Robust and efficient Fourier-Mellin transform approximations for invariant grey-level image description and reconstruction

Stéphane Derrode, F. Ghorbel

HAL Id: hal-00325095

https://hal.science/hal-00325095

Submitted on 26 Sep 2008

Robust and efficient Fourier-Mellin transform approximations for gray-level image reconstruction and complete invariant description

Stéphane Derrode

Corresponding author,
Département Image et Traitement de l'Information,
École Nationale Supérieure des Télécommunications de Bretagne,
Technopôle Brest-Iroise, BP 832, 29285 Brest Cedex, France.
Phone: int+ 33 2 29 001 057; Fax: int+ 33 2 29 001 098

E-mail: stephane.derrode@enst-bretagne.fr

and

Faouzi Ghorbel

Laboratoire Cristal, pôle GRIFT,
École Nationale des Sciences de l'Informatique,
Domaine universitaire de La Manuba,
2010 La Manuba, Tunisia.
Phone: int+ 216 1 600 444; Fax: int+ 216 1 600 449

E-mail: faouzi.ghorbel@ensi.rnu.tn

This paper addresses the gray-level image representation ability of the Fourier-Mellin Transform (FMT) for pattern recognition, reconstruction and image database retrieval. The main practical difficulty of the FMT lies in the accuracy and efficiency of its numerical approximation and we propose three estimations of its analytical extension. Comparison of these approximations is performed from discrete and finite-extent sets of Fourier-Mellin harmonics by means of experiments in: (i) image reconstruction via both visual inspection and the computation of a reconstruction error; and (ii) pattern recognition and discrimination by using a complete and convergent set of features invariant under planar similarities.

Experimental results on real gray-level images show that it is possible to recover an image to within a specified degree of accuracy and to classify objects reliably even when a large set of descriptors is used. Finally, an example will be given, illustrating both theoretical and numerical results in the context of content-based image retrieval.

1

1. INTRODUCTION

Recent applications in the field of image analysis such as object-oriented video coding, content-based retrieval and digital image watermarking, induced a renewal of interest for scenes representation. The latter requires to precise the shape of objects in the scene and the geometrical transformations that act globally on the objects [1, 2]. A lot of works to date were therefore motivated for the search of global features invariant to some geometrical transformations. In the 2D case, planar closed contours and gray-level objects have been widely studied.

In the case of planar contours, several similarity-invariant and starting-point independent Fourier descriptors have been proposed for closed-curve description under planar Euclidean motions [3]. Refinements of such sets have then been defined using completeness and stability properties. A stable set of invariants insures that small shape variations do not induce a noticeable change in the feature values [4], whereas a complete set guarantees that an object is fully identified up to a certain transformation [5]. Completeness is recognized as an important criterion for full shape discrimination and reconstruction from features [6].

In the case of 2D gray-level objects, similarity-invariant description can be achieved with moment invariants. Hu, in a landmark paper [7], proposed seven invariants based on combinations of low-order geometric moments. In order to extract a complete set of rotation-invariant features, research has focused on moment invariants computed from images expressed in polar rather than rectangular coordinates. Reddi [8] proposed using radial and angular moments, but this attempt was not sufficient to attain completeness. Abu-Mostafa and Psaltis [9] suggested the use of complex moments. Since a simple relation connects complex moments of similar objects, a normalization process [10] may be used to derive a complete set of rotation- and scale-invariant moments. However, it is well known that the recovery of an image from geometric moments is numerically unstable and computationally expensive [11]. Image reconstruction from a finite set of moment descriptors has been addressed successfully only with orthogonal moments because of their straightforward inverse transform and their good reconstruction capabilities (see works on Legendre, Zernike, and related moments [12, 13, 14], and disk harmonic coefficients [15]).

Late on the 70's, the optical research community introduced the Fourier-Mellin Transform (FMT) for pattern recognition [16, 17] and it was later used in digital signal and image processing [18, 19, 20]. Several sets of rotation- and scale-invariant features based on the FMT modulus have been designed, but completeness was not reached since the descriptors do not hold phase information. Furthermore, it was also pointed out that numerical estimation of the Mellin integral brings up crucial difficulties [21]. Recently, a solution for the convergence of the integral has been given by using the Analytical Fourier-Mellin Transform (AFMT), and a complete set of similarity-invariant features for planar gray-level images was proposed [22]. But, it is essentially impossible to realize an exact estimation since image data originates as sample measurements on a 2D grid. Therefore the accuracy and efficiency of different approximations is of practical interest and can benefit to recent applications of the FMT such as content-based image retrieval [23] and digital image watermarking [24].

In this paper, emphasis is placed on the numerical image representation ability of the AFMT for pattern recognition and gray-level image reconstruction purposes. The general lay-out is as follows: The direct and inverse AFMT and some of its relevant properties are briefly reviewed in Section 2. Three equivalent expressions of AFMT integrals are used for deriving three numerical approximations of the Fourier-Mellin transform in Section 3. The approximations, namely, the direct-, the fast-, and the Cartesian-AFMT estimations, differ in resampling methods of the discrete image on polar, log-polar and cartesian grids and in numerical integration performed on these grids. Then, the comparison between discrete and finite-extent Fourier-Mellin spectrums is conducted in two ways. In Section 4, the inverse problem of reconstructing an image from its numerical spectrum is examined via both visual inspection and the computation of a reconstruction error. In Section 5, pattern recognition and content-based image retrieval regardless of the orientation and scale of gray-level images is tested with the computation of a real distance derived from a complete and convergent set of AFMT similarity invariant descriptors.

2. THE ANALYTICAL FOURIER-MELLIN TRANSFORM

The Fourier-Mellin transform is mainly concerns with the study of similarity transformations (direct product of the rotation and scale groups). Harmonic analysis on group provides the basis for the definition of other Fourier transforms suited to more general groups of geometrical transformations. Interesting results have been brought to the image-processing field by several authors. For examples, Gauthier et al [25] worked on the Fourier transform in the case of 2D and 3D motion groups (semi-direct product of the translation and rotation groups in 2 and 3 dimensions), Ghorbel [2] focus in particular on Fourier transforms defined on 2D and 3D parameterizations for affine transformations, and, in a recent paper, Turski [26] proposed some Fourier transforms for sub-groups of the affine group (including shearing transformations). Other references include [27, 28, 29]. It was often pointed out that there is still a lot of works to do regarding practical considerations for image analysis (approximation, fast algorithms, ...). These considerations remain crucial in the case of the well-known FMT.

Throughout this paper, we shall denote by $\mathbb{Z}^1$ the additive group of integers, $\mathbb{R}$ the additive group of the real line, $\mathbb{R}_+^*$ the multiplicative group of positive and non-zero real numbers, and $\mathbb{S}^1$ the unit circle of the plane $\mathbb{R}^2$. All these groups are locally compact and commutative. The direct product $\mathbb{R}_+^* \times \mathbb{S}^1$ forms a locally compact and commutative group under the following law : $(\alpha, \theta) \circ (\rho, \psi) = (\alpha\rho, \theta + \psi)$.

2.1. The standard Fourier-Mellin transform

Let f denote a function representing a gray-level image defined over a compact set of $\mathbb{R}^2$. The standard Fourier-Mellin transform of f is given by:

$$\forall (k, v) \in \mathbb{Z} \times \mathbb{R}, \quad \mathcal{M}_f(k, v) = \frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} f(r, \theta) \, r^{-iv} \, e^{-ik\theta} \, d\theta \frac{dr}{r}. \qquad (1)$$

f is assumed to be summable over $\mathbb{R}_+^* \times \mathbb{S}^1$ under the measure $d\theta \frac{dr}{r}$, i.e.

$$\int_0^\infty \int_0^{2\pi} \left| f(r,\theta)\, r^{-iv}\, e^{-ik\theta} \right| d\theta \frac{dr}{r} = \int_0^\infty \int_0^{2\pi} \frac{1}{r}\, f(r,\theta)\, d\theta dr < \infty, \qquad (2)$$

since f is positive.

The FMT is a global transform and applies on all pixels in the same way. Textured images cannot be taken into account directly and objects must first be localized and isolated from the scene to match one of the requirements regarding the existence of the integral in Eq. (1).

To be computed, there is also a need to define a center of coordinates. This center shall be independent to translation in order to make the FMT representation shift-invariant. Translation invariant positioning procedures, such as the auto-correlation function or the modulus of Fourier transform, make the representation incomplete, which is not satisfactory regarding the objectives, i.e. reconstruction and complete invariant description. For isolated objects, the center of mass is certainly the easiest way to set the origin of coordinates, even if others specific shape points can be used.

