un photon et que l'on couple avec une caméra TV.

#### 1.1.6 La réduction photométrique.

Les mesures données par tout récepteur n'ont qu'un sens relatif. On doit en premier lieu les réduire de manière à obtenir des informations quantitatives ayant un sens physique. Les réductions à effectuer proviennent du récepteur (non linéarité, variation de sensibilité, fond sous-jacent, pixels défectueux connus ou non) et de l'ensemble instrumental, atmosphère, télescope et analyseur.

Pour cela il faut toujours des étalonnages. L'instrument observe des sources connues, de laboratoire et célestes. On utilise les mesures obtenues pour réduire les observations.

En ce qui concerne les effets atmosphériques, il est souvent facile de supprimer la lumière du ciel nocturne, par contre les corrections d'absorption sont plus délicates.

L'élément le plus difficile dans la réduction est celui de l'établissement des sources étalons. De très nombreuses observations sont nécessaires, avec une réduction tenant compte des différentes conditions. Malgré cela des écarts très importants existent entre les différents observateurs, souvent en raison d'un choix différent des étalons ou de réductions incorrectes.

### 1.1.7 L'analyse des flux.

Les flux lumineux observés permettent d'obtenir diverses quantités physiques. Dans le cas de récepteurs de type image, l'analyse conduit à l'obtention de positions, de magnitudes, de paramètres de forme ou de texture. L'analyse de mesures temporelles conduit à des déterminations de périodes et des amplitudes d'harmoniques. Dans le cas des sources non stationnaires, l'analyse temps-fréquence permet de préciser leur comportement photométrique.

Cette analyse s'effectue par comparaison à une modélisation de l'objet (analyse déterministe) soit avec des méthodes d'analyse structurale d'usage générale en reconnaissance de forme. A ce niveau, l'astronome possède une série de paramètres physiques significatifs sur les sources étudiées. Il doit les analyser en tenant compte des connaissances *a priori* sur ces sources. Les méthodes dont ils disposent sont souvent très simple :

**Comparaison par rapport à une prédiction :** il s'agit de voir, si compte tenu des erreurs les valeurs des quantités physiques obtenues sont compatibles avec les observations antérieures ou les prédictions d'un modèle.

**La distribution des paramètres :** à partir du catalogue des mesures, on effectue des histogrammes, et on les ajuste par des lois adéquates. L'existence de différents modes permet de déduire l'existence de groupes.

**L'étude des corrélations entre paramètres** permet de montrer l'existence d'un lien statistique entre différentes quantités. L'analyse en composantes principales, sous différentes formes, permet de généraliser l'étude des corrélations à  $n$  variables. L'analyse des corrélations peut permettre de déduire des *lois* empiriques entre les variables.

**L'analyse des diagrammes** bi ou  $n$ -dimensionnels permet de séparer les différentes classes d'objet, grâce en particulier à *l'analyse des nuées*. L'introduction des réseaux artificiels de neurones dans les années 80 ont permis de faire des progrès très sensibles dans ce domaine.

Avec des méthodes empiriques, souvent non informatisées, on compare les diagrammes obtenus à ceux que l'on déduit d'une modélisation, ou de connaissances antérieures du problème.

L'analyse des données conduit ainsi à guider l'astronome dans sa recherche sur la structure de l'univers. Son rôle est de pouvoir sélectionner les interprétations pertinentes des faits observés et de le guider vers une nouvelle approche du problème.

## 1.2 L'image numérique.

Un ordinateur ne peut stocker dans ses mémoire que des quantités dénombrables, et en nombre nécessairement limité, le transfert d'une image dans un ordinateur est soumis à 2 règles essentielles : l'échantillonnage et la discrétisation.

### 1.2.1 Le théorème de l'échantillonnage.

Il s'agit d'extraire du signal continu (et n-dimensionnel) que constitue l'image l'ensemble des échantillons de travail. Cette sélection pourrait s'effectuer de manière arbitraire, mais un théorème fondamental donne l'assise sur laquelle elle doit reposer. Il s'agit du théorème de l'échantillonnage de Shannon, dont l'énoncé pour un signal 1D est le suivant :

**Théorème 1 (Shannon)** *Soit une fonction  $f(x)$ , définie sur  $]-\infty, +\infty[$ , admettant une transformée de Fourier  $F(\nu)$ . Si  $f(x)$  est à support bornée, c'est-à-dire si  $F(\nu) = 0$  pour  $\nu > \nu_c$ , alors la fonction  $f(x)$  est parfaitement connue en tout point si on possède un échantillon de ses valeurs, avec un pas régulier  $p$  entre chacun valant :  $p \leq 1/2\nu_c$ .*

Ce théorème admet un complément indispensable, celui qui permet de recalculer en chaque point la fonction  $f(x)$  à partir des échantillons  $f(kp)$  :

$$f(x) = \sum_{-\infty}^{+\infty} f(kp) \frac{\sin \pi \left( \frac{x}{p} - k \right)}{\pi \left( \frac{x}{p} - k \right)}.$$

On généralise aisément ce théorème d'échantillonnage et d'interpolation aux fonctions à plusieurs dimensions, par séparation des variables. A 2 dimensions l'échantillonnage conduit à un tableau régulier à 2 dimensions. A 3D, on a un tableau aussi 3D, etc. La fréquence de coupure peut ne pas être la même dans chaque dimension, le pas n'est alors pas le même dans chaque dimension.

### 1.2.2 La formation de l'image.

Le théorème de l'échantillonnage fait intervenir la fréquence de coupure de la fonction  $f(x)$ . Pour cela, il faut, en principe, en étudier la transformation de Fourier, ce qui nécessite un premier échantillonnage! Dans le cas d'une image de synthèse ou obtenue par des moyens introduisant un échantillonnage interne, le problème ne se pose pas, par contre pour les images naturelles sur support photographique par exemple, on doit procéder à un choix. Celui-ci est guidé par la nature du système ayant conduit à l'image à traiter. En effet, toute image naturelle est obtenue à partir d'une chaîne complexe d'acquisition dont on peut distinguer différents éléments.

**L'objet de l'étude** : il peut se trouver à distance finie ou à l'infini. Il est caractérisé par une distribution de luminosité  $O(X, Y)$ , que l'on cherche à étudier.

**Le collecteur** : il s'agit du système optique permettant de former l'image sur le récepteur.

**Le récepteur d'images** : l'image est formée par le collecteur sur un récepteur. Dans le plan du récepteur nous obtenons un éclaircissement de l'image  $I(x, y)$ .

### 1.2.3 Relation objet-image.

Un système imageur permet donc de passer d'une fonction objet  $O(X, Y)$  à celle image  $I(x, y)$ . Souvent ce passage, moyennant d'éventuelles corrections que nous étudierons ultérieurement, est linéaire c'est-à-dire qu'il satisfait aux conditions suivantes :

**Additivité** : si  $O_1$  et  $O_2$  sont deux distributions objets conduisant aux images  $I_1$  et  $I_2$ , l'image  $I$  formée à partir de  $O_1 + O_2$  est la somme  $I_1 + I_2$ .

**Multiplication par un scalaire** : si  $k$  est une constante positive quelconque, l'image formée à partir de  $kO$  est  $kI$ .

Décomposons l'objet  $O$  en une infinité d'objets n'émettant chacun que dans un domaine d'aire  $dXdY$  avec une luminosité  $O(X, Y)$ , dans ce domaine. En vertu de la linéarité, nous obtenons pour la luminosité en chaque point de l'image par :

$$I(x, y) = \int_{-\infty}^{+\infty} O(X, Y)p(x, y; X, Y)dXdY \quad (1)$$

$p(x, y; X, Y)$  est la part de la luminosité de l'objet en  $(X, Y)$  qui est formée au point  $(x, y)$  de l'image. La relation (1) n'est pas élémentaire, on la simplifie généralement en admettant que le système obéit à une autre condition : l'invariance par translation.

#### 1.2.4 Systèmes convolutifs.

La plus souvent, le système optique est tel que lorsqu'on effectue une translation sur l'objet, l'image ne subit elle aussi qu'une translation. Dans ce cas l'image d'un point est une tache dont la forme est aussi invariante. En adoptant pour l'objet et l'image des coordonnées telles que le grandissement soit de 1 en tout point, la relation (1) se simplifie sous la forme :

$$I(x, y) = \int_{-\infty}^{+\infty} O(X, Y)p(x - X, y - Y)dXdY \quad (2)$$

#### 1.2.5 Profil instrumental.

La fonction  $p(x, y)$  décrit la manière dont un point est étalée par le système optique. On la désigne soit par fonction d'étalement, soit par profil instrumental. En anglais, elle est appelée *Point Spread Function* (PSF).

Ce profil instrumental est le produit des différentes aberrations du système et des effets de diffraction :

**Aberrations géométriques** : elles sont induites par les lois de l'optique géométrique. *L'aberration de sphéricité* décrit la largeur centrale de l'image. *La courbure de champ* est relative à l'élargissement de la tache hors de l'axe. *L'astigmatisme* est lié au rapport des allongements de l'image dans le sens tangentiel et sagittal. *La coma* se rencontre lorsque la tache d'un point devient de plus en plus dissymétrique au fur et à mesure que l'on s'éloigne du centre. D'autres aberrations géométriques surviennent lorsque l'on s'écarte trop des conditions de Gauss, comme par exemple *l'aberration de décentrement*.

**Les effets chromatiques** : une image optique est rarement observée dans une bande spectrale très restreinte. Le plus souvent, des effets chromatiques importants existent, conduisant à élargir la tache image d'un point.

**Effet de la diffraction** : on ne considère en général que ceux de la diffraction de Fraunhofer, à l'infini. En raison des propriétés ondulatoires de la lumière, même avec un système parfait du point de vue géométrique (miroir parabolique par exemple), l'image d'un point ne peut être un point, mais une tache (tache d'Airy) dont la structure est conditionnée par la forme de la pupille du système optique. La diffraction forme la limite fondamentale de tout système imageur.

### 1.2.6 La fonction de transfert.

En désignant par  $\hat{O}(\nu_x, \nu_y)$  et  $\hat{I}(\nu_x, \nu_y)$ , les transformées de Fourier de  $O(x, y)$  et  $I(x, y)$ , et par  $\hat{p}(\nu_x, \nu_y)$  celle de  $p(x, y)$ , la relation (2) s'écrit sous la forme :

$$\hat{I}(\nu_x, \nu_y) = \hat{O}(\nu_x, \nu_y)\hat{p}(\nu_x, \nu_y) \quad (3)$$

$\hat{p}(\nu_x, \nu_y)$  est appelée *fonction de transfert optique* du système imageur. Il peut s'agir d'une fonction complexe si le profil instrumental n'est pas symétrique. On considère souvent alors les fonctions de modulation et de phase.

### 1.2.7 Echantillonnage d'une image.

Quel que soit  $O(x, y)$ ,  $I(x, y)$  possède la fréquence de coupure de  $p(x, y)$  en raison de la relation (3). Or le profil instrumental admet nécessairement une fréquence de coupure, celle liée à la diffraction de Fraunhofer, déterminée par la taille de la pupille.

La connaissance du système imageur détermine donc celle de la fréquence de coupure associée conduisant au pas d'échantillonnage de l'image quel que soit l'objet. Souvent, l'échantillonnage s'effectue simplement en examinant la taille des plus petits détails et en prenant un pas moitié.

### 1.2.8 La discrétisation.

La discussion précédente concerne la manière dont on sélectionne les échantillons discrets de l'image. Mais on n'a considéré que des valeurs dans l'ensemble des réels  $\mathbf{R}$ . Or un ordinateur ne peut travailler que sur des éléments discrets. Une conversion est indispensable. La transformation en valeurs discrètes est associée à une limitation physique de la connaissance dans la précision des mesures : le bruit. En effet, comme tout signal, la mesure de  $I(x, y)$  est bruitée pour diverses raisons :

**Bruit de la chaîne d'acquisition :** consistant dans la fluctuation du signal présent même en l'absence d'un signal réel ;

**Bruit intrinsèque à l'image :**  $I(x, y)$  est une variable aléatoire en raison d'une part de la nature de la lumière (bruit de photons), et d'autre part, aux mécanismes associés à la détection des photons.

En raison du bruit,  $I(x, y)$  a une loi distribution, avec une espérance  $\bar{I}(x, y)$  et une dispersion  $\sigma(I)$ . La discrétisation consiste à définir des niveaux de valeurs discrets  $I_n$ . Si  $I(x, y)$  est compris entre  $I_n$  et  $I_{n+1}$ , on affecte la mesure  $n$ . La distance  $D_n$  entre les niveaux  $n$  et  $n+1$  est déterminée par  $\sigma(I_n)$ . L'image est bien discrétisée si :

$$D_n < \sigma(I_n)$$

Dans le cas contraire, on a une perte d'informations sur la connaissance de  $I(x, y)$ , car la mesure du niveau serait bien moins précise que celle qu'on aurait en l'absence de quantification. On peut remarquer qu'une bonne quantification peut justifier une largeur variable du pas codeur. On peut résoudre ce problème en effectuant une transformation préalable des données  $I_n$  de manière à ce que cette nouvelle variable ait une dispersion constante. Par exemple pour un comptage de photons  $n$ , si  $n$  est suffisamment grand (plus que 30) il suffit d'appliquer la

transformation d'Anscombe :

$$x = 2\sqrt{n + \frac{3}{8}}$$

A chaque loi  $\sigma(I)$  on peut associer une transformation stabilisant la variance.

### 1.2.9 Echantillonnage en présence de bruit.

Si une fonction a une fréquence de coupure  $\nu_c$  nous pouvons parfaitement restituer le signal si le pas entre deux échantillons est  $1/2\nu_c$ . Lorsque le signal est bruité, le bruit peut avoir une densité spectrale ayant une fréquence de coupure bien plus élevée. L'échantillonnage au pas de Shannon n'est pas valable pour le bruit. En raison des répliques dans la transformée de Fourier, liées à l'échantillonnage, cela se traduit par une augmentation du bruit dans une proportion d'autant plus grande que sa fréquence de coupure  $\nu_b$  est plus élevée.

L'échantillonnage doit donc tenir compte de cet aspect. On devrait donc être amené à échantillonner avec un pas de  $1/2\nu_b$ . Cela conduirait à un sur-échantillonnage par rapport au signal très superflu, dont l'intérêt est de décrire correctement le bruit, ce qu'on ne recherche pas en général. La voie logique consiste à filtrer le bruit de manière à réduire sa fréquence de coupure, afin de retrouver un nombre d'échantillons en harmonie avec le signal étudié. Pour cela on effectue une intégration analogique liée à une aire d'analyse : spot du scanner (photographie), surface d'intégration (comptage de photons, dispositifs à transfert de charges, etc.).

Si l'aire d'intégration est uniforme, le résultat correspond à une convolution par une porte. Considérons le problème à une dimension. Soit  $e(x)$  le signal d'entrée,  $s(x)$  le signal de sortie et  $L$  la largeur de la fente. Nous avons la relation :

$$s(x) = \int_{-\infty}^{+\infty} e(u)\text{Rect}_L(x-u)du$$

Dans l'espace de Fourier cela conduit à :

$$\hat{s}(\nu) = L\hat{e}(\nu)\frac{\sin \pi\nu L}{\pi\nu L}$$

La fente agit donc comme un filtre, mais en laissant des hautes fréquences spatiales, puisque la fonction  $\text{sinc}(x)$  converge lentement. Si on choisit une largeur  $L = 1/\nu_c$ , le filtre s'annule pour  $\nu = \nu_c$ , mais il subsiste du bruit haute fréquence. Pour  $\nu < \nu_c$ , on atténue le bruit et le signal. L'atténuation est significative pour  $\nu > \frac{3}{4}\nu_c$ . C'est pourquoi, on préfère échantillonner avec une fente plus étroite  $L = \frac{1}{2\nu_c}$ , pour laquelle l'atténuation haute fréquence reste faible.

## 1.3 Image numérique et Image réelle.

Dans la section précédente nous avons discuté des bases sur lesquelles sont fondée la mise en mémoire d'une image. Examinons maintenant la relation réelle existant entre l'image numérique et l'image réelle.

### 1.3.1 Structure de l'image numérique.

Une image numérique est formée de 2 parties :



1. Un tableau de valeurs échantillonnées ayant autant de dimensions que l'image réelle correspondante ;
2. Une série d'attributs permettant de pouvoir déterminer en tout point la valeur de l'image.

Éventuellement d'autres descripteurs de l'image, n'ayant pour but que de faciliter son identification ou son traitement sont définis. Les attributs fondamentaux sont :

- Le nombre de dimensions de l'image ;
- Pour chaque dimension, le nombre d'éléments discrets, la coordonnée du premier élément et le pas d'échantillonnage.
- Le mode de codage des données.

Souvent on ne traite que des images 2D, avec un codage des données bien défini. On ne garde alors que les 6 attributs liés à l'échantillonnage :

1. Nombre de Colonnes de l'image  $N_c$
2. Nombre de lignes  $N_l$
3. Abscisse de départ  $X_d$
4. Ordonnée de départ  $Y_d$
5. Pas d'échantillonnage en X  $D_x$
6. Pas d'échantillonnage en Y  $D_y$

Le nombre d'éléments de l'image ou nombre de pixels est donc  $N_c \times N_l$ .

Les descripteurs de l'image peuvent être nombreux : date de création, historique, valeur des extréma d'intensité, paramètres du profil instrumental, etc. Dans le cas des images astronomiques la convention FITS d'écriture des fichiers permet un échange aisé entre les laboratoires.

### 1.3.2 Valeur en un point.

Le problème essentiel de l'imagerie numérique s'énonce de la manière suivante : soit une image numérique avec ses attributs et son tableau de valeurs échantillonnées. Déterminer la valeur de l'image réelle correspondante à un point M de coordonnées données.

Considérons le cas 2D, pour simplifier. Soit  $(x, y)$  les coordonnées de M. Dans le tableau des valeurs, M correspond aux colonne et ligne suivantes :

$$C = (x - X_d)/D_x + 1$$

$$L = (y - Y_d)/D_y + 1$$

Mais  $C$  et  $L$  ne sont pas nécessairement des entiers : M n'est pas alors un point de la grille d'échantillonnage. Il faut donc interpoler. On déduit une colonne  $I_c$  et ligne  $I_l$  en prenant les parties entières de  $C$  et  $L$ , et on calcule les résidus :

$$D_c = C - I_c \quad \text{et} \quad D_l = L - I_l$$

Plusieurs interpolations sont possibles :

**Plus proche voisin** en prenant la partie entière de  $C + 0.5$  et de  $L + 0.5$ . On obtient ainsi une représentation en marches d'escalier. Pour une image à une dimension, la fonction image est restituée par une expression du type :

$$\tilde{f}(x) = \sum_{-\infty}^{+\infty} V(k)R(x - k)$$

où  $R(x)$  désigne la fonction porte, nulle à l'extérieur de l'intervalle  $[-\frac{1}{2}, \frac{1}{2}]$  et valant 1 à l'intérieur de cet intervalle. Les écarts entre  $f(x)$  et  $\tilde{f}(x)$  peuvent être importants si le signal contient des hautes fréquences.

**Linéaire** par :

$$I(x, y) = (1 - Dc)(1 - Dl)I1 + (1 - Dc)Dl.I2 + Dc(1 - Dl)I3 + Dc.Dl.I4$$

où  $I1, I2, I3, I4$  ont pour coordonnées  $(Ic, Il), (Ic, Il + 1), (Ic + 1, Il), (Ic + 1, Il + 1)$ . On évite les discontinuités obtenues avec l'interpolation précédente, mais on reste discontinue pour les dérivations partielles. A une dimension l'interpolation linéaire restitue une fonction qu'on peut exprimer sous la forme :

$$\tilde{f}(x) = \sum_{-\infty}^{+\infty} V(k)T(x - k)$$

où  $T(x)$  désigne la fonction triangle, nulle à l'extérieur de l'intervalle  $[-1, 1]$ , valant 1 en 0, et linéaire de part et d'autre de l'origine. La fonction triangle est l'autoconvolution de la fonction rectangle. Dans l'espace de Fourier on obtient  $\text{sinc}^2(\nu)$ , qui converge plus vite que  $\text{sinc}(\nu)$ . Les écarts entre  $f(x)$  et  $\tilde{f}(x)$  sont moins importants.

**du type Spline** l'interpolation est plus complexe, mais permet d'obtenir une fonction continue et dérivable. On utilise le plus souvent une approximation B-spline, telle que :

$$\tilde{f}(x) = \sum_{-\infty}^{+\infty} V(k)B_l(x - k)$$

où  $B_l(x)$  désigne la fonction ayant pour transformée de Fourier  $\text{sinc}^{l+1}(\nu)$ . Cette fonction est composée de morceaux de polynômes se raccordant sur les entiers sans discontinuités jusqu'à l'ordre  $l$  exclu. Malheureusement on n'obtient pas une vraie interpolation car  $f(k) \neq \tilde{f}(k)$ . Il est possible d'obtenir une interpolation en déterminant des  $\tilde{V}(k)$  qui résolve une équation de convolution discrète. On obtient ainsi une interpolation spline, qui donne une excellente restitution de la fonction, mais au prix d'un temps de calcul important.

**Pseudo-spline** On peut obtenir sur 4 points successifs une interpolation continue et dérivable, avec des morceaux de polynômes de degré 3. Les calculs sont aisés à mettre en œuvre.

**l'interpolation de Shannon** Elle est très rarement utilisée en 2D. Par contre, elle est parfois employée en 1D, pour des problèmes très précis (doublement du nombre d'échantillons par exemple).

**Interpolation dyadique** Il arrive souvent, par exemple pour des zooms, que l'on ait besoin d'avoir besoin d'une interpolation au milieu d'un segment. La méthode naturelle consiste à prendre la demi-somme. Mais on peut raffiner en prenant des termes plus lointains. Par exemple pour 4 points la meilleure combinaison est (Deslaurier-Dubuc) :

$$V(k + \frac{1}{2}) = -\frac{1}{16}(V(k-1) + V(k+2)) + \frac{9}{16}(V(k) + V(k+1))$$

Plus le nombre de points est élevé plus l'interpolation donne d'excellents résultats au fur et à mesure que l'on divise l'intervalle en deux.

Nous venons de voir comment on peut estimer en un point la valeur de l'image. Nous n'avons qu'à peine mentionné l'interpolation de Shannon alors qu'elle est à la base de l'échantillonnage. Les interpolations linéaire ou spline ne sont pas associées au théorème d'échantillonnage. Leur choix vient de la facilité de calcul. Mais il ne faut pas non plus considérer comme un acte de foi le théorème de Shannon, car en fait il est très rare qu'on remplisse les bonnes conditions :

- L'image est de dimension finie. Cela entraîne l'existence de hautes fréquences. Il y a toujours des effets de bord, réductibles par apodisation.
- La fréquence de coupure peut être si élevée qu'on doit obtenir un trop grand nombre d'échantillons. Divisé par deux le pas correspond en 2D à multiplier par quatre les pixels et souvent par huit les temps de calcul. On recherche donc à limiter les temps de traitements.
- La fonction  $\sin x/x$  converge trop lentement pour qu'on puisse utiliser l'interpolation de Shannon sur une image de petite dimension.



## Chapitre 2

# Généralités sur l'analyse des images

### 2.1 Objectifs de l'Analyse des images

#### 2.1.1 Analyse des images et information

Analyser une image peut avoir un sens très différent selon l'état de connaissance du domaine. Il est bien connu que deux personnes différentes ne perçoivent pas les mêmes choses dans une photographie ou pis encore dans un tableau. Il n'y a pas de raison *a priori* qu'il n'en soit pas de même du point de vue numérique. La machine n'analysera une image qu'à partir d'une stratégie définie par son utilisateur.

L'analyse d'une image, comme de tout ensemble de données, consiste d'une manière ou une autre à indiquer dans quel état se trouve un système associé à l'image. Par exemple, le système considéré est celui consistant dans l'ensemble des valeurs en chaque pixel. S'il y a  $N$  pixels et que la valeur est codée sur  $a$  bits, le nombre d'états possibles est  $2^{Na}$ . On dit aussi que sa capacité d'information est de  $aN$  bits. C'est aussi le nombre de bits qu'il faut pour pouvoir stocker l'ensemble des valeurs de l'image, indication de l'état dans lequel se trouve le système.

On peut associer d'autres systèmes descriptifs à une image, cela dépend essentiellement de l'état de connaissance *a priori* sur l'image que l'on a. Par exemple, en Astronomie l'image d'une zone céleste peut n'être formée que par un ensemble de taches ponctuelles. Seuls les positions et flux émis par les astres sont porteurs d'information. La quantité d'information est alors beaucoup plus faible que dans l'exemple précédent.

Le choix des méthodes employées pour l'analyse de l'image est conduit par le principe suivant :

*On doit toujours choisir la méthode d'analyse conduisant au minimum de quantité d'information recueillie, tout en restant compatible avec les données enregistrées, et en rendant compte de la totalité de l'information reçue.*

La capacité d'information n'est qu'une borne supérieure. L'analyse doit, grâce à une connaissance apportée par l'utilisateur, réduire de manière importante cette valeur.

#### 2.1.2 Analyse des images en Astronomie.

Les astronomes ont des objectifs qui s'expriment en termes plus concrets. La plupart des analyses se situent dans le cadre de grands relevés de l'Univers pour lesquels des milliards de pixels vont être enregistrés et analysés. L'astronome ne peut pas faire une étude très précise de

toute l'information reçue, il veut en obtenir, aussi rapidement que possible la quintessence.

Généralement il s'agit d'identifier, de mesurer et de classer les sources astronomiques. Grâce à des logiciels spécialisés l'astronome constitue des catalogues de sources, avec leurs positions sur le ciel et leur flux exprimé en magnitude. Il cherchera au moins à séparer les étoiles des galaxies, ou de tout autre type d'objet nébulaire.

La simple mise sur catalogue n'est pas un objectif scientifique très ambitieux. Tout astronome souhaite faire des découvertes. En premier lieu il peut s'agir de détecter des sources nouvelles. Cela va de soi lorsqu'on arrive à observer avec un niveau de détection plus faible que les observateurs précédents. Mais, on peut détecter des sources nouvelles en appliquant une nouvelle approche d'analyse des images.

Il est encore plus excitant de découvrir une nouvelle classe d'objet. Cela n'est généralement pas possible en se basant uniquement sur une image, il faut combiner des informations provenant de différents télescopes à des longueurs d'onde variées pour comprendre qu'on est en présence d'un objet différent des sources connues.

L'analyse d'un lot d'images permet la détection et la mesure des variations de position et de magnitude des objets. Cela peut conduire à des découvertes importantes, astéroïdes géocroiseurs, étoile proche de faible luminosité, supernova lointaine, ...

