

2. Les moments de Legendre comme descripteur de forme

Dans ce paragraphe, nous considérons la représentation d'une forme par sa fonction caractéristique. Cette fonction assigne la valeur 1 aux points contenus dans la forme et la valeur 0 à ceux situés à l'extérieur de celle-ci. Nous noterons Ω_{int} la région support de l'objet et Ω_{ext} le reste du domaine image, Ω (figure 1). Cette représentation est évidemment peu parcimonieuse et il est souhaitable de définir une caractérisation plus compacte de la forme. Cela revient à « encoder » la géométrie de l'objet en un nombre plus réduit de paramètres, formant un vecteur que nous appellerons descripteur de l'objet. Le principe général que nous adoptons consiste à projeter la fonction caractéristique de l'objet sur une base de fonctions analytiques, à support global sur le domaine image. Nous développons ici le cas des fonctions de bases polynomiales réelles, ce qui revient à décrire la forme par ses moments géométriques. En effet, cela permet de définir facilement une représentation invariante par translation et facteur d'échelle, propriété fortement souhaitable pour un descripteur. D'autre part, l'utilisation de fonctions de bases orthogonales limite la redondance d'information et offre donc une description « hiérarchique » de la forme, comme nous l'illustrerons à la fin de ce paragraphe. Nous utilisons les polynômes de Legendre, orthogonaux pour le produit scalaire associé à la fonction de poids unité. Notons que d'autres bases de polynômes, notamment complexes, sont envisageables. Citons par exemple les polynômes de Zernike [44] ou encore la base ART (*Angular Radial Transform*) introduite pour la norme de compression MPEG-7 [2, 39]. Ces bases, également orthogonales, présentent l'intérêt d'être naturellement invariantes par rotation, mais sont définies en coordonnées polaires dans le plan complexe, ce qui impose des traitements supplémentaires dans les schémas algorithmiques.

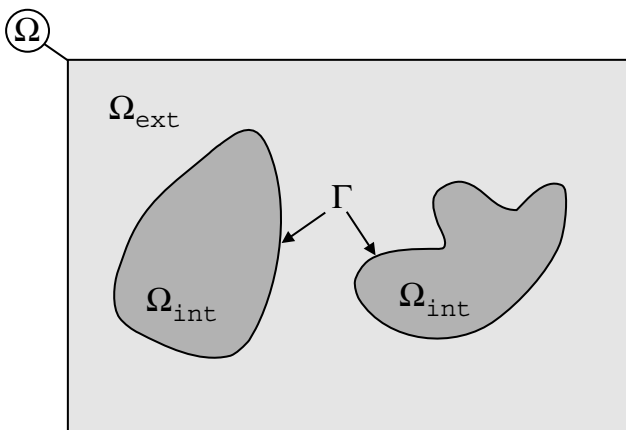


Figure 1. Partition de l'image.

2.1 Définition générale

En toute généralité une image peut être considérée comme une distribution bidimensionnelle $f(x, y)$ dont les valeurs sont l'intensité des pixels aux positions (x, y) . L'ensemble des moments géométriques associés, noté $\{M_{p,q}, (p, q) \in \mathbb{N}^2\}$, est défini par :

$$M_{p,q} = \iint_{\Omega} x^p y^q f(x, y) dx dy. \quad (1)$$

Il s'agit de la projection de la fonction f sur les monômes, $p + q$ représentant l'ordre d'un moment. La fonction étant bornée, continue par morceaux, à support fini et l'ensemble $\{x^p y^q\}$ constituant une base complète de polynômes (théorème d'approximation de Weierstrass [10]), l'ensemble des moments géométriques décrit de manière unique l'image f et réciproquement. Ces moments ont donc naturellement été utilisés comme descripteurs d'image dans des applications diverses comme la reconnaissance, la classification, l'estimation de pose ou encore la reconstruction.

Nous nous intéressons à la description d'une forme et, dans ce cas, comme nous l'avons mentionné en introduction, l'image $f(x, y)$ est la fonction caractéristique, binaire, de l'objet. La définition (1) s'écrit alors plus simplement :

$$M_{p,q} = \iint_{\Omega_{int}} x^p y^q dx dy. \quad (2)$$

2.2 Invariance par changement d'échelle et translation

Une propriété souhaitable pour un descripteur de forme est l'invariance lorsque celle-ci est soumise à des transformations géométriques. Nous nous intéressons ici à l'invariance par translation et changement d'échelle des moments géométriques. L'invariance par translation s'obtient simplement en effectuant un changement de repère. Le nouveau repère a son origine au centre de gravité (\bar{x}, \bar{y}) de la forme. On obtient alors ce qu'on appelle des moments centrés qui s'écrivent :