Another problem associated with the computation of the FMT, that has been less addressed, concerns the existence of the integral too. From Eq. (2), the FMT exists for functions which are equivalent to Kr^σ in the neighborhood of the origin, for some constants K and $\sigma > 0$. However, an image function does not meet this requirement since, in the vicinity of the center of coordinates, f is equivalent to the gray value of the centroid, which is generally different from 0. As a result, f is not summable and the integrals in (1) diverge [22].

2.2.   The analytical Fourier-Mellin transform and inverse transform

Due to the singularity at the origin of coordinates, a solution generally adopted is to cancel the image over a small disk around the origin [21]. However, this approximation has serious effects on the numerical computation of the FMT because of the following reasons:

• Image values nearer the origin have a larger effect on the FMT than image values remote from the centroid because of the $1/r$ weighting in the measure of the Fourier-Mellin integrals. Hence, a significant information content of the image is lost in addition to removing a small disk in the image centroid.

• It may cause stretching problems when images are enlarged : how large must the disk be if the image is scaled by an unknown factor ? By cancelling a disk of constant radius for every image, different amounts of information are removed.

More recently, a rigorous approach has been introduced to tackle the difficulties described above. According to Eq. (2), Ghorbel [22] suggested computing the standard FMT of $f_\sigma(r,\theta) = r^\sigma f(r,\theta)$ instead of $f(r,\theta)$, where σ is a fixed and strictly positive real number. Hence, the integral (1) exists and is called the Analytical Fourier-Mellin Transform (AFMT) of f:

$$\forall (k,v) \in \mathbb{Z} \times \mathbb{R}, \quad \mathcal{M}_{f_\sigma}(k,v) = \frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} f(r,\theta)\, r^{\sigma-iv}\, e^{-ik\theta}\, d\theta \frac{dr}{r}, \qquad (3)$$

with $\sigma > 0$. The AFMT gives a unique representation of images and the Inverse AFMT (IAFMT) enables f to be retrieved as:

$$\forall (r, \theta) \in \mathbb{R}_+^* \times \mathbb{S}^1, \quad f(r, \theta) = \int_{-\infty}^{+\infty} \sum_{k \in \mathbb{Z}} \mathcal{M}_{f_\sigma}(k, v) \, r^{-\sigma + iv} \, e^{ik\theta} dv. \tag{4}$$

$\mathcal{M}_{f_\sigma}$ is assumed to be summable over $\mathbb{Z} \times \mathbb{R}$.

It is worth noting that this general form of the FMT can be found in mathematical literature on harmonic analysis [30], but less attention has been devoted to its use in image processing. The AFMT is close to the Laplace transform defined under the planar similarity group in the group theory and harmonic analysis viewpoint [2]. In the following, we recall some basic properties of the AFMT that will be used latter.

2.3. Gray-level shape and shift theorem for planar similarities

For pattern recognition, the interest in the standard FMT comes essentially from its shift theorem under planar similarities. Let derive now the transformation law in the AFMT case. Let g denote the rotation and size change of an object f through angle $\beta \in \mathbb{S}^1$ and scale factor $\alpha \in \mathbb{R}_+^*$, i.e. $g(r, \theta) = f(\alpha r, \theta + \beta)$. The two objects f and g have the same shape and will be termed similar objects. The AFMT of g is

$$\forall (k, v) \in \mathbb{Z} \times \mathbb{R}, \quad \mathcal{M}_{g_\sigma}(k, v) = \frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} f(\alpha r, \theta + \beta) \, r^{\sigma - iv} \, e^{-ik\theta} \, d\theta \frac{dr}{r},$$

and performing a simple change of the variables r and θ gives the following relations:

$$\forall (k, v) \in \mathbb{Z} \times \mathbb{R}, \quad \mathcal{M}_{g_\sigma}(k, v) = \alpha^{-\sigma + iv} \, e^{ik\beta} \, \mathcal{M}_{f_\sigma}(k, v). \tag{5}$$

While the classical Fourier transform converts translation into a pure phase change, Eq. (5) shows that the AFMT converts a similarity transformation in the original domain into a complex multiplication in the Fourier-Mellin domain. These relations can be seen as the shift theorem for the planar similarity group, and make the AFMT appropriate for extracting features that are invariant to scale and rotation changes. However, the usual modulus-based FMT descriptors are no longer invariant to scale because of the $\alpha^{-\sigma}$ term. A normalization process is used in Section 5 to derive a complete set of AFMT invariant descriptors.

3. THREE APPROXIMATIONS TO THE AFMT

While the analytical extension of the AFMT can be used to solve the crucial problem of the convergence of the FMT integrals for gray-level images, it is still of practical interest to propose approximation methods to the AFMT since no exact transform can be derive. Indeed, the direct and inverse AFMT are valid for functions with real-valued parameters ($r \in \mathbb{R}_+^*$, $\theta \in \mathbb{S}^1$, and $v \in \mathbb{R}$). As no discrete version exists, we propose three methods to estimate the direct and inverse AFMT pair. The approximations differ (i) in discrete image resampling into polar, log-polar, or rectangular grids and, (ii) in the way numerical integration of the gray-level function is performed. The performance of the methods is evaluated by means of reconstruction and pattern discrimination experiments in Sections 4 and 5.

3.1. Discrete and finite-extent AFMT

An object $f(p, q)$ is considered as a discrete image defined on a rectangular grid with finite spatial extent (Fig. 1). Since f vanishes outside a bounded domain of the 2D plane, its AFMT is not band-limited and requires an infinite number of harmonics to fully describe the object. However, we can only deal with a finite set of harmonics in practice, with the result that a part of the original image content is lost. According to the Riemann-Lebesgue theorem, the FMT tends towards zero as $|v|$ and $|k|$ tend to infinity, and the information lost by numerical truncation can be made as weak as desired for a given application.

The discrete and finite-extent AFMT $\widehat{\mathcal{M}}_{f_\sigma}(k, v)$ is computed for $k \in [-\mathrm{K}, \mathrm{K}]$ and $v \in [-\mathrm{V}, \mathrm{V}]$, with a sampling step value over axis v set to 1. Following the recommendation of Goh [31], σ is set to 0.5. For real-valued functions, the AFMT is symmetrical:

$$\mathcal{M}_{f_\sigma}(-k, -v) = \overline{\mathcal{M}_{f_\sigma}(k, v)},$$

where the overline denotes complex conjugation. The effective size $S_{\mathrm{K,V}}$ of this representation is one half of the size of the AFMT due to the symmetry property. Finally, $\widehat{f}_{\mathrm{K,V}}(p, q)$ indicates the reconstructed version of $f(p, q)$ from a finite set of Fourier-Mellin harmonics. The selection of the ranks K and V may be guided in terms of image spectral content and application, but no general rule can be defined.
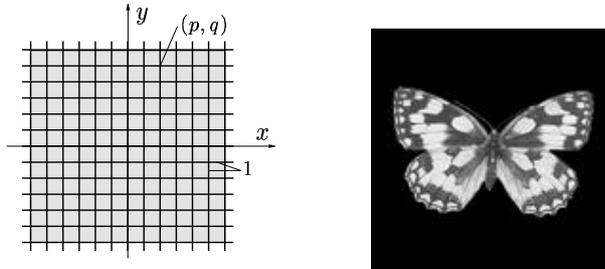

FIG. 1.   Cartesian grid and butterfly image used to illustrate the three approximation methods of the AFMT.

3.2. The direct algorithm

The direct AFMT approximation (D-Afmt) consists of resampling $f(p, q)$ in discrete polar coordinates and estimating the Fourier-Mellin integrals (3) and (4) (see Fig. 2). The polar sampling grid is built from the intersection between M concentric circles with increasing radii of fixed spacing and N radial lines originating from the image centroid. The angular and radial sampling steps are, respectively,

$$\triangle\theta = \frac{2\pi}{\mathrm{M}}, \quad \triangle\rho = \frac{\mathrm{R}}{\mathrm{N}},$$

where R denotes the radius of the smallest disk required to contain the object.