### 2.1.3 Possibilités offertes par l'analyse des images.

Avec la constitution de grandes archives d'images accessibles par Internet, il est possible d'interroger ces bases de manière à extraire les objets recherchés.

Un des problèmes intéressants réside dans la possibilité de paramétrer automatiquement l'image de manière à pouvoir facilement l'indexer. Ceci est bien sûr d'une très grande utilité pour les images destinées au grand public, mais il y a des retombées d'une grande utilité pour la recherche pure compte tenu du nombre croissant d'informations disponibles.

Il est clair que la capacité de paramétrer des images pour les indexer ouvre la porte à la possibilité d'interprétation automatique des images. Il ne s'agit pas simplement de reconnaître les objets mais aussi de comprendre les liens qui peuvent exister entre toutes les composantes de la scène.

À ce jour, seuls les astrophysiciens peuvent donner l'interprétation d'une image du ciel. Les mécanismes de vision artificielle sont encore trop rudimentaires pour être appliqués à l'analyse d'une scène astronomique.

## 2.2 Les différents types d'analyse.

On considère plusieurs types d'analyse, parmi lesquelles :

**L'analyse interactive** qui transfère à l'astronome le soin de décider des quantités à évaluer, la machine n'ayant qu'un rôle de visualisation et d'estimation des quantités recherchées ;

**L'analyse globale** qui consiste à calculer sur l'image des quantités globales, soit statistiques, soit liées à une transformation de l'image. Elle conduit à l'analyse texturale ;

**L'analyse déterministe** dans laquelle l'image est réduite en divers objets, pour lesquels on détermine pour chacun d'entre eux une suite définie de paramètres ;

**La vision par ordinateur** la machine détecte, mesure, classe, reconnaît à partir d'une stratégie pré-définie ;

**L'analyse morphologique** Les objets étudiés ont des formes qui peuvent être décrites par des outils variés. Ceux issus de la Morphologie Mathématique constitue un domaine important de l'analyse des images ;

**L'analyse multi-échelle** l'image est examinée à différentes échelles. Les variations de l'une à l'autre permettent d'identifier les objets de taille différente.

**L'analyse multi-images** l'analyse tient compte de l'ensemble des images conduisant à classer ces images, mais aussi les pixels, à étudier des variations temporelles ou des déplacements.





## Chapitre 3

# L'analyse interactive

### 3.1 Objectifs.

Nous verrons qu'il est difficile d'obtenir une analyse complète des images sans l'aide de l'astronome. Son expérience, ses connaissances, ses objectifs sont indispensables pour donner un sens aux traitements à réaliser.

Dans de nombreuses expériences, le flot d'images arrive sur le terminal de l'équipe responsable du projet (*Principal Investigator*). Les chercheurs doivent rapidement identifier les éléments nouveaux et porteurs de découverte. Pour cela ils effectuent un *Quick look* sur les images, ne retenant ensuite qu'une faible fraction de l'information reçue.

Les astronomes demandent donc à la machine de les aider au mieux dans la tâche de sélection, celle-ci pouvant souvent être la seule phase d'analyse des images.

### 3.2 Matériel et logiciels associés.

#### 3.2.1 Le matériel.

L'analyse interactive des images nécessite une station de travail aussi rapide que possible. Pour qu'il y ait interaction il faut que le temps d'exécution des tâches soient du même ordre de grandeur que l'écriture des paramètres des commandes. Aujourd'hui, les PC permettent d'avoir cette puissance, même pour des images de dimension  $2000 \times 2000$ . Pour les dimensions plus grandes, le chercheur est limité par la taille des écrans de visualisation. Des stations plus rapides, comme les stations Alpha, permettent une meilleure interaction.

Pour accélérer on fait appel de plus en plus au parallélisme. Soit plusieurs processeurs sont associés à l'unité centrale, soit plusieurs machines partagent le traitement des données, des logiciels spécialisés effectuant plus ou moins efficacement ce partage.

La visualisation des images s'effectuent le plus souvent directement sur l'écran de la station. La carte graphique varie selon le nombre de points d'écran visualisables et le nombre de bits par pixels. Aujourd'hui on trouve sur des PC à des prix abordables des résolutions de  $1260 \times 1024$  sur 24bits permettant une visualisation d'excellente qualité.

Des auxiliaires graphiques existent : imprimantes et traceurs en couleur, tablettes de saisie graphique. Ce dernier périphérique permet une interaction d'excellente qualité avec les images.

En général, la souris de la station est le moyen privilégié d'interaction. Le traitement des images bidimensionnelles ne nécessitent pas de moyens plus puissants. Pour celles à trois di-

mensions des salles spécialisées permettent une visualisation optimale. Des lunettes spécifiques conduisent à une vision stéréoscopique très utile pour explorer ces images.

Les images numériques occupent une place considérable sur les supports de stockage. On peut les compresser, mais cela conduit à une perte sérieuse d'information. Il faut prévoir des disques de taille importante. L'archivage s'effectue plutôt sur des bandes magnétiques, comme les DAT ou les DLT. Le CD-ROM est idéal pour l'archivage de petites images, le DVD semble être le support le mieux adapté pour l'archivage des images de plus grande taille.

Aujourd'hui de nombreuses images sont stockées sur des supports accessibles via Internet. La capacité des réseaux en permet un transfert assez aisé.

### 3.2.2 Les logiciels généraux.

Quel que soit le système d'exploitation Unix/Linux, Windows ou MacOS, il existe des logiciels permettant de visualiser une image et d'effectuer un minimum de retouches. Certains disposent d'un nombre considérable d'outils de traitement et d'analyse. Les plus connus sont :

- Photoshop, de la société ADOBE. Il est considéré comme la Rolls des logiciels pour les images. Il est conçu pour le monde des imprimeurs et des photographes. Il accepte les images FITS. Son coût est assez élevé. Avec Illustrator, on peut intégrer des fonctions graphiques et de dessin.
- Paint Shop Pro, un excellent logiciel de visualisation et de traitement, dont le coût est inférieur au précédent
- CorelDraw, un intégré contenant les fonctions images et graphiques. Parfaitement adapté à la réalisation d'affiches.

Pour les systèmes Unix/Linux d'autres outils similaires existent :

- xv : c'est l'outil standard de visualisation, sans traitement important;
- Image\_Magick, un logiciel plus riche de traitement.

### 3.2.3 Les logiciels complets.

Plusieurs compagnies ont développé des logiciels permettant le traitement complet des images grâce à un langage de commande adapté :

- MATLAB : c'est devenu l'outil de développement le plus populaire dans les milieux du traitement de signal et des images, et chez les mathématiciens. Des boîtes à outils permettent de très nombreuses fonctions de traitement et d'analyse.
- Mathematica : ce n'est pas vraiment un concurrent. Il s'agit plutôt d'un outil de calcul formel permettant de visualiser une image.
- IDL : développé d'abord par des astronomes, c'est devenu un outil de traitement et d'analyse très populaire dans le monde médical. La NASA diffuse des outils spécialisés pour les fichiers FITS.
- KHOROS : au départ il s'agissait d'une plate-forme diffusée gratuitement, aujourd'hui c'est le logiciel commercial probablement le plus riche. De nombreux outils existent pour la télédétection.

L'existence de ces logiciels rend inutile le développement de programmes faisant les fonctions qui y sont contenues. Le prix d'achat est rapidement amorti par leur usage.

### 3.2.4 Les logiciels spécialisés.

Les astronomes ont construit des logiciels spécialisés à l'analyse interactive des images. Il y en a plutôt à profusion :

- MIDAS : c'est le logiciel de l'ESO. D'une structure ancienne, il n'est plus maintenu. Il n'existe que sur Unix/Linux.
- IRAF : c'est le logiciel diffusé par le NRAO américain et utilisé par le STScI. Il n'existe que sous Unix/Linux. C'est la référence en matière de traitement des images. La partie interactive n'est pas la mieux conçue. Pour cela on préfère en général saoiimage. NRAO diffuse aussi fitsview pour les stations Windows.

Les astronomes anglais diffusent, via l'organisation STARLINK, de très nombreux logiciels. A chaque projet spatial est en général attaché un logiciel d'exploitation, comme xanadu pour les images X.

Des amateurs diffusent aussi des logiciels adaptés au traitement des images astronomiques (QMIPS32, Pise Atlas par exemple).

## 3.3 L'interface et son contenu.

### 3.3.1 Commandes ou actions graphiques.

Les systèmes anciens de traitement des images étaient basés sur le mode de commandes. L'utilisateur passait une partie notable de son temps à entrer les commandes ou à répondre à la machine. Avec le développement des interfaces graphiques le concept des *Graphic User Interface (GUI)* a pris corps. Aujourd'hui les grands logiciels sont basés sur ce concept.

Le mode commande avait l'avantage de permettre facilement le traitement par lot, supprimant l'interaction avec l'utilisateur. L'utilisation des GUI peut lasser son utilisateur quand il s'agit de traiter des centaines d'images.

Un mouchard est souvent créé dans le cas des logiciels spécialisés. Toutes les opérations sont inscrites de manière à pouvoir comprendre la nature des résultats ultérieurement.

## 3.4 Commandes principales.

Un logiciel interactif comprend en général les outils suivants :

- Des commandes relatives aux fichiers à ouvrir et à sauvegarder. En astronomie, la gestion des descripteurs FITS des images est très utile.
- Des commandes permettant d'extraire et de copier une zone de l'image.
- Des commandes d'agrandissement de l'image, avec ou sans interpolation.
- Des commandes d'amélioration du contraste, que nous examinerons ultérieurement.
- Des commandes de réduction du bruit et d'amélioration de la résolution.
- Une saisie de valeurs en des points marqués avec la souris

Certains logiciels permettent de définir des objets avec la souris et d'en déterminer divers paramètres.

### 3.5 Débruitage.

Lorsqu'on mesure en un point  $M$  une valeur image  $I$ , même si l'on sait qu'il s'agit d'une valeur bruitée, il faut, pour la modifier, admettre que les autres mesures, voisines de préférence, contiennent aussi une information sur la valeur réelle de l'image au point  $M$ . C'est sur cette hypothèse que repose toutes les méthodes de réduction.

#### 3.5.1 Filtre de Shannon

Nous avons vu que l'image reçue était, au bruit près, la convolution de l'objet par le profil instrumental  $P$ . Or  $P$  est à support borné en fréquence. Toutes les valeurs au delà de la fréquence de coupure ne peuvent être due qu'au bruit dans l'image. Il est donc facile de supprimer dans l'espace de Fourier cette contribution, en multipliant par une porte centrée sur 0 et de largeur  $2\nu_c$ . Il s'agit du filtrage de Shannon, dont l'effet est nul si l'échantillonnage a été correctement effectué.

Dans l'espace direct, ce filtrage correspond à une convolution avec une fonction en  $\sin x/x$  dont la convergence est très lente, conduisant donc à des effets de bord. C'est pourquoi on préfère utiliser des profils apodisés, c'est-à-dire qui ne présentent pas de discontinuité, jusqu'à un certain ordre de dérivation. La convergence est alors plus rapide et les effets de bords limités.

#### 3.5.2 Filtre de Wiener-Kolmogorov.

Le filtre de Shannon n'offre donc que peu d'intérêt pratique, par contre le filtre de Wiener-Kolmogorov est essentiel. Il se déduit des considérations suivantes :

- Bruit gaussien, avec une densité spectrale  $B(\nu_x, \nu_y)$ .
- Signal que l'on peut associer à un processus stochastique, lui aussi gaussien de densité spectrale  $S(\nu_x, \nu_y)$ .
- Moyenne locale stationnaire : le résultat en chaque point est une moyenne pondérée des mesures voisines. La pondération est la même sur toute l'image.
- Estimation des poids par moindres carrés : puisque le processus de bruit est gaussien.

Celui conduit au filtre  $W(\nu_x, \nu_y)$  suivant :

$$W(\nu_x, \nu_y) = S(\nu_x, \nu_y) / (S(\nu_x, \nu_y) + B(\nu_x, \nu_y))$$

Ce filtre tient donc compte à la fois des propriétés du signal, mais aussi de celles du bruit. Par exemple, si le bruit est négligeable devant le signal (basse fréquence)  $W = 1$ . On n'a donc pas d'altération. Par contre, si le signal est nul, ou négligeable devant le bruit, on a :  $W = 0$ . On supprime donc ce qui ne contient plus d'information. Compte tenu de la fonction de transfert optique, le filtre de Wiener ressemble généralement à un filtre de Shannon apodisé.

#### 3.5.3 Lissages.

Au lieu de se placer dans l'espace de Fourier, on peut chercher une réduction du bruit par lissage, en ne considérant que des valeurs voisines. Plusieurs masques sont utilisées :

**Moyenne glissante** on effectue la moyenne de  $(2n+1) \times (2n+1)$  pixels autour de chaque pixel.

On admet implicitement que les différences entre ces valeurs ne sont pas statistiquement significatives. On montre facilement que cela revient toujours à perdre en résolution.

**Lissage de Sheppard** on admet que localement l'image est représentée par un polynôme de degré  $k$ . Ce lissage donne d'excellents résultats.

### 3.5.4 Méthodes non linéaires.

Les méthodes précédentes sont stationnaires et linéaires. Si le calcul du filtre de Wiener dépend de l'image, l'opération de filtrage ou de lissage est elle-même linéaire. Il existe de nombreuses méthodes non linéaires ou/et non stationnaires employées couramment :

**Le filtre de la médiane** consiste dans la détermination en chaque point de la médiane des valeurs voisines. Il permet de supprimer le bruit *poivre et sel*, non gaussien.

**L'équilibrage du bruit** consiste dans l'intégration autour de chaque pixel dans des domaines de plus en plus vastes, jusqu'à obtention d'un flux supérieur à un seuil donné. Le résultat est le rapport du flux mesuré sur l'aire nécessaire. Cela est adapté aux images de comptages, avec un bruit poissonnien.

**Le filtrage-compression** consiste en 2 phases : compression de l'image par un facteur  $a$  puis réextension par interpolation. La méthode est linéaire, mais non stationnaire. Elle est très bien adaptée aux fortes réductions de bruit.

**Lissage dépendant du niveau** en chaque point on effectue un lissage avec un masque dépendant du niveau de l'image. Généralement le lissage est d'autant plus énergique que le signal est faible. Cette méthode est bien adaptée à tout les processus où le bruit dépend du signal.

**Méthodes associées à la déconvolution non linéaire** il suffit d'effectuer la convolution du résultat trouvé par le profil instrumental. C'est avec cette méthode que l'on obtient les résultats les plus intéressants.

Il existe de très nombreuses autres méthodes de réduction du bruit, mais les paragraphes précédents contiennent les plus employées.

## 3.6 Amélioration de la résolution.

Nous avons vu que la relation image s'écrit, dans le cas d'un système imageur linéaire et stationnaire :

$$I(x, y) = O(x, y) \otimes P(x, y) \quad (1)$$

Soit par transformation de Fourier :

$$\hat{I}(\nu_x, \nu_y) = \hat{O}(\nu_x, \nu_y) \times \hat{P}(\nu_x, \nu_y) \quad (2)$$

La connaissance, qui est le plus souvent possible, du profil instrumental  $P(x, y)$  peut conduire par solution de l'équation (1) (ou (2)), à obtenir la distribution de l'objet  $O$  à partir de celle de  $I$ . Dans l'espace de Fourier cela consiste à calculer  $\hat{O}$  par :

$$\hat{O} = \hat{I} / \hat{P} \quad (3)$$

Cette opération, appelée *déconvolution* est limitée par plusieurs effets :

- En raison de la diffraction, nécessairement  $\hat{P}$  est nul pour toute fréquence supérieure à celle de coupure. La division est alors impossible. L'information sur  $O$  n'est pas accessible par  $I$  pour ces fréquences.

- La mesure de  $I$  est bruitée. Il faut tenir compte du bruit, donc introduire un débruitage, par exemple par filtrage dans le cas stationnaire.
- La connaissance du profil n'est jamais parfaite, le champ est limité, il y a des défauts sur l'image, etc. Le problème posé ne correspond donc pas strictement à la réalité.

Malgré ces difficultés de très nombreuses méthodes ont été mises au point. Elles sont classées entre méthode linéaires et non linéaires.

**Inversion itérative de l'opérateur.** Si la solution n'est pas possible sous la forme (3), on montre que la méthode suivante, due à Van Cittert permet d'obtenir, en l'absence de bruit en quelques itérations :

$$O_n(x, y) = I(x, y) - O_{n-1}(x, y) \otimes P(x, y) + O_{n-1}$$

On pose  $O_0(x, y) = I(x, y)$ . La convergence s'effectue en 3-4 itérations, après quoi la solution devient de plus en plus bruitée. On montre facilement que l'itération converge effectivement vers la solution de l'équation (1) en l'absence de bruit et en l'absence de trous fréquentiels.

**Méthodes de gradient.** On peut stabiliser la solution par différentes manières. On aboutit par exemple à la méthode itérative du gradient à pas fixe :

$$O_n(x, y) = O_{n-1}(x, y) + aP(-x, -y) \otimes (I(x, y) - O_{n-1}(x, y) \otimes P(x, y))$$

On montre facilement que la limite est bien solution de (1). Le facteur  $a$  joue un grand rôle dans la vitesse de convergence. Il est facile de déterminer la valeur optimale de  $a$ .

**Déconvolution avec le filtrage de Wiener.** Le filtre de Wiener est :

$$W = S/(S + B)$$

Or la densité spectrale du signal image est égale à celle de l'objet multipliée par le carré du module  $M$  de la fonction de transfert. La déconvolution s'effectue par division par  $P$ . Or  $M$  est sûrement nul lorsque  $P$  l'est,  $W$  l'est aussi et le filtre résultant  $D$  l'est également. Il n'y a donc plus de problèmes pour les hautes fréquences spatiales. On simplifie souvent, en prenant aussi un profil symétrique, avec une fonction de transfert réelle, ce qui donne l'expression du filtre :

$$D = rP^*/(r|P|^2 + 1)$$

où  $r$  est le rapport Signal/Bruit, le signal étant relatif à l'énergie de l'objet et non de l'image.

**Déconvolution par convolution directe.** On montre qu'il est possible de construire des masques permettant une amélioration de la résolution. Ils résultent de la déconvolution d'une représentation locale, pour tout pixel, sur une base linéaire, avec un support limité et le critère des moindres carrés.

**Méthodes non linéaires.** Les méthodes précédentes étaient toutes linéaires. Elles peuvent améliorer la résolution mais présentent quelques défauts, dont celui de conduire à des valeurs négatives. Les méthodes non-linéaires utilisent le plus souvent cette contrainte de positivité.

La Méthode du Maximum d'Entropie (MEM) est basé sur le critère du maximum d'entropie. Les moindres carrés servent de critère de vraisemblance de la solution. On obtient un excellent résultat en prenant la restriction positive de  $O(x, y)$  dans la méthode du gradient.

De nombreuses méthodes ont été développées avec d'autres contraintes de régularisation.

**Ondelettes et bases optimales.** Depuis une vingtaine d'années, de nombreuses bases de décomposition d'un signal ou d'une image ont été introduites. Elles permettent d'obtenir une représentation telle que l'information est concentrée dans un minimum de paramètres. Il est donc naturel de les utiliser pour comprimer numériquement les images. Mais cette compacité de représentation permet aussi de réduire de manière optimal le bruit et par conséquent de déconvoluer efficacement les images.

La transformation en ondelettes est l'outil qui a été en premier exploité. D'autres bases, comme les paquets d'ondelettes ou les curvelets sont également très utilisés.

### 3.7 Amélioration du Contraste.

Le contraste d'une image numérique dépend des conditions de visualisation. Celles-ci sont souvent définies par une valeur minimum et une maximum, correspondant aux intensités minimum et maximum du moniteur (ou des couleurs extrêmes). En jouant sur ces valeurs on peut donc augmenter ou diminuer le contraste. Mais les consoles images permettent aussi de modifier la répartition des contrastes. Par exemple, une transformation logarithmique permet de les adoucir, alors qu'une transformation exponentielle conduit à obtenir des contrastes plus forts.

Ainsi par transformation de la valeur de l'image, on peut obtenir une image ayant, pour une intensité donnée le contraste souhaité. La limite provient du bruit (ou des niveaux de discrétisation). On ne peut plus, en effet, percevoir de variations entre des intensités dont la différence est plus faible que la valeur du bruit. Les transformations peuvent n'être que locales, tenant compte des variations d'un fond par exemple.

L'image stockée est visualisée sous la forme de demi-teintes, chaque niveau de gris étant liée à la valeur image. Le mode pseudo-couleur est très souvent employé, on associe à chaque valeur une couleur grâce à une table de correspondance entre une échelle de couleur et la valeur image (*Look Up Table* ou LUT).

Les LUTs sont soit définies une fois pour toutes et stockées dans une mémoire morte, soit construites de manière interactive ou par modélisation.

Le mode vraie couleur est utilisé pour visualiser simultanément plusieurs plans graphiques, jusqu'à un maximum de 3. A chaque plan, une couleur particulière est affectée.





# Chapitre 4

## Analyse globale de l'image

### 4.1 Principes de l'analyse globale.

L'analyse globale consiste dans la caractérisation d'une image simplement à l'aide de quelques paramètres déduits d'une analyse statistique des valeurs de l'image. Cette analyse n'a pas de sens de manière isolée. Elle suppose qu'on étudie un ensemble d'images  $\{V_k(i, j)\}$  et que sur cet ensemble on a défini une classification basée sur ces paramètres. Leur simple détermination conduit à indiquer l'état de l'image.

Ce mécanisme est très utile pour la sélection rapide des images à traiter, celles qui sont susceptibles de contenir une information importante. Nous ne cherchons pas à mesurer précisément mais à alerter. Les calculs doivent donc être simple et rapide, en accord avec l'objectif de sélection.

Il y a deux approches possibles du problème soit à partir la physique des caractères à détecter, soit selon l'apparence de l'image (texture). Pour cela les outils sont toujours à peu près les mêmes : histogrammes, transformations de Fourier, corrélation. La sélection du meilleur outil est effectué à partir de l'analyse physique ou de manière purement empirique.

La notion de texture a été développée pour mieux quantifier l'apparence. Il ne s'agit pas d'un outil mathématique, mais de critères basés sur différentes notions. La notion de texture s'applique plutôt localement que globalement. Dans une image il existe, en général, des zones de différentes textures.

### 4.2 Utilisation de l'histogramme.

L'histogramme est l'outil privilégié pour effectuer des sélections d'image. Si  $\{V_k(i, j)\}$  est la collection d'images étudiées, leurs histogrammes  $\{H_k(p)\}$  permet une sélection très efficace lorsqu'il reste stable pour les images à rejeter, alors qu'il présente des variations détectables pour celles que l'on veut retenir, ou inversement. L'histogramme n'est pas nécessairement utilisé dans son ensemble. On extrait, le plus souvent, des paramètres plus faciles à interpréter : la moyenne, la dispersion, des quantiles, des modes, des paramètres d'ajustement, etc.. On pourrait les déduire directement, sans calculer l'histogramme, mais la présentation de l'analyse avec l'histogramme permet une meilleure vue d'ensemble.

### 4.2.1 Détermination de l'histogramme.

L'histogramme consiste dans le nombre de pixels correspondant à un niveau donné. Pour le déterminer il faut définir :

**Une valeur minimum**  $V_i$  : si l'intensité d'un pixel est en dessous de  $V_i$  soit on élimine le pixel, soit on attribue cette valeur. Ceci permet d'éviter d'avoir des indices négatifs ;

**Une valeur maximum**  $V_m$  : on élimine les pixels ayant une intensité supérieure à  $V_m$ , ou on lui attribue cette valeur ;

**Un intervalle de comptage**  $D_v$  permettant de définir un nombre  $N$  d'intervalles tel que :

$$N = E\left(\frac{V_m - V_i}{D_v}\right) + 1$$

où  $E(x)$  désigne la partie entière de  $x$ .

Pour une valeur donnée  $V$  d'un pixel, on incrémente l'histogramme d'une unité, pour l'indice  $I$  donné par :

$$I = E\left(\frac{V - V_i}{D_v}\right) + 1$$

L'histogramme est donc un tableau facile à calculer. Le choix des seuils et du pas peut résulter des mesures. Il est facile de déterminer préalablement le minimum et le maximum des données. L'intervalle de comptage résulte de la dispersion des mesures, qu'on peut déterminer, sur un échantillon de l'image. Dans de nombreux systèmes, les données sont codées en entier sur 8 bits, soit sur 256 niveaux, ce qui est compatible avec une détermination sur tous les niveaux possibles.

Une autre méthode pour construire l'histogramme consiste dans un classement préalable pour ordre croissant des intensités. Cela conduit à un temps de calcul beaucoup plus élevé, même avec l'algorithme basé sur les arbres binaires. Néanmoins, cette méthode permet d'éviter de discrétiser les valeurs des intensités.

## 4.3 Paramètres fondamentaux.

### 4.3.1 Paramètres déduits des moments.

A partir de l'histogramme  $H(i)$  on calcule facilement les moments  $m_n$  des intensités par :

$$m_n = \sum_{i=1, N} i^n \times H(i)$$

En raison des seuils haut et bas, il se peut que ces moments soient légèrement différents de ceux que l'on déduirait directement des intensités. Le calcul est plus rapide que le calcul direct, puisqu'on effectue une fois pour tous les pixels de même niveau les produits avec les puissances de  $i$ .

A partir de ces moments, on déduit immédiatement le niveau moyen  $i_m$  :

$$i_m = \frac{m_1}{m_0}$$

La dispersion (ou écart-type) exprimée en niveau se calcule par :

$$\sigma_i^2 = \frac{m_2}{m_0} - i_m^2$$

En principe l'expression est positive, mais à cause des erreurs d'arrondi, il arrive souvent qu'on obtienne une valeur négative. On préfère alors calculer les moments centrés  $\mu_n$  :

$$\mu_n = \sum_{i=1,N} (i - i_m)^n \times \frac{H(i)}{m_0}$$

Ces moments ont été normalisés. La dispersion  $\sigma_i$  est une quantité jouant un rôle très important dans les traitements. Le moment centré d'ordre 3,  $\mu_3$ , est souvent réduit par :

$$\mu'_3 = \frac{\mu_3}{\sigma_i^3}$$

On réduit de même le moment d'ordre 4 de la manière suivante :

$$\mu'_4 = \frac{\mu_4}{\sigma_i^4} - 3$$

$\mu'_4$  est égal à 0 pour une distribution gaussienne. Si  $\mu'_4$  est positif la distribution est plus plate qu'une gaussienne, sinon elle est plus piquée.  $\mu'_4$  est appelé le cumulatif d'ordre 4 de la distribution.