$$\iint_{\Omega_{int}} (x - \bar{x})^p (y - \bar{y})^q dx dy. \quad (3)$$

Considérons maintenant le cas du changement d'échelle. Si l'on dispose d'un objet binaire f et du même objet dans un rapport de taille γ , f' , alors on a la relation $f'(x, y) = f(x/\gamma, y/\gamma)$. En effectuant un changement de variable on a :

$$M'_{p,q} = \iint_{\Omega} x^p y^q f\left(\frac{x}{\gamma}, \frac{y}{\gamma}\right) dx dy = \gamma^{p+q+2} M_{p,q}. \quad (4)$$

À partir de cette équation on peut, par exemple, exprimer γ en fonction des moments $M'_{0,0}$ et $M_{0,0}$ (respectivement les aires des deux objets binaires f' et f) : $\gamma = (M'_{0,0}/M_{0,0})^{1/2}$. En remplaçant γ par cette valeur dans (4) on obtient :

$$\frac{M'_{p,q}}{(M'_{0,0})^{\frac{p+q+2}{2}}} = \frac{M_{p,q}}{(M_{0,0})^{\frac{p+q+2}{2}}}. \quad (5)$$

Nous obtenons ainsi des moments invariants par changement d'échelle, que l'on appelle *moments normés* par rapport à l'aire de l'objet, notée $|\Omega_{int}|$. On peut maintenant définir les moments géométriques centrés et normés de la fonction caractéristique d'une forme :

$$\eta_{p,q} = \iint_{\Omega_{int}} \frac{(x - \bar{x})^p (y - \bar{y})^q}{|\Omega_{int}|^{(p+q+2)/2}} dx dy. \quad (6)$$

2.3 Les moments de Legendre

Nous disposons à ce stade d'un descripteur de forme constitué par l'ensemble des moments géométriques centrés et normés, calculés à partir de la projection de la fonction caractéristique sur la base des polynômes $x^p y^q, (p, q) \in \mathbb{N}^2$. Les problèmes de redondance d'information dus à la non orthogonalité de cette base de projection sont bien connus [30, 44, 45]. Pour cette raison, nous utilisons la base des polynômes de Legendre, orthogonaux pour le produit scalaire associé à la fonction de poids la plus simple, la fonction unité. Ces polynômes, notés $P_n(x)$, définis sur l'intervalle $[-1, 1]$, sont donnés par la relation :

$$P_n(x) = \frac{1}{2^n n!} \times \frac{d^n}{dx^n} (x^2 - 1)^n, \quad (7)$$

et la condition d'orthogonalité est traduite par :

$$\langle P_n(x), P_m(x) \rangle = \int_{-1}^1 P_n(x) P_m(x) dx = \frac{2}{2m+1} \delta_{mn}. \quad (8)$$

Ceci permet de définir les moments de Legendre, en admettant en toute généralité que $\Omega \subset [-1, 1] \times [-1, 1]$:

$$\lambda_{p,q} = \frac{(2p+1)(2q+1)}{4} \iint_{\Omega_{int}} P_p(x) P_q(y) dx dy. \quad (9)$$

Dans la suite de cet article, nous noterons C_{pq} la constante de normalisation $(2p+1)(2q+1)/4$. Si l'on écrit $P_n(x) = \sum_{i=0}^n a_{ni} x^i$, il existe une relation linéaire entre les moments classiques et ceux de Legendre. On souhaite, par ailleurs, former un descripteur invariant par translation et changement d'échelle. Il suffit pour cela de calculer les moments de Legendre à partir des moments géométriques centrés et normés. La relation avec ces derniers s'écrit alors :

$$\lambda_{p,q} = C_{pq} \sum_{i=0}^p \sum_{j=0}^q a_{pi} a_{qj} \eta_{i,j}. \quad (10)$$

La propriété d'orthogonalité offre, d'une part, une représentation non redondante en information et moins sensible aux instabili-

tés numériques. D'autre part, elle permet une reconstruction analytique simple de la fonction image f :

$$f(x, y) = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \lambda_{p,q} P_p(x) P_q(y). \quad (11)$$