Wherever the polar sampling point does not correspond to a grid location, the gray-level value is estimated from the four nearest pixels using bi-linear interpolation. Hence, the polar representation of an image is an $[\mathrm{N}, \mathrm{M}]$-matrix whose values correspond to

$$\widehat{f}(\hat{\rho}_n, \hat{\theta}_m), \quad n \in [0, \mathrm{N}-1], m \in [0, \mathrm{M}-1],$$

where $\hat{\rho}_n = n\triangle\rho$ and $\hat{\theta}_m = m\triangle\theta$ are respectively the beam with index n and the circle with index m. Replacing integrals over circles and beams by sums in (3), we get the D-Afmt approximation $\widehat{\mathcal{M}}_{f_\sigma}^D$ of f :

$$\forall k \in [-\mathrm{K}, \mathrm{K}], \forall v \in [-\mathrm{V}, \mathrm{V}], \quad \widehat{\mathcal{M}}_{f_\sigma}^D (k, v) = \triangle\rho\triangle\theta \sum_{n=0}^{\mathrm{N}-1} \widehat{F}_k (\hat{\rho}_n) (\hat{\rho}_n)^{\sigma - iv - 1} ,$$

where

$$\widehat{F}_k (\hat{\rho}_n) = \sum_{m=0}^{\mathrm{M}-1} \widehat{f} \left(\hat{\rho}_n, \hat{\theta}_m \right) e^{-i\frac{km}{\mathrm{M}}}.$$

From (4), the discrete polar image $\widehat{f}_{\mathrm{K},\mathrm{V}}^D(\hat{\rho}_n, \hat{\theta}_m)$ is recovered with the following approximation of the inverse AFMT :

$$\forall n \in [0, \mathrm{N}-1], \forall m \in [0, \mathrm{M}-1], \quad \widehat{f}_{\mathrm{K},\mathrm{V}}^D \left(\hat{\rho}_n, \hat{\theta}_m \right) = (\hat{\rho}_n)^{-\sigma} \sum_{v=-\mathrm{V}}^{\mathrm{V}} \widehat{G}_k \left(\hat{\theta}_m, v \right) (\hat{\rho}_n)^{iv} ,$$

where

$$\widehat{G}_k \left(\hat{\theta}_m, v \right) = \sum_{k=-\mathrm{K}}^{\mathrm{K}} \widehat{\mathcal{M}}_{f_\sigma}^D (k, v) e^{i\frac{km}{\mathrm{M}}}.$$

The discrete Fourier series may be computed with a fast Fourier transform algorithm (FFT) if M and $(2\mathrm{K} + 1)$ are integral powers of 2. The reconstructed image is retrieved with a polar to Cartesian conversion.
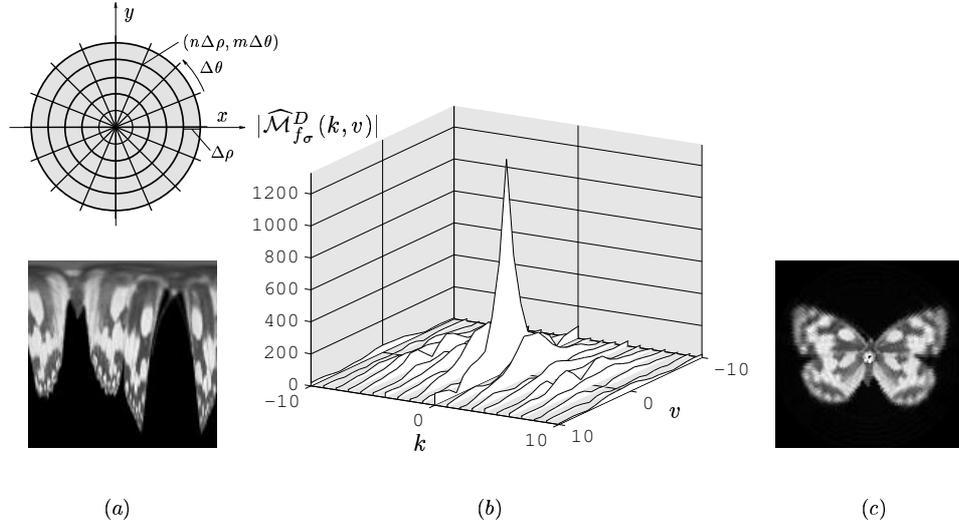

FIG. 2. Illustration of the direct approximation of the AFMT (D-Afmt) for the butterfly in Fig. 1. (a) Polar sampling grid and representation of the butterfly (N = M = 128). (b) Magnitude of central Fourier-Mellin harmonics (K = V = 10). (c) Cartesian reconstruction obtained with K = V = 41.

3.3. The fast algorithm

With a change of the integration variable from r to $q = \ln(r)$, Eq. (3) and (4) can be rewritten as the Fourier transforms of the deformed object $e^{t\sigma} f(e^t, \theta)$:

$$
\begin{cases}
\mathcal{M}_{f_\sigma}(k,v) = \dfrac{1}{2\pi} \displaystyle\int_{-\infty}^{+\infty} \int_0^{2\pi} e^{t\sigma} f(e^t, \theta) \, e^{-i(k\theta+tv)} d\theta dt. \\[12pt]
f(e^t, \theta) = e^{-t\sigma} \displaystyle\int_{-\infty}^{+\infty} \sum_{k\in\mathbb{Z}} \mathcal{M}_{f_\sigma}(k,v) \, e^{i(k\theta+tv)} dv.
\end{cases}
\tag{6}
$$

The fast AFMT approximation (F-Afmt) consists of resampling $f(p,q)$ in discrete log-polar coordinates and estimating the Fourier integrals (see Fig. 3). The numerical conversion from Cartesian to log-polar coordinates of the discrete image is performed using an algorithm similar to the previous one except that circles are exponentially spaced. The AFMT is computed using a 2D Discrete Fourier Transform (DFT), with log-polar coordinates :

$$
\forall k \in [-K, K], \forall v \in [-V, V], \quad \widehat{\mathcal{M}}_{f_\sigma}^F(k,v) = \triangle\theta \sum_{n=0}^{N-1} e^{\sigma\hat\rho_n} \sum_{m=0}^{M-1} \widehat{f}\left(e^{\hat\rho_n}, \hat\theta_m\right) e^{-i\left(k\hat\theta_m + v\hat\rho_n\right)}.
$$

It is worth noting that this algorithm assumes the image to be periodic in the radial direction.

From (6), the discrete log-polar image $\widehat{f}\left(e^{\hat\rho_n}, \hat\theta_m\right)$ is recovered by computing a 2D DFT over the Fourier-Mellin coordinates:

$$
\forall n \in [0, N-1], \forall m \in [0, M-1], \quad \widehat{f}\left(e^{\hat\rho_n}, \hat\theta_m\right) = e^{-\sigma\hat\rho_n} \sum_{v=-V}^{V} \sum_{k=-K}^{K} \widehat{\mathcal{M}}_{f_\sigma}^F(k,v) e^{i\left(k\hat\theta_m + v\hat\rho_n\right)}.
$$

If M and N, $(2K+1)$ and $(2V+1)$ are integral powers of 2, the FFT algorithm can be used for reducing the computation time. The Cartesian image $\widehat{f}_{K,V}^F(p,q)$ is retrieved with a log-polar to Cartesian conversion.

3.4.   The Cartesian algorithm

Another expression of the AFMT pair can be obtained by performing a polar to rectangular change of variable in (3) and (4). Hence the AFMT is expressed according to the Cartesian coordinates of f:

$$
\begin{cases}
\mathcal{M}_{f_\sigma}(k,v) = \dfrac{1}{2\pi} \displaystyle\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y) \, (x+iy)^{-k} \, (x^2+y^2)^{\frac{k-2+\sigma-iv}{2}} \, dxdy. \\[12pt]
f(x,y) = \displaystyle\int_{-\infty}^{+\infty} \sum_{k\in\mathbb{Z}} \mathcal{M}_{f_\sigma}(k,v)(x+iy)^k \, (x^2+y^2)^{\frac{-k-\sigma+iv}{2}} \, dv.
\end{cases}
\tag{7}
$$

In this case, no resampling of the discrete image is necessary and the AFMT can be estimated directly from the rectangular grid (see Fig. 4). The Cartesian AFMT
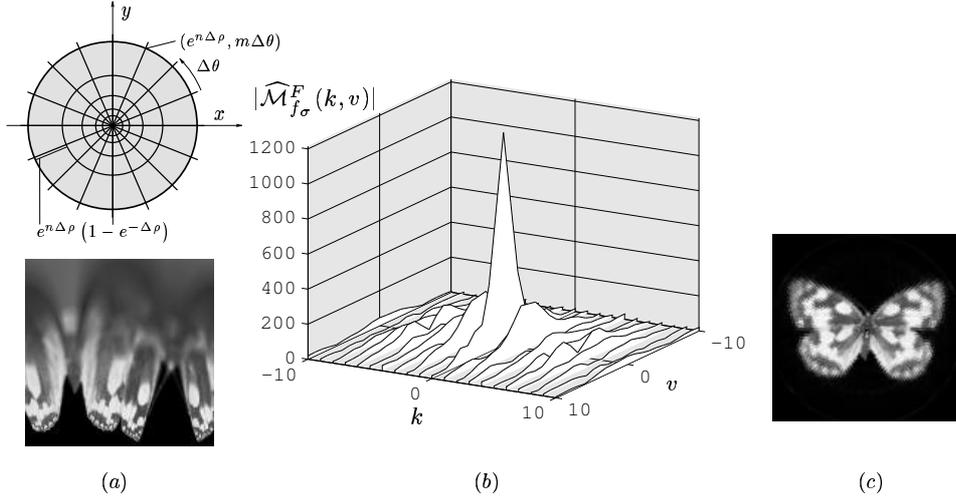
(a)   (b)   (c)

FIG. 3. Illustration of the fast approximation of the AFMT (F-Afmt) for the butterfly in Fig. 1. (a) Log-Polar sampling grid and representation of the butterfly (N = M = 128). (b) Magnitude of central Fourier-Mellin harmonics (K = V = 10). (c) Cartesian reconstruction obtained with K = V = 41.