La moyenne, l'écart-type, les moments centrés réduits d'ordre 3 et 4 contiennent une information très utile, mais souvent insuffisante pour sélectionner une image. Une des raisons provient de la non *robustesse* de ces quantités. La robustesse est une propriété statistique importante pour l'estimation d'une quantité. Elle est relative à la réponse de l'estimateur à la présence de données aberrantes. Tous les moments sont d'autant moins robustes que le degré est élevé. Pour pallier ce grave défaut, on a le choix entre des estimations avec rejet, ou à des statistiques d'ordre.

### 4.3.2 Estimation avec rejet.

Pour stabiliser l'estimation des moments, on rejette une partie des données. Il y a plusieurs stratégies possibles.

**Limitation à k-sigma** On calcule un premier ensemble  $(i_m, \sigma_i)$ . On élimine tous les éléments à l'extérieur de l'intervalle  $[i_m - k\sigma_i, i_m + k\sigma_i]$ . Ceci permet de recalculer un nouvel ensemble  $(i_m, \sigma_i)$ . On itère jusqu'à convergence.

Le résultat dépend de la valeur  $k$  choisi. Si  $k$  est très grand (10 par exemple), la deuxième estimation sera très peu différente de la première valeur. Par contre, si  $k$  est trop petit (2 par exemple), on rejette toujours des valeurs, ce qui conduit à des résultats aberrants. La convergence dépend de la loi de distribution. Dans le cas d'une loi de Gauss,  $k$  doit être au minimum de 3.

**Estimation à partir des statistiques d'ordre** On enlève systématiquement  $x$  pourcents des premières et des dernières données. Le calcul est plus rapide, puisqu'il n'y a pas d'itération. On choisit le plus souvent  $x = 10$ .

### 4.3.3 Statistiques d'ordre.

La médiane  $m_d$  est un des paramètres extraits de l'histogramme des plus utilisés. Son calcul est immédiat après formation de l'histogramme. On somme les valeurs de ce tableau, en commençant par le premier élément jusqu'à obtention de la moitié du nombre de pixels. La médiane est robuste, peu sensible aux valeurs aberrantes. C'est la meilleure quantité permettant de donner une indication de l'intensité *moyenne* d'une image. Sur une série d'images on peut ainsi contrôler le niveau moyen, afin de sélectionner celles correspondant à un niveau trop fort ou trop faible.

Les quartiles  $\{q_k, k = 1, 3\}$  permettent la division de l'histogramme en 4 intervalles contenant autant de pixels. Bien sûr,  $q_2$  est égale à  $m_d$ . Les deux autres quartiles donnent une idée de la dispersion des mesures et de la dissymétrie de l'histogramme, de manière plus robuste qu'avec les moments. On utilise aussi les déciles, permettant de partager l'histogramme en 10 parties contenant chacune autant de pixels. Nous avons vu que les déciles étaient utilisés pour obtenir une estimation plus robuste de la moyenne et de la dispersion. Dans le cas de la sélection des images, les déciles peuvent être très utiles.

On définit d'autres quantiles, comme les centiles, conduisant à des critères de plus sensible à la fluctuation statistique. Les intensités minimum et maximum sont très souvent utilisées dans l'analyse des images. Elles sont non seulement essentielles pour la visualisation et le calcul de l'histogramme, mais elles permettent aussi de facilement sélectionner des images.

## 4.4 Utilisation des modes.

### 4.4.1 Les modes de l'histogramme.

Lorsqu'on trace l'histogramme d'une image, on voit toujours des maxima, plus ou moins prononcés. Chaque maximum est appelé mode. La position du mode indique la zone dans laquelle il y a accumulation de pixels. Ceci peut être très différent de la moyenne, ou de la médiane. Par exemple, si nous avons une image comportant de l'écriture noire sur un papier blanc. L'histogramme présente deux modes, les positions de la moyenne et de la médiane sont intermédiaires, et sont donc moins porteurs d'information.

### 4.4.2 Recherche des modes.

Cette recherche est en principe très simple, puisqu'il s'agit de détecter les niveaux pour lesquels on a un nombre plus grand de pixels que pour les niveaux voisins. Avec ce principe, on détecte toujours un très grand nombre de maxima, dont la plupart ne sont pas significatifs. En outre, lorsqu'on fait varier l'intervalle de comptage, le nombre et la position des modes varient de manière très significative. On doit donc procéder différemment, pour obtenir des résultats exploitables.

### 4.4.3 Utilisation du lissage.

La méthode la plus classique pour détecter les modes consiste dans le lissage préalable. En général, on effectue une moyenne glissante, avec un masque simple (gaussien, constant, etc.). On ajuste le lissage de manière à voir disparaître les fluctuations dépourvues de sens. On localise facilement les maxima résiduels.

La procédure précédente, utilisée assez systématiquement, est totalement impropre du point de vue statistique. En effet, le signal correspond à un comptage et obéit donc à une statistique de Bernoulli, qu'on peut simplifier en statistique de Poisson. L'application d'un lissage est justifiée pour un processus gaussien. Des biais s'introduisent dans le cas d'un processus de Poisson. Ces biais sont souvent du second ordre par rapport à la fluctuation et souvent négligés.

Pour être rigoureux il faut donner une représentation analytique la loi de probabilité associée à l'histogramme. Les modes, entre autres, se déduisent alors immédiatement.

#### 4.4.4 Somme de gaussiennes.

Une des méthodes d'ajustement les plus riches consiste dans l'approximation par une somme de distributions gaussiennes. Sur le plan formel, le problème est simple à écrire :

$$H(i) = \sum_j a_j G\left(\frac{i - m_j}{\sigma_j}\right)$$

où  $G$  est la fonction de Gauss  $\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ . Le problème le plus délicat réside dans la première valeur des différents  $(a_j, m_j, \sigma_j)$ . Des algorithmes basés sur l'analyse des nuées permettent d'obtenir une bonne approximation.

La somme de gaussiennes est facile à comprendre du point de vue physique : chaque gaussienne correspond à une population statistique. Du point de vue de l'image, on obtient ainsi une excellente information sur le contenu, sans avoir effectué l'analyse complète.

## 4.5 Seuils et informations extraites.

### 4.5.1 Cas d'un histogramme monomode.

Il arrive souvent que l'histogramme ne présente qu'un mode significatif. Dans ce cas nous avons sur un fond, typiquement gaussien, un éclaircissement moyen, typiquement exponentiel. L'histogramme peut être ajusté par la convolution d'une loi de Gauss par une loi exponentielle. Ceci permet d'obtenir les 3 paramètres suivants :

- $f$  : intensité du fond ;
- $\sigma$  : écart de l'intensité ;
- $a$  : paramètre de la loi exponentielle.

Une des premières applications de cet ajustement est le calcul du seuil de détection  $s$  :

$$s = f + k.\sigma$$

Le coefficient  $k$  est typiquement de 3 - 5.

Ce calcul permet de déduire de l'histogramme le flux émis au dessus du fond, soit à partir d'une intensité donnée, soit pour l'ensemble des pixels.

La détermination de  $f$  et  $\sigma$  permet d'obtenir l'histogramme corrigé. Il faut, en principe, effectuer la déconvolution de l'histogramme avec la gaussienne associée au bruit. On peut aussi procéder en cherchant à estimer pour chaque niveau le nombre de pixels appartenant au fond.

### 4.5.2 Cas d'un histogramme à deux modes.

Dans ce cas, le seuil le plus important concerne la séparation entre les deux classes d'éléments. Le plus souvent, on prend pour valeur le minimum entre ces modes.

Dans le cas de distributions gaussiennes  $p_1$  et  $p_2$ , on peut analyser plus finement le problème en examinant les probabilités. Soit  $i_m$  est l'intensité mesurée, on peut estimer la probabilité  $P_1$  pour que  $i \geq i_0$ , pour le premier mode, et  $P_2$  pour que  $i \leq i_0$ , pour le second mode. Le seuil est calculé par la valeur  $i_s$  pour laquelle  $P_1 = P_2$ . On calcule alors facilement le nombre de pixels et l'intensité correspondant à chaque mode.

### 4.5.3 Cas général.

Lorsqu'on effectue une décomposition en modes de l'histogramme il est facile ensuite de déterminer d'une part les seuils permettant de localiser les pixels correspondant à chacun, et d'autre part les propriétés globales de ces pixels (nombres de pixels, intensités correspondantes).

## 4.6 Histogrammes de fonctions dérivées.

Dans les paragraphes précédents nous avons analysé l'utilisation de l'histogramme des intensités. Une grande partie de ces méthodes sont applicables à d'autres variables que l'intensité. L'intérêt est parfois académique, mais souvent réel.

### 4.6.1 Utilisation des dérivées premières.

On calcule parfois les histogrammes des différences en  $x$  ou  $y$ . Cela permet de déterminer des seuils pour détecter des pixels de variation trop rapide. On préfère plutôt l'histogramme des valeurs absolues des différences en  $x$  et  $y$ , ou de leur somme. Cela conduit à l'estimation des pixels anormaux. Le calcul du gradient algébrique  $g(i, j)$  s'effectue par :

$$g(i, j) = \sqrt{(I(i, j) - I(i - 1, j))^2 + (I(i, j) - I(i, j - 1))^2}$$

L'histogramme du gradient permet d'obtenir la meilleure estimation de l'importance des contrastes dans l'image. Le calcul est plus long que précédemment, mais on gagne en rigueur. On considère parfois l'histogramme des gradients relatifs  $g/I$ . Cela permet de tenir compte des variations locales de luminosité.

### 4.6.2 Histogrammes de fonctions de second ordre.

C'est surtout le laplacien  $L$  d'intensité qui est utilisé. On le calcule par :

$$L(i, j) = \begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix}$$

On déduit ainsi un seuil de détection des pixels ayant un Laplacien significativement négatif (pic) ou positif (trou). On peut sélectionner ainsi, sans détection des objets, les images riches en objets, ou en trous. Le calcul du Laplacien s'effectue parfois avec d'autres masques.

## 4.7 Histogrammes doubles, Cooccurrence.

### 4.7.1 Histogramme de fonctions locales.

L'étude de l'histogramme de la mesure simultanée de deux quantités associées à chaque pixel permet souvent de donner une information globale sur l'image très importante. On peut, par exemple, étudier la distribution simultanée de l'intensité et du gradient. On délimite ainsi une région de pixels de fort contraste et de faible intensité. Ceci permet la sélection d'images basée sur la présence de trous ou de taches.

### 4.7.2 Matrice de Cooccurrence.

D'autres fonctions dérivées de l'intensité sont utilisables, mais les études les plus importantes en ce domaine sont basées sur l'histogramme des intensités en 2 points distants d'un vecteur donné  $\vec{v}$ . On appelle matrice de cooccurrence  $S(i, j, \vec{v})$  la loi de distribution associée.  $i$  et  $j$  sont deux niveaux d'intensité donnés.

L'établissement de la matrice de cooccurrence est aussi simple que celle de l'histogramme. En général on rediscrétise de manière à avoir un nombre de niveaux moyens (32 – 64), en harmonie avec le nombre de pixels à traiter.

La transformation des niveaux peut se faire de manière linéaire, ou bien en tenant compte de la loi de distribution des intensités. On peut convenir ainsi de choisir une discrétisation conduisant à un nombre similaire de pixels dans chaque intervalle.

### 4.7.3 Choix du vecteur.

La matrice dépend du vecteur  $\vec{v}$  considéré. Si l'échantillonnage est très serré, il existe une très grande relation entre les intensités  $i$  et  $j$ , pour un vecteur de norme unitaire. Aucune information utile ne peut être extraite, en dehors de celle contenue dans l'histogramme de l'intensité. Si l'on prend un vecteur trop grand, il y a totale indépendance entre les intensités. Aucune information ne pourra aussi être extraite.

On comprend alors le lien entre le vecteur  $\vec{v}$  et la fonction d'autocorrélation. Si la distribution des intensités est un champ aléatoire homogène gaussien, toute l'information est effectivement contenue dans l'histogramme de l'intensité et la fonction d'autocorrélation.

Le choix de  $\vec{v}$  s'effectue à partir de l'étude de la fonction d'autocorrélation. Plusieurs vecteurs de longueur et d'orientation différentes sont utilisées.

### 4.7.4 Paramètres extraits de la matrice de cooccurrence.

On extrait plusieurs paramètres de cette matrice (selon Connors) :

#### L'inertie

$$I(\vec{v}) = \sum_{i,j} (i - j)^2 S(i, j, \vec{v})$$

#### La disymétrie

$$A(\vec{v}) = \sum_{i,j} (i - i_m + j - j_m)^3 S(i, j, \vec{v})$$

**La proéminence**

$$B(\vec{v}) = \sum_{i,j} (i - i_m + j - j_m)^4 S(i, j, \vec{v})$$

**L'homogénéité locale**

$$L(\vec{v}) = \sum_{i,j} \frac{1}{1 + (i - j)^2} S(i, j, \vec{v})$$

**L'énergie**

$$E(\vec{v}) = \sum_{i,j} S^2(i, j, \vec{v})$$

**L'entropie**

$$H(\vec{v}) = \sum_{i,j} -S(i, j, \vec{v}) \log S(i, j, \vec{v})$$

Avec :

$$i_m = \sum_{i,j} i S(i, j, \vec{v})$$

et

$$j_m = \sum_{i,j} j S(i, j, \vec{v})$$

Pour un vecteur  $\vec{v}$  donné, on obtient ainsi 6 paramètres. D'autres paramètres ont été définis. Ces paramètres appliquées à un ensemble d'images permettent de les classer selon leur apparence (texture). Cette classification nécessite un apprentissage préalable sur des échantillons.

## 4.8 Utilisation de la fonction d'autocorrélation.

**Estimation de la fonction d'autocorrélation.** Cette fonction  $C(\delta x, \delta y)$  se détermine aisément par sommation :

$$C(\delta x, \delta y) = \frac{\sum I(x, y) I(x + \delta x, y + \delta y)}{N(\delta x, \delta y)}$$

où la somme s'effectue sur tous les pixels pour lesquels il y a des couples possibles.  $N(\delta x, \delta y)$  correspond au nombre de couples.

**Longueur de corrélation.** Dans le cas d'un signal monodimensionnel la longueur de corrélation s'estime à partir de la mesure de  $\delta x$  correspondant à la moitié de la corrélation d'ordre 0.

Pour une image  $2D$ , on détermine une courbe de corrélation définie par la relation :

$$C(\delta x, \delta y) = 0.5C(0, 0)$$

Cette courbe peut être ajustée par une ellipse, indiquant :

**une direction de l'allongement** correspondant à la direction du grand axe ;

**une valeur de l'allongement** rapport du grand axe sur le petit axe ;

**une longueur de corrélation** valeur du grand axe.



**Autres informations extraites.** Les pics de la fonction d'autocorrélation donnent une information sur l'existence d'une répétition de motif. En principe, à un pic primaire, on obtient une infinité de pics secondaires.

En fait, en dehors de la mesure de la longueur de corrélation, la fonction d'autocorrélation est moins facile à exploiter que sa duale par la transformation de Fourier, la densité spectrale de l'énergie.

## 4.9 Utilisation de la Densité spectrale de l'énergie.

**Calcul de la densité spectrale de l'énergie.** On utilise bien sûr la transformée de Fourier discrète, conduisant aux parties réelle  $a(\nu_x, \nu_y)$  et imaginaire  $b(\nu_x, \nu_y)$ . La densité spectrale est :

$$D(\nu_x, \nu_y) = a^2(\nu_x, \nu_y) + b^2(\nu_x, \nu_y)$$

Rappelons que cette densité spectrale est associée à la fonction d'autocorrélation cyclique, par la transformation de Fourier discrète.

**Fréquence de coupure.** Il s'agit de la fréquence pour laquelle :

$$D(\nu) = 0 \quad \text{si} \quad \nu \geq \nu_c$$

Dans le cas d'une image à deux dimensions, on peut avoir non pas une fréquence de coupure, mais un contour délimitant le domaine où la fonction n'est pas nulle.

**Fréquence caractéristique.** On introduit la courbe définie par :

$$D(\nu_x, \nu_y) = 0.5D(0, 0)$$

L'ajustement par une ellipse permet d'obtenir une direction, une dimension et un allongement caractéristiques.

**Pic central.** Le pic central peut être analysé à partir des quantités précédentes, en déterminant un profil radial dans des couronnes elliptiques. D'une part ce profil permet d'obtenir une information caractéristique, d'autre part cela permet, par soustraction de faire apparaître les pics mineurs, associés à l'existence de fréquences privilégiées.

Pour chaque pic on extrait :

- L'énergie contenue dans le pic
- Le laplacien pour le centre du pic
- L'aire du pic
- La distance du pic à l'origine
- L'orientation du pic

On peut extraire d'autres paramètres sur les pics, ou inter-pics (angle entre eux).

## 4.10 Champ de Markov.

Dans les années 70, la notion de champ de Markov a été introduite afin de modéliser les images ayant pour caractéristique des zones homogènes séparées par des discontinuités (bords des objets). L'idée repose sur la dépendance de la valeur d'un pixel avec les seuls pixels voisins.

Plusieurs modèles ont été introduits pour représenter différents types d'image. Les paramètres associés peuvent être exploités pour identifier la nature d'une image.

En astronomie, les champs de Markov ne sont utilisés que pour l'analyse des surfaces planétaires, pour lesquelles ce modèle peut être pertinent. Par contre, il est inadapté pour modéliser les images de galaxies ou de nébuleuse gazeuse pour lesquelles les dépendances à longue distance sont importantes.

## 4.11 Analyse de texture.

L'analyse de texture permet d'informer sur l'apparence d'une image. Cette analyse utilise les critères que nous avons définies dans ce qui précède. Le choix d'un paramètre particulier pour indiquer la texture d'une image ne résulte pas, en général, d'une analyse théorique de l'image.

La méthode générale consiste dans un apprentissage sur un échantillon d'images, ou de sous-images. Une analyse des nuées, ou mieux encore, une analyse discriminante linéaire, permet de déduire les fonctions permettant de caractériser au mieux la texture examinée.

Cette méthode n'est pas la seule employée. Nous verrons que d'autres outils, basés sur la segmentation de l'image, permettent d'obtenir aussi une bonne description de l'apparence d'une image.

# Chapitre 5

## Analyse paramétrique de l'image

### 5.1 Modélisation de l'image.

Il arrive souvent que l'on puisse modéliser complètement ou partiellement l'image de travail. Cela provient par exemple :

**D'une analyse de la scène complète** cette analyse introduit différents objets dont la forme est connue, mais pour lesquels la position, le niveau d'éclairement, le niveau de fond et éventuellement l'échelle sont à déterminer ;

**De l'étude morphologique d'un objet présent** on cherche à représenter d'une manière analytique un objet particulier présent dans l'image. La forme de cet objet dépend de plusieurs paramètres. La manière dont l'objet est éclairé est aussi pris en compte ;

**D'une interprétation physique** complète de l'image d'un objet. L'image observée dépend de quelques paramètres physiques.

L'analyse paramétrique consiste donc d'une manière générale à la recherche de la valeur des paramètres permettant d'ajuster l'image au modèle.

### 5.2 Méthodes d'estimation.

Il n'existe pas de critère parfait d'estimation des paramètres. Plusieurs estimateurs sont applicables. Leur emploi dépend de la statistique de bruit.

#### 5.2.1 Estimateur des Moindres Carrés.

C'est l'estimateur naturel pour un processus gaussien. Soit  $V(i, j)$  l'intensité de l'image au pixel  $(x, y)$ .  $F(i, j, a_k)$  correspond à la modélisation. Le principe énonce que le meilleur choix des  $a_k$  est tel :

$$R = \sum_{i,j} (V(i, j) - F(i, j, a_k))^2$$

doit être maximum.

On pondère souvent la sommation pour tenir compte d'une importance variable des différents pixels pour la modélisation. Nous verrons l'intérêt de cette pondération pour tenir compte de

l'influence du voisinage. Nous avons alors :

$$R = \sum_{i,j} p(i,j)(V(i,j) - F(i,j, a_k))^2$$

doit être maximum.

La sommation n'est pas nécessairement complètement faite dans l'ensemble de l'image. Nous verrons quelles sont les différentes stratégies de limitation de la fenêtre.

### 5.2.2 Principe du maximum de vraisemblance.

Dans le cas d'un processus de Poisson (comptage de photons) on ne peut utiliser les moindres carrés sans avoir des biais importants. On applique alors le principe du maximum de vraisemblance :

$$V = \sum_{i,j} V(i,j) \log F(i,j, a_k) - F(i,j, a_k)$$

doit être maximum.

Les résultats sont similaires à ceux obtenus avec les moindres carrés avec une pondération proportionnelle à  $1/V(i,j)$ .

### 5.2.3 Autres méthodes.

Il arrive souvent qu'on utilise d'autres méthodes, moins rigoureuses sur le plan statistique, mais conduisant à des calculs plus simples.

Par exemple, on emploie souvent les moments pour l'analyse des images. Nous verrons dans le prochain chapitre qu'ils sont faciles à déterminer simultanément pour un ensemble d'objets.

### 5.2.4 Intervalle de confiance.

L'estimation conduit toujours à un résultat. Pour pouvoir comparer les valeurs trouvées à d'autres mesures, il faut connaître l'erreur sur la mesure effectuée. L'intervalle de confiance est défini à partir d'un seuil d'incertitude  $\epsilon$ . Si  $\epsilon = 0,05$ , l'intervalle de confiance est de 0,95. A l'aide des lois de probabilités, on estime les valeurs  $a_{\min}$  et  $a_{\max}$  entre lesquelles on a une probabilité 0,95 de trouver un résultat de l'ajustement  $a$ .

L'estimation de l'intervalle de confiance est très difficile à réaliser. On simplifie souvent avec des lois normales.

## 5.3 Tests des modèles.

Le test des modèles est indispensable pour assurer que l'on a pas réalisé une hypothèse non fondée.

### 5.3.1 Cas gaussien.

La méthode générale consiste dans la distribution du résidu  $R$  si l'hypothèse est vraie.  $R$  doit suivre une loi de Pearson à  $k$  degrés de liberté.  $k$  est égal au nombre de pixels utilisés, moins le nombre de paramètres utilisés dans le modèle.

En utilisant les tables du  $\chi^2$ , il est facile de déduire à partir du résidu, réduit de la variance du bruit, la probabilité pour que le modèle corresponde bien aux données.

### 5.3.2 Test du $X^2$ .

Dans le cas d'un processus de Poisson, le test du  $\chi^2$  défini ci-dessus n'est pas correct, en raison de la dépendance de la variance du bruit avec le signal. On pondère le  $\chi^2$  en en tenant compte :

$$X^2 = \sum_{i,j} \frac{(V - F)^2}{F}$$

On appelle aussi souvent  $\chi^2$  ce test.

Ce n'est pas le seul test disponible dans le cas d'un processus de Poisson. D'autres méthodes existent. Elles conduisent à peu près aux mêmes résultats que le  $X^2$ .

Le  $X^2$  ne suit pas tout à fait une loi de Pearson. Dans le cas d'un processus de Poisson, son espérance est de  $k - 1$ , où  $k$  est défini de la même manière que ci-dessus.

Sa variance  $v$  est plus complexe :

$$v = 2k + \sum_{i,j} \frac{1}{F(i, j, a_k)}$$

## 5.4 L'utilisation du rejet.

Lorsqu'on effectue une estimation, on doit définir une stratégie de choix des pixels concernés. Une fois l'estimation effectuée nous disposons des résidus individuels

$$r(i, j) = V(i, j) - F(i, j, a_k)$$

Dans l'hypothèse où notre modèle est correct, les  $r$  doivent être distribués selon la loi du processus statistique concerné. Admettons qu'il s'agisse d'un processus gaussien.

Il est facile alors d'estimer les paramètres de la distribution des résidus  $r_m$  et  $\sigma_r$ .  $r_m$  doit être, en général, voisin de 0.  $\sigma_r$  est proche de l'écart-type de la loi de distribution des  $V$ .

On compare chaque  $r(i, j)$  à  $\sigma_r$ . La loi de Gauss conduit à une probabilité  $p(x)$  pour que  $y \geq x$ , où  $y$  est la mesure effectuée. Pour que  $|r(i, j)| \geq 3 \times \sigma_r$ , cette probabilité est de 0,005, donc extrêmement faible. Si on constate qu'on trouve plusieurs résidus satisfaisant à cette condition on peut considérer qu'il y a une anomalie. On a le choix entre 2 stratégies :

**Eliminer le modèle pour insuffisance** cela veut dire qu'on a une confiance absolue dans la statistique des mesures et que tout écart ne peut être dû qu'à une impossibilité d'ajustement avec le modèle concerné.

**Rejeter les pixels concernés de l'ajustement** Pour une raison plus ou moins bien connue, le processus statistique n'est pas normal en ce point (défaut), à moins que ce ne soit une insuffisance notoire du modèle (influence de l'environnement par exemple).

Le rejet des pixels incriminés conduit à une nouvelle estimation de l'écart-type des résidus, plus faible que la valeur précédente. On peut donc itérer pour rejeter de nouveaux pixels.

Il ne faut pas, bien sûr, que l'on aboutisse à la suppression progressive de tous les pixels. Avec un rejet à  $3\sigma_r$ , on se stabilise assez vite. Ce n'est pas le cas d'un rejet à  $2\sigma_r$ , où on peut converger vers un seul pixel restant. Le coefficient de rejet joue un rôle important. Il est le plus souvent pris entre 3 et 5. Au delà, on n'élimine plus que les gros défauts.

## 5.5 Choix des pixels.

L'ajustement ne s'effectue presque jamais sur l'ensemble complet de l'image. Si celle-ci est très riche, cela signifie un très grand nombre de paramètres simultanés à déterminer, alors qu'en procédant localement, l'image peut être modélisée par beaucoup moins de paramètres.

Cela introduit deux aspects importants de l'ajustement : le choix des pixels pour un ajustement donné et l'influence du voisinage sur une estimation. Nous examinerons ultérieurement ce second point.

**Fenêtre constante.** Il s'agit de la manière la plus simple de procéder. Les objets composant l'image ont une taille voisine. On extrait autour de chacun d'eux une fenêtre constante. L'ajustement s'effectue donc ensuite dans le cadre d'une modélisation d'une image complète.

**Délimination par un niveau d'intensité.** On estime que les éléments de trop faible intensité ne sont pas porteurs d'information sur l'objet. On choisit un niveau d'éclairement  $V_s$  permettant de délimiter les pixels. L'opération d'extraction des pixels sera décrite au chapitre suivant. On obtient une table d'étiquette qui permet d'affecter à chaque pixel un domaine. L'ajustement dépend du seuil  $V_s$  choisit. Si ce seuil est très bas, des objets proches se rejoignent, s'il est trop haut on perd des pixels et la précision est plus faible.