2.4 Capacité de description des moments de Legendre

Nous nous intéressons aux moments en tant que descripteurs de forme. Pour juger de leur capacité de représentation, nous examinons maintenant le problème de la reconstruction. Nous avons vu précédemment qu'il était possible théoriquement de reconstruire de manière univoque une image à partir de ses moments (11). Comme on peut le constater, en théorie, l'image reconstruite s'écrit comme une combinaison linéaire infinie de moments. Dans la pratique, on doit évidemment effectuer une troncature. Le problème qui se pose alors est le suivant : quel est l'effet de cette troncature sur la qualité de reconstruction, autrement dit, sur la capacité de description des moments de Legendre ? Si l'on note N l'ordre jusqu'auquel sont calculés les moments de Legendre, l'image reconstruite \hat{f}_N s'écrit :

$$\hat{f}_N(x, y) = \sum_{p=0}^N \sum_{q=0}^P \lambda_{p,q} P_p(x) P_q(y). \quad (12)$$

La figure 2 montre des résultats de reconstruction pour 4 objets binaires de référence, à des ordres N de troncature de plus en plus grands. Comme prévu, on vérifie que, quelles que soient la position et la taille de l'objet de référence, l'objet reconstruit est centré et son aire représente la même proportion du domaine image. Par inspection visuelle, on constate que l'ordre 40 est suffisant pour identifier les objets de références. Cependant pour l'image « main », une reconstruction à un ordre plus élevé semble nécessaire pour bien distinguer certains détails, comme l'index et le majeur. L'erreur commise lors de la reconstruction peut s'écrire comme la somme de deux termes [30]. L'un provient de la troncature, l'autre de l'approximation discrète des moments continus. Les comportements asymptotiques de ces termes sont antagonistes : pour un ordre de reconstruction tendant vers l'infini, l'erreur due à la troncature tend vers zéro tandis que celle due aux approximations augmente. Sur la figure 2, les reconstructions des images « carré » et « lady » à l'ordre 50 mettent en évidence les instabilités numériques liées au cumul des erreurs de discrétisation et à la précision arithmétique du calculateur. Ceci est encore plus flagrant si l'on calcule l'erreur quadratique moyenne normalisée de reconstruction en fonction de l'ordre de troncature. L'expression de cette grandeur est donnée par :

$$e^2(N) = \frac{\iint (\hat{f}_N(x, y) - f(x, y))^2 dx dy}{\iint f(x, y)^2 dx dy}. \quad (13)$$

Naturellement, avant de calculer cette quantité, la forme de référence est centrée et son aire est normalisée de la même façon