(C-Afmt) approximation $\widehat{\mathcal{M}}_{f_\sigma}^C$ is computed by using sums in place of integrals:

$$
\begin{cases}
\widehat{\mathcal{M}}_{f_\sigma}^C(k,v) = \dfrac{1}{2\pi} \sum_{q=\mathrm{Q}_{min}}^{\mathrm{Q}_{max}} \sum_{p=\mathrm{P}_{min}}^{\mathrm{P}_{max}} f(p,q)\,(p+iq)^{-k}\,(p^2+q^2)^{\frac{k-2+\sigma-iv}{2}}. \\[2em]
\widehat{f}_{\mathrm{K,V}}^C(p,q) = (p^2+q^2)^{\frac{-\sigma}{2}} \sum_{v=-\mathrm{V}}^{\mathrm{V}} \sum_{k=-\mathrm{K}}^{\mathrm{K}} \widehat{\mathcal{M}}_{f_\sigma}^C(k,v)\,(p+iq)^{k}\,(p^2+q^2)^{\frac{-k+iv}{2}}.
\end{cases}
$$

The coordinates m and n correspond to a pixel position from the object centroid. P_{min}, P_{max}, Q_{min}, and Q_{max} indicate the coordinates, with respect to the image centroid, of the smallest rectangle that fully contains the object. For the sake of compatibility with other approximations, we used the trapezoidal integration rule. The discrete image is recovered directly in rectangular coordinates. It is worth noting that no fast algorithm is available for this version and the computation cost is higher than that of the other approximations.

A careful visual inspection shows differences between the three reconstructions. The question of which approximation gives the best reconstruction for a given finite-extent spectrum is addressed in the next section.

4. RECONSTRUCTION EXPERIMENTS

An image can be fully reconstructed from its Fourier-Mellin transform. In the discrete case, it must be shown that a finite number of Fourier-Mellin harmonics can reconstruct the original image to within a specified degree of accuracy. For this purpose, the reconstruction error between an image $f(p,q)$ and its reconstructed version from a finite number of its Fourier-Mellin harmonics is used as a measure of the image reconstruction ability of the AFMT. According to works done on reconstruction from orthogonal moments [13, 14], we use the discrete normalized
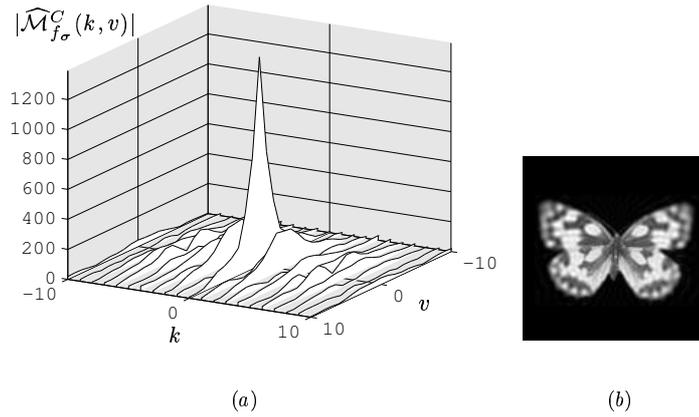
(a)   (b)

FIG. 4. Illustration of the Cartesian approximation of the AFMT (C-Afmt) for the butterfly in Fig. 1 using the original grid. (a) Magnitude of central Fourier-Mellin harmonics (K = V = 10). (b) Cartesian reconstruction obtained with K = V = 41.

mean square reconstruction error (NMSRE) defined by

$$\varepsilon^2\left(\widehat{f}_{K,V}\right) = \frac{\displaystyle\sum_{p=P_{min}}^{P_{max}} \sum_{q=Q_{min}}^{Q_{max}} \left| f(p,q) - \widehat{f}_{K,V}(p,q) \right|^2}{\displaystyle\sum_{p=P_{min}}^{P_{max}} \sum_{q=Q_{min}}^{Q_{max}} f(p,q)^2}, \tag{8}$$

where $\widehat{f}_{K,V}$ is one of the three reconstructions obtained from the inverse D-Afmt, F-Afmt, or C-Afmt algorithms with K and V harmonics. It is well-known that an image with a high NMSRE can be of better visual quality than an image with a small one. In our case, it is used to measure the improvement when the number of harmonics involved in the reconstruction process increases and to characterize the individual parts of the Fourier and Mellin transforms.

Experiments that are reported in the following used a binary image of the letter F and the image of a butterfly (Fig. 5). For increasing values of K and V, the results show the reconstructed images $\widehat{f}_{K,V}$, the difference between the original and the reconstructed images: $f - \widehat{f}_{K,V} + 128$, and the evolution of the NMSRE for the reconstructed images $\varepsilon\left(\widehat{f}_{K,V}\right)$ (8).
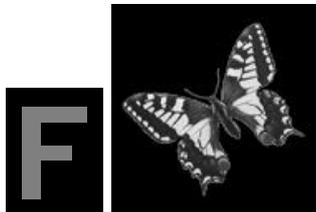

FIG. 5. Images used for reconstruction experiments. Left: Binary image of the letter F (original size 36×49). Right: *Euplagia Quadripunctaria* (original size 112×73).

4.1.   Binary image of a typographic letter

Fig. 6 shows the evolution of the reconstruction and the NMSRE of a binary image with the typographic letter F for an increasing number of harmonics along both frequency axes corresponding to the Fourier and Mellin transforms. The harmonics were computed with the C-Afmt algorithm.

As expected, the reconstruction error is reduced when the spectral dimensions increase, confirming the visual improvement of the reconstructed image. By including increasingly higher-order harmonics in both frequency directions, we get a visual illustration of the contribution of each dimension independently. For lower dimensions, results show respectively the radial and the angular influence of the Mellin and the Fourier parts of the AFMT. For larger dimensions, these influences are combined and finally give a high quality reconstruction with a low NMSRE of about 0.05. From the NMSRE, it can be also pointed out that the two dimensions carry different information of nearly equal importance.

4.2.   Gray-level image of a butterfly

Fig. 7, 8 and 9 show the reconstruction and the NMSRE of the butterfly image for the three approximations of the direct and inverse AFMTs. This image presents high contrast texture variations and is relevant in order to compare the three algorithms.

From the reconstruction results, it becomes apparent that low-order Fourier-Mellin harmonics contain gross shape information and higher-order harmonics fill the fine details of an image. An object of finite spatial extent will possess Fourier-Mellin harmonics of infinite order. When the AFMT is computed to a finite order, the inverse AFMT approximation results in a blurred and corrupted reconstruction (Gibb's phenomenon) [31]. These degradations are reduced with the contribution of additional increasing order Fourier-Mellin harmonics, confirming the visual improvement of the reconstructed image.

However, a careful inspection of Fig. 7 shows that the improvement in the reconstruction from the D-Afmt algorithm is counter-balanced by a divergent hole in the image centroid where numerical values exceed the gray-level limit. From the reconstructed images, it becomes apparent that the hole size increases as the number of Mellin harmonics grows. Further experiments have shown that using a smaller radial step in the polar representation reduces the hole and improves the NMSRE, but requires an additional computational effort. For the F-Afmt and C-Afmt approximations, the difference between the original and the reconstructed images shows the effect of the sampling grids and outlines the existence of circular waves due to the spectrum truncation. These waves are reduced with the contribution of higher order Fourier-Mellin harmonics and finally the difference becomes small.