**Utilisation du Laplacien.** Le Laplacien de l'intensité de l'image est un bon estimateur du bord des objets. Le fond est considéré positif et les pixels des objets ont un Laplacien négatif. La ligne de Laplacien nul délimite l'objet. Un calcul de seuil permet d'obtenir un niveau pour lequel le pixel ne peut plus appartenir au fond. Avec un algorithme d'étiquetage on délimite les pixels de chaque domaine où le Laplacien est négatif. L'ajustement s'effectue avec ces pixels. La différence profonde avec l'utilisation des pixels au dessus d'un seuil, c'est l'absence de détermination préalable du niveau du fond, nécessaire dans les ajustements avec délimitation des niveaux.

D'autres fonctions de bord permettent d'extraire les pixels d'ajustement.

Une fois choisie la méthode d'ajustement ainsi que les pixels concernés, l'analyse statistique peut être réalisée. Néanmoins, il est rare que le modèle puisse tenir compte de l'ensemble des objets présents. On doit tenir compte de l'environnement pour l'ajustement d'un objet donné. Pour cela il y a deux moyens, qu'on peut utiliser simultanément : introduire une pondération qui permette de mieux tenir compte de l'influence des voisins dans la détermination des paramètres, et procéder par itération en soustrayant, à chaque étape, l'image calculée par modélisation.

## 5.6 Modification de la pondération.

Le plus souvent l'analyse paramétrique concerne l'image d'un objet brillant sur un fond uniforme. Admettons que nous avons déterminé puis soustrait cette valeur du fond. L'ajustement conduit, dans le cas des moindres carrés pondérés à :

$$R = \sum_{i,j} p(i,j)(V(i,j) - F(i,j, a_k))^2$$

D'après ce qui précède,  $F$  est quasi nul à l'extérieur de la zone de l'objet. La présence proche d'un autre objet conduit à mesurer une intensité  $V$  très forte. L'ajustement conduit à modifier

les paramètres  $a_k$  de manière à en tenir compte. Non seulement cela fausse la détermination de ces valeurs, mais cela augmente fortement l'écart-type des résidus. On peut rejeter les pixels de résidu élevé, avec le procédé exposé précédemment. Cela permet d'éliminer l'influence de pixels d'intensité très éloignée de la prédiction du modèle. L'influence des autres objets reste néanmoins d'une manière plus discrète.

Nous avons vu différentes stratégies de délimitation des objets. La plupart d'entr'elles conduisent à supprimer des pixels qui pourraient être porteur d'une information, mais susceptible aussi de perturber la mesure. Une manière de tenir compte de tous les pixels choisis, en tenant compte de l'information qu'ils transmettent sur les paramètres  $a_k$  consiste dans la modification de la pondération. On choisit :

$$p(i, j) = F(i, j, a_k)$$

On doit donc procéder par itération, partant d'une pondération unitaire. A chaque étape on obtient un jeu de paramètres  $a_k$  et une nouvelle pondération. Ce sont les pixels pour lesquels  $F$  est le plus élevé qui contribue le plus à l'ajustement.

On pondère parfois avec d'autres fonctions (gradient, laplacien, etc.).

## 5.7 Correction de l'environnement.

On a  $N$  objets. L'image est représentée sous la forme :

$$V(i, j) = \sum_n F(i, j, a_{k,n})$$

Pour un objet donné, on a à faire un ajustement avec des pixels et une pondération déterminés. A une étape  $l$  du processus itératif, on dispose d'une série de paramètres  $a_{k,n}^{(l)}$ . Ceci permet de calculer pour un objet  $m$  l'influence des autres objets en calculant pour tous les pixels concernés pour l'ajustement de cet objet la somme :

$$B_m(i, j) = \sum_{n \neq m} F(i, j, a_{k,n}^{(l)})$$

On soustrait cette valeur et on effectue l'ajustement en pondérant éventuellement par la fonction  $F(i, j, a_{k,n}^{(l)})$ .

## 5.8 La modélisation restreinte.

Nous avons ainsi une procédure qui permet d'une part d'ajuster avec un modèle une partie de l'image, et d'autre part de tenir compte au mieux de l'influence des zones voisines. Cette procédure est générale, les paramètres  $a_k$  n'étant que des variables dans une expression analytique. Le plus souvent, les modélisations sont beaucoup plus physiques, les paramètres ayant alors une valeur sémantique claire.

Une modélisation très classique consiste à décrire l'image sous la forme :

$$V(i, j) = \sum_n a_n F\left(\frac{i - x_n}{c_n}, \frac{j - y_n}{c_n}\right) + b(i, j)$$

Dans cette expression  $b(i, j)$  correspond à la carte du fond sur lequel se superpose l'objet ;  $(x_n, y_n)$  sont les coordonnées du centre de l'objet  $n$ ,  $a_n$  en représente le facteur d'amplitude et  $c_n$

le facteur d'échelle. On pourrait aussi introduire une rotation  $\theta_n$  de l'objet  $n$ , mais cela alourdit l'écriture.

Le problème de l'analyse de l'image consiste dans l'estimation des paramètres  $\{x_n, y_n, a_n, c_n\}$  des différents objets. La carte du fond n'est pas, généralement, porteuse d'information, mais sa détermination est indispensable pour estimer les autres paramètres. Il y a une ambiguïté apparente dans l'expression ci-dessus, on pourrait écrire que les  $a_n$  sont nuls et que  $V(i, j) = b(i, j)$ , autrement dit tout peut être du fond. Il nous faut donc approfondir cette notion très importante.

## 5.9 La cartographie du fond.

### 5.9.1 Concept général.

L'intensité en un point d'une image est nécessairement une quantité positive. L'existence d'un fond signifie :

- Le signal est la somme des objets et d'un fond ;
- L'échelle de variation du fond est très grande par rapport à celle des objets ;
- Il existe des pixels appartenant aux objets d'intensité proche du fond

La recherche des points du fond et la cartographie résultante provient de ces constatations. Selon leur utilisation, on a des méthodes un peu différentes.

### 5.9.2 Cas d'un fond constant.

Pour bien comprendre les méthodes qui vont suivre, plaçons nous dans le cas où on admet que le fond est constant. Soit  $b$  sa valeur. Il s'agit d'une information globale sur l'image et nous avons vu au chapitre précédent l'utilisation de l'histogramme des intensités pour extraire cette quantité.

Du point de vue technique, on construit l'histogramme  $H(n)$  des niveaux de l'image. Si l'histogramme est multimode, on ne s'intéresse qu'au premier, car le signal doit être nécessairement positif. On détecte donc la position du premier mode significatif, le plus souvent il s'agit du mode de plus grande population. S'il n'est pas ainsi, cela veut dire que le fond est dans sa plus grande partie occulté par un très grand objet. Sa cartographie est donc quasi impossible.

La valeur du fond est ensuite déterminée. Plusieurs valeurs sont utilisées :

**La médiane** c'est une quantité facile à déterminer, mais cela suppose que la population du fond est beaucoup plus importante que celle des objets.

**La moyenne avec rejet** cela permet d'éliminer une grande partie des pixels polluant le mode du fond ;

**L'abscisse du mode** on utilise un petit lissage local, ou on effectue un ajustement autour du mode ;

**Le fond déterminé par ajustement** Le modèle le meilleur correspond à la convolution d'une gaussienne avec une distribution de Laplace. L'ajustement doit se faire dans la région du pic.

Outre la position du fond, on en déduit aussi l'écart-type des intensités.



### 5.9.3 Cartographie à partir de l'histogramme local.

Si on considère un point de l'image  $(i_0, j_0)$ , on extrait autour de ce point une sous-image  $V_0(i, j)$ . L'histogramme de cette sous-image contient une information locale pour le pixel  $(i_0, j_0)$ . En particulier, on peut déterminer le fond  $b_0$ .

Pour obtenir la carte complète, on peut donc :

1. définir une taille de fenêtre  $L$  ;
2. calculer l'histogramme dans cette fenêtre ;
3. en déduire le paramètre de fond point par point.

Cette procédure a quelques variantes selon le choix du paramètre extrait de l'histogramme, et la méthode de cartographie, car souvent on ne calcule pas tous les fonds point par point, mais plutôt bloc par bloc, on se ramène à la carte complète par une interpolation spline.

Le choix de l'échelle est guidé par d'une part l'échelle supposée des variations de fond, d'autre part la taille des objets physiques.

### 5.9.4 Recherche des points de fond.

La méthode précédente ne contient pas la recherche explicite des points du fond. Dans une certaine mesure, on peut même dire qu'il n'est pas nécessaire qu'il en existe réellement, le fond n'étant qu'un paramètre de la modélisation. Une autre voie consiste dans la représentation numérique du fond par ajustement sur des points reconnus comme appartenant au fond.

L'idée est simple. On admet que le fond  $b(i, j)$  est une fonction  $g(i, j, a_k)$ , typiquement un polynôme de faible degré. On ajuste donc les intensités de l'image  $V(i, j)$  avec cette fonction. Pour les points n'appartenant pas au fond, on a un très net écart, ce qui conduit à les rejeter. On itère le calcul jusqu'à convergence.

La carte ainsi obtenue correspond à un ajustement sur les pixels du plus grand mode. On peut accélérer le processus en sélectionnant localement avec l'histogramme local. Cette méthode permet d'obtenir des cartes très lisses, mais ne permet pas de restituer les variations rapides du fond. En outre, avec les polynômes on est vite limité en degré, si on veut éviter des grandes oscillations sans réalités physiques.

### 5.9.5 Recherche des points du fond et cicatrisation.

On pose un nouveau critère pour le fond : pour les pixels correspondants les gradients sont les plus faibles. La recherche des points est basée sur les contrastes locaux. On élimine les pixels de plus forts contrastes.

Pour obtenir la valeur du fond dans les points éliminés, on effectue une cicatrisation, c'est-à-dire qu'on remplace des bords vers le centre les intensités par une valeur interpolée linéairement à partir des pixels du contour. Le mécanisme de cicatrisation utilise l'algorithme d'érosion que nous verrons au prochain chapitre.

### 5.9.6 Ajustement fonctionnel.

Il est assez courant de déterminer la valeur du fond simultanément avec les paramètres de l'objet, soit par exemple :

$$V(i, j) = a_n f\left(\frac{i - x_n}{c_n}, \frac{j - y_n}{c_n}\right) + b_n$$

Le fond  $b_n$  est donc constant dans l'environnement de l'objet  $n$ .

## 5.10 Méthodes de détermination de la position.

### 5.10.1 Ajustement.

C'est l'application naturelle de ce qui précède. On pose, pour simplifier que :

$$V(i, j) = af(i - x_0, j - y_0)$$

La fonction  $f$  est connue à une translation près. Il peut s'agir d'une gaussienne, ou du profil instrumental, ou de la forme d'un objet à étudier. Les paramètres  $a$ ,  $x_0$  et  $y_0$  sont inconnus. L'application des moindres carrés avec une pondération  $p(i, j)$  conduisent à la minimisation de :

$$R = \sum_{i,j} p(i, j)(V(i, j) - af(i - x_0, j - y_0))^2$$

On élimine facilement le terme  $a$  en obtenant son équation normale associée :

$$\sum_{i,j} p(i, j)f(i - x_0, j - y_0)V(i, j) = a \sum_{i,j} p(i, j)f^2(i - x_0, j - y_0)$$

On peut tenir compte de cette relation pour redeterminer le résidu :

$$R = \sum_{i,j} p(V - af)^2 = \sum_{i,j} pV(V - af) - a \sum_{i,j} p.f(V - af)$$

Le second terme est nul en raison de l'équation normale, ce qui conduit au résidu :

$$R = \sum_{i,j} pV^2 - a \sum_{i,j} pVf$$

soit :

$$R = \sum_{i,j} pV^2 - \frac{(\sum_{i,j} pVf)^2}{\sum_{i,j} pf^2}$$

La minimisation du résidu est équivalent à la maximisation du second terme. Au terme de pondération près on reconnaît un terme de corrélation. Les équations normales donnant les positions s'écrivent :

$$\frac{\partial R}{\partial x_0} = -2 \frac{\Sigma_1 \Sigma_1^x}{S_2} + 2 \frac{\Sigma_1^2 S_1^x}{S_2^2} = 0 \quad (5.1)$$

$$\frac{\partial R}{\partial y_0} = -2 \frac{\Sigma_1 \Sigma_1^y}{S_2} + 2 \frac{\Sigma_1^2 S_1^y}{S_2^2} = 0 \quad (5.2)$$

où :

$$\Sigma_1 = \sum_{i,j} pVf \quad (5.3)$$

$$\Sigma_1^x = \sum_{i,j} pV \frac{\partial f}{\partial x} \quad (5.4)$$

$$\Sigma_1^y = \sum_{i,j} pV \frac{\partial f}{\partial y} \quad (5.5)$$

$$S_2 = \sum_{i,j} pf^2 \quad (5.6)$$

$$S_1^x = \sum_{i,j} pf \frac{\partial f}{\partial x} \quad (5.7)$$

$$S_1^y = \sum_{i,j} pf \frac{\partial f}{\partial y} \quad (5.8)$$

En remarquant que :

$$a = \frac{\Sigma_1}{S_2}$$

on obtient :

$$\frac{\partial R}{\partial x_0} = -2a(\Sigma_1^x - aS_1^x) \quad (5.9)$$

$$\frac{\partial R}{\partial y_0} = -2a(\Sigma_1^y - aS_1^y) \quad (5.10)$$

On élimine le facteur  $-2a$  dans les équations. Ce qui conduit à deux équations normales assez simples :

$$\Sigma_1^x - aS_1^x = 0 \quad (5.11)$$

$$\Sigma_1^y - aS_1^y = 0 \quad (5.12)$$

En procédant de manière à ajuster sa fenêtre de travail de manière symétrique, les termes  $S_1^x$  et  $S_1^y$  sont nuls dans le cas où la fonction  $f$  est elle-même symétrique. Ce qui conduit à une simple recherche de l'annulation des termes  $\Sigma_1^x$  et  $\Sigma_1^y$ .

La résolution complète s'effectue de manière itérative en différenciant les termes  $\Sigma_1^x$  et  $\Sigma_1^y$  par rapport à  $x_0$  et  $y_0$ . On se place toujours dans l'hypothèse où on peut supprimer les sommations des termes de dérivées premières pour des raisons de symétries. Ceci conduit à l'estimation suivante des variations de  $x_0$  et  $y_0$  :

$$\delta x_0 = \frac{\sum_{i,j} pV \frac{\partial f}{\partial x}}{\sum_{i,j} pV \frac{\partial^2 f}{\partial x^2}} \quad (5.13)$$

$$\delta y_0 = \frac{\sum_{i,j} pV \frac{\partial f}{\partial y}}{\sum_{i,j} pV \frac{\partial^2 f}{\partial y^2}} \quad (5.14)$$

La détermination des positions par ajustement conduit donc à plusieurs méthodes :

- Maximum de corrélation, pour minimiser le résidu ;
- Annulation d'un terme de corrélation pour satisfaire les équations normales ;
- Détermination itérative de la position pour résoudre ces équations.

### 5.10.2 Maximum de corrélation.

La procédure est simple, on doit corréler l'image avec le profil de l'objet  $f$ . Techniquement, on peut utiliser un lissage avec un masque représentant la fonction, ou un filtrage, en passant par la transformation de Fourier. Le filtrage est plus adéquat lorsque l'image est grande et que le profil est complexe, non réductible à un produit de deux fonctions en  $x$  et  $y$ .

Du fait de l'échantillonnage, la recherche des maxima de corrélation s'effectue avec une résolution du pixel. On améliore la position avec une interpolation locale du second ou du troisième degré.

### 5.10.3 Position par le maximum.

La recherche de la position d'un objet pourrait s'effectuer en recherchant sur l'image les maxima d'intensité brute. Cette position est sensible au bruit. Un filtrage est nécessaire pour réduire ce bruit. Lorsque l'objet a un profil donné le meilleur lissage correspond à la corrélation croisée avec la forme. On retrouve ainsi l'utilisation du maximum de corrélation.

Si cherche à positionner un objet quelconque, il est évident que la position d'un maximum local ne peut correspondre qu'à un vague indication, l'objet pouvant d'ailleurs avoir plusieurs maxima. On doit procéder tout autrement.

### 5.10.4 Barycentre.

A l'aide d'une règle définie plus haut, nous avons délimité les pixels de l'objet. La position est définie par :

$$x_0 = \frac{\sum_{i,j} pV_i}{\sum_{i,j} pV} \quad (5.15)$$

$$y_0 = \frac{\sum_{i,j} pV_j}{\sum_{i,j} pV} \quad (5.16)$$

Le calcul dépend des pixels sélectionnés. La prise en considération du fond est aussi un facteur important.

Le barycentre peut être estimé avec d'autres fonctions que l'intensité de l'image. Dans le cas d'une analyse de scène, ce qui comptera beaucoup plus que l'éclairage, c'est l'appartenance à l'objet. La position ainsi définie est plus stable par rapport aux conditions d'éclairage.

Le barycentre peut être calculé avec les points du contour des objets.

### 5.10.5 Centre de symétrie.

Une autre manière, moins classique de définir la position d'un objet consiste dans la recherche d'un centre de symétrie.

Nous avons délimité les pixels appartenant à l'objet, et nous avons soustrait le fond éventuel. On considère le résidu  $S$  :

$$S(i_0, j_0) = \frac{\sum_{i,j} p(i,j) |V(i,j) - V(2i_0 - i, 2j_0 - j)|}{\sum_{i,j} p(i,j)}$$

On détermine par itération la position  $(i_0, j_0)$  pour laquelle  $S$  est minimum. La position peut être améliorée en utilisant une approximation locale polynômiale du second ou du troisième degré. Cette méthode dépend de plusieurs facteurs :

- la délimitation des pixels de l'objet ;
- la soustraction du fond ;
- la pondération  $p$  choisie ;
- la distance utilisée.

De très bon résultats peuvent être obtenus, même quand les objets n'ont pas réellement un centre de symétrie.

## 5.11 Calcul des flux.

**Ajustement.** Nous avons vu que si l'on connaît la forme  $f$  de l'objet, le facteur d'intensité  $a$  est estimé par :

$$a = \frac{\sum_{i,j} pVf}{pf^2}$$

Le calcul du flux  $\Phi$  associé à l'objet en résulte par intégration :

$$\Phi = a \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy$$

**Maximum de corrélation.** L'expression nous donnant  $a$  correspond à la corrélation de l'image avec le profil, avec éventuellement une pondération et un facteur constant de normalisation. Si la forme des objets à étudier est toujours la même, il suffit pour obtenir leur flux de connaître l'intensité aux maxima de corrélation. La procédure est donc très facile à mettre en oeuvre.

**Intégration directe.** Le calcul du flux peut s'effectuer de manière directe, sans ajustement, par sommation des intensités des pixels, soit :

$$\Phi = \sum_{i,j} V(i, j)$$

Le choix des pixels concernés, et la suppression du fond jouent les principaux rôles dans cette méthode. En particulier, l'intégration s'effectue le plus souvent dans des fenêtres de taille discrète, ce qui conduit à des erreurs systématiques dues à la géométrie. On compense cette erreur en introduisant une pondération sur les bords de la fenêtre (bords flous). Les résultats sont alors stables par rapport aux décentrement.

## 5.12 Estimation du paramètre d'échelle.

### 5.12.1 Ajustement.

Nous avons introduit dans la modélisation restreinte un paramètre d'échelle  $c_n$ . Sa détermination peut s'effectuer par l'ajustement. Le calcul simultané de la position et du facteur d'échelle conduit à des calculs un peu compliqué. Souvent on estime le facteur d'échelle après avoir déterminé la position. On en revient toujours à la maximisation de :

$$R' = \frac{\Sigma_1^2}{S_2}$$

où  $\Sigma_1$  et  $S_2$  ont les mêmes significations que ci-dessus. Compte-tenu que la somme  $S_2$  dépend de l'échelle on ne peut pas procéder par maximum de corrélation. On procède par itérations, à partir d'une méthode classique de recherche du maximum d'une fonction. L'équation normale associée laisse apparaître le terme  $-2a$  en facteur (comme pour la position). On en revient à annuler :

$$T = \sum_{i,j} (V - af) \left( (i - x_0) \frac{\partial f}{\partial x} + (j - y_0) \frac{\partial f}{\partial y} \right)$$

Pour obtenir une variation de  $c$  on doit différencier à nouveau. Mais la présence du terme  $af$  complique fortement. Contrairement au cas de l'estimation de la position on ne peut le négliger.

### 5.12.2 Estimation à partir des moments.

Si les pixels de l'objet sont délimités, il est facile d'en calculer l'échelle, grâce à la détermination de l'aire, simple nombre de pixels de l'objet. Les moments d'ordre 2 sont souvent utilisés pour estimer l'échelle. Si  $m_{xx}, m_{xy}, m_{yy}$ , sont ces moments réduits au barycentre déterminés avec les moments d'ordre 1, on en déduit les longueur des grand et petit axes ( $a, b$ ) et l'orientation  $\theta$  d'une ellipse ayant mêmes moments. Les moments sont calculés sans pondérer avec l'intensité.

## 5.13 Autres paramètres de forme.

Les méthodes d'ajustement permettent d'obtenir toujours les paramètres d'un modèle. Toutefois, il est rare que l'on dépasse le cas de la modélisation restreinte. En général, les paramètres de forme résultent plutôt d'autres approches :

- les moments des domaines associés aux objets ;
- des paramètres associés au contour ;
- des quantités plus ou moins empiriques sur des fonctions de l'image (moments du profil radial, résidu d'un ajustement, etc.).

L'étude de la morphologie des objets aura pour but de dégager des règles pour sélectionner ces paramètres.

# Chapitre 6

## La vision par ordinateur

### 6.1 La vision par une machine.

#### 6.1.1 Les éléments du problème.

L'analyse paramétrique d'une image n'a de sens que si on est capable de la décrire, au moins localement, sous une forme analytique. Ceci se produit couramment dans le cas d'images provenant d'expériences ou d'observations scientifiques. C'est d'ailleurs l'un des objectifs majeurs de la science de pouvoir modéliser avec assez de précision l'objet des recherches de manière à prédire les observations. Cette modélisation peut être possible dans le cas d'images moins techniques, mais conduit à un formalisme très lourd, donc difficilement utilisable.

C'est pourquoi on a construit des méthodes d'analyse de l'image se passant de la modélisation formelle, basées sur l'analyse des structures. Ces méthodes forment un domaine appelé souvent *vision par ordinateur*.

La notion de vision introduit l'idée d'une description de la scène correspondant à l'image. En particulier, l'analyse doit conduire à une reconnaissance des objets présents dans l'image. Au niveau supérieur, elle implique qu'on décrit leurs relations afin de comprendre la scène.

#### 6.1.2 Les objets.

L'analyse structurale est basée sur l'existence d'objets identifiables formant la scène à décrire. Cette hypothèse est triviale lorsqu'on examine une image de la vie courante. Notre cerveau cherche toujours à décomposer en éléments ce que nous voyons, mais souvent au prix d'un effort d'interprétation.

Si j'observe une table de bois, le plus souvent je ne m'intéresse pas aux détails des noeuds et des anneaux, ce que je retiens c'est l'existence à un certain endroit d'un volume que j'ai identifié à du bois. C'est essentiellement la fonction que je vais associer à l'objet détecté qui me fait penser à une table. La localisation dans la scène de la table m'indique ensuite son rôle.

Avec cet exemple, trois aspects importants de la vision artificielle sont abordés :

**La texture** On ne cherche pas à décomposer l'image de la table, on regroupe les éléments de l'image qui appartiennent au même objet car l'apparence est la même. C'est donc la texture qui est la caractéristique de l'appartenance.

**La connexité** j'ai localisé *un* volume. Un objet est toujours considéré comme bien individualisé. On doit être capable de faire un chemin dans l'espace entre deux molécules quelconques

d'un objet en ne passant que par des molécules lui appartenant. Cette propriété forme ce qu'on appelle en topologie mathématique la connexité.

**Reconnaître c'est interpréter** La reconnaissance de la table ne vient qu'à partir d'un petit raisonnement tenant compte de la scène. Sans cela, il peut y avoir ambiguïté.

Les phases de la vision sont donc claires : détection des objets, en se basant sur des critères texturaux, caractérisation des objets en vue de la reconnaissance, classification des objets, attribution des classes et interprétation.

Ce schéma peut subir quelques variantes, selon la nature des objets et la connaissance préalable de la scène.

### 6.1.3 Domaines et Objets.

Pour l'instant, nous ne traitons que des images à 2 dimensions, provenant le plus souvent de l'observation d'une scène à 3 dimensions. Dans cette projection, il y a une perte importante de l'information. En particulier, les objets, qui occupent dans l'espace une position bien spécifique, se retrouvent parfois superposés.

La connexité est une propriété se conservant par projection, la démonstration en est assez évidente lorsqu'on considère la définition. Mais les superpositions conduisent dans l'éclatement de l'objet en différents domaines de l'image. C'est le processus de vision qui conduit à regrouper, par l'interprétation, les différents morceaux.

L'analyse structurale ne peut donc se baser directement sur les objets, mais elle commence par l'identification des domaines pouvant former, seuls ou en groupes, les projections des objets formant la scène à reconnaître.

## 6.2 Appartenance d'un pixel à un domaine.

### 6.2.1 Aspect multiple de la vision.

Nous allons chercher à caractériser les pixels appartenant à un domaine de l'image. Nous allons voir que plusieurs voies sont possibles, c'est-à-dire qu'il n'existe pas une mais plusieurs visions. Il suffit pour s'en rendre compte de supprimer rapidement la couleur sur son récepteur de télévision, on voit immédiatement que la couleur joue un rôle capital dans l'analyse de la scène. L'analyse d'une image *Noir et Blanc* est nettement moins riche et souvent ambiguë.

Puisque nous procédons avec une machine, il n'est pas essentiel que le processus de vision suive notre comportement physiologique. Nous pouvons étendre les capacités d'analyse de la machine par l'introduction de nouveaux critères qui n'auraient pas de sens du point de vue physiologique.

### 6.2.2 Critère de couleur.

C'est le critère classique pour la recherche d'une vision proche de la vision humaine. Les aspects principaux de l'analyse multispectrale sont décrits dans le chapitre suivant, car l'affectation d'une couleur n'est possible que si l'on possède une image dans chacun des canaux Rouge, Vert et Bleu (pour avoir la terminologie de la télévision en couleur).

On peut attribuer à chaque pixel  $(i, j)$  une *couleur*  $c(i, j)$  correspondant à une position définie dans le diagramme des couleurs. Les 3 images sont donc réduites à une seule. Les différents



niveaux de couleur proviennent du découpage dans l'espace de couleur, généralement effectué une fois pour toutes.

Il existe deux types de découpage, selon que l'on tienne compte ou pas de la luminosité intrinsèque du pixel. Dans le premier cas, le découpage est à 3 dimensions, il est à 2 dimensions dans le cas contraire. L'oeil effectue un découpage à 3 dimensions.

