

Transformée de Fourier-Mellin numérique - Reconstruction et estimation de mouvement d'objets à niveaux de gris

Stéphane Derrode, Faouzi Ghorbel

Groupe de Recherche Images et Formes
Ecole Nouvelle d'Ingénieurs en Communication
Institut National des Télécommunications
Cité scientifique - rue G. Marconi, 59658 Villeneuve d'Ascq
Tél : 03 20 33 55 16, Fax : 03 20 33 55 98, e-mail : derrode@enic.fr

RÉSUMÉ

La transformée de Fourier-Mellin est utilisée pour estimer les paramètres de mouvement relatifs entre des objets à niveaux de gris de même forme mais de taille et d'orientation différentes. Nous définissons une énergie basée sur la distance euclidienne exprimée sur l'espace de Fourier-Mellin et le théorème du retard appliqué au groupe des similitudes vectorielles. La localisation du minimum de cette énergie fournit une estimation des paramètres de mouvement. Les résultats expérimentaux confirment la robustesse des différentes approximations numériques du spectre de Fourier-Mellin et la fiabilité des paramètres de mouvement estimés.

1 Introduction

La description de formes planes à niveaux de gris est un problème important et difficile de l'analyse d'images. Elle trouve notamment application en reconnaissance de formes, en codage orienté-objet de séquences animées, en indexation d'objets dans une base de données multimédia.

Dans cette étude, nous nous intéressons à l'estimation de mouvement d'objets à niveaux de gris extraits d'une scène complexe par prétraitements. Nous supposons que le mouvement est approximativement décrit par une similitude (transformation géométrique composée de translation, rotation et homothétie).

Dans ce cas, la transformée de Fourier-Mellin (TFM) fournit une description pertinente des objets puisqu'elle convertit une similitude dans le domaine spatial en un déphasage dans le domaine de Fourier-Mellin. De nombreuses études portent d'ailleurs sur son utilisation en reconnaissance de formes [3]. Nous utilisons la TFM pour estimer les paramètres de mouvement entre deux objets de même forme mais de taille et d'orientation différentes.

Il est cependant bien connu que l'approximation de la TFM soulève des problèmes cruciaux (notamment de divergence de l'intégrale [4]). Nous proposons donc au préalable, 3 méthodes numériques d'estimation du spectre de Fourier-Mellin. Nous les comparons en termes de temps de calcul, de fiabilité de reconstruction de l'objet original et de robustesse de l'estimation des paramètres de mouvement.

2 La transformée de Fourier-Mellin analytique

La théorie des groupes et de l'analyse harmonique [1] permettent de généraliser la transformée de Fourier classique. Il a en effet été démontré que cette dernière est liée au groupe additif $(R,+)$. L'objet de cette généralisation

ABSTRACT

In order to assess motion parameters between gray-level objects having the same shape with distinct scale and orientation, we use the Fourier-Mellin transform. We define an energy based on the Euclidean distance expressed in the Fourier-Mellin domain and the shift theorem applied to the group of planar similarities. The location of the minimum energy gives the motion parameters. Experimental results are presented, confirming the robustness of the discrete estimations of the Fourier-Mellin spectrum and the reliability of the estimated motion parameters.

consiste en la définition d'une transformée de Fourier sur un groupe suffisamment régulier (groupe commutatif, groupe localement compact, ...).

2.1 La transformée de Fourier-Mellin

La transformée de Fourier-Mellin (TFM) est la transformée de Fourier sur le groupe des similitudes planes vectorielles $G = R_+^x \times S^1$, où S^1 est le cercle unité de R^2 . Ce groupe, muni de la loi de composition $(\alpha, \beta) \cdot (\rho, \theta) = (\alpha\rho, \beta + \theta)$ est commutatif et localement compact. Posons $d\mu_G = d\theta \cdot dr / r$, la mesure normalisée, invariante et positive du groupe. Notons $L^1(G, d\mu_G)$, l'espace vectoriel normé des fonctions à valeurs réelles définies sur G . Un élément de $L^1(G, d\mu_G)$ est une fonction définie en coordonnées polaires vérifiant :

$$\|f\|_{L^1(G, d\mu_G)} = \int_0^\infty \int_0^{2\pi} |f(r, \theta)| \frac{dr}{r} d\theta < \infty$$