In order to compare the reconstruction results for the three AFMT approximation methods, Fig. 10 shows plots of NMSRE of the butterfly image for K = V. For the D-Afmt algorithm, the NMSRE reaches a minimum at K = V = 45, and then starts to increase as the number of Fourier-Mellin harmonics involved in the reconstruction process increases. This illustrates the impact of the divergent hole on the reconstruction. For the F-Afmt and C-Afmt algorithms, the NMSRE always decreases as the spectrum dimensions increase. For larger spectra (K = V > 20),
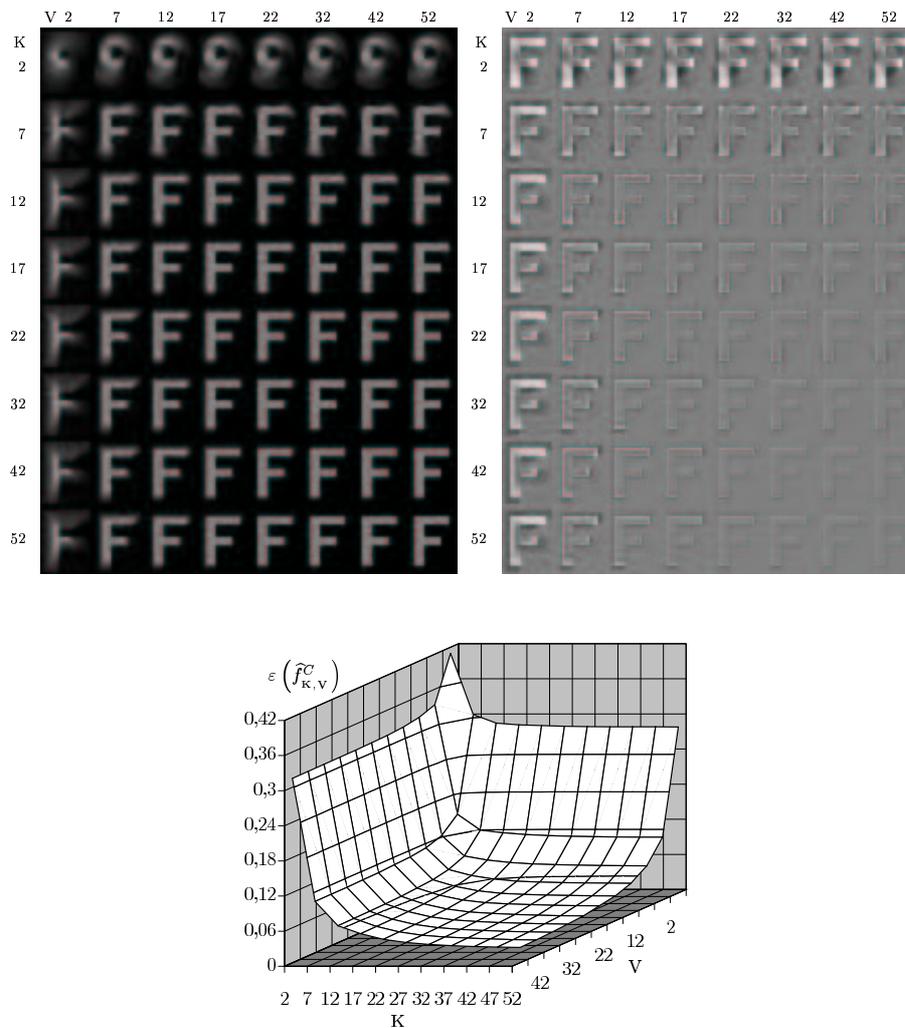
FIG. 6.   Evolution of reconstruction of a binary image of the lettre F with the C-Afmt algorithm for varying dimensions K (lines) and V (columns) and corresponding mean square reconstruction error.

the C-Afmt algorithm provides a lower reconstruction error than the F-Afmt algorithm. The total error introduced in estimating the integral over rectangular coordinates for a sufficiently large spectrum produced less error than the log-polar resampling.

Each Fourier-Mellin harmonic is computed over the entire image so that a too small polar or log-polar resampling rate results in a poor Fourier-Mellin representation. In the case of log-polar description, the radial and angular steps M and N set the maximum number of Fourier-Mellin harmonics available in both dimensions, i.e. $K \leq 2M + 1$ and $V \leq 2N + 1$ (see section 3.3). So that a low resampling rate results in a truncated FM representation with higher frequencies discarded, and the inverse F-Afmt algorithm reconstructs a smoothed image. The effects of resampling become apparent whatever the degree of approximation used (bi-linear or higher).
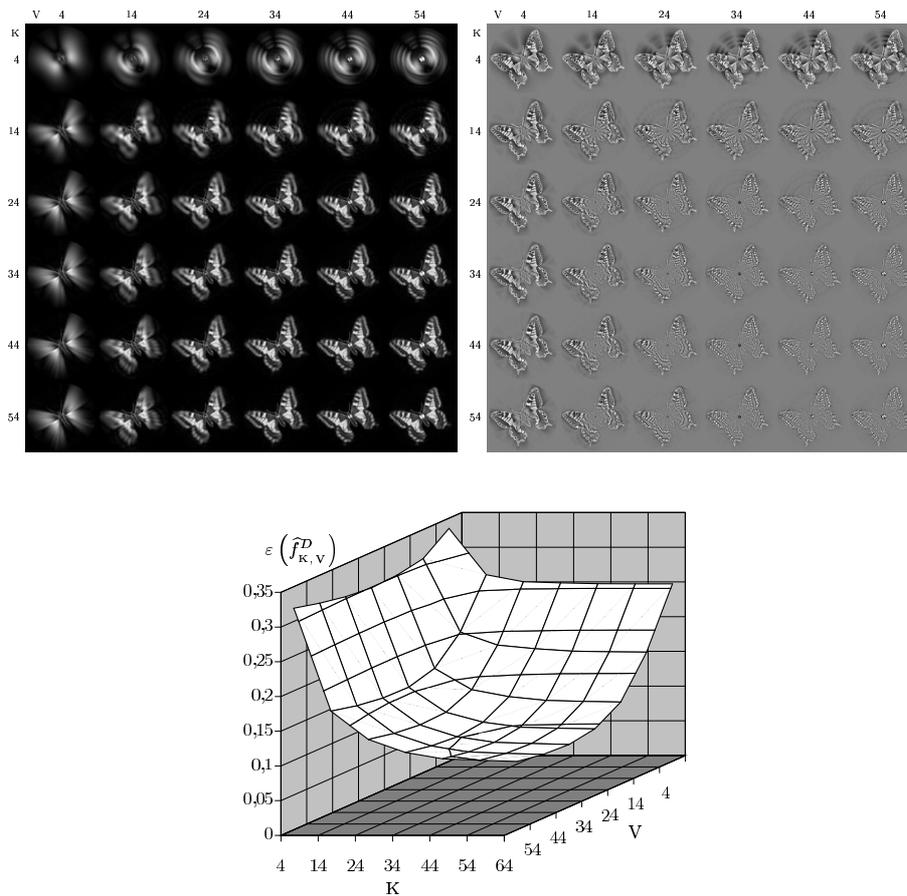
FIG. 7. Evolution of reconstruction of the image of a butterfly with the D-Afmt algorithm for varying dimensions K (lines) and V (columns) and corresponding mean square reconstruction error.

In the case of a high resampling rate, the degree of approximation has a small impact on precision and we verified that the reconstruction error improvement and the visual gain are negligible.

The next section goes on with the comparisons for pattern recognition and image retrieval from a complete and convergent set of invariants computed from the AFMT.