### 6.2.3 Critère de contraste.

Dans notre vision, on utilise aussi l'existence des bords des objets, conduisant à un très grand gradient lorsqu'on passe d'un pixel d'un domaine d'un objet à celui correspondant à un autre objet. Les pixels de fort gradient donc constituent les lignes de contour des objets. Ces lignes devraient donc entourer les différents domaines de l'image.

Ce schéma n'est que très approximatif pour diverses raisons :

**Les contours des objets sont souvent flous** Les variations de luminosité peuvent être faibles, et la fonction d'étalement de la chaîne d'acquisition des images suffit à atténuer les bords, pour les rendre difficilement détectables.

**Les objets ne sont pas nécessairement de luminosité homogène** Il existe des gradients de luminosité qui peuvent être supérieurs à ceux des bords.

**L'existence d'un bruit peut conduire à des faux contours** Si le bruit de l'image est élevé on a des gradients locaux importants, dus uniquement aux fluctuations de la luminosité. Ils peuvent complètement noyer la perception des vrais contours.

L'utilisation concrète du critère de contraste s'appuie sur la détermination des lignes de bord. Nous en décrirons les principales méthodes de calcul ci-dessous.

### 6.2.4 Critère de luminosité.

L'observation du ciel étoilé la nuit est un bon exemple de la vision à partir de la luminosité. Les étoiles sont vues parce qu'elles ont une intensité lumineuse supérieure au seuil de détection de l'oeil. La différence de détection entre un ciel bien noir et celui d'une ville est en partie due au changement de mode de détection. Lorsque le ciel blanchit, le contraste diminue, de même que le seuil de détection.

Le critère de luminosité est utilisé dans tous les cas où il s'agit de voir des sources faibles sur un fond bruité.

### 6.2.5 Critère de texture.

L'analyse d'une image de paysage ne conduit pas à discerner dans le détail l'ensemble des objets, mais à séparer les zones de différente texture. On distingue ainsi les pelouses, les bois, etc.. Un critère spécifique caractérise les éléments correspondant au gazon, ceux liés aux arbres, etc.. L'analyse de texture n'est pas globale, mais locale, en décomposant l'image en sous-images.

L'application du critère de texture n'est possible que si l'on a, dans un premier temps, classé les différentes textures présentes dans l'image. L'extraction des sous-images de plusieurs variables texturales, comme celles décrites au chapitre 3, permet une analyse multivariée, conduisant à la classification des textures présentes. Une connaissance préalable de la nature des textures permet d'aller plus loin dans l'identification des textures. Le critère de texture s'emploie le plus souvent pour l'analyse d'une collection d'images.

La décomposition en sous-images pour déterminer la texture associée à un pixel, conduit à une perte en résolution. Il est assez difficile de déterminer ainsi des bords très nets. Pourtant, les objets texturés peuvent avoir dans la réalité des bords bien tranchés. On examinera l'utilisation de méthodes itératives pour la recherche des bords.

### 6.2.6 Critères du second ordre.

Avec le critère précédent on ne détecte pas nécessairement un domaine parce qu'il correspond à un objet bien individualisé, connexe. La pelouse est formée de brins d'herbe que l'on peut chacun individualiser, pourtant notre œil voit un ensemble bien défini associé au groupe des herbes.

On peut généraliser ce concept en introduisant des critères de *second ordre*. Dans une première analyse, basée sur un des critères ci-dessus, on détecte les différents domaines. Puis en utilisant des règles d'association des domaines (aire moyenne locale, orientation moyenne, corrélation entre type de domaine, etc.) on localise les différentes structures présentes.

Ce critère se situe entre celui associée à la texture, basée sur l'environnement et liée au pixel, et la reconnaissance de forme.

### 6.2.7 Autres critères.

Il est possible de créer d'autres modes de vision en se basant sur diverses lignes que l'on peut tracer sur la surface de l'image.

Les lignes de vallée joue un rôle important pour l'analyse des images formées de pics. Il s'agit des lignes de plus grandes pentes partant des cols de l'image en direction des fonds. Les lignes de crête, lignes de plus grande pente allant vers les sommets, appelées aussi lignes de bassin versant ou de partage des eaux, ne forment pas un ensemble très utile.

On peut définir aussi de manière interactive sur l'image les bords des domaines. On ne peut parler alors de vision artificielle. Enfin, on considère parfois dans l'étude d'un ensemble d'images les lignes de bord formées à partir d'une image de référence, qu'on transfère sur les autres images.

## 6.3 Bords et fonctions de bord.

### 6.3.1 Bords d'un tracé à une dimension.

Dans le cas d'une image à une dimension, la notion de bord d'un domaine est immédiate : il s'agit des points d'inflexion du profil. Dans le cas d'une fonction purement mathématique continue  $f(x)$ , l'estimation de la position des points d'inflexion s'obtient simplement par double dérivation. Mais s'il s'agit d'un signal numérique bruité et échantillonné, on ne peut se baser sur cette simple définition. La dérivation suppose une représentation locale analytique, et l'existence du bruit conduit à lisser.

Pour l'obtention des bords d'une image à 2 dimensions, on essaie de revenir au cas à une dimension, en choisissant la bonne orientation. On va donc retrouver les différentes méthodes présentées ci-dessous. Le choix de l'orientation complique bien sûr un peu.

### 6.3.2 Bords par les maxima des gradients.

**Les gradients de Prewitt.** On calcule les gradients en  $x$  et  $y$  avec les masques :

$$G_x = \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} \quad G_y = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{pmatrix}$$

L'intensité du gradient  $G$  est donné soit par  $\sqrt{G_x^2 + G_y^2}$  ( $L_2$ ) soit par  $|G_x| + |G_y|$  ( $L_1$ ). Les points de contour sont ceux pour lesquels cette intensité dépasse un seuil. Celui-ci est fixé par la distribution statistique de  $G$ .

**Les gradients cardinaux.** On détermine les variations dans chacune des 8 directions cardinales : Est, Nord-Est, Nord, Nord-Ouest, Ouest, Sud-Ouest, Sud et Sud-Est. Pour cela on procède par lissage avec 8 masques.

**Autres méthodes de calcul du gradient.** On peut évidemment avoir une approximation en prenant les différences des intensités des pixels successifs (Roberts) :

$$G_x = \begin{pmatrix} 0 & 0 \\ -1 & 1 \end{pmatrix} \quad G_y = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}$$

Le gradient  $G$  calculé dans  $L_1$  ou  $L_2$  est très sensible au bruit.

Sobel utilise les masques :

$$G_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \quad G_y = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix}$$

Ces masques permettent de lisser un peu le bruit par rapport à ceux de Prewitt.

### 6.3.3 Gradients et fonctions de bord.

Avec les méthodes de gradients, on détermine les pixels pour lesquels l'intensité subit un saut significatif. L'image ainsi obtenue n'est pas formée de lignes, mais de domaines de faible épaisseur. Il faut ensuite appliquer des algorithmes de retrécissement (Cf. ci-dessous) pour obtenir des lignes. Mais celles-ci peuvent ne pas se fermer, ce qui pose des problèmes considérables pour l'identification des domaines.

Il est préférable de déterminer directement les lignes de bord à partir d'un contour nul d'une fonction soit  $B(x) = 0$ . Les conditions ci-dessus sont alors respectées *de facto*. Le problème réside alors dans le choix de la meilleure fonction de bord.

**Laplacien de l'intensité.** A une dimension, les points d'inflexion s'obtiennent en annulant la dérivée seconde. De nombreux auteurs ont proposé l'utilisation du Laplacien de l'intensité à 2 dimensions :

$$L(x, y) = \frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2}$$

Les pics d'intensité ont des valeurs négatives, alors que les creux ont des valeurs positives. Mais la réciproque n'est pas nécessairement vraie. Dans le cas des objets très contrastés le laplacien donne toutefois des résultats très valables.

La détermination du laplacien s'effectue le plus souvent par convolution. 3 masques sont principalement utilisés (on calcule en fait l'opposé du laplacien) :

$$L_1 = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{pmatrix} \quad L_2 = \begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix} \quad L_3 = \begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix}$$

Le premier masque provient de l'application de la définition en remplaçant les dérivées par les différences. Le second est une approximation très employée, même si sa justification théorique n'est pas évidente. Le troisième provient de l'ajustement local polynômial de degré 2.

La meilleure estimation du Laplacien résulte de l'ajustement par pavés de l'intensité avec des fonctions splines. Une fois obtenue la matrice des intensités sur une maille éventuellement comprimée, on dérive dans chaque pavé les polynômes du 3ième degré en  $x$  et  $y$ .

On considère les lignes de Laplacien nul, ou on seuille à partir de la distribution statistique de cette quantité.

**Lignes de gradient maximum.** Les méthodes liés au gradient présentées ci-dessus sont toutes basées sur l'obtention d'une valeur supérieure à un seuil, or le bord à une dimension est lié à la maximisation de la pente, c'est-à-dire du gradient. Si on examine le problème à deux dimensions, il faut trouver les points correspondants aux maxima. Si  $M$  de coordonnées  $(x_0, y_0)$  est un point quelconque du plan, on trouve toujours une direction pour laquelle le gradient est maximum. La recherche de la maximisation du gradient doit se compléter par la définition de la direction pour laquelle le gradient est maximal. Il est raisonnable de choisir celle de la normale aux isoniveaux (isophotes), soit le long de la ligne de plus grande pente, c'est d'ailleurs celle du gradient vectoriel.

Les lignes de bord associées au gradient maximum sont donc celle pour lesquelles le gradient le long de la ligne de plus grande pente est maximum.

Considérons la ligne isophote passant par  $M$ . Il est préférable de changer de repère en choisissant les axes  $(X, Y)$  tangent et normal à l'isophote en  $M$ .  $F(x, y)$  désigne l'intensité au point  $(x, y)$ . Nous allons déterminer l'expression de  $F$  autour de  $M$  dans le coordonnées  $(X, Y)$ .

La différentielle totale de  $F$  au second ordre est :

$$dF = \left(\frac{\partial F}{\partial x}\right)_0 dx + \left(\frac{\partial F}{\partial y}\right)_0 dy + \frac{1}{2} \left[ \left(\frac{\partial^2 F}{\partial x^2}\right)_0 dx^2 + 2 \left(\frac{\partial^2 F}{\partial x \partial y}\right)_0 dx dy + \left(\frac{\partial^2 F}{\partial y^2}\right)_0 dy^2 \right]$$

La ligne isophote est définie localement par  $dF = 0$ . Ceci permet d'obtenir l'équation de la tangente :

$$\left(\frac{\partial F}{\partial x}\right)_0 dx + \left(\frac{\partial F}{\partial y}\right)_0 dy = 0$$

D'où la direction  $\tan \theta = -(\frac{\partial F}{\partial x})_0 / (\frac{\partial F}{\partial y})_0$ , ce qui permet de passer des coordonnées  $(x, y)$  aux  $(X, Y)$  :

$$X = (x - x_0) \cos \theta + (y - y_0) \sin \theta \quad (6.1)$$

$$Y = -(x - x_0) \sin \theta + (y - y_0) \cos \theta \quad (6.2)$$

et

$$x = (X - X_0) \cos \theta - (Y - Y_0) \sin \theta \quad (6.3)$$

$$y = (X - X_0) \sin \theta + (Y - Y_0) \cos \theta \quad (6.4)$$

On déduit l'ensemble des dérivées partielles de  $F$  par rapport aux  $X$  et  $Y$  :

$$\left(\frac{\partial F}{\partial X}\right) = 0 \quad (6.5)$$

$$\left(\frac{\partial F}{\partial Y}\right) = \sqrt{\left(\frac{\partial F}{\partial x}\right)_0^2 + \left(\frac{\partial F}{\partial y}\right)_0^2} \quad (6.6)$$

$$\left(\frac{\partial^2 F}{\partial X^2}\right) = \left(\frac{\partial^2 F}{\partial x^2}\right)_0 \cos^2 \theta + \left(\frac{\partial^2 F}{\partial y^2}\right)_0 \sin^2 \theta + 2\left(\frac{\partial^2 F}{\partial x \partial y}\right)_0 \cos \theta \sin \theta \quad (6.7)$$

$$\left(\frac{\partial^2 F}{\partial Y^2}\right) = \left(\frac{\partial^2 F}{\partial x^2}\right)_0 \sin^2 \theta + \left(\frac{\partial^2 F}{\partial y^2}\right)_0 \cos^2 \theta - 2\left(\frac{\partial^2 F}{\partial x \partial y}\right)_0 \cos \theta \sin \theta \quad (6.8)$$

$$\left(\frac{\partial^2 F}{\partial X \partial Y}\right) = \left[\left(\frac{\partial^2 F}{\partial x^2}\right)_0 - \left(\frac{\partial^2 F}{\partial y^2}\right)_0\right] \cos \theta \sin \theta + \left(\frac{\partial^2 F}{\partial x \partial y}\right)_0 (\cos^2 \theta - \sin^2 \theta) \quad (6.9)$$

En ce qui concerne les bords c'est la fonction  $B(x, y) = (\partial^2 F / \partial Y^2)$  qui nous intéresse. En effet, dire que le gradient est maximum le long de la plus grande pente est équivalent à avoir  $B = 0$ .

Cette fonction dépend des dérivées partielles secondes, comme le Laplacien, mais elle n'est pas linéaire, et ne peut donc se calculer par convolution. Néanmoins, de même que le Laplacien, elle est invariante par rotation.

Son calcul nécessite une bonne réduction du bruit. La meilleure manière d'estimer  $B(x, y)$  consiste dans l'utilisation d'une interpolation spline, après réduction ou non de la matrice de l'image. Le coût d'obtention (temps CPU, mémoires) est beaucoup plus important que pour le Laplacien.

Si on prend la forme assez caractéristique d'une gaussienne  $\exp(-(x^2 + y^2)/2\sigma^2)$ , alors que le Laplacien donne un bord pour :

$$x^2 + y^2 = 2\sigma^2$$

la fonction  $B(x, y)$  conduit à la courbe d'équation :

$$x^2 + y^2 = \sigma^2$$

C'est  $B(x, y)$  qui correspond à la position des points d'inflexion de la gaussienne unidimensionnelle, qui engendre par rotation la gaussienne 2D. Le Laplacien élargit le domaine d'une manière assez importante.

On peut également seuiller la fonction de bord, pour tenir compte des fluctuations de cette fonction.

#### 6.3.4 Points de bord déterminés de manière dyadique.

Le critère d'appartenance au bord déterminé localement est parfois insuffisant en raison du bruit, des faibles contrastes et de la texture. Rosenfeld a introduit une détermination dyadique

des points de bord. On détermine un gradient à partir d'une matrice  $2 \times 2$  de pixels successifs, puis d'une matrice  $4 \times 4$ , puis  $8 \times 8$ , etc.. Avec une telle procédure un point de bord non vu à la première étape, parce que trop peu contrasté, apparaît dans une phase suivante, s'il correspond à un changement significatif de la texture ou de la luminosité, associées à cette échelle.

La transformation en ondelettes a permis de généraliser cette procédure dyadique.

## 6.4 Appartenance à un domaine.

### 6.4.1 Critère basé sur la détection par fonction.

La première méthode pour indiquer l'appartenance d'un pixel à un domaine est basée sur la valeur d'une fonction  $G$  associée à l'image : intensité, laplacien, fonction de bord, fonction liée à la couleur, paramètre textural, etc.. Pour cela il faut déterminer les seuils haut et bas d'appartenance du pixel. C'est l'histogramme de  $G$  qui conduit, le plus souvent, à la détermination de ces seuils.

Nous avons examiné les différentes méthodes d'analyse de l'histogramme permettant la détermination des différentes populations de pixels. Dans le cas de la vision artificielle l'objectif réside dans la classification des pixels, de manière à identifier les domaines contenant une population homogène.

La procédure est donc simple :

- Choix de la fonction discriminante ;
- Calcul de cette fonction ;
- Etude de la distribution des valeurs. Cette étude s'effectue soit sur les données de l'image, soit sur une partie de ces données, soit sur des données venant d'image de même catégorie. On en déduit les seuils de discrimination des différentes classes de pixels.
- Affectation d'une classe à chaque pixel.

### 6.4.2 Critère basé sur la délimitation des domaines.

Nous avons examiné ci-dessus différentes méthodes de détermination des bords conduisant à délimiter les domaines par des lignes. On ne possède pas de fonction permettant de déduire l'appartenance d'un pixel à une classe.

Dans ce cas, l'affectation s'effectue de proche en proche, selon le principe de la connexité : deux points appartiennent au même domaine s'ils peuvent être reliés par une ligne à l'intérieur de ce domaine.

### 6.4.3 Appartenance floue.

L'étude de la distribution des fonctions d'appartenance et celle des lignes délimitant les domaines nous montrent que l'appartenance d'un point à un point domaine n'est pas toujours assurée : il existe une probabilité  $p$  pour que ce pixel satisfasse la condition d'appartenance, et  $q = 1 - p$  pour qu'il n'y appartienne pas.

L'appartenance peut être considérée comme floue, avec le sens de la théorie des sous-ensembles flous. Le calcul des  $p$  résulte de la distribution de la fonction servant de critère d'appartenance, ou bien de considérations plus complexes sur les bords.

#### 6.4.4 Relaxation dans la fonction d'appartenance.

Le calcul des  $p$  est d'abord local. Mais il est évident que l'on est tenté de tenir compte des probabilités voisines pour obtenir un critère plus sûr. Le lissage de la fonction conduit un peu à améliorer les probabilités, mais c'est souvent au prix d'une perte élevée en résolution. On préfère plutôt opérer de la manière itérative suivante :

1. On considère des poids  $w_{kl}$  caractéristiques des liens entre un pixel et celui situé à la distance vectorielle  $(k, l)$ .  $k$  et  $l$  varient dans l'intervalle  $(-n, +n)$ .
2. Soit  $p_0(i, j)$  et  $q_0(i, j)$  les probabilités d'appartenance ou non du pixel  $(i, j)$  déterminées initialement. On a bien sûr  $p_0 + q_0 = 1$ .
3. On effectue le calcul itératif :

$$p_m(i, j) = \sum_{k,l} p_{m-1}(i, j) w_{kl} \quad (6.10)$$

$$q_m(i, j) = \sum_{k,l} q_{m-1}(i, j) w_{kl} \quad (6.11)$$

4. On normalise les poids de manière à ce que la somme fasse 1 :

$$p_m(i, j) = \sum_{k,l} p_m(i, j) / (p_m(i, j) + q_m(i, j)) \quad (6.12)$$

$$q_m(i, j) = \sum_{k,l} q_m(i, j) / (p_m(i, j) + q_m(i, j)) \quad (6.13)$$

## 6.5 Segmentation de l'image.

### 6.5.1 Etiquetage des pixels d'une image.

Après avoir défini un critère d'appartenance d'un pixel à un domaine, la seconde étape de la vision artificielle consiste dans la constitution des domaines. Il s'agit d'affecter à chaque pixel un nombre (étiquette) caractéristique du domaine auquel il appartient.

Tous les pixels appartenant à un même domaine connexe doivent avoir la même étiquette. Ceci constitue la base de la segmentation de l'image. Il faut donc donner la même étiquette à tous les pixels qui peuvent être joints par une ligne de pixel ayant le critère d'appartenance. Tous les algorithmes de segmentation sont basés sur ce principe.

### 6.5.2 Algorithme associé à une lecture vidéo de l'image.

On considère une image numérique  $F(i, j)$ . On va la segmenter en procédant ligne par ligne, et colonne par colonne (lecture vidéo). Soit  $M(i, j)$  un point courant de l'image.  $G(i-1, j)$  est son voisin de gauche, et  $H(i, j-1)$  est celui d'en haut. Leurs étiquettes sont déterminées nécessairement avant celle de  $M$ .

On dit que  $M \in P$  si  $M$  a le critère d'appartenance au domaine. Dans le cas contraire  $M \in \bar{P}$ . Pour un pixel donné nous pouvons avoir l'un des 5 cas suivants :

**Non appartenance au domaine** Si  $M \in \bar{P}$ , le pixel  $M$  n'appartient pas à un domaine défini par le critère choisi. On lui affecte une étiquette nulle.

**Création d'un nouveau domaine** Si  $M \in P$ , mais  $(G, H) \in \overline{P}$ , aucun segment ne peut être joint avec les voisins haut et gauche, on doit créer un nouveau domaine pour  $M$ , c'est-à-dire que l'on donne à  $M$  une nouvelle étiquette. Cette création ne signifie pas que  $M$  n'est pas relié à un pixel déjà étiqueté. C'est le rôle de la fusion qui permet de tenir compte des chemins qui ne passent pas par  $G$  ou  $H$ .

**Continuation en ligne** Si  $(M, G) \in P$  et  $H \in \overline{P}$ , on peut faire un chemin de  $G$  à  $M$ , ils appartiennent donc au même domaine. On donne à  $M$  l'étiquette de  $G$ .

**Continuation en bas** Si  $(M, H) \in P$  et  $G \in \overline{P}$ , on peut faire un chemin de  $H$  à  $M$ . On donne à  $M$  l'étiquette de  $H$ .

**Fusion des domaines** Si  $(M, G, H) \in P$ , on devrait donner à la fois à  $M$  l'étiquette de  $G$  et celle de  $H$ . Ceci est bien sûr impossible si elles sont différentes. Cela veut dire qu'entre  $G$  et  $H$  on a pu trouver un chemin en passant par  $M$ . Les domaines correspondant sont fusionnés.

Pour gérer les fusions des domaines, on affecte à chacun d'entre eux à leur création un indice  $K$ . Cet indice est au départ égal au rang de création du domaine. Lors d'une fusion, on modifie les indices des domaines en leur affectant le plus petit indice associé.

Ainsi si on fusionne les domaines  $K_1$  et  $K_2$ , on examine les indices  $I(K_1)$  et  $I(K_2)$ . Si  $k$  est la valeur minimale, on pourrait affecter à  $I(K_1)$  et  $I(K_2)$  ces valeurs. Ceci conduirait très vite à des domaines connexes morcelés, parce que des informations de fusion sont perdues. Ces informations ne concernent pas uniquement les domaines  $K_1$  et  $K_2$ , mais aussi tous les domaines qui leur ont été fusionnés. Il faut rechercher *les ancêtres* des domaines, c'est-à-dire effectuer l'itération suivante :

$$J = I(K) \quad (1)$$

Si  $J \neq K$  on remplace  $K$  par  $J$  pour déterminer un nouvel indice  $J$ . L'ancêtre  $A$  est obtenu lorsque  $K = J$ .

Lorsqu'on a obtenu les ancêtres  $A_1$  et  $A_2$ , on leur affecte leur minimum comme nouvel indice. Tous les domaines ayant pour ancêtre  $A_1$  et  $A_2$  vont donc de ce fait se trouver fusionner.

Cet algorithme est facilement mis en oeuvre. On peut écrire ligne par ligne les valeurs intermédiaires. Lorsque le dernier pixel est lu, on relit cette table intermédiaire d'étiquettes. Si la valeur lue est 0, on la garde, dans le cas contraire, si  $K$  est la valeur lue on affecte au pixel l'ancêtre du domaine. On prend soin de donner à chaque ancêtre des indices consécutifs afin de ne pas avoir des vides inutiles dans la table d'étiquettes.

### 6.5.3 Variante vectorielle.

Avec l'algorithme précédent, on étiquette pixel par pixel, or sur une ligne on peut avoir un très grand nombre de pixels successifs  $\in P$ , ayant donc la même étiquette. On peut accélérer l'étiquetage en procédant vectoriellement.

On décompose la ligne en une suite de vecteurs  $V_k$  ayant deux composantes  $(d_k, f_k)$  indiquant le début et la fin des pixels  $\in P$ . On compare cette suite à la ligne précédente : si un vecteur  $W_l$  à une intersection commune avec  $V_k$ , ils doivent avoir même étiquette.

On retrouve le concept de la fusion des domaines, lorsque deux vecteurs  $W_l$  et  $W_m$  de la ligne précédente peuvent être liés à un même vecteur  $V_k$ . La gestion des fusions s'effectue de la même manière qu'avec l'algorithme pixel par pixel.



On gagne sur l'étiquetage, en groupant par blocs de pixel. Mais cela n'a d'intérêt que si l'image est formée de grandes plages de pixels  $\in P$ , dans le cas contraire, on doit effectuer trop d'intercomparaison ligne à ligne.

#### 6.5.4 Procédures par dilatations successives.

On lit l'image de manière vidéo. Lorsqu'on détecte un pixel  $\in P$  et n'ayant pas été étiqueté on crée une nouvelle étiquette  $K$ . On effectue une opération de dilation (voir ci-dessous pour plus de détails), c'est-à-dire qu'on ajoute autour de ce pixel ses 4 voisins qui  $\in P$ , formant un nouveau domaine  $D_1$ . Tous les points ont la même étiquette  $K$ .

On dilate  $D_1$ , en ajoutant tous les pixels  $\in P$  voisins d'un pixel de  $D_1$ . On obtient un domaine  $D_2$  connexe, pour lequel tous les pixels ont forcément la même étiquette.

On itère le processus de dilatation jusqu'à invariance, tous les voisins du domaine formé  $\in \bar{P}$ .

Ce processus est simple, mais nécessite une lecture répétée de l'image pour chaque dilatation. L'image doit être impérativement stockée en mémoire vive. Si les domaines sont petits, peu de dilatations sont nécessaires pour chaque domaine, ce qui permet un travail rapide.

Cet algorithme s'adapte assez bien aux machines à parallélisme massif, où chaque pixel est associé à un processeur.

De nombreux algorithmes ont été proposés pour l'étiquetage des domaines. Il s'agit souvent d'une variante des algorithmes précédents.

On utilise parfois les contours : on détecte les points de contour (voir ci-dessous). Ces points ont bien sûr même étiquette. On érode le domaine, on détecte le nouveau contour, et on affecte la même étiquette à ces pixels. Le domaine érodé n'est plus nécessairement connexe, ce qui complique l'algorithme.

L'utilisation d'une décomposition *Quadtree* de l'image, a conduit aussi à un algorithme procédant sur cette structure.

## 6.6 Caractères structuraux associés aux domaines.

### 6.6.1 Caractères accessibles pendant l'étiquetage.

L'intérêt essentiel de l'étiquetage réside dans l'identification des pixels appartenant à un même domaine connexe. Cette opération effectuée, il est facile d'extraire sur les domaines diverses quantités servant de base à la reconnaissance des domaines.

Les quantités les plus souvent considérées sont celles qui peuvent être extraites pendant l'étiquetage, essentiellement les moments et les valeurs extrêmes.

En effet, lorsqu'on affecte à un pixel une étiquette, il est facile soit d'ajouter dans un tableau associé à cette étiquette une quantité dépendant du pixel, ou de prendre la valeur minimale ou maximale de cette quantité avec celle préalablement déterminée pour cette étiquette. En fin d'étiquetage, on regroupe les morceaux attachés au même ancêtre en additionnant ou en prenant le minimum ou maximum des quantités affectées à chaque domaine ayant cet ancêtre. Il existe des petites variantes dans la procédure (ligne par ligne ou en fin de lecture).