La TFM d'une telle fonction s'écrit :

$$M_f(k, iv) = \frac{1}{2\pi} \int_0^{+\infty} \int_0^{2\pi} f(r, \theta) \cdot e^{-ik\theta} r^{-iv} \cdot \frac{dr}{r} d\theta \quad (1)$$

avec $k \in Z$ et $v \in R$

Il est bien connu en analyse harmonique que la TFM inverse est définie comme la transformée de Fourier sur le groupe dual de G , $\hat{G} = R \times Z$ où Z désigne l'ensemble des entiers relatifs. La mesure invariante de ce groupe est donnée par $d\mu_{\hat{G}} = dv \otimes d\delta$ où $d\delta$ représente la mesure de comptage sur Z . Si $M_f \in L^1(\hat{G}, d\mu_{\hat{G}})$, la TFM inverse existe et s'écrit :

$$f(r, \theta) = \int_{-\infty}^{+\infty} \sum_{k \in \mathbb{Z}} M_f(k, iv) \cdot e^{ik\theta} r^{iv} \cdot dv \quad (2)$$

avec $r \in \mathbb{R}_+^\times$ et $\theta \in [0; 2\pi[$

2.2 Le prolongement analytique de la TFM

Pour réaliser une description invariante par rapport aux translations, l'objet est souvent décrit en coordonnées polaires par rapport à son centre d'inertie. Or, dans ce cas, le support de l'objet n'est généralement pas contenu dans un compact de G , i.e. une couronne fermée, bornée, centrée à l'origine des coordonnées et privée de celle-ci. L'intégrale (1) diverge en général. Deux solutions ont été proposées pour résoudre ce problème. Soit annuler la fonction sur un disque de rayon ε où ε est suffisamment petit pour minimiser la perte d'information [2]. Soit modifier la fonction $f(r, \theta)$ par $g(r, \theta) = r^{\sigma_0} f(r, \theta)$ avec $\sigma_0 > 0$ [4]. Ainsi définie, g est intégrable sur G et sa TFM est appelée TFM analytique de f :

$$M_f(k, -\sigma_0 + iv) = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{+\infty} r^{\sigma_0} f(r, \theta) \cdot r^{-iv} e^{-ik\theta} \cdot d\theta dr / r \quad (3)$$

Cette transformée peut être vue comme la transformée de Laplace sur G . Elle admet une transformée inverse qui s'exprime par la formule suivante :

$$f(r, \theta) = \int_{-\infty}^{+\infty} \sum_{k \in \mathbb{Z}} M_f(k, s) \cdot r^s e^{ik\theta} \cdot ds \quad (4)$$

avec $s = -\sigma_0 + iv$, $\sigma_0 > 0$

La transformée de Fourier-Mellin analytique a permis de construire une famille de paramètres invariants par rapport aux translation, rotation et homothétie [4]. Cette famille est convergente et complète (à l'instar des descripteurs de Fourier-Mellin [3]). Ces deux propriétés ont pour conséquence de définir une vraie distance entre les objets.

3 Approximations numériques de la TFM et reconstruction des objets

Il est bien connu que l'approximation de la TFM soulève des problèmes numériques cruciaux. La TFM analytique a été introduite pour résoudre le problème de divergence de l'intégrale. Nous présentons 3 méthodes d'estimation numériques de cette transformée basées sur des ré-échantillonnages différents de l'image initialement décrite en coordonnées cartésiennes. Ces méthodes nécessitent une interpolation de la surface image [5].