5. APPLICATION TO INVARIANT PATTERN RECOGNITION AND CONTENT-BASED IMAGE RETRIEVAL

In the following, we examine the extent to which a numerical spectrum of an object is a reliable estimation of its AFMT by mean of experiments on pattern recognition and Content-Based Image Retrieval (CBIR) from a complete set of similarity invariant descriptors. The investigation gives an idea of the robustness of the three algorithms and their ability to perform pattern classification.
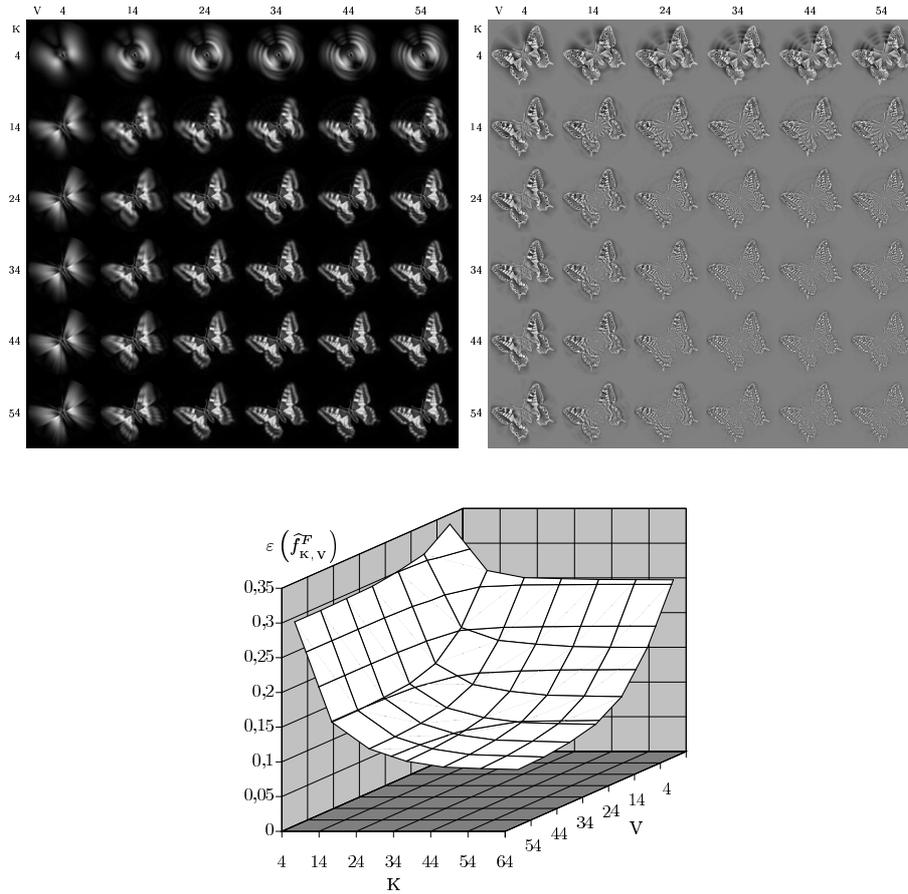
FIG. 8. Idem to Fig. 7 with the F-Afmt algorithm.

5.1. Standard Fourier-Mellin descriptors

Since the usual FMTs of two similar objects only differ by a phase factor, a standard set of descriptors invariant to object position, orientation, and size is generally obtained by taking the magnitude of selected Fourier-Mellin harmonics [16]. However, the phase of the AFMT representation contains an important amount of information on the object shape, as illustrated in Fig. 11. In this experiment, we reconstructed an object from the combination of the AFMT phase of a first image with the AFMT magnitude of a second one. It is clear from images (c) and (d) that we first recognized the object whose phase has been used for reconstruction.

Furthermore, most experiments on pattern recognition only deal with a small set of low-order standard descriptors because of the numerical difficulties encountered in estimating the FMT. Therefore, the corresponding set of invariant features is dramatically reduced and only represents an outline of the shape of the object. For pattern recognition purposes, the classification of an unknown object as one of a set of reference patterns is achieved with a comparison method such as: the computation of an error between the features [32], neural networks [33, 34], or statistical classifiers [19, 35, 20]. These methods are turned out to be efficient when the models to compare have simple and distinct shapes (typographic symbols or
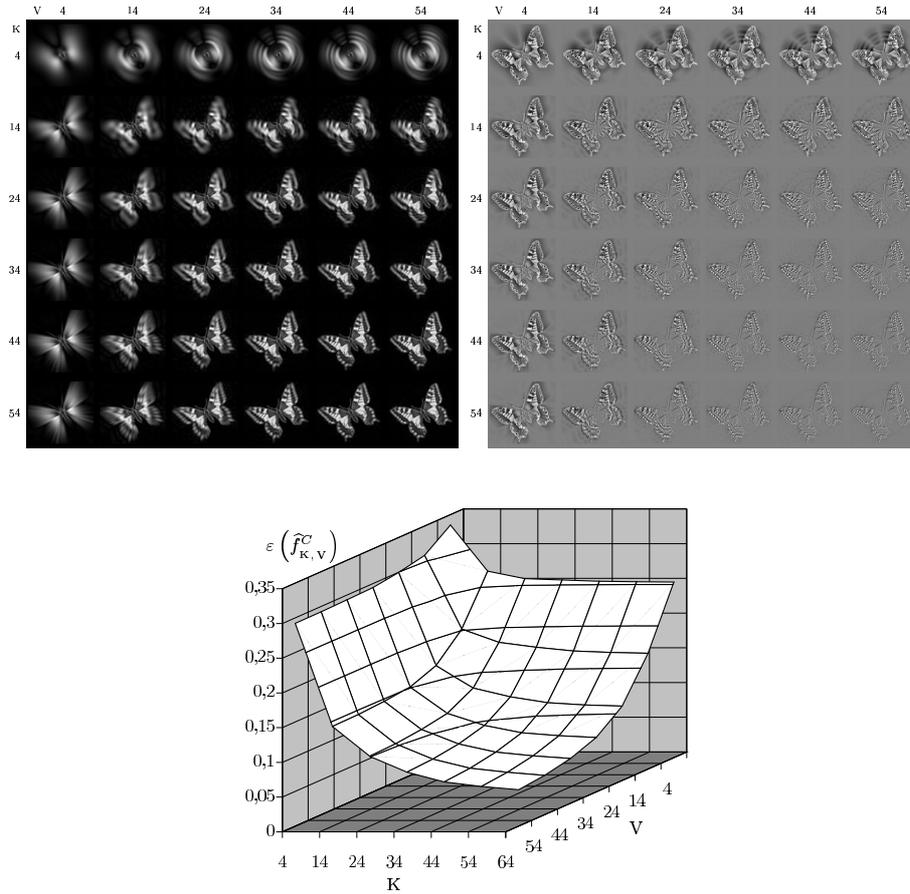
FIG. 9. Idem to Fig. 7 with the C-Afmt algorithm.

letters). Nevertheless, image databases become more and more voluminous and heterogeneous. Due to the lack of completeness, one can find distinct objects with identical descriptor values, and the classification process may mix up objects.

5.2. A complete family of AFMT descriptors

In order to overcome the difficulties described above, a complete family of similarity invariant features based on the AFMT has been suggested in [22]. This family can be easily rewritten and applied to any strictly positive σ value in the following way:

$$\forall (k,v) \in \mathbb{Z} \times \mathbb{R}, \quad \mathcal{I}_{f_\sigma}(k,v) = \mathcal{M}_{f_\sigma}(0,0)^{\frac{-\sigma+iv}{\sigma}} \, e^{ik \arg(\mathcal{M}_{f_\sigma}(1,0))} \, \mathcal{M}_{f_\sigma}(k,v), \quad (9)$$

with $\sigma > 0$. Each feature $\mathcal{I}_{f_\sigma}(k,v)$ is constructed in order to compensate for the $\alpha^{-\sigma+iv} e^{ik\beta}$ term that appears in the shift theorem (5). The compensation is achieved via the two Fourier-Mellin harmonics $\mathcal{M}_{f_\sigma}(0,0)$ and $\mathcal{M}_{f_\sigma}(1,0)$, which act as normalization parameters. Fig. 12 shows the magnitude of the central AFMT invariant descriptors of the butterfly in Fig. 5. We used the F-Afmt approximation
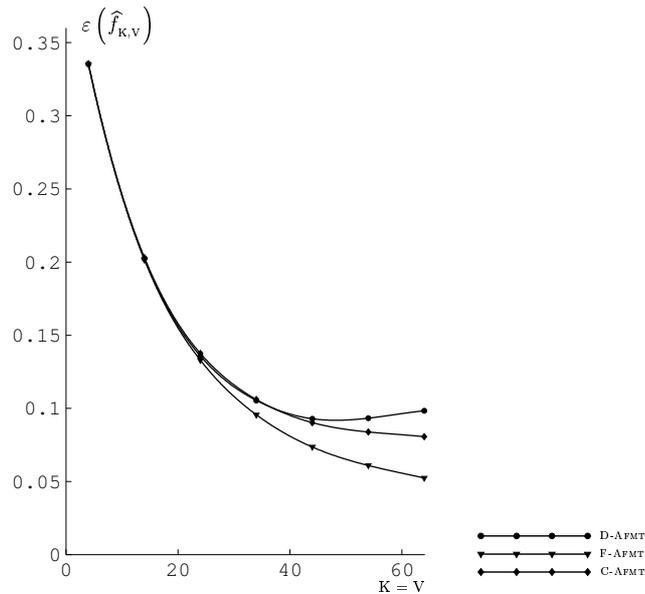
FIG. 10.   Plots of the NMSRE of the butterfly image in Fig. 5 versus spectral dimensions (K=V) for the three algorithms.