### 6.6.2 Moments et moments réduits.

Soit  $F(i, j)$  l'intensité du pixel courant. Les moments sont les sommes sur les pixels du domaine d'une fonction de  $F$ ,  $i$ ,  $j$ , et des fonctions associées à l'intensité (dérivées, différences,

primitives).

Les moments primitifs sont ceux qui sont déterminés sans connaissance préalable sur le domaine. Par exemple, l'aire du domaine, moment de la quantité 1 sur les pixels du domaine, ne dépend que l'identification des pixels du domaine. Par contre, si  $(i_c, j_c)$  est le centre du domaine,  $\sigma$  une dimension caractéristique du domaine,  $G(i, j) = \exp(-(i - i_c)^2 + (j - j_c)^2 / 2\sigma^2)$ , les moments pondérés avec  $G(i, j)$  ne sont pas des moments primitifs. Ils ne peuvent pas non plus être déterminés à partir d'autres moments primitifs.

**Aire A.** C'est le moment obtenu avec un poids de 1.

**Moments d'ordre 1 en  $(i, j)$ .**

$$m_x = \sum_D i \quad (6.14)$$

$$m_y = \sum_D j \quad (6.15)$$

La sommation s'effectue bien sûr sur les éléments du domaine. On en déduit immédiatement les coordonnées pixel du centre du domaine :

$$i_c = m_x / A \quad (6.16)$$

$$j_c = m_y / A \quad (6.17)$$

**Aire équivalente, allongement, orientation.** On détermine les moments d'ordre 2 par :

$$m_{x^2} = \sum_D i^2 \quad (6.18)$$

$$m_{xy} = \sum_D ij \quad (6.19)$$

$$m_{y^2} = \sum_D j^2 \quad (6.20)$$

La connaissance du centre  $(i_c, j_c)$  et de l'aire  $A$  permet de déduire des moments d'ordre 2 centrés :

$$M_{x^2} = m_{x^2} - A \times i_c^2 \quad (6.21)$$

$$M_{xy} = m_{xy} - A \times i_c j_c \quad (6.22)$$

$$M_{y^2} = m_{y^2} - A \times j_c^2 \quad (6.23)$$

On n'utilise pas directement ces moments, mais on en déduit des quantités plus invariantes, en identifiant les moments réduits à ceux d'une ellipse. Celle-ci est caractérisée par ses axes  $a$  (grand) et  $b$  (petit), et l'orientation  $\theta$  de son grand axe.

On a les relations suivantes :

$$M_{x^2} = \frac{\pi}{4} ab (a^2 \cos^2 \theta + b^2 \sin^2 \theta) \quad (6.24)$$

$$M_{y^2} = \frac{\pi}{4} ab (a^2 \sin^2 \theta + b^2 \cos^2 \theta) \quad (6.25)$$

$$M_{xy} = -\frac{\pi}{4} ab \cos \theta \sin \theta (a^2 - b^2) \quad (6.26)$$

Ces relations permettent de déduire  $a$ ,  $b$ ,  $\theta$ . En particulier :

$$\tan 2\theta = \frac{2M_{xy}}{M_{y^2} - M_{x^2}}$$

Les valeurs des axes se calculent à partir des sommes et différences.

Ainsi à partir des moments d'ordre 2, on obtient une aire équivalente  $\pi ab$ , un allongement  $a/b$  et une orientation  $\theta$ . Ces éléments peuvent jouer un rôle dans la classification du domaine.

**Utilisation des moments d'ordre supérieur.** L'extraction des moments d'ordre 3, 4 ou supérieur conduit, après réduction, à des moments invariants assez caractéristiques de la forme du domaine.

La réduction tient compte de la position du centre  $(i_c, j_c)$ , de l'orientation  $\theta$  et de la dimension du grand axe, servant d'échelle du domaine. Dans une première phase on réduit de la position du centre en tenant compte des relations :

$$M_{x^\alpha y^\beta} = \sum_D (i - i_c)^\alpha (j - j_c)^\beta$$

Le développement des puissances conduit à une expression dépendant des  $m_{x^n y^m}$ . La réduction de l'orientation s'effectue en tenant compte des relations de transformations des axes entre le repère de base et celui des axes de l'ellipse définie par les moments d'ordre 3 :

$$X = x \cos \theta + y \sin \theta \quad (6.27)$$

$$Y = -x \sin \theta + y \cos \theta \quad (6.28)$$

On déduit aisément les moments  $M_{X^\alpha Y^\beta}$ , en développant des expressions du type :

$$M_{X^\alpha Y^\beta} = \sum_D (x \cos \theta + y \sin \theta)^\alpha (-x \sin \theta + y \cos \theta)^\beta$$

On remplace symboliquement dans le calcul les sommes  $\sum_D x^n y^m$  par les moments  $M_{x^n y^m}$ .

La réduction de l'échelle est plus immédiate. Elle ne consiste qu'à diviser les moments obtenus  $M_{x^\alpha y^\beta}$  par  $a^{\alpha+\beta}$ .

Les moments d'ordre 3 donnent une information sur la dissymétrie du domaine alors que ceux d'ordre 4 indique si le domaine est suffisamment proche de la forme elliptique. On les prend relativement rarement en considération.

**Moments pondérés avec l'intensité.** On préfère souvent dans des images en demi-teintes utiliser les moments pondérés avec l'intensité du pixel :

$$m'_{x^n y^m} = \sum_D F(i, j) i^n j^m$$

On obtient ainsi, à l'ordre 0, le flux  $\Phi$  émis par le domaine :

$$\Phi = \sum_D F(i, j)$$

Les moments d'ordre 1 conduisent à la position du barycentre du domaine. Cette position peut être significativement différente du centre précédent. Les moments d'ordre 2 permettent

de déterminer les éléments d'une distribution elliptique gaussienne ayant même moment d'ordre 2. Au facteur  $\frac{\pi}{4}ab$  près on retrouve les mêmes relations pour calculer les axes, en divisant les moments pondérés par le flux  $\Phi$  :

$$\frac{M'_{x_2}}{\Phi} = a^2 \cos^2 \theta + b^2 \sin^2 \theta \quad (6.29)$$

$$\frac{M'_{y_2}}{\Phi} = a^2 \sin^2 \theta + b^2 \cos^2 \theta \quad (6.30)$$

$$\frac{M'_{xy}}{\Phi} = -\cos \theta \sin \theta (a^2 - b^2) \quad (6.31)$$

On déduit aisément les axes et l'orientation. Les différences avec les valeurs non pondérées sont caractéristiques de la distribution de la luminosité dans le domaine. L'extension à des moments d'ordre supérieur ne pose pas de problème particuliers. On obtient ainsi des invariants par translation, rotation, échelle et luminosité totale. Ces invariants peuvent être utilisés pour la classification ultérieure.

Les moments d'ordre 0 en  $(i, j)$  conduisent à des quantités très importantes pour la classification : flux intégral  $\Phi$ , dispersion de l'intensité dans le domaine  $\sigma_F$ , dissymétrie, aplatissement de la distribution, etc..

D'autres pondérations fonction de l'intensité du pixel sont parfois utilisées pour limiter la dynamique, ou pour tenir compte de la non-linéarité du récepteur.

**Autres pondérations.** On extraie parfois des moments pondérés avec d'autres fonctions : intensité du gradient, valeur du laplacien et d'une autre fonction du bord, etc.. Ces moments peuvent donner une information nouvelle par rapport aux autres quantités, principalement les moments d'ordre 0 en  $(i, j)$  : gradient, laplacien moyen, dispersion du gradient ou du laplacien, etc..

**Amélioration des valeurs des moments.** Les valeurs des moments sont calculés à l'intérieur des domaines. Ils sont sensibles à la délimitation du domaine. Les moments correspondent au calcul de l'intégrale par les rectangles intérieurs (si les pixels de la frontière ne sont pas inclus) ou extérieurs (dans le cas contraire). Ce sont donc des valeurs nécessairement minimisés ou maximisés. Une meilleure approximation résulte de l'utilisation de la méthode des trapèzes, c'est-à-dire en ajoutant la moitié de la contribution des pixels de la frontière.

Nous verrons, plus loin, une méthode plus précise, basée sur la compression et la dilatation du domaine.

### 6.6.3 Valeurs extrêmes.

Au cours de l'étiquetage des domaines, il est facile d'effectuer des comparaisons de valeurs correspondant aux pixels de même étiquette. On en déduit ainsi les extrema de diverses quantités définies sur les pixels du domaine. Cela concerne essentiellement :

**Les coordonnées des pixels.** On détermine ainsi le *rectangle d'encadrement*, correspondant aux points de coordonnées avec les couples  $(i_{min}, i_{max})$  et  $(j_{min}, j_{max})$ . Les valeurs minimum et maximum sont considérées sur les pixels du domaine.

La comparaison de la surface réelle à celle du rectangle d'encadrement conduit au *taux de remplissage* du domaine.

**L'intensité.** Les valeurs maximales de l'intensité peuvent donner une information sur sa classe. Ceci est utilisé en astronomie pour classer les galaxies.

**Autres fonctions.** D'autres fonctions peuvent être aussi utiles : valeur maximale du gradient, du Laplacien, ou d'une fonction de bord.

#### 6.6.4 Quantités déduites ultérieurement à l'étiquetage.

Une fois l'étiquetage terminé, il est facile d'estimer sur le domaine diverses expressions en n'utilisant que les pixels associés.

Ceci concerne en premier lieu les quantités associées à un ajustement paramétrique du domaine. Nous avons examiné en détail dans le chapitre précédent quels étaient les procédés employés à cette fin. Le calcul des paramètres du modèle nécessite une sommation sur les pixels, avec éventuellement une valeur approximative des paramètres. Le processus est alors itératif, et ne peut en aucun cas s'effectuer pendant l'étiquetage.

L'ajustement avec des tâches de distribution gaussienne est assez classique. Des modèles plus complexes peuvent être employés.

La segmentation sert ici simplement pour déterminer les pixels sur lesquels l'ajustement est effectué.

D'autres quantités peuvent aussi être déduites comme la corrélation interne. Leur rôle est assez limité.

## 6.7 Morphologie et Paramétrisation.

Dans ce qui précède le domaine est décrit soit par des moments ou des extrema, soit avec des paramètres de modèle. Cette analyse est très restrictive par rapport à ce qu'on peut déterminer comme critères associés au domaine.

On va considérer, en premier lieu, le cas d'une image binaire. L'étude de la forme est liée biunivoquement à celle de son contour. Nous allons donc examiner en détail ce que l'étude du contour apporte par rapport à celle du domaine.

L'étude de la forme introduit en outre la notion de *squelette*. On peut l'aborder de plusieurs manières. Celle associée aux opérations d'érosion et de dilatation, permet d'introduire ces opérateurs morphologiques et leurs applications.

L'étude de la topologie du domaine nous conduit au graphe d'inclusion.

La description du domaine avec les extrema (maxima, minima et cols) introduit la décomposition à l'aide des lignes de vallée et des lignes de crêtes. D'autres lignes, associées à la surface image, jouent parfois un rôle dans l'analyse morphologique.

L'étude des profils, radiaux ou angulaires permet une autre approche de la forme des domaines, et nous amène à la décomposition en pics.

On voit donc que l'analyse morphologique forme un vaste domaine, aux outils mathématiques très variés.

## 6.8 Analyse avec le Contour.

### 6.8.1 Extraction du contour du domaine.

L'extraction des points de contour peut résulter de l'étiquetage : pour un domaine particulier, on extrait les points pour lesquels un des voisins au moins ne possède pas la propriété discriminante.

D'autres algorithmes permettent d'extraire le contour sans avoir préalablement étiqueté le domaine. Prenons par exemple l'algorithme de Pavlidis :

**Détection d'un premier point** avec une lecture type vidéo on détecte le premier point d'un nouveau contour ;

**Recherche du point suivant** on choisit la direction SO. Celles O, NO, N, et NE sont exclues car on aurait dû détecter le point de contour avant s'il existait. Si la direction SO n'a pas la propriété discriminante on teste dans l'ordre S, SE et E. Si aucun point n'est adéquat on est en présence d'un domaine formé d'un seul domaine.

**Itération sur l'ensemble du contour** à chaque nouveau point, on recherche le point suivant dans la direction la plus proche du point précédent. On tourne toujours dans le sens direct jusqu'à obtention de ce point.

**Arrêt** on arrête bien sûr lorsqu'on retrouve le premier point.

L'avantage de l'algorithme précédent par rapport à la simple détermination après étiquetage réside dans la constitution en temps réel de son codage. En effet, on voit qu'il suffit de connaître les coordonnées du point initial et la suite des différentes orientations pour retracer le contour.

Avec le code de Freeman, le contour est codé avec une suite de valeurs entre 0 et 7 occupant chacune 3 bits.

### 6.8.2 Analyse du contour.

L'étude du contour s'effectue avec des moyens très variés : approximation par morceaux, séries de Fourier, ajustement fonctionnel, analyse topologique, etc..

**Approximation par morceaux et points essentiels.** La représentation du contour peut être simplifiée par une approximation par morceaux : interpolation linéaire, interpolation spline ou pseudospline, etc.. L'approximation peut s'effectuer avec une réduction constante de la maille, avec un algorithme basé sur les moindres carrés ( $L_2$ ).

La meilleure réduction s'obtient avec une approximation dans  $L_\infty$ , avec un maillage non régulier. L'algorithme de choix des points de la maille et de la détermination des coordonnées correspondantes est très complexe.

Les points obtenus forment un ensemble minimum à partir duquel le contour complet peut être redéterminé. On les appelle *points essentiels* du contour.

**Séries de Fourier :** En fait les coordonnées des points de contour en fonction de l'ordre forment des séries périodiques, puisqu'en fin de tracé on retrouve le premier point. On peut donc développer en séries de Fourier ces coordonnées.

Pour avoir une phase physiquement significative, on exprime les coordonnées en fonction de l'ordre, mais plutôt en fonction de la longueur  $l$ . Celle-ci varie de 1 pour les codes E, N, O, et

S et de  $\sqrt{2}$  pour les directions diagonales. Les développements conduisent à des expressions du type :

$$x(l) = a_0 + \sum_{n=1,N} a_n \cdot \cos(2\pi nl/L) + b_n \sin(2\pi nl/L) \quad (6.32)$$

$$y(l) = c_0 + \sum_{n=1,N} c_n \cdot \cos(2\pi nl/L) + d_n \sin(2\pi nl/L) \quad (6.33)$$

où  $L$  désigne la longueur totale du contour, et  $(x(l), y(l))$  les coordonnées courantes d'un point.

Les paramètres  $(a_n, b_n, c_n, d_n)$  dépendent d'une part du choix du premier point choisi pour déterminer le contour (effet de phase), et d'autre part de l'orientation des axes. On s'affranchit aisément de la position du centre, contenue dans les valeurs  $(a_0, c_0)$ . En éliminant les effets de phase et d'orientation des axes, on obtient de nouvelles expressions du type :

$$x(l) = a_1 \cos(2\pi l/L) + \sum_{n=2,N} a_n \cdot \cos(2\pi nl/L) + b_n \sin(2\pi nl/L) \quad (6.34)$$

$$y(l) = c_1 \sin(2\pi l/L) + \sum_{n=2,N} c_n \cdot \cos(2\pi nl/L) + d_n \sin(2\pi nl/L) \quad (6.35)$$

L'échelle de l'objet est déterminée par le nouveau coefficient  $a_1$ . On réduit donc de l'échelle en divisant par cette valeur. Le coefficient  $c_1$  est relatif à l'ellipticité  $e$  de l'objet. La réduction finale conduit donc aux relations :

$$x(l) = \cos(2\pi l/L) + \sum_{n=2,N} a_n \cdot \cos(2\pi nl/L) + b_n \sin(2\pi nl/L) \quad (6.36)$$

$$y(l) = e \sin(2\pi l/L) + \sum_{n=2,N} c_n \cdot \cos(2\pi nl/L) + d_n \sin(2\pi nl/L) \quad (6.37)$$

Les coefficients réduits sont caractéristiques de la forme du contour et peuvent donc être utilisés pour la reconnaissance.

**Représentation polaire.** Avec la méthode précédente, nous avons exprimé séparément les abscisses et ordonnées. Il peut sembler plus naturel d'obtenir une représentation en série de Fourier du contour exprimant la manière avec laquelle la distance au centre varie avec l'angle au centre.

Cette représentation se déduit des expressions précédentes, en tenant compte des expressions classiques :

$$\rho(l) = \sqrt{x^2(l) + y^2(l)} \quad (6.38)$$

$$\theta(l) = \text{Arctan}\left(\frac{y(l)}{x(l)}\right) \quad (6.39)$$

On déduit les coefficients  $(e_n, f_n)$  du développement :

$$\rho(\theta) = e_0 + \sum_{n=1,N} e_n \cos(n\theta) + f_n \sin(n\theta)$$

En pratique, on doit procéder par itérations. La représentation polaire contient une information plus comprimée que celle paramétrique en  $(x, y)$ , mais elle n'est possible que sur des contours sans valeur multiple de rayon pour un arc donné.

### 6.8.3 Courbure et équation intrinsèque.

La courbure en un point est une notion essentielle de l'étude d'une courbe. Si nous avons une représentation analytique du contour, nous pouvons calculer cette courbure  $c$ , ou son inverse le rayon de courbure  $R$  en chaque point, soit :

$$c = \frac{x'y'' - y'x''}{(x'^2 + y'^2)^{\frac{3}{2}}}$$

Si  $c$  est nul, localement le contour est linéaire. La ligne de contour présente un point d'inflexion. Il y a changement de convexité de la courbe. Le paragraphe suivant est consacré à l'étude des convexités.

Les extrema de  $c$  sont aussi très utile à étudier. Leur ensemble forme à la fois les lignes de crête et les lignes de vallée. Nous verrons plus loin comment on peut également les étudier.

La relation entre la courbure en chaque point et la longueur parcourue  $s$  le long du contour forme l'équation intrinsèque du contour. Très peu d'études d'imagerie numérique impliquent cette fonction bien qu'il s'agisse d'une caractéristique complète et invariante du domaine.

### 6.8.4 Convexité.

L'étude du contour permet d'étudier de manière précise la convexité du domaine. Un ensemble est convexe si les points de tout segment de droite tracé à partir de 2 points quelconques de cet ensemble est contenu dans cet ensemble. La convexité implique la connexité, mais non l'inverse.

Le contour d'un ensemble convexe est une courbe fermée convexe, c'est-à-dire dont la courbure est dirigée vers l'intérieur. Dans le cas d'un ensemble non convexe, le contour est formé d'arcs convexes (courbure vers l'intérieur) et d'arcs concaves (courbure vers l'extérieur). L'étude de la courbure du contour permet de déterminer si l'ensemble est convexe. Dans le cas contraires, on détermine les arcs convexes et concaves.

Les arcs convexes entourent les caps du domaine, tandis que les arcs concaves correspondent aux baies. L'analyse du contour conduit donc à une décomposition topologique en caps et baies. On dit aussi, dans le cas d'un domaine non convexe, que l'objet a plusieurs branches, chaque branche correspondant à une partie convexe entourée de 2 parties concaves.

Au sens strict, l'étude de la convexité n'a de sens que pour une topologie continue. En effet, dans le cas d'une image échantillonnée, les points de tout segment de droite n'appartiennent pas à cette image. Cette étude sous-tend une représentation continue de l'image, avec une délimitation continue des domaines.

En se basant sur une interpolation spline avec les points de contour, on dérive la courbure en chaque point, ce qui permet de déterminer les arcs convexes et concaves.

Une autre détermination peut résulter de la décomposition du contour avec les séries de Fourier en  $(x, y)$ .

Cette description est très sensible à la représentation choisie, ainsi qu'au bruit. Les contours d'un domaine est beaucoup plus régulier après lissage de l'image. Une description hiérarchique de la convexité est indispensable pour obtenir des éléments stables. L'analyse de Fourier permet en procédant harmonique par harmonique, d'obtenir une telle description. D'autres mécanismes sont possibles.

Il faut donc être très prudent si l'on construit un système de reconnaissance basé sur la structure du contour.



### 6.8.5 Paramètres associés au contour.

Nous avons vu ci-dessus la manière d'obtenir à partir du contour divers paramètres de forme. Ceux-ci étaient associés à sa représentation mathématique. L'analyse du contour permet, en outre, l'extraction de plusieurs paramètres intégraux, conduisant à des caractéristiques utiles du domaine étudié :

**Le périmètre.** Il s'agit en première approximation du nombre de points de contour  $N_c$  du domaine. Une valeur plus proche de la réalité physique s'obtient, dans le cas d'un voisinage à 4 points, en déterminant le nombre  $N_v$  de voisins hors du domaine par pixel. Nous verrons qu'il s'agit du contour du domaine obtenu par une extension de 1 pixel. Cette approximation est meilleure car elle tient compte du fait que le contour est formé de points appartenant au domaine, donc est forcément plus petit que celui qu'on déterminerait à partir d'une représentation continue de la fonction image. On s'en rend compte très bien de cet aspect lorsqu'on examine un domaine formé d'un pixel.  $N_c$  est égal à 1, ce qui donne un périmètre absurde, alors que  $N_v$  vaut 4, ce qui correspond au périmètre du carré entourant le pixel considéré.

Une meilleure approximation se situe entre  $N_c$  et  $N_v$ . La demi somme peut être envisagée comme un bon approximant. La valeur "exacte" dépend de la représentation analytique et est forcément très difficile à obtenir. Une bonne stratégie repose sur la variation du nombre de points de contour en fonction de la distance au contour, ou d'un autre paramètre associé au domaine.

Considérons le nombre de points de contour  $N_c(k)$  obtenus sur le domaine après  $i$  dilatations ( $k$  positif), ou  $-k$  érosions ( $k$  négatif).  $N_c(0)$  est le nombre de points de contour du domaine ( $= N_c$ ). L'étude de  $N_c(k)$  en fonction de  $k$  peut conduire à une estimation. Mais elle est *a priori* sans grand intérêt puisqu'elle conduit soit à  $N_c$ , soit à  $N_v$ , soit à la demi somme, selon le choix de l'indice  $k$  à considérer.

Soit  $P(i, j)$  la valeur de la fonction servant à délimiter le domaine : cela peut être le Laplacien, la fonction de bord associée au gradient maximum, le flux au dessus du seuil, ou une fonction de la texture. On intègre cette fonction sur le contour des domaines dilatés, érodés ou non, ce qui conduit à une suite  $P_c(k)$  de valeurs.  $P_c(0)$  n'est pas nécessairement nul, en raison de la géométrie discrète. Le vrai contour correspond à une valeur nulle de  $P_c(k)$ . On obtient aisément la racine  $k_0$  de l'équation :

$$P_c(k) = 0$$

Elle est comprise nécessairement entre 0 et 1. On en déduit le périmètre en calculant par une interpolation adéquate la valeur  $N_c(k_0)$ .

**Compacité.** La valeur du périmètre  $L_p$  permet de dériver un paramètre de forme important, la compacité  $C$ . Soit  $A$  l'aire du domaine. Nous avons l'inégalité :

$$C = \frac{L_p^2}{4\pi A} \geq 1$$

L'égalité n'a lieu que pour des domaines circulaires. Si  $C$  est très grand cela est dû à l'existence de nombreuses branches, allongeant le contour sans augmenter l'aire. La géométrie discrète influe sur la valeur de la compacité. Il est recommandé d'utiliser la procédure décrite ci-dessus pour le calcul du périmètre et de la surface du domaine.

**Moments et extrema sur le contour.** On peut considérer les moments et les extrema sur les points du contour. De ces valeurs on dérive un centre, des longueurs d'axes, une orientation, une aire équivalente, un allongement, etc..

On préfère utiliser les valeurs obtenues avec l'ensemble du domaine.

## 6.9 Reconnaissance par corrélation.

### 6.9.1 Principe.

Dans de nombreux problèmes de reconnaissance de forme, les objets à détecter forment un ensemble défini  $\{F_k(i, j)\}$ . L'intensité du pixel en un point est de la forme :

$$F(i, j) = aF_k(i - i_0, j - j_0)$$

où  $(i_0, j_0)$  sont les coordonnées du centre de l'objet, et  $a$  le facteur d'amplitude.  $k$  est l'indice de la forme de l'objet à reconnaître.

Cette expression n'est valable que dans une petite fenêtre centrée sur l'objet et on admet que les objets sont suffisamment loin les uns des autres pour qu'il n'y ait pas superposition. Nous avons vu dans le chapitre précédent comment on peut déterminer les coordonnées du centre de l'objet, si l'on forme l'hypothèse que l'objet est l'une des formes  $k$ . Nous admettons que le fond est nul, nous avons déjà examiné la manière d'éliminer ce paramètre.

Nous obtenons aisément :

$$a_k = \frac{\sum_{i,j} F_k(i, j)F(i, j)}{\sum_{i,j} F_k^2(i, j)}.$$

Le résidu  $R$  est :

$$R_k = \sum F^2 - \frac{(\sum F_k F)^2}{\sum F_k^2}.$$

Si l'objet observé a une forme  $l$ , on teste toutes les formes, et on détermine celle  $m$  correspond au minimum du résidu. Or la minimisation de  $R_k$  est équivalent à la maximisation du terme  $A_k$  :

$$A_k = \frac{(\sum F_k F)^2}{\sum F_k^2}.$$

On peut normaliser les fonctions de manière à ce que la somme  $\sum F_k^2$  fasse 1, ce qui ne conduit qu'à l'examen de la corrélation croisée entre l'image (dans une fenêtre à définir) et les différentes formes :

$$C_k = \sum F_k F.$$

### 6.9.2 Procédure par lissage.

On peut déterminer une taille typique pour effectuer une série de lissage correspondant à chaque forme. On obtient ainsi un ensemble de grandes images  $G_k(i, j)$ . En raison du bruit, on doit définir un seuil de détection  $s$  : si le signal dans une image lissée est supérieur à  $s$ , il y a détection. La position de l'objet est défini par le maximum de corrélation. On se contente donc de détecter les maxima significatifs de chaque image.

Pour un maxima significatif détecté, l'identification de l'objet est effectué par la détermination de l'image pour laquelle la corrélation est maximum. Un test sur le résidu permet, en outre, de confirmer si l'objet a bien la forme identifiée.

Cette procédure n'est possible que si le nombre de formes est faible ( $< 10$ ), sinon le temps de calcul et le nombre d'images à manipuler deviennent trop élevés. Elle s'adapte parfaitement à des méthodes optiques.

### 6.9.3 Utilisation de la segmentation.

On évite le lissage avec l'ensemble des formes, si on détecte dans un premier temps l'existence d'un objet centré en un point assez précis, puis on effectue la corrélation avec les formes centrées en ce point.