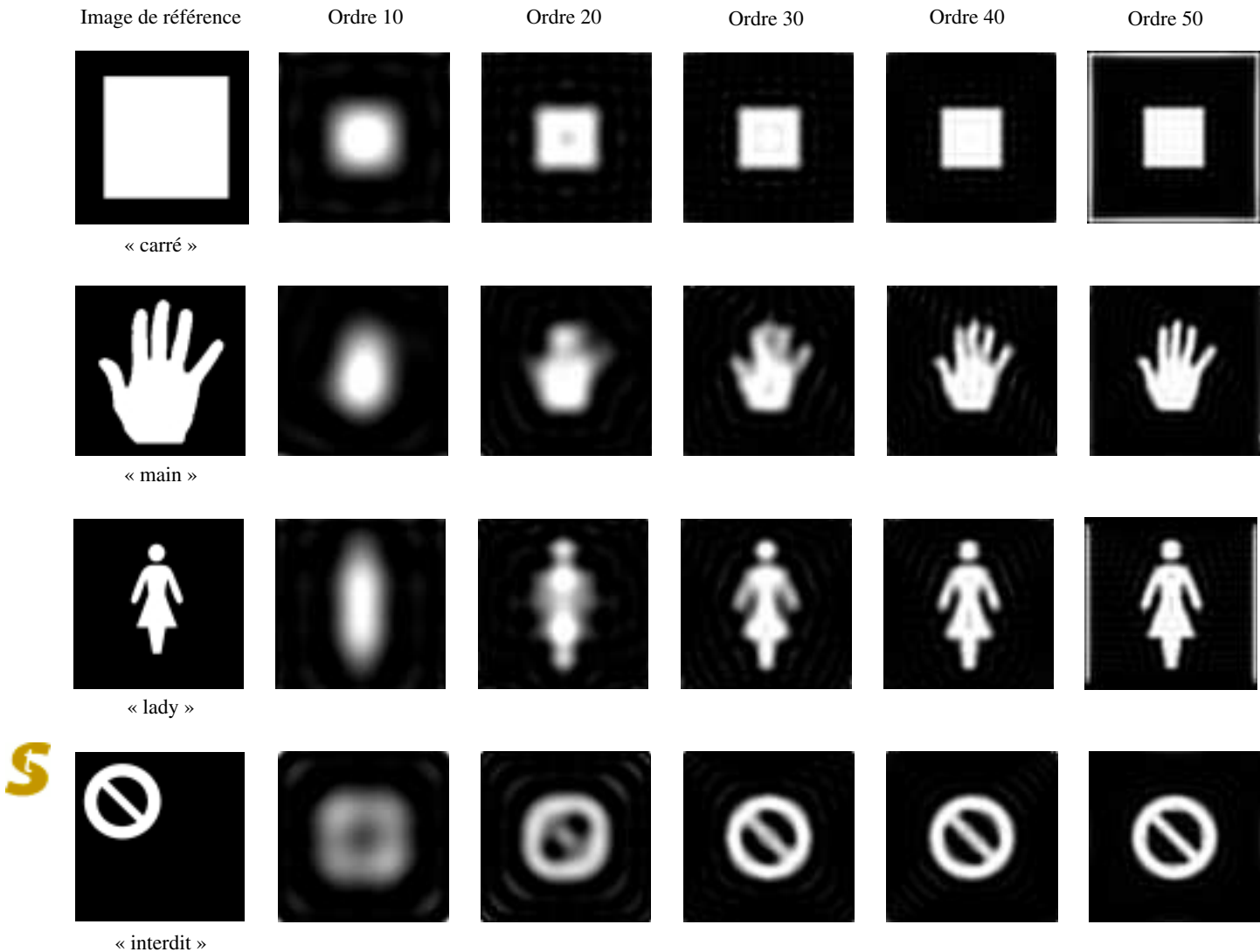


Figure 2. Reconstruction de formes à différents ordres. On peut constater l'invariance de la reconstruction à la position et à l'échelle de la forme de référence.

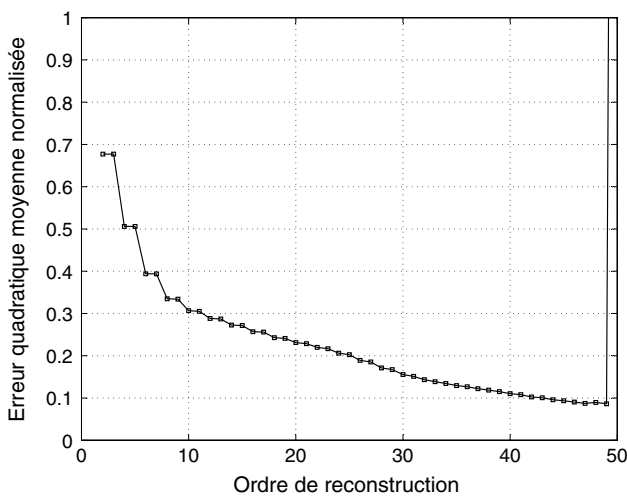


Figure 3. Erreur de reconstruction de l'image « lady ». À l'ordre 50, l'erreur (non visible sur le graphique avec cette échelle) vaut 5.13.

que l'image reconstruite. La figure 3 confirme pour l'objet « lady » l'amélioration constante de la qualité de la reconstruction avec l'augmentation du nombre de moments pris en compte, jusqu'à l'ordre 49. Au-delà de cet ordre, des instabilités apparaissent. Dans la pratique, nous nous servons de telles courbes pour déterminer l'ordre maximal pour la description d'un objet donné : pour tous ceux que nous avons étudiés, un ordre inférieur à 50 est suffisant. Des résultats similaires sont observés, par exemple, dans [30] pour des objets très complexes (caractères chinois).

2.5 Le descripteur de forme

Comme nous venons de le voir, on peut construire une description globale d'un objet à partir d'un nombre restreint de ses moments de Legendre. Nous « encoderons » donc une forme par les $((N + 1)(N + 2))/2$ premiers moments de sa fonction

caractéristique, calculables à partir des moments géométriques classiques centrés et normés à l'aide des relations (10) et (6). Nous les rangerons dans un vecteur appelé descripteur et noté λ_N ou plus simplement λ , l'ordre de troncature, N , n'apparaissant dans les calculs développés que si nécessaire.