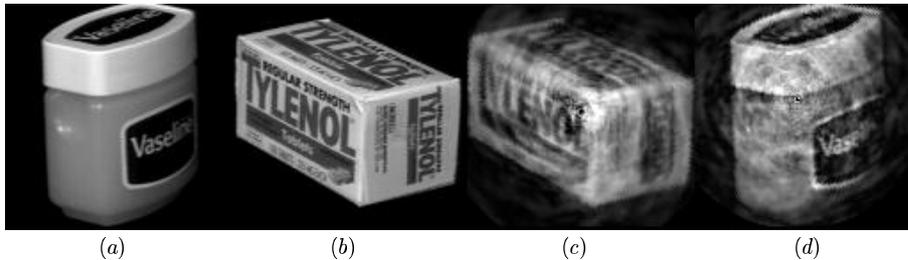|			
(a)	(b)	(c)	(d)

FIG. 11.   Illustration of the importance of phase in the AFMT representation. (a) and (b) Two gray-level objects from the COIL database (Columbia university). (c) image reconstructed from the inverse AFMT (C-Afmt algorithm) with the modulus of (a) and the phase of (b). (d) idem with the modulus of (b) and the phase of (a).

with $K = V = 10$. The central harmonic ($k = v = 0$) has a value of 1, and the magnitude of the set converges towards zero as k and v increase.

The set of features in Eq. (9) is complete since it is possible to recover the AFMT of an object from all of its invariant features and the two normalization parameters. Since the family is also convergent for square integrable functions defined on $\mathbb{R}_+^* \times \mathbb{S}^1$, it is possible to define a mathematical distance between shapes [22]:

$$d_2(\mathcal{I}_{f_\sigma}, \mathcal{I}_{g_\sigma}) = \left(\int_{-\infty}^{+\infty} \sum_{k \in \mathbb{Z}} |\mathcal{I}_{f_\sigma}(k, v) - \mathcal{I}_{g_\sigma}(k, v)|^2 \, dv \right)^{\frac{1}{2}}. \tag{10}$$

The invariant distance (10) is the well-known d_2 Euclidean distance expressed in the Fourier-Mellin domain. The distance is zero if and only if the objects are similar, regardless of translation, rotation and scale. When objects are distinct, the distance is used for the quantification of the dissimilarity between the objects.
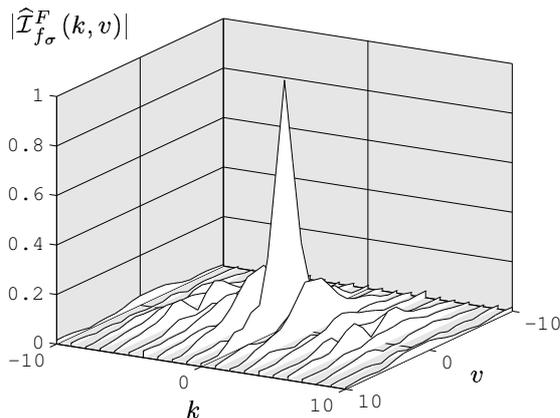
FIG. 12. Magnitude of central AFMT invariant descriptors for the butterfly image in Fig. 5 (F-Afmt approximation, K = V = 10).

The completeness property of the feature set shows an infinite number of invariants, which appears to be incompatible with numerical constraints. However, the expression in Eq. 9 allows to extract as many descriptors as needed for a given application. By including higher-order invariants, theoretical completeness is best attained and we get a more and more accurate representation of shapes. Furthermore, no additional error is introduced in computing the invariants and the three AFMT approximations can be compared by means of pattern recognition experiments.

5.3. Pattern recognition and discrimination experiments

We now present some results of discrimination experiments with the images of butterflies in Fig. 13. These images were chosen in order to illustrate the numerical behavior of the different approximations. It must be pointed out that comparing butterflies as in Fig. 13 is rather difficult since their global aspect shows obvious shape characteristics, independent of size and orientation, that allow the images to be understood by the human eye as being butterflies. The main difference is in the texture of the wings and contour-based invariants such as Fourier descriptors will probably fail.
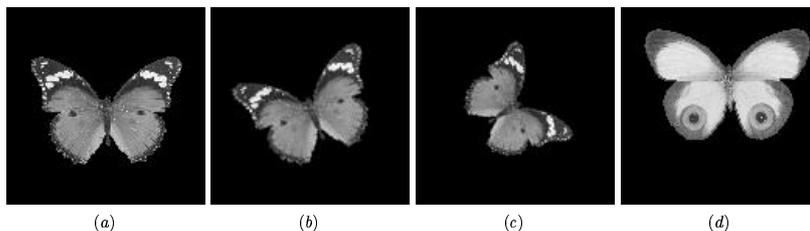

FIG. 13. Four test images used in the pattern recognition experiments (original size 128×128). Fig. (a), (b) and (c) represent the same butterfly but with different size and orientation. Fig. (d) is an image of a butterfly with a different shape. (a) Graphium Agamemnon. (b) image (a) scaled by 0.9 and rotated by 20. (c) image (a) scaled by 0.75 and rotated by -60. (d) Taenaris Macrops.

The computation of the invariant distance is performed in three steps. Firstly, we estimate the AFMT of the two objects to compare with one of the three approximations. Secondly, we compute the invariant descriptor values for each object according to Eq. (9). We obtain as many invariants as there are Fourier-Mellin harmonics. Thirdly, we substitute summation for the integration in (10) and get the invariant distance between the two objects. When objects are similar, the numerical approximation of the distance is a measure of the error between invariant values. This question is examined in the following experiments by investigating the numerical behavior of the distance when an increasing number of descriptors is used for its computation.

The two graphs in Fig. 14 show the invariant distance between images with the same shape. As expected, the distance increases when increasing portions of the Fourier-Mellin spectrum are used. There are two main reasons for not obtaining exact invariance (zero distance) in case of similar objects. Firstly, the Cartesian grid is not well suited for similarity transforms, and objects may present small differences due to interpolation. Secondly, the approximation of the AFMT introduces numerical errors in the computation of the invariant features. Note that the distance values are different for the same pair of images but different approximation of the AFMT.

In our experiments, the evolution of the distance computed from the D-Afmt approximation shows a scale factor sensibility when the set of invariant features grows. The direct algorithm does not provide a reliable estimation of higher-order Fourier-Mellin harmonics. This point confirms the results of the reconstruction experiments reported in Section 4. Further experiments have also shown that the rotation angle has lesser effect on the results as compared to the scale factor. For the C-Afmt and F-Afmt approximations, the distances are more stable. The fast algorithm gives always a lower distance, whatever the size of the feature set. The additional error coming from higher-order invariants is low, and we may conclude that the F-Afmt and C-Afmt algorithms give a better estimate of higher-order Fourier-Mellin harmonics.
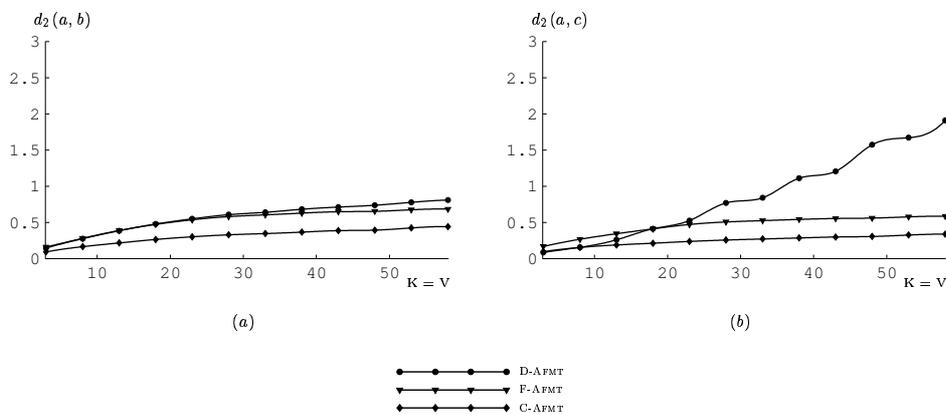

FIG. 14.   Plot of the invariant distance between similar images versus the spectral dimensions involved in its computation. (a): images in Fig. 13(a) and 13(b). (b): images in Fig. 13(a) and 13(c).