La détection peut provenir d'une segmentation avec un caractère textural adapté aux formes à reconnaître. La localisation est effectuée, par exemple avec les moments du domaine associé à l'objet.

La reconnaissance par corrélation est sensible à l'échelle linéaire de l'objet et à son orientation. La détection par segmentation permet d'en tenir compte. En effet, en utilisant les moments d'ordre 2, il est facile de calculer le facteur d'échelle et l'angle de rotation permettant la superposition des formes sur l'image à étudier.

On peut aller plus loin en calculant pour la série de formes différents moments et paramètres structuraux ce qui conduit à une reconnaissance basée sur ces paramètres. Ce qui revient aux méthodes étudiées ci-dessus.

### 6.9.4 Corrélation et contour.

Les méthodes de corrélation peuvent aussi s'appliquer à d'autres paramètres que les intensités des pixels. Si la forme d'un objet est identique à celle d'un objet de référence, on peut comparer tous les paramètres résultants.

Le contour est l'un des ensembles intéressants. La manière la plus correcte de l'utiliser consiste dans ses développements réduits en série de Fourier. Ces développements sont invariants par translation, rotation, et homothétie, donc bien caractéristiques de la forme. On peut corréler les valeurs obtenues avec les développements des formes de référence et choisir celle qui donne le maximum de corrélation.

D'autres paramètres structuraux, comme le squelette, peuvent aussi être utilisés dans des méthodes de corrélation.



# Chapitre 7

## Etudes morphologiques.

### 7.1 Morphologie mathématique.

#### 7.1.1 Erosion et Dilation.

Les opérations d'érosion et de dilatation forment les outils de base de la morphologie mathématique.

L'érosion d'un domaine consiste dans l'élimination de tous les pixels voisins d'un point extérieur au domaine. Le voisinage à 4 points éliminent moins de pixels que celui à 8 points. C'est ce dernier qui conduit aux meilleurs résultats pour les opérations composées.

Au contraire, la dilatation consiste dans l'opération d'adjonction au domaine de tous les points voisins d'un des points du domaine.

Les opérations d'érosion  $E$  et de dilatation  $D$  forment des opérateurs transformant une image binaire en une autre image binaire. La transformation est non linéaire.

#### 7.1.2 Composition des opérations.

La composition d'une érosion puis d'une dilatation ( $DE$ ) est appelé ouverture. Ce n'est pas l'opérateur identité. En effet, des pixels disparus à l'érosion ne sont pas restaurés à la dilatation. Soit  $A$  un domaine, désignons par  $B$   $E(A)$ , par  $C$   $D(A)$ , par  $D$   $D(B)$  et par  $E$   $E(C)$ . La différence  $F$  entre  $A$  et  $B$  représente les points de contour. Celle  $G$  entre  $A$  et  $D$  forme l'ensemble des points du contour qui ne peuvent être restauré après dilatation. Ils forment des caps de largeur de 1 ou 2 pixels.

De même la composition d'une dilatation puis d'une érosion ( $ED$ ) est appelé fermeture.

La différence  $H$  entre  $C$  et  $A$  forme le contour extérieur du domaine. Celle  $I$  entre  $E$  et  $A$  est l'ensemble des points du contour extérieur qui ne peut être restitué après érosion. Il s'agit des points des baies étroites d'un ou de deux pixels.

La suppression des pixels de  $H$  et l'ajout de ceux de  $I$  conduisent à obtenir un contour plus régulier.

#### 7.1.3 Distance au contour.

L'itération des opérations d'érosion permet de déterminer pour chaque pixel du domaine la distance au contour. Avec la notation du paragraphe précédent, on obtient ainsi une suite d'ensemble  $C_k$ , où  $k$  désigne l'ordre d'érosion.  $k - 1$  est aussi la distance du pixel au contour.

On obtient de la même manière une distance pour les pixels externes avec la dilatation, avec les ensembles  $H_k$ .

Des algorithmes procédant en lecture ligne par ligne conduisent à la détermination de ces distances, souvent utiles pour l'étude du domaine. Nous en avons vu une application pour l'amélioration des estimations des quantités associées au contour et au domaine.

#### 7.1.4 Squelette d'un domaine.

Le squelette d'un domaine est l'ensemble des points situés à égale distance des points du bord.

Dans le cas d'une image échantillonnée sur une grille carrée, le problème de la détermination des points du contour appartenant au squelette est résolu sans difficulté : il s'agit des points de  $G$ . L'ensemble du squelette est déterminé par l'itération du processus  $DE$ , conduisant à une suite d'ensembles  $G_k$ .

Avec les voisinages à 4 points, le squelette d'un domaine est connexe, avec de nombreuses branches, même pour un ensemble apparemment assez régulier. Cela est dû aux effets de la géométrie discrète.

Par contre le squelette déterminée par itération avec un voisinage à 8 points n'est pas nécessairement connexe : les parties ne correspondant pas à une même étape d'itération ne peuvent être voisines. Par contre, il est beaucoup plus simple.

Chaque pixel du squelette est affecté d'un indice correspondant à l'étape du processus itératif à laquelle il est apparu. Le domaine peut être restauré parfaitement à partir de la connaissance des pixels du squelette, et de leur indice. Il y a donc biunivocité entre la connaissance du squelette et celle du contour. On le rend connexe, en dilatant une fois chacune des parties.

Le squelette contient une information importante sur la morphologie. Cette information n'est pas restaurable avec les moments et extrema calculés sur le domaine. L'étude du contour permet par contre de l'obtenir, mais au prix de grandes difficultés. Cette information consiste dans la décomposition en branches de l'objet. L'existence de plusieurs parties convexes et concaves du contour se traduit par un squelette fait de plusieurs branches. En tenant compte de l'étape à laquelle un pixel du squelette a été formé dans le processus itératif, il est possible par dilatation d'affecter à chaque pixel une étiquette d'appartenance à un des morceaux du domaine.

Ceci permet sur chaque partie d'obtenir la surface, le flux, la position, ou divers paramètres de forme, conduisant à une description détaillée du domaine.

## 7.2 Filtre morphologique.

Considérons une ligne d'une image discrétisée. Son tracé  $V(i)$  est le contour d'un domaine discret.

On applique le processus d'érosion puis de dilatation. On supprime du profil les pics de largeur d'un ou de deux pixels. Inversement, si on applique le processus de dilatation puis d'érosion on remplit les trous de largeur d'un ou de deux pixels. Cela a pour résultat d'adoucir le profil. On appelle filtre morphologique cette opérateur sur l'image.

On peut itérer et obtenir ainsi des profils de plus en plus doux. Ceci constitue une opération non linéaire utile pour éliminer des défauts ou pour déterminer un fond.

## 7.3 Morphologie en niveaux de gris.

### 7.3.1 Germes de la ligne.

Appliquons sur le domaine-ligne le processus de détermination du squelette indiqué ci-dessus. On constate que le squelette consiste dans une succession d'ajout d'intensité qu'on introduit étape par étape d'itération.

Pour reconstituer l'intensité d'une ligne en chaque point, il suffit de partir de ces ajouts en connaissant l'étape correspondante. On les appelle germes de la ligne-image.

La description du profil est biunivoque avec la connaissance des germes.

### 7.3.2 Volume image et filtre morphologique.

On généralise aisément ce qui précède au cas d'une image complète. On considère dans un espace discret à trois dimensions le volume formée par l'image. L'érosion consiste toujours dans la suppression des pixels voisins avec un point extérieur, alors que la dilatation rajoute des points extérieurs voisins au domaine.

On considère soit le voisinage à 8 points, soit celui à 26 points, selon qu'on prend les pixels diagonaux ou non. Le voisinage à 26 pixels conduit à minimiser le nombre de germes.

La composition érosion-dilatation supprime les pics étroits, alors que la dilatation suivie de l'érosion remplit les trous.

### 7.3.3 Germes d'une image.

Le squelette obtenue sur le volume de l'image forme l'ensemble des *germes* de l'images. La connaissance des différents germes avec l'indice de l'étape à laquelle il a été formé permet la reconstruction complète de l'image.

Alors que le squelette permet seulement la reconstruction du contour du domaine, les germes contiennent toute la morphologie du domaine, et peuvent, de ce fait, contribuer à sa classification et à sa reconnaissance.

## 7.4 Structures de la surface image.

### 7.4.1 La surface image.

Nous avons souvent introduit la surface correspondant à la relation, dans un espace à trois dimensions, entre l'intensité du pixel et les coordonnées. Cette surface image n'a aucun sens physique mais c'est un moyen commode de se représenter les intensités de l'image. Les zones brillantes correspondent à des pics, alors que les zones faibles sont des trous de la surface.

L'étude de la surface image conduit à définir des éléments importants de l'image.

### 7.4.2 Les extrema.

Ce sont les points où les dérivées premières en  $x$  et  $y$  s'annulent. Nous avons 3 types d'extrema : les maxima, les minima et les cols (ou points de selle) :

$$\left(\frac{\partial^2 F}{\partial x \partial y}\right)^2 - \frac{\partial^2 F}{\partial x^2} \cdot \frac{\partial^2 F}{\partial y^2} > 0$$

$$\left(\frac{\partial^2 F}{\partial x \partial y}\right)^2 - \frac{\partial^2 F}{\partial x^2} \cdot \frac{\partial^2 F}{\partial y^2} < 0 \quad \text{Col} \quad (7.1)$$

$$\text{et } \frac{\partial^2 F}{\partial x^2} > 0 \quad \text{Minimum} \quad (7.2)$$

$$\left(\frac{\partial^2 F}{\partial x \partial y}\right)^2 - \frac{\partial^2 F}{\partial x^2} \cdot \frac{\partial^2 F}{\partial y^2} < 0$$

$$\text{et } \frac{\partial^2 F}{\partial x^2} < 0 \quad \text{Maximum} \quad (7.3)$$

La fonction  $F(x, y)$  correspond à l'intensité de l'image. La recherche des points extrêmes par dérivation est très délicate. On préfère des algorithmes plus rapides basés simplement sur le voisinage.

Ainsi, la recherche des maxima s'effectue par comparaison de l'intensité du pixel à celle de ces voisins. Le voisinage peut être à 4 ou 8 points. On procède de même pour les minima.

La recherche des points de selle sur une image discrète est plus délicate. On montre qu'il en existe de deux types :

**Axes normaux** il s'agit des points pour lesquels la densité croît dans le sens des  $x$ , et décroît dans le sens des  $y$ , ou inversement ;

**Axes à 45°** dans le cas d'une représentation bilinéaire dans un carré les isophotes sont des hyperboles équilatères. S'il y a dégénérescence dans ce carré d'une isophote en ses asymptotes, on a un col avec une décroissance de l'intensité dans une direction diagonale, et une croissance dans l'autre sens.

### 7.4.3 Les lignes associées aux dérivées partielles premières.

Les lignes de dérivées partielles nulles sont parfois utilisées pour l'analyse de l'image. Souvent on utilise une méthode basée sur les différences entre les points voisins. Cela suppose un filtrage préalable de l'image. Ces lignes présentent le défaut majeur de ne pas être invariante par rotation.

### 7.4.4 Graphe de liaison entre extrema.

Les lignes de dérivée partielles permettent de déterminer un graphe de liaison entre les différents extrema. Ce graphe est une caractéristique de l'image, ou de parties de l'image. Bien que les lignes de dérivées partielles nulles dépendent de l'orientation, le graphe de liaison est invariant.

### 7.4.5 Lignes de vallée.

Les lignes de vallée et les lignes de crête sont les lignes naturelles de liaison entre les différents extrema d'une surface. Plusieurs définitions sont possibles :

**L'ensemble des points de l'isophote où la courbure est extrémale** la courbure d'une isophote en un point est déterminée et on déduit aisément les pixels pour lesquels cette courbure est maximale le long de l'isophote. Nous avons vu ci-dessus que l'intensité du pixel pouvait s'exprimer localement dans un repère orienté sur la tangente et la normale à l'isophote. Nous avons, en particulier calculer les dérivées partielles d'ordre 1 et 2 par



rapport à  $(X, Y)$ . La dérivée partielle  $\partial^2 F / \partial Y^2$  a permis ainsi de définir une ligne de bord associée au gradient maximum.

Considérons maintenant les pixels pour lesquels la dérivée  $\partial^2 F / \partial X \partial Y$  est nulle. Dans le sens de la tangente à l'isophote ( $Y = 0$ ), l'intensité présente un extremum. Le rayon de courbure est extrémal. D'après les calculs ci-dessus, cela correspond à l'équation :

$$\sin \theta \cos \theta \left( \frac{\partial^2 F}{\partial y^2} - \frac{\partial^2 F}{\partial x^2} + (\cos^2 \theta - \sin^2 \theta) \frac{\partial^2 F}{\partial x \partial y} \right) = 0$$

Avec :

$$\tan \theta = - \frac{\partial F}{\partial X} / \frac{\partial F}{\partial Y}$$

Le réseau ainsi déterminé contient à la fois les lignes de vallée et les lignes de crête. La séparation n'est pas évidente.

**Les lignes qui séparent les bassins ou les massifs** elles séparent les points de la surface pour lesquels si une goutte tombe en ce point elle descendra au même minimum (bassin), ou, fictivement, elle remontera au même maximum (massif).

Les lignes se déterminent par le contour après segmentation.

**Les lignes de plus grande pente passant par les cols** Le réseau des lignes de plus grandes pentes joue un rôle important dans l'étude d'une surface. Ces lignes sont orthogonales au réseau des lignes isophotes.

Par un point de la surface, il ne passe qu'une seule ligne de plus grande pente, sauf pour les extréma. Par un minimum ou un maximum, il passe une infinité de lignes. Dans l'analogie hydrologique, les gouttes d'eau peuvent venir au fond dans toutes les directions, elle peuvent de même s'éloigner du maximum quel que soit l'orientation. Par un col, il ne passe que deux lignes de plus grande pente, orthogonales entre elle. Soit, la goutte descend vers un fond dans un sens ou l'autre, soit elle remonte vers un sommet, dans un sens ou l'autre.

Les lignes de vallée sont les lignes de plus grande pente partant des cols et se dirigeant vers les fonds, alors que les lignes de crête sont celles qui se dirigent vers les sommets.

Le réseau des lignes de vallée permet de séparer les domaines contenant un maximum et un seul (pic), alors les lignes de crêtes permettent de séparer ceux contenant un minimum et un seul (trou). Chacun des réseaux peut s'employer indépendamment pour l'étude des pics ou des trous d'intensité. Ils peuvent être utiles pour l'analyse complète de l'image, détection des objets inclus, ou être utilisé localement, pour décomposer un domaine.

Par construction ce réseau est invariant par rotation des axes. Il permet enfin de construire le graphe de liaison entre les extrema.

Ce réseau de lignes de l'image est très sensible au bruit. Leur tracé suppose un lissage préalable. En outre, dans les régions du fond de l'image, où seul le bruit est présent, ce réseau conduit à une multitude de petits pics ou trous non significatifs. Il est facile de les éliminer par seuillage ou par cicatrisation.

#### 7.4.6 Les lignes de points d'inflexion.

On peut également s'intéresser sur la surface image aux points correspondant aux inflexions des isophotes. Ils permettent de séparer les branches des domaines, dans le cas où ils ne sont

pas convexes. On peut les obtenir, en reprenant la représentation locale de l'intensité autour d'une isophote. Ce sont les lignes pour lesquels la dérivée partielle en  $X^2$  est nulle. La courbe isophote est alors du troisième degré en  $X$ , et elle présente donc une inflexion. Les lignes des points d'inflexion s'obtiennent donc par la relation :

$$\frac{\partial^2 F}{\partial x^2} \cos^2 \theta + \frac{\partial^2 F}{\partial y^2} \sin^2 \theta + 2 \frac{\partial^2 F}{\partial x \partial y} \sin \theta \cos \theta = 0$$

La détermination de ces lignes exige une représentation de type spline, pour une bonne estimation des dérivées partielles. Ce réseau s'utilise plutôt sur un domaine que sur l'ensemble de l'image. On préfère toutefois étudié les branches des objets à partir des opérateurs morphologiques.

#### 7.4.7 Autres lignes.

D'autres lignes de la surfaces images peuvent bien sûr être définies et utilisées pour l'étude des domaines : géodésiques, lignes de courbure moyenne ou totale nulle, etc..

### 7.5 Profils associés aux extrema.

#### 7.5.1 Profils radiaux.

Dans de nombreuses applications de l'imagerie numérique, l'analyse consiste dans une étude de pics. Toute l'information est contenue dans la localisation et la caractérisation de ces pics.

On effectue une détection des maxima (ou des minima s'il s'agit de trous), pour la localisation. Si le pic admet une symétrie radiale cela peut suffire. L'information sur la forme est alors contenue dans le profil radial du pic. Ce profil est déterminé en faisant la moyenne des intensités des pixels situés à une distance donnée du sommet du pic.

Un processus de rejet permet l'élimination des pixels dont la valeur est trop éloignée de la moyenne (contamination par un autre pic).

La structure discrète de l'image numérique complique la détermination précise du profil radial, un pixel ne se trouvant pas, en général, à une distance entière du centre. On peut raffiner la détermination du profil en utilisant le concept d'appartenance floue à une couronne d'un pixel. Le poids d'appartenance est déterminée en fonction de l'écart au rayon.

#### 7.5.2 Cas d'un objet de forme elliptique.

Lorsque les contours d'un objet sont de formes elliptiques homothétiques, on améliore sensiblement le profil en prenant la moyenne sur des couronnes elliptiques.

L'objet est complètement décrit par sa position, le rapport des axes des ellipses des contours, leur orientation et le profil radial.

#### 7.5.3 Décomposition à partir des profils.

L'analyse morphologique peut se déduire des profils. On détecte les maxima et on détermine le rapport des axes et l'orientation des contours dans leur voisinage. Ceci permet de déterminer les profils dans les couronnes elliptiques associés à chaque maximum.

La connaissance des positions, du rapport des axes, de l'orientation et du profil radial permet la reconstitution d'une image *a priori* similaire à l'image de travail. On soustrait l'image de synthèse et on obtient ainsi les écarts au modèle.

On peut itérer sur la différence, et ainsi de suite jusqu'à ce que les écarts soient négligeables.

Cette méthode d'analyse est bien adaptée aux images constituées de pics, ayant éventuellement un peu de superposition.

#### 7.5.4 Profil angulaire.

On peut déterminer pour chaque pic, le profil angulaire, c'est-à-dire la variation de l'intensité en fonction de l'angle.

L'apport d'information de ce profil est généralement relativement faible.



## Chapitre 8

# La Vision Multiéchelle.

### 8.1 Les transformations Multi-échelles

Du fait du théorème d'échantillonnage, il est traditionnel de découper l'image en éléments, les pixels, et par conséquent de considérer ces éléments comme les atomes de l'image. De ce fait on admet que chaque pixel apporte une information indépendante sur l'image. En fait, ce découpage en pixels est purement fonctionnel, et ne présume pas que l'image soit constituée d'éléments juxtaposés. D'autres découpages, comme celui lié au contour des objets, peuvent être plus porteurs pour l'interprétation.

Les images du ciel ont la caractéristique d'être formées de taches diffuses ou très ponctuelles superposées sur un fond lentement variable. La valeur d'un pixel dépend donc d'au moins deux informations observées à deux échelles différentes. Il est donc tout à fait naturel d'introduire des transformations qui tiennent compte de cette caractéristique. Les approches sont nombreuses. Plusieurs d'entre elles ont été introduites de manière empirique, comme une bonne recette, d'autres, comme la transformation en ondelettes, dérivent d'une approche mathématique plus profonde. Parmi les plus intéressantes sur le plan opérationnel nous avons :

**Le laplacien pyramidal** il est naturel de réduire une image en faisant une moyenne des quatre valeurs contenues dans une zone de  $2 \times 2$  pixels. On perd nécessairement de l'information. Pour la récupérer, on interpole le tableau et on effectue les différences entre les valeurs initiales et interpolées. Il existe diverses variantes selon la moyenne et l'interpolation effectuée. En itérant, on obtient une pyramide de coefficients contenant l'information aux différentes échelles ;

**La transformation de Haar** Si on considère la somme et la différence de deux valeurs, on les restitue en effectant la demi-somme et la demi-différence. Sur ce constat, en examinant par groupe de deux on obtient d'une part un tableau lissé, d'autre part des valeurs significatives des variations locales. On itère sur le tableau lissé, ce qui permet d'obtenir une transformation de même dimension que le tableau initial, avec une inversion aisée. A deux dimensions, on procède séparément ligne par ligne et colonne par colonne ;

**L'analyse multirésolution** c'est la généralisation de la transformation de Haar, mais on effectue des combinaisons un peu plus compliquées. L'intérêt principal réside dans l'association à deux fonctions plus régulières que les marches d'escalier qui sont introduites avec la transformation de Haar.

**L'algorithme à trous** les transformations précédentes réduisent l'échantillonnage, ce qui pose

beaucoup de problèmes ultérieurs, pour le filtrage, la déconvolution ou la vision. L'algorithme à trous peut se déduire de la transformation en ondelettes continue, mais aussi du laplacien pyramidal ou de l'analyse multirésolution. On effectue une transformation sans décimation, c'est-à-dire sans réduction de l'échantillonnage. Nous en verrons une description précise ultérieurement ;

**Les pyramides morphologiques** au lieu d'appliquer des opérateurs linéaires, on peut aussi mettre en œuvre des opérateurs non linéaires, comme ceux issus de la Morphologie Mathématique. Par exemple, on détermine dans un pavé  $3 \times 3$  le minimum, et on ne prend qu'un point sur deux (lignes et colonnes). On interpole morphologiquement en prenant le maximum dans l'environnement  $3 \times 3$  du point considéré et on effectue la différence. On extrait ainsi les pics à l'échelle 1. On itère, ce qui permet d'obtenir les pics à toutes les échelles. On peut aussi prendre les maxima, ou utiliser un autre élément structurant.

**La pyramide de médianes** une autre variante non linéaire consiste dans la détermination de la valeur médiane dans l'environnement  $3 \times 3$ . La réduction de l'échantillonnage et la réinterpolation sont suivies d'une différence pour déterminer les coefficients à l'échelle 1. On itère sur le tableau décimé pour obtenir les coefficients à toutes les échelles. La pyramide de médianes permet de mieux compacter les structures d'une échelle à l'autre.

Nous avons vu toute une série de transformations multiéchelles qui sont caractérisées par les éléments suivants :

**Une règle de lissage** pouvant être linéaire ou non ;

**Une décimation des données** cette décimation est en fait virtuelle dans l'algorithme à trous : on ne décime pas, mais on tient compte par des entrelacements de l'absence de décimation. Cette règle s'applique aussi au cas non linéaire ;

**Un calcul de coefficients** cela peut résulter d'une différence entre l'image originale et l'image lissée interpolée, mais des calculs plus subtils permettent la réduction du tableau de données ;

**Une itération sur le tableau lissé** permettant d'obtenir les détails à des échelles de plus en plus grandes ;

**Une règle de reconstruction** triviale pour les détails provenant des différences, plus subtile dans le cas des méthodes non redondantes comme l'analyse multirésolution.

Il serait présomptueux de vouloir développer dans le cadre de ce cours toutes les applications associées à ces transformations. On se limitera à l'utilisation de l'algorithme à trous.

## 8.2 L'Algorithme à trous

Le choix de l'algorithme est associé à différents critères. En premier lieu nous souhaitons une vision isotrope, il n'y a pas de raison de privilégier une direction particulière. Or ce n'est pas le cas de l'analyse multirésolution. Nous voulons suivre d'échelle en échelle un objet, car sa perception ne s'effectue jamais sur une seule échelle. Enfin, l'algorithme doit être rapide. Ceci nous a amené à l'application de l'algorithme à trous.

Sans entrer dans des détails mathématiques subtils, on peut montrer son origine. On introduit la notion d'approximation par une suite translatée :

$$c(0, k) = \langle f(x), \phi(x - k) \rangle$$

$\phi(x)$  est appelée fonction d'échelle. Elle doit satisfaire l'équation de dilatation :

$$\frac{1}{2}\phi\left(\frac{x}{2}\right) = \sum_n h(n)\phi(x-n)$$

Le pavé de résolution est formé des coefficients définis de la manière suivante :

$$c(i, k) = \left\langle f(x), \frac{1}{2^i}\phi\left(\frac{x-k}{2^i}\right) \right\rangle$$

On obtient la récurrence :

$$c(i+1, k) = \sum_n h(n)c(i, k+n2^i)$$

Il est donc facile d'obtenir le pavé de résolution à partir de l'approximation initiale. On doit spécifier le filtre  $h$ . On a choisi celui associé à l'interpolation B-spline cubique centrée :

$$h(0) = \frac{3}{8} \quad h(\pm 1) = \frac{1}{4} \quad h(\pm 2) = \frac{1}{16} \quad h(n) = 0 \text{ si } |n| > 2$$

Ceci permet d'obtenir pour les approximations successives une très grande régularité ( jusqu'à l'ordre 3 des dérivées). Les coefficients (en ondelettes) sont obtenues simplement en effectuant les différences entre plans successifs.

La transformation se présente sous la forme d'une série d'images de même dimension  $N \times N$  que l'image initiale. Le nombre maximum d'échelles est  $\log_2 N$ .

### 8.3 Visualisation

Il y a différentes manières de représenter les données produites par l'algorithme à trous. La plus simple consiste dans la représentation de la transformée en ondelettes comme une série d'images. Chacune d'entre elles peut alors être visualisée indépendamment des autres, en utilisant les outils classiques des logiciels d'imagerie (table des couleurs, histogramme, pointé, coupe, ...).

Les plans d'ondelettes correspondent à des bandes de fréquence différentes, et les images en ondelettes montrent ainsi des structures de différente taille. Dans la première image, on perçoit surtout le bruit et les étoiles, tandis que dans la dernière, les grandes structures comme les bras de la galaxie apparaissent.

Une image de synthèse peut aussi être créée en binarisant chaque plan aux seuils  $+3\sigma_i$  déterminés, en affectant une couleur à chaque échelle, et en superposant l'ensemble des plans.

### 8.4 Coefficients Significatifs

Si l'image est localement uniforme, il est facile de voir qu'on obtient un coefficient nul. Mais pour des images astronomiques l'image n'est jamais tout-à-fait uniforme en raison du bruit. Le coefficient  $w$  ne sera pas nul. Sa loi de distribution  $Prob(w)$  peut être estimée aisément pour un processus gaussien, soit analytiquement soit expérimentalement. Pour d'autres bruits le calcul est plus délicat.