3.1 Approximations de la TFM Analytique

1^{ère} méthode : elle consiste en une approximation directe de l'intégrale de Fourier-Mellin analytique (3) par échantillonnage de l'image en coordonnées polaires. L'intégration numérique est réalisée en utilisant une transformée de Fourier discrète (TFD) sur chaque cercle suivie d'une approximation de l'intégrale de Mellin :

$$\hat{F}_k(q\xi) = TFD(\hat{f}(q\xi, m\beta)) \text{ pour } k \in \{0, 1, \dots, N-1\}$$

$$\hat{M}_f(k, s) = \frac{\xi}{2\pi} \cdot \sum_{q=1}^{R \max/\xi} \hat{F}_k(q\xi) \cdot (q\xi)^{s-1}, \quad v \in \{-V..V\}$$

avec $s = -\sigma_0 + iv$, $V \in \mathbb{R}_+^\times$

ξ et β représentent respectivement les pas d'échantillonnage radial et angulaire.

2^{ème} méthode : Il est possible d'écrire l'intégrale de Mellin en une transformée de Fourier sur R par changement de variable log-polaire ($q = \ln r$). Ainsi, on peut écrire M_f comme :

$$M_f(k, s) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_0^{2\pi} e^{q\sigma_0} f(e^q, \theta) \cdot e^{-i(qv+k\theta)} \cdot d\theta dq$$

La TFM est alors estimée par une TFD bidimensionnelle de l'image en coordonnées log-polaires. Elle a été décrite dans [2].

3^{ème} méthode : Il est possible d'exprimer la TFM directement en coordonnées cartésiennes par changement de variable polaires - cartésiennes :

$$M_f(k, s) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \cdot (x+iy)^{-k} \cdot (x^2 + y^2)^{\frac{k-s}{2}-1} dx dy$$

Cette méthode ne nécessite pas de ré-échantillonnage de l'image. Par contre, elle requiert l'approximation numérique d'une double intégrale.

3.2 Comparaison des différentes méthodes d'estimations

Nous avons estimé le spectre de l'objet de la figure 1 par les trois méthodes présentées ci-dessus. Les figures 2, 3, 4

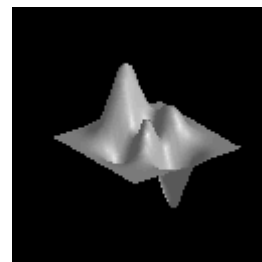


Figure 1 : Image représentant un objet à niveaux de gris

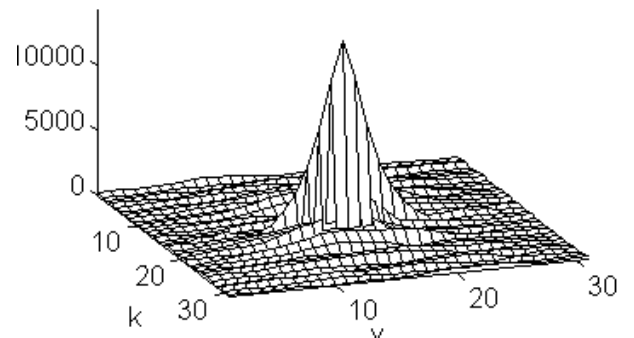


Figure 2 : Amplitude du spectre de Fourier-Mellin de l'objet de la fig. 1 estimé par la 1^{ère} méthode.

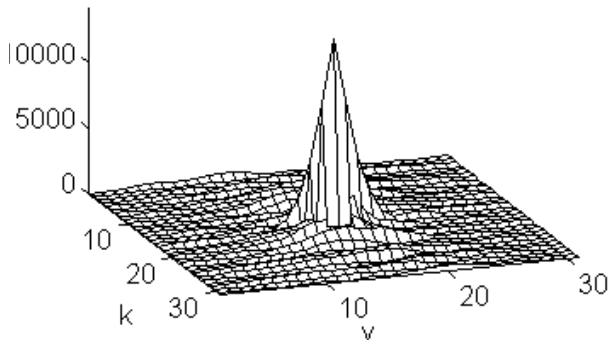


Figure 3 : Amplitude du spectre de Fourier-Mellin de l'objet de la fig. 1 estimé par la 2^{ème} méthode.

représentent l'amplitude de la partie centrale du spectre obtenu (32*32 composantes).