3. Une contrainte géométrique globale pour les contours actifs orientés région

3.1 Définition d'un *a priori* de forme à partir des moments de Legendre

Nous sommes capables « d'encoder » une forme, même à topologie complexe, à partir des moments de Legendre de sa fonction caractéristique. Notre objectif est d'imposer à un contour actif une contrainte de forme géométrique. Nous allons donc considérer le *snake* comme délimitant une forme (le domaine Ω_{int} , cf. figure 1) et faire porter la contrainte sur celle-ci. Les moments de Legendre de la fonction caractéristique de la région intérieure au contour sont rangés dans le vecteur $\lambda(\Omega_{int})$ que l'on notera plus simplement λ . Si l'on se donne une forme de référence dont le descripteur, noté λ^{ref} , est connu, on peut définir une énergie *d'a priori* de forme de la façon suivante :

$$J_{forme}(\Omega_{int}) = d(\lambda, \lambda^{ref}). \quad (14)$$

Dans cette définition, l'énergie associée à la région Ω_{int} , donc au contour actif, est une distance d entre ses moments et ceux de la forme de référence. Minimiser cette énergie, c'est contraindre la forme définie par le *snake* à ressembler à la forme de référence. L'intérêt est que cette distance est basée sur un nombre réduit de paramètres, ne nécessite pas de mise en correspondance entre formes et bénéficie directement des invariances géométriques des descripteurs. Enfin, on en dérive relativement facilement une équation d'évolution pour le contour actif, comme nous allons le voir au paragraphe suivant.

3.2 Dérivation de l'équation d'évolution du contour actif

Nous avons défini un *a priori* de forme pour un contour actif au moyen d'une fonctionnelle d'énergie. Classiquement, l'étape qui suit est le calcul des variations de cette fonctionnelle pour obtenir une équation d'évolution portant sur le contour. Nous nous sommes placés dans une approche région : l'énergie dépend de

la partition de l'image, en d'autres termes, on voit apparaître dans sa formulation des intégrales doubles au travers des moments $\lambda_{p,q}(\Omega_{int})$. Plusieurs techniques ont été développées pour résoudre ce problème de calcul variationnel. L'une des premières présentées fut celle de Zhu et Yuille [49] : les intégrales de surface sont transformées en intégrales portant sur le contour à l'aide du théorème de Stokes. Ensuite sont déduites les équations d'Euler Lagrange associées à cette énergie. Enfin, ces équations sont plongées dans un schéma dynamique du type *descente de gradient* et l'on obtient une équation d'évolution pour le contour. D'autres auteurs [4, 35] ont utilisé une approche basée sur les ensembles de niveaux. Dans ce cas, l'implantation algorithmique est fixée dès le départ.

3.2.1 Définition et utilisation de la dérivée eulérienne

La méthode que nous nous proposons d'utiliser est équivalente à la première approche, comme démontré dans [1], mais plus simple à mettre en œuvre. Elle est basée sur la notion de dérivée de domaine, aussi appelée dérivée eulérienne, couramment utilisée en mécanique des milieux continus. L'introduction de ce formalisme est en grande partie liée aux travaux de Zolésio [15, 40]. En traitement d'images, il a été utilisé par Schnörr [36], pour le calcul du flot optique. Plus récemment, Jehan-Besson *et al.* [23] l'ont mis à profit pour définir un formalisme rigoureux de dérivation des équations d'évolution pour les contours actifs orientés région. Le principe général est simple : on cherche à dériver directement la fonctionnelle par rapport au domaine Ω_{int} . Cette approche permet d'une part d'éviter les problèmes liés à la transformation des intégrales régions en intégrales frontières ; d'autre part, elle reste générique dans le sens où elle est indépendante du mode d'implantation choisi. Voici le théorème fondamental que nous allons utiliser :

Théorème : La dérivée eulérienne de la fonctionnelle $F(\Omega(\tau)) = \iint_{\Omega(\tau)} k(\mathbf{x}, \Omega(\tau)) d\mathbf{x}$, dans la direction \vec{V} est la suivante :

$$\delta_E(F(\Omega(\tau))) = \iint_{\Omega(\tau)} \frac{\partial k}{\partial \tau}(\mathbf{x}, \Omega(\tau)) d\mathbf{x} - \int_{\Gamma(\tau)} k(\mathbf{x}, \Omega(\tau)) \langle \vec{V}, \vec{N} \rangle ds, \quad (15)$$

où \mathbf{x} est la position dans le plan, $\Omega(\tau)$ le domaine du plan qui évolue, $\Gamma(\tau)$ la frontière de ce domaine, \vec{V} le champ de vecteur vitesse dans lequel $\Omega(\tau)$ évolue et \vec{N} la normale unitaire à $\Gamma(\tau)$ dirigée vers l'intérieur du domaine.

Le théorème traduit le principe de conservation pour un domaine matériel (que l'on suit dans son mouvement). Le domaine d'intégration variant avec le temps τ , la dérivée de l'intégrale prend en compte le mouvement du domaine matériel, ce qui est exprimé par l'intégrale sur le contour qui apparaît dans le second membre de l'égalité, en plus de la variation de k sur le domaine.