Si  $w$  est observé, on peut s'interroger sur l'hypothèse d'une uniformité locale ou non. Soit  $\mathcal{H}_0$  l'hypothèse selon laquelle l'image est localement constante. On rejette  $\mathcal{H}_0$  ( $w > 0$ ) si :

$$Prob[W > w(i, k, l)] < \epsilon$$

Ce rejet dépend d'un seuil d'incertitude  $\epsilon$ . Si ce seuil est par exemple de 0.01, cela signifie que sur 100 pixels en moyenne 1 sera détecté supérieur au seuil en raison du bruit. Ceci conduit à un taux de fausses alarmes très important. Si on réduit trop fortement le seuil, on rate des détections de structures réelles. Un compromis est nécessaire. Souvent on choisit 0.001 ou 0.0001. Dans certains cas, on introduit un flou entre ces deux seuils.

Si dans le cas gaussien il est assez aisé d'obtenir un bon seuillage, le problème est plus délicat pour d'autres bruits. La détermination des coefficients significatifs s'effectue assez bien dans le cas des transformations linéaires. Dans le cas non-linéaire, on procède par simulation, mais le seuillage est instable car les lois dépendent de l'environnement.

## 8.5 Le filtrage adaptatif multiéchelle

Grâce à la transformation multiéchelle et au seuillage, on a isolé les éléments significatifs avec lesquels on est capable de reconstruire l'image. Cette reconstruction se fait directement avec l'analyse multirésolution qui est non redondante. Mais des artefacts importants apparaissent dans la procédure. Dans le cadre de l'algorithme à trous, transformation redondante, la reconstruction ne s'effectue pas par une simple inversion et il est alors possible de mieux contrôler la reconstruction et de réduire les artefacts.

On procède par itération obtenant une image  $I^{(n)}$  à l'étape  $n$ . La différence :

$$R^{(n)} = I - I^{(n)}$$

forme le résidu à cette étape. Si le bruit est éliminé, le résidu doit être équivalent à une séquence de bruit, sinon on détectera dans le bruit des coefficients significatifs, avec lesquels, par reconstruction on en déduira un résidu significatif. Ce résidu doit être ajouté à l'image reconstruite pour obtenir une nouvelle image plus correcte.

Différentes variantes sont applicables selon la méthode de reconstruction et d'utilisation du résidu significatif, mais l'idée générale reste similaire.

## 8.6 La restauration Multiéchelle.

L'équation entre l'objet  $O$  et l'image  $I$  peut s'écrire :

$$I = O \star P + B$$

où  $P$  est la fonction d'étalement connue et  $B$  le bruit inconnu. On peut réécrire cette équation sous la forme :

$$B = I - O \star P = R$$

On se ramène ainsi au problème du filtrage. Le résidu entre l'image initiale  $I$  et l'image reconstruite  $I^{(n)} = O^{(n)} \star P$  ne doit comporter de terme significatif. Une fois le résidu significatif déterminé on reconstruit une nouvelle approximation de l'objet avec une expression de la forme :

$$O^{(n+1)} = \mathcal{K}(O^{(n)}, \tilde{R}^{(n)})$$



où  $\mathcal{K}$  peut être, par exemple :

**Van Cittert**  $\mathcal{K}(O, R) = O + R$ ;

**Gradient à pas fixe**  $\mathcal{K}(O, R) = O + \alpha \tilde{P} \otimes R$ , où  $\tilde{P}$  est l'opérateur adjoint ( $\tilde{P}(k, l) = P(-k, -l)$ )  
et  $\alpha$  un paramètre facile à estimer ;

**Lucy**  $\mathcal{K}(O, R) = O + (\frac{R}{O \otimes P}) \otimes \tilde{P}$ .

## 8.7 La vision multiéchelle

Un modèle plus complet de vision a été développé. Grâce au déploiement en échelle et au seuillage des coefficients significatifs, les objets physiques apparaissent comme des volumes dans l'espace de la transformée. Une procédure de segmentation a été mise au point dans cet espace. Les objets sont identifiés comme des sous-arbres associés au graphe d'inclusion des domaines d'une échelle à l'autre. Il est donc essentiel de pouvoir définir un voisinage dans l'espace de la transformée, ce qui est plus aisé à réaliser avec l'algorithme à trous.

Les images des objets sont reconstruites à partir des coefficients significatifs dans le volume associé, par résolution itérative du problème inverse. Les mesures sont effectuées sur les images reconstruites. Les tests effectués conduisent à une bonne précision.

Pour des objets isolés de petite taille, les méthodes classiques d'imagerie astronomique donnent des résultats similaires. Mais le modèle multiéchelle permet une description correcte des structures plus complexes. La notion de sous-objet permet de décrire complètement la hiérarchie des structures.

## 8.8 Conclusions.

L'utilisation de transformations multiéchelle permet un très grand nombre d'applications en imagerie astronomique. Les outils associés sont simples, même si des difficultés mathématiques peuvent apparaître pour des problèmes précis.

D'autres applications ont été aussi mises en œuvre avec succès, comme la mise en correspondance automatique d'images. Des problèmes de fusion de données ont eu aussi une solution originale et exploitable.

Les transformations multiéchelles permettent en fait de mieux décrire les objets qu'une simple analyse pixel par pixel, parce qu'elles séparent, par définition, les informations à différentes échelles. Elles conduisent aussi à une description de l'image basée sur des objets, décrits par leurs coefficients et le volume correspondant. Ceci va dans le sens d'une description d'images par leur contenu qui forme l'une des voies les plus prospectives pour la vision par ordinateur.



## Chapitre 9

# Analyse d'un ensemble d'Images

### 9.1 Introduction.

Il arrive fréquemment en astronomie qu'un même objet, ou une même zone, soit observée plusieurs fois, mais dans des conditions un peu différentes. Il peut s'agir soit :

**D'analyse multispectrale** si l'on observe avec différents filtres colorés.

**De cartographie de la polarisation** si les images sont obtenues à travers des polariseurs.

**De détermination de mouvements** s'il s'agit d'observations prises à différents instants.

Malgré la diversité des problèmes on retrouve toujours divers points communs :

**Le rééchantillonnage des images** dans un même système de coordonnées ;

**La mise en évidence des pixels, ou des zones**, pour lesquels des différences apparaissent ;

**Une réduction** nécessaire de l'information.

Par contre, nous verrons que chaque problème amène en plus quelques particularités.

### 9.2 Le Rééchantillonnage.

Les images pour être comparées doivent être ramenées dans un même repère. Il se peut qu'elles le soient, pour des observations successives avec une caméra CCD, par exemple. Dans le cas contraire la procédure de rééchantillonnage est la suivante :

**Définition de plusieurs points de repère** sur l'image. Ces points doivent être bien définis.

Des étoiles forment le plus souvent de bons repères. Dans le cas d'une série d'images d'une structure variable, il peut être très difficile de définir un bon repère. Celui-ci doit alors être défini de manière instrumentale.

**Saisie sur les images des positions des repères** Elle peut être interactive ou automatique.

Dans ce dernier cas on utilise souvent la corrélation croisée entre les repères et une forme stellaire.

**Détermination des équations de transformation** entre le repère de l'image et celui de référence. Pour des raisons liées à l'algorithme de rééchantillonnage on utilise toujours une relation de la forme :

$$\text{Repère Image} = \text{Fonction} (\text{Repère Référence})$$

La transformation est le plus souvent une similitude. S'il existe des distorsions, on utilise des polynômes du 2nd ou du 3ième degré, homogènes en  $x,y$ . Il arrive aussi qu'on rééchantillonne en secondes d'arc sur le ciel dans un repère EO/NS par rapport à un centre de coordonnées définies préalablement.

**On recalcule** les valeurs de l'image rééchantillonnée en se donnant des coordonnées de départ, un pas d'échantillonnage dans chaque direction et une taille de la matrice.

Chaque pixel du rééchantillonnage a des coordonnées  $X,Y$  dans le repère de référence, ce qui permet d'obtenir les coordonnées  $x,y$  dans l'image de travail. On détermine, par interpolation, la valeur de l'image au point  $x,y$  et l'on affecte cette valeur dans l'image rééchantillonnée au point  $X,Y$ .

Si le principe est très simple, la réalisation peut être compliquée par des problèmes d'accès aux valeurs de l'image.

### 9.3 Réduction Photométrique.

De même que pour le traitement d'une image individuelle la réduction photométrique est une opération essentielle si l'on veut comparer 2 ou plusieurs images. Cela concerne la suppression du fond, la transformation en éclaircissement et les corrections de sensibilité.

Il est souvent préférable de transformer en magnitude au dessus du fond de ciel pour une analyse multispectale. Il arrive parfois qu'au lieu d'utiliser directement la magnitude, ou l'éclaircissement, on considère le gradient ou le laplacien. Cela renforce les intercomparaisons pour les pixels correspondant à une structure complexe.

Nous admettrons dans toute la suite que tous les problèmes de réduction géométrique comme photométrique ont été résolus.

### 9.4 Comparaison des Images 2 à 2.

La comparaison de 2 images peut s'effectuer de différentes manières :

**Rapport des 2 images** : c'est l'opération la plus simple. Si les images sont exprimées en magnitude, il faut faire seulement la différence.

La visualisation du rapport fait apparaître différentes zones. Une analyse structurale permet d'obtenir un catalogue de ces zones avec les positions, les aires, etc.

L'histogramme, ou divers paramètres statistiques peuvent être aussi très utiles.

**La différence** contient une information souvent différente de celle du rapport, en particulier dans les zones proches du fond.

**Le diagramme** obtenu à partir des valeurs pixels homologues dans les deux images est aussi très riche. Il permet :

- de déterminer la relation  $v_2 = f(v_1)$  entre les valeurs pixels des 2 images ;
- de mettre en évidence des groupes de pixels.

Nous examinerons plus en détail ce problème des groupes dans un paragraphe ultérieur.

A partir de la relation moyenne  $v_2 = f(v_1)$  on déduit aisément une carte des écarts  $v_2 - f(v_1)$ , ou des rapports  $v_2/f(v_1)$  qui permet de voir les différences par rapport à la loi moyenne des images.

**La visualisation colorée** permet de percevoir immédiatement les différences entre les images. **L'alternance** sur une console image est aussi très instructive sur les différences.

## 9.5 La comparaison de 3 images.

On peut étudier 3 images, en étudiant séparément les couples 2 à 2. Mais il existe des particularités à ce cas de 3 images dont la *Vision en vraie couleur* consistant à associer une couleur différente (Rouge, Vert, Bleu) sur une console image.

En astronomie on considère souvent les diagrammes de couleur. Si une des images est exprimée en magnitude V, la seconde en B et la troisième en U, on forme le diagramme des pixels ayant pour coordonnées (U-B),(B-V). Chaque zone de ce diagramme peut correspondre à un type d'objets. On visualise les cartes obtenues avec les pixels d'une zone particulière du diagramme.

## 9.6 Analyse d'une collection importante d'images.

Lorsque le nombre d'images dépasse 3, il devient impossible de pouvoir effectuer une comparaison globale, il faut introduire de nouvelles méthodes d'analyse. Plusieurs stratégies sont possibles :

**L'extraction de paramètres** en chaque point. Si  $V_n(x, y)$  désigne la valeur en chaque point de l'image  $n$ , on peut posséder une modélisation de  $V_n$  de telle sorte que l'on extrait en chaque point les paramètres correspondants par ajustement. Cette situation se rencontre dans l'analyse des données obtenues avec le Pérot-Fabry à balayage, où l'on extrait, par exemple le flux dans la raie étudiée, le continu et la vitesse radiale.

**L'analyse en composantes principales** des mesures pour réduire la dimension. Il s'agit de considérer pour l'image  $n$   $V_n(x, y)$  comme une variable aléatoire. On réduit la dimension en utilisant les corrélations entre les différents  $V_n$  (au sens statistique du terme). Ceci conduit à la décomposition de Karhunen-Loève avec la détermination d'une base orthogonale. On utilise la corrélation croisée des  $V_n$ .

**L'analyse des nuées** permet d'étudier le diagramme formé par tous les points de coordonnées  $V_n$ , pour un pixel donné. Il conduit à déterminer des classes dans ce diagramme.

**L'analyse non-linéaire** consiste dans la recherche de lois exprimant la valeur  $V_n$  en fonction des autres valeurs. En principe on travaille sur une classe donnée d'objets.

## 9.7 L'analyse en composantes principales.

Si nous considérons les  $V_n(x, y)$  comme une variable aléatoire pour laquelle chaque pixel correspond à une réalisation, on déduit aisément les paramètres statistiques suivants ( $Np$  est le nombre total de pixels) :

**Moyennes** :  $M_n = \sum_{x,y} V_n(x, y)/Np$

**Variances** :  $\sigma_n^2 = \sum_{x,y} (V_n - M_n)^2/Np$

**Variances croisées**  $\sigma_{nm}^2 = \sum_{x,y} (V_n - M_n)(V_m - M_m)/Np$

**Coefficient de corrélation** :  $c_{nm} = \frac{\sigma_{nm}}{\sigma_n \sigma_m}$

L'analyse en composantes principales consiste dans la déduction de paramètres non corrélés à partir de la connaissance de la matrice de corrélation. On montre que ces coefficients sont de la forme :  $W_k = \sum_n (V_n - M_n) C_{nk}$  où  $C_{nk}$  désigne la composante  $n$  du  $k$ ème vecteur propre de la matrice de corrélation. Pour obtenir les paramètres les plus porteurs de l'information on montre qu'il faut procéder par valeur propre décroissante.

Cette transformation des valeurs de l'image ne serait qu'un amusement mathématique si cela ne permettait pas de réduire la dimensionnalité. En effet, du fait de l'existence d'une corrélation, les vecteurs propres de rang élevé ne contiennent pratiquement plus d'informations. On ne considère le plus souvent que les 2 ou 3 premiers vecteurs, afin de pouvoir effectuer une comparaison 2 à 2 ou 3 à 3 des images obtenues à partir des  $W_k$ .

L'analyse en composantes principales réduit la dimension, mais aussi permet de localiser des pixels insolites. En effet, si la valeur propre diminue il peut subsister une fraction faible, mais significative de pixels anormaux.

L'histogramme des  $W_k(x, y)$  permet de déterminer le seuil à partir duquel les valeurs sont anormales. Il devient alors intéressants de visualiser l'image ainsi obtenue.

Dans une variante de l'analyse en composantes principales on cherche à exprimer l'ensemble des images  $V_n(x, y)$  sur une base orthogonale  $W_k(x, y)$  telle que la distance moyenne entre  $V_n(x, y)$  et la reconstitution sur la base  $U_n(x, y)$  soit minimum. Nous avons :

$$U_n(x, y) = \sum_k \lambda_k A_{nk} \cdot W_k(x, y)$$

où  $A_{nk}$  sont les composantes de  $V_n$  sur  $W_k$ .  $A_k$  et  $W_k$  sont orthonormées. Nous posons donc :

$$\sum_{n,x,y} (U_n(x, y) - V_n(x, y))^2 \text{ minimum}$$

Si l'on désigne par  $Cor(n, m)$  la matrice des intercorrélations entre  $V_n(x, y)$  et  $V_m(x, y)$ , soit

$$Cor(n, m) = \sum_{x,y} V_n(x, y) \cdot V_m(x, y)$$

Les coefficients  $C_{nk}$  doivent être les vecteurs propres de cette matrice, et  $\lambda_k^2$  représente ses valeurs propres.

De même que pour l'analyse en composantes principales on montre facilement qu'il faut procéder par valeur propre décroissante. La limitation du nombre de vecteurs de base provient du test sur le résidu entre les  $V_n$  et les  $U_n$ . Cette valeur peut s'obtenir aisément dans le processus de diagonalisation.

Par construction cette méthode tient compte du bruit des mesures et de la répartition du signal. Elle donc mieux adaptée au problème.

## 9.8 Détermination de la Base.

On montre que la base  $W_k(x, y)$  est de la forme :

$$W_k(x, y) = \frac{1}{\lambda_k} \sum_k A_{nk} \cdot V_n(x, y)$$

C'est-à-dire que c'est la moyenne pondérée des images initiales avec les composantes. De même on montre aussi que :

$$A_{nk} = \frac{1}{\lambda_k} \sum_{x,y} W_k(x,y) \cdot V_n(x,y)$$

C'est-à-dire que les composantes  $A_{nk}$  représentent bien les composantes de  $V_n(x,y)$  sur la base  $W_k(x,y)$ .  $A_{nk}$  et  $W_k(x,y)$  forment des vecteurs propres associés.

## 9.9 Diagramme des Images.

Si l'on possède de nombreuses images, il peut être utile de les positionner dans un diagramme associé (*espace des images*). Pour cela on considère le point de coordonnées des composantes de l'image  $n$ ,  $A_{nk}$ . Il y a autant de coordonnées qu'il y a de vecteurs de base.

Ce diagramme permet de mettre en évidence les diverses classes d'image, que l'on peut séparer avec l'analyse des nuées.

## 9.10 Diagramme des pixels.

De même on peut associer à tout pixel  $(x,y)$  un point dans *l'espace des pixels* dont les coordonnées sont les valeurs des composantes  $W_k(x,y)$ . Il y a autant de dimensions que de vecteurs de base et autant de points que de pixels.

On peut mettre ainsi en évidence des classes de pixels, que l'on peut ensuite visualiser.

Si le nombre de vecteurs de base est de 2, il est facile de déterminer les classes interactivement. Mais on doit, le plus souvent faire appel à l'analyse des nuées. L'étude des résidus  $V_n(x,y) - U_n(x,y)$  peut être un moyen de mettre en évidence les pixels anormaux.

## 9.11 Analyse des Nuées.

Considérons l'espace à  $N$  dimensions  $R_N$ ,  $N$  étant le nombre total d'images, dans lequel chaque pixel correspond à un point de coordonnées des valeurs de l'image  $V_n(x,y)$ . L'ensemble des points forme un nuage. Nous cherchons à décomposer ce nuage en plusieurs groupes ou classes. Plusieurs stratégies sont possibles :

**Pavage** : On estime que les valeurs  $V_n$  sont entachées d'erreur et l'on divise leur intervalle de variation en quelques segments. L'espace  $R_N$  sera donc divisé en un certain nombre de pavés. Tous les pavés ne contiennent pas nécessairement des pixels. Le problème réside donc dans la détermination des pavés remplis.

C'est la seule méthode qui permet de déterminer les classes d'objets qui au bruit près sont similaires. Le seul défaut consiste dans l'importance éventuelle du nombre de classes. Ce pavage est souvent utilisé comme première étape pour les autres méthodes. Celles-ci sont basées sur une notion de groupes physiques, de distribution plutôt gaussienne, monomodale de toute façon. On regroupe donc des pixels différents du point de vue du bruit, mais dont on pense qu'ils appartiennent à une même classe.

**La Classification Hiérarchique Ascendante** : elle consiste dans un regroupement progressif des points de  $R_N$  en se basant sur la distance entre classes.

Pour une étape donnée de l'algorithme nous avons un ensemble de classes  $C_i$ , définies d'une part par une position dans  $R_N$   $P_n(i)$  et un nombre de pixels associés  $M(i)$ . On calcule les distances interclasses  $D_{ij}$  et on regroupe les classes  $i$  et  $j$  correspondant à la plus faible distance. La classe nouvelle a pour position le barycentre des 2 classes, et pour poids la somme des pixels associés  $M(i)$  et  $M(j)$ .

On procède ainsi jusqu'à l'obtention d'une distance interclasse que l'on estime trop grande pour justifier un regroupement.

Après obtention des classes, on détermine pour chacune d'entre elles :

Les Centres  $C_n(i)$ ;

Les axes d'inerties conduisant à une matrice par classe  $Cor_{nm}(i)$ ;

Ceci permet de déduire pour un point de coordonnées  $X_n$  les probabilités d'appartenance à une classe  $P_i(X_n)$ . On peut attribuer la classe correspondant à la probabilité maximum, ou déduire une nouvelle probabilité d'appartenance à chaque classe de manière à avoir une somme égale à 1.

**Distance de Mahalanobis** : nous avons admis implicitement que la distance utilisée était euclidienne. D'autres distances peuvent être utilisées ( $L_0$ ,  $L_1$ ,  $L^p$ ), mais l'une des distances les plus utilisées est celle de Mahalanobis.

Nous avons introduit dans l'analyse factorielle une matrice de corrélation des valeurs  $V_n$ . Si on utilise les variables réduites :

$$R_n = \frac{V_n - M_n}{\sigma_n}$$

et si l'on désigne par  $C_{nm}^{-1}$  la matrice inverse de corrélation, la distance de Mahalanobis entre deux points de coordonnées réduites  $R_n$  et  $S_n$  est :

$$D = \sum_{nm} (R_n - S_n) C_{nm}^{-1} (R_m - S_m)$$

L'utilisation de cette distance évite la réduction préalable de la dimensionnalité avec l'analyse factorielle.

**Classification par rejet** : Avec l'ensemble des pixels, on déduit un centre  $M_n$  et une matrice de corrélation permettant de calculer une probabilité dans le cas d'une distribution gaussienne. Des pixels peuvent avoir une probabilité très faible. On est donc amené à les rejeter du nuage. Ce qui conduit à redéterminer un centre, une matrice de corrélation et une nouvelle probabilité d'appartenance pour les points restants. On itère jusqu'à convergence. La première classe est ainsi obtenue.

On reprend les pixels éliminés et on redétermine une nouvelle classe avec le même processus. Et cela jusqu'à utilisation de tous les pixels. Ce processus est lent, mais permet une bonne classification, si l'on a bien des nuages gaussiens.

**Analyse des nuées dynamiques** : dans les processus précédents la convergence peut être très longue. Avec l'analyse des nuées dynamiques, les choses peuvent aller beaucoup plus vite. La méthode est basée sur un processus itératif d'affectation de classe à chaque point du nuage. Au départ on choisit au hasard la classe à affecter à chaque point. Puis on calcule les centres et la matrice de corrélation, ce qui conduit aux probabilités d'appartenance à chaque classe. On modifie les affectations et on itère jusqu'à convergence. Le processus



marche d'autant mieux que le choix des classes initiales est correct. Un large pavage est une bonne technique.

**L'ajustement de la fonction de répartition** : la situation idéale consiste dans la détermination de la fonction de répartition des points du nuage sous la forme d'une somme de gaussiennes.

Le problème est déjà délicat à 1 dimension. Il faut avoir des paramètres assez proches des valeurs finales pour effectuer l'estimation correcte avec le principe du maximum de vraisemblance. Mais le problème est quasi inextricable à  $N$  dimensions, et c'est la raison pour laquelle seule des méthodes assez empiriques permettent cette détermination.

## 9.12 Présentation des résultats.

Après utilisation de ces méthodes on est capable d'associer à tout pixel une classe (éventuellement avec une probabilité d'appartenance). Ceci peut être visualisé sur une console en affectant à chaque classe une couleur différente. L'image obtenue contient alors toute l'information de la série d'images.

A partir de chaque classe une analyse morphologique peut être effectuée : domaines, aires, flux, etc. Une interprétation physique des classes peut résulter d'un modèle, agissant soit sur les centres des classes, soit cherchant à interpréter l'ensemble des paramètres descriptifs. Ce dernier type d'analyse est très difficile.

## 9.13 Réseaux de neurones artificiels.

Dans les années 70 le concept de réseau de neurones artificiels (RNA) a été introduit. Cela a permis de développer des méthodes très efficaces pour reconnaître l'appartenance à une classe à partir d'un apprentissage.

Un neurone  $N_i$  consiste dans un opérateur linéaire agissant sur  $N$  paramètres  $p_n$ . Il effectue une combinaison linéaire :

$$r_i = \sum_{n=1, N} \varpi_{in} p_n \quad (9.1)$$

$\varpi_{in}$  est le poids synapsique du réseau  $i$  pour le paramètre  $n$ . Après calcul de  $r_i$  le RNA prend la décision 0 ou 1 par seuillage. Si  $r_i > s_i$  on décide 1 et 0 dans le cas contraire. La décision est relative à l'appartenance à une classe.

Cette décision  $d_i$  peut être plus nuancée en introduisant une fonction d'activation, comme la fonction sigmoïde :

$$d_i = f(r_i) = \frac{1}{1 + \exp^{-\frac{r_i - s_i}{a_i}}} \quad (9.2)$$

Dans un réseau à une couche, ou perceptron, chacun donne une réponse selon ses poids. Le but consiste dans l'obtention d'une décision correcte de classification. Les poids et les seuils doivent être adaptés à cela. Leur optimisation s'effectue par une méthode de gradient. Si  $\delta_i$  est la décision correcte on montre qu'un procédant échantillon par échantillon, la correction des poids est liée aux écarts entre  $d_i$  et  $\delta_i$ .

Les perceptrons ne peuvent classer que dans le cas de structures simples. C'est pourquoi on a introduit des réseaux multi-couches. Après une première décision, on dispose de nouveau

paramètres sur lequel on effectue une nouvelle série de décisions avec des poids synaptiques adaptés.

Avec ce type de réseaux, il est possible, après apprentissage, de reconnaître des structures ayant une classification complexe.

## 9.14 Analyse des Mouvements.

Lorsqu'on compare des images pris à des instants différents, l'étude des mouvements est l'un des objectifs essentiels. L'analyse discriminante linéaire permet d'obtenir une approche dans le cas de mouvements de faible amplitude, indiquant parfaitement les zones variant en phase. Par contre dès que l'amplitude des mouvements augmente, en raison de la linéarité de la décomposition, trop de vecteurs propres sont nécessaires et l'interprétation des résultats n'est plus immédiat.

L'étude des mouvements s'effectue en utilisant plusieurs mécanismes, parfois simultanément :

**La détermination de la corrélation maximum** entre l'image de travail et une des images de référence. On détermine le décalage entre les images de manière à obtenir un maximum de corrélation.

L'image de référence peut être déterminée par la moyenne des images de travail. On peut alors procéder par itération, en redéterminant une nouvelle image de référence en tenant compte des décalages mesurés. Une précision remarquable peut être obtenue.

Cette méthode est globale. On peut la régionaliser de manière à déterminer un mouvement local.

**Recherche de traits caractéristiques** : sur chaque image on observe divers points caractéristiques, et on étudie séparément leur mouvement. Les points caractéristiques peuvent résulter de diverses approches : maxima, centres de domaines, points singulier d'un contour, centre de symétrie.

**Expression de l'image en fonction du temps** : on peut chercher à exprimer chaque image  $V(x,y,t)$  sous la forme :

$$V(x, y, t) = \sum_k Z_k(x, y) F_k(t)$$

où  $F_k(t)$  représente une base de fonctions (polynômes, trigonométriques, etc). De même que pour l'analyse discriminante linéaire, la base  $Z_k(x, y)$  est déterminée par moindres carrés.

L'analyse des mouvements se réduit donc en une analyse des images  $Z_k$ . On se ramène aux méthodes discutées précédemment.

La notion de *Flot Optique* a été introduite pour caractériser le champ des vecteurs de déplacement entre deux images.