La figure 5 présente les reconstructions de l'objet de la figure 1 obtenues à l'aide des 3 approximations de la TFMA inverse.

L'approximation numérique de la TFM analytique par changement de variable log-polaire fournit la reconstruction la plus fidèle des 3 méthodes (figure 5b). Cela s'explique par l'utilisation de transformées discrètes exactes (TFD et TFD Inverse). Elle est également plus rapide que les 2 autres pour les mêmes raisons (FFT 2D). Cependant, des expériences en comparaison de formes [4] et en estimation des paramètres de mouvement (paragraphe suivant)

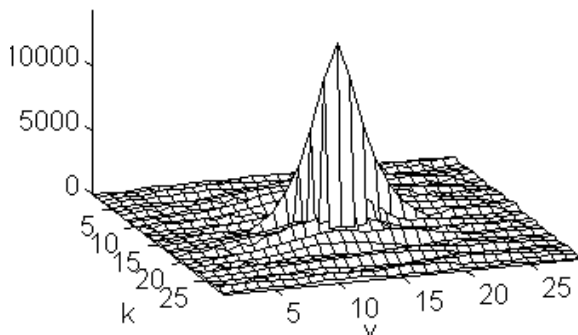
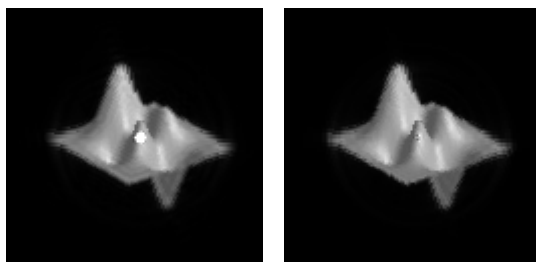
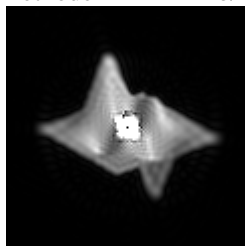


Figure 4 : Amplitude du spectre de Fourier-Mellin de l'objet de la fig. 1 estimé par la 3^{ème} méthode.



a/ 1^{ère} méthode

b/ 2^{ème} méthode



c/ 3^{ème} méthode

Figure 5 a, b, c : Reconstruction de l'objet de la fig. 1 par les 3 méthodes d'estimation de la TFM analytique en utilisant seulement 32*32 composantes.

montrent que le spectre est moins bien estimé que celui obtenu par la 1^{ère} méthode (figure 2). La 3^{ème} méthode s'avère, quand à elle, peu pertinente puisque ni le spectre ni l'objet reconstruit ne sont bien estimés (figures 4 et 5c). Elle est, de plus, particulièrement coûteuse en temps de calcul.

Aux vues des résultats, la méthode d'estimation basée sur la représentation polaire à pas radial constant de l'image (1^{ère} méthode) est la plus robuste. Elle semble en effet fournir le meilleur compromis entre stabilité numérique et temps de calcul.

4 Estimation des paramètres de mouvement

Dans ce qui suit, nous décrivons une méthode d'estimation originale des paramètres de mouvement. Cette méthode est basée sur le calcul de la distance euclidienne exprimée dans l'espace de Fourier-Mellin. Elle présente un certain nombre de propriétés souhaitables : rapidité, robustesse...

4.1 Distance euclidienne entre objets

D'après le théorème de Plancherel-Parseval appliqué au groupe des similitudes planes, la TFM analytique est une isométrie de $L^2(G, d\mu_G)$ dans $L^2(\hat{G}, d\mu_{\hat{G}})$. Par conséquent, la distance entre deux objets f et g supposés dans $L^1(G, d\mu_G) \cap L^2(G, d\mu_G)$ s'exprime de manière strictement équivalente sur ces deux espaces :

$$\begin{aligned} d_2(f, g) &= \sqrt{\int_0^{+\infty} \int_0^{2\pi} |r^{\sigma_0} g(r, \theta) - r^{\sigma_0} f(r, \theta)|^2 \cdot dr/r \cdot d\theta} \\ &= \sqrt{\sum_{k \in \mathbb{Z}} \int_{-\infty}^{+\infty} |M_g(k, s) - M_f(k, s)|^2 \cdot ds} \end{aligned} \quad (5)$$

D'un point de vue numérique, l'estimation de la distance sur l'espace de Fourier-Mellin s'avère plus rapide et mieux adaptée. En effet, l'énergie du spectre est concentrée au voisinage de l'origine (théorème de Riemann-Lebesgue). La troncature du spectre de Fourier-Mellin (inévitable d'un point de vue numérique) a donc une influence modérée sur la qualité de l'estimation de cette distance, contrairement au calcul de cette même distance sur l'espace original où tous les pixels des objets doivent être considérés.

4.2 Notion de forme et théorème du retard

On dit que deux objets f_1 et f_2 de $L^1(G) \cap L^2(G)$ ont la même forme au sens de la transformation géométrique similitude si et seulement si :

$$\exists(\alpha, \beta) \in G, \quad \forall(r, \theta) \in G, \quad f_2(r, \theta) = f_1(\alpha r, \theta + \beta)$$

Une forme est donc définie comme la classe des objets se différenciant uniquement par une similitude vectorielle. Pour de tels objets, le théorème du retard s'écrit :

$$M_{f_2}(k, iv) = \alpha^{iv} e^{ik\beta} M_{f_1}(k, iv)$$

En tenant compte des modifications apportées à f et g

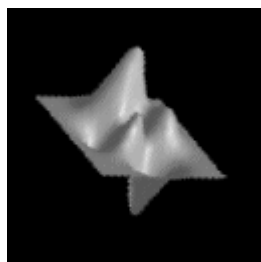


Figure 6 : Objet identique à l'objet de la figure 1 mais agrandi d'un facteur $\alpha = 1.1$ et tourné d'un angle $\beta = 20^\circ$.

pour les raisons de convergence justifiées dans le paragraphe 2.2, les équations (6) et (7) deviennent :

$$\begin{cases} r^{\sigma_0} f_2(r, \theta) = r^{\sigma_0} f_1(\alpha r, \theta + \beta) \\ M_{f_2}(k, s) = \alpha^s e^{ik\beta} M_{f_1}(k, s) \end{cases} \quad (6)$$

Une rotation d'angle β et un facteur d'échelle α appliqués à l'image dans le domaine spatial se traduisent par un déphasage d'angle $k\beta + v \ln \alpha$ et un facteur d'amplitude $\alpha^{-\sigma_0}$ dans le domaine fréquentiel.

4.3 Estimation des paramètres de mouvement

Estimer le mouvement relatif de deux objets f et g de même forme revient à déterminer les paramètres α et β de la similitude qui les distingue. Un algorithme consiste à :

- appliquer une similitude de paramètres (ρ, ψ) sur f ,
- comparer l'objet ainsi obtenu avec le second objet.

Si les objets sont identiques alors $(\rho, \psi) = (\alpha, \beta)$, sinon on réitère l'opération pour d'autres couples (ρ, ψ) .

La difficulté réside dans le choix d'une méthode de comparaison des objets qui soit suffisamment robuste sans être trop coûteuse en temps de calcul. Pour cela, on propose de modifier la distance euclidienne entre objets (5) pour tenir compte de la transformation appliquée à f . On définit alors une énergie dépendant de (ρ, ψ) :

$$E_{fg}(\rho, \psi) = \int_0^{+\infty} \int_0^{2\pi} |r^{\sigma_0} g(r, \theta) - r^{\sigma_0} f(\rho r, \theta + \psi)|^2 \frac{dr}{r} d\theta$$

Les paramètres de mouvement sont obtenus en minimisant cette énergie :

$$(\alpha, \beta) = \text{Arg min}_{\rho, \psi} (E_{fg}), \quad \rho \in \mathbb{R}_+^\times \text{ et } \psi \in [0; 2\pi[\quad (7)$$

En tirant parti de la distance euclidienne (Eq. 5) et du théorème du retard (Eq. 6), cette énergie se réécrit :

$$E_{fg}(\rho, \psi) = \sqrt{\int_{-\infty}^{+\infty} \sum_{k \in \mathbb{Z}} |M_g(k, s) - \rho^s e^{ik\psi} M_f(k, s)|^2 \cdot dv} \quad (8)$$

4.4 Exemple numérique

La figure 7 représente la surface d'énergie entre les objets des figures 1 et 6. Les spectres des objets sont estimés par la 1^{ère} méthode numérique. Il apparaît très nettement que l'énergie présente un minimum dont les coordonnées

coïncident avec les paramètres de mouvement fixés pour l'expérience, à $\alpha = 1.1$ et $\beta = 20^\circ$.

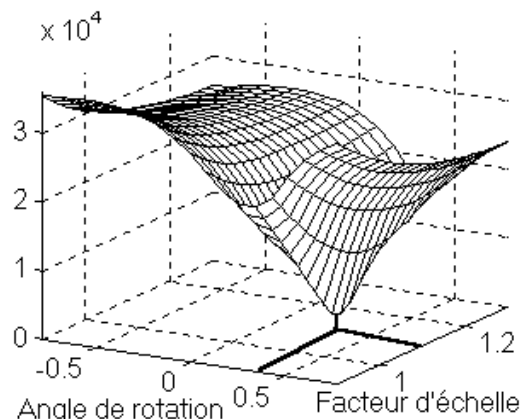


Figure 6 : Surface d'énergie entre les objets des figures 1 et 6. Les spectres sont estimés par la 1^{ère} méthode. Le minimum de la surface se situe au point $\rho = 1.1$ et $\psi = 20^\circ$.

5 Conclusion

Nous avons utilisé la TFM analytique pour estimer les paramètres de mouvement entre objets de même forme mais de taille et d'orientation différentes. Nous avons construit une énergie qui, lorsqu'elle est minimisée, fournit les paramètres de mouvement. Parmi les 3 méthodes d'approximation de la TFM, nous avons retenu celle qui offrent le meilleur compromis entre temps de calcul, fiabilité d'estimation du spectre et qualité de reconstruction.

Les résultats numériques permettent d'envisager son utilisation dans un schéma de codage orienté-objet de séquences animées, et en indexation, par le contenu, d'objets à niveaux de gris.

Les auteurs tiennent à remercier G. Eude, O. Avaro (CNET - Issy Les Moulineaux) et H. Sanson (CCETT - Cesson-Sévigné) pour leur soutien financier.

6 Références

- [1] J. Dieudonné. "Eléments d'Analyse Moderne", Gauthier-Villars, 1974.
- [2] PE. Zwicke, Z. Kiss. "A new Implementation of the Mellin Transform and Its Application to Radar Classification", IEEE Trans. on PAMI, Vol. 5 N°2, pp 191-199, March 1983.
- [3] Y. Sheng, HH. Arsenault. "Experiments on Pattern Recognition using Invariant Fourier-Mellin Descriptors", J. of Opt. Soc. Am. A, Vol. 3, N° 6, pp 771-776, June 1986.
- [4] F. Ghorbel. "A Complete Invariant Description for Grey-Level Images by the Harmonic Analysis Approach", PRL, Vol. 15, pp 1043-1051, Oct. 1994.
- [5] HE. Bez. "On Analysis of Symmetry for Plane Curves and Plane Curve Algorithms", CAGD, Vol. 9, pp 125-142, 1992.