Theoretical completeness of the invariant set is best attained with a large set of descriptors. On the other hand, previous experiments show that invariant distance

increases as the spectrum is enlarged. To get an idea of the numerical discrimination performance of the invariant family, we divided the distances obtained with a pair of distinct objects with that obtained with a pair of similar objects. The farther the ratio is from 1, the greater the discrimination is. Fig. 15(b) illustrates the ratio between the images in Fig. 13(a) and (c) and the images in Fig. 13(a) and (d) (the distance is reported in Fig. 15(a)). The D-Afmt approximation provides the best discrimination until the dimensions of the invariant set reach 8×8. For larger set, the ratio is close to 1, and discrimination between objects becomes impossible. The C-Afmt and F-Afmt algorithms provides a discrimination ratio that is greater than 3, even when the feature set is large, indicating that it is still possible to distinguish shapes.
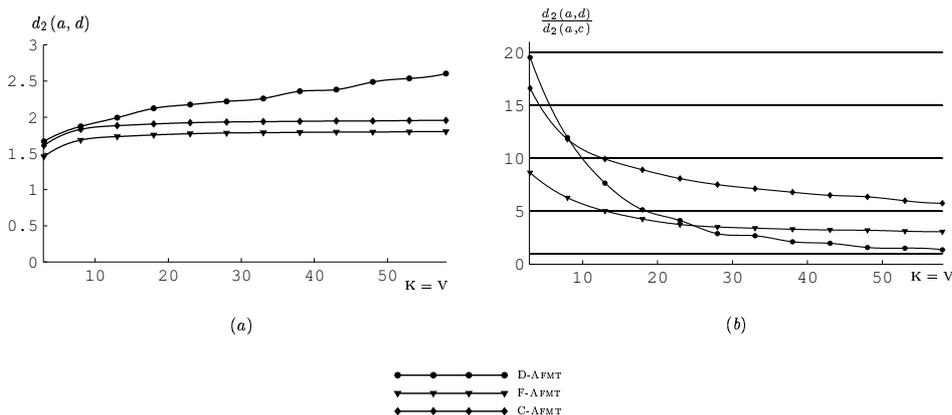

FIG. 15. (a) Invariant distance between images in Fig. 13(a) and (d), and (b) discrimination ratio between images in Fig. 13(a), (c) and (d) versus the spectral dimensions involved in its computation.

5.4. Content-based image retrieval

In the following, we examine the application of the complete set of AFMT descriptors and the invariant distance to perform content-based image retrieval. The experiments presented here don't cover the many aspects of CBIR systems (relevance feedback, combination of descriptors, fast search, ...), but illustrate the FMT in this context.

At archival time (off-line), each model f^i, $i \in [1, \text{N}]$ of the image database is indexed by its set of invariant descriptors $\mathcal{I}_{f_\sigma^i}(k, v)$, with $k \in [-\text{K}, \text{K}]$ and $v \in [-\text{V}, \text{V}]$. At query time (on-line), the retrieval of a query object g, regardless of the objects pose, orientation and size, is achieved by :

1. computing its invariant features from Eq. (9),

2. estimating the distance between the input object g and the N models f^i, by using Eq. (10),

3. sorting and selecting the M models f^j which give the smallest distance to g.

In this way, all images from the database can be compared with any query image, and ranked by the value of d_2. The search through the database is exhaustive and future work will include an effective access method such as tree indexing and

hashing as pointed out in the conclusion. Since invariants are used as indexes, a new model can be added to the database without modifying the models already stored and no voting algorithm for selecting the best models is required.

Concerning to the experiments, we used the F-Afmt approximation since it appears more robust and faster than the two other estimations. We computed about one thousand invariant descriptors for each gray-level images in the database. The computation time for deriving the invariant representation of an image is about 3 seconds on a 300 MHz PC (including file I/O, log-polar resampling (128×128), 2D FFT computation and invariant extraction). The computation time only depends on the size of the image, not on its complexity. At query time, the search through the database is performed sequentially.

Fig. 16 shows the top M = 9 ranked images retrieved for three query image (from up to bottom) by a larger set of invariants (from left to right, S =13, 181 and 761). The first retrieval (R_1) has been conducted on a collection of N = 98 images of butterflies. 69 images represent distinct butterflies and 29 images show translation, rotation and scale changes. The two objects with a shape similar to the query have been first selected, whatever the size of the feature set. This result confirms the invariance and the discrimination of our set of descriptors.

The next two retrievals in Fig. 16 (R_2 and R_3) have been conducted on the Columbia Object Image Library [36]. This database contains N = 1440 images of 20 different 3D objects: 72 images per object taken at 5 degrees in pose. In this collection, the camera motion clearly infringes the similarity transformation model underlying the image representation of the AFMT. These experiments intend to illustrate the robustness of the descriptors to partial occlusions and shape distortions (non-similarity transformations).

In the first retrieval, the number of invariants shows no influence on the search, while the second result becomes better as the dimension of the invariant features set enlarges, since more and more images of the query object are retrieved. These results have been chosen amongst the worst retrievals we obtained since most of the time the process retrieves more than 10 images of the query object.

These interesting results illustrate the stability of the complete invariant family. This stability is due to the normalization process and, above all, to the robustness of the AFMT approximations, since each invariant descriptor is computed from the corresponding Fourier-Mellin harmonic. The small difference of texture between two consecutive images of the same object is coded by a small difference in the Fourier-Mellin harmonics values so that the invariant descriptors are also closed to and the invariant distance is low.

6. CONCLUSION

In this paper, three approximations of the AFMT were proposed and investigated through experiments with real gray-level images by means of image reconstruction and pattern recognition. From both theoretical and numerical improvements, results show that it is possible to recover an image to within a given degree of accuracy and to classify objects reliably even when a large set of invariant descriptors is used. Actually, it has been pointed out that a relatively small set of AFMT harmonics may characterize an image adequately, but that the fine details can be recreated only by including higher-order harmonics. From reconstruction experiments, the

$S_{K,V}$	13 (K = V = 2)	181 (K = V = 9)	761 (K = V = 19)
R_1	pap13_1 (0.00), pap13_2 (0.10), pap13_3 (0.12); pap18_1 (0.18), pap18_2 (0.20), pap10_1 (0.20); pap10_2 (0.22), pap10_3 (0.25), pap44_1 (0.32)	pap13_1 (0.00), pap13_2 (0.20), pap13_3 (0.20); pap10_1 (0.48), pap10_2 (0.50), pap18_1 (0.51); pap18_2 (0.51), pap10_3 (0.53), pap44_1 (0.60)	pap13_1 (0.00), pap13_3 (0.21), pap13_2 (0.24); pap18_1 (0.60), pap18_2 (0.60), pap10_1 (0.65); pap10_2 (0.65), pap10_3 (0.68), pap44_1 (0.74)
R_2	obj8__1 (0.00), obj8__0 (0.10), obj8__2 (0.16); obj8__71 (0.21), obj8__70 (0.26), obj8__69 (0.32); obj8__66 (0.39), obj7__43 (0.41), obj7__42 (0.41)	obj8__1 (0.00), obj8__0 (0.23), obj8__2 (0.35); obj8__71 (0.43), obj8__70 (0.49), obj8__69 (0.60); obj8__66 (0.71), obj7__42 (0.75), obj8__68 (0.76)	obj8__1 (0.00), obj8__0 (0.31), obj8__2 (0.46); obj8__71 (0.54), obj8__70 (0.62), obj8__69 (0.71); obj8__66 (0.81), obj7__42 (0.83), obj4__50 (0.85)
R_3	obj19__48 (0.00), obj19__47 (0.12), obj19__49 (0.15); obj5__14 (0.25), obj19__50 (0.26), obj5__15 (0.27); obj19__51 (0.30), obj5__13 (0.31), obj5__16 (0.31)	obj19__48 (0.00), obj19__47 (0.29), obj19__49 (0.30); obj19__50 (0.50), obj19__51 (0.56), obj5__14 (0.57); obj5__15 (0.60), obj5__16 (0.64), obj19__15 (0.65)	obj19__48 (0.00), obj19__47 (0.37), obj19__49 (0.41); obj19__50 (0.61), obj19__51 (0.65), obj5__14 (0.67); obj5__15 (0.70), obj19__15 (0.73), obj19__16 (0.75)

FIG. 16. Retrieval results for three query images (top-left corner) for larger sets of AFMT descriptors ($S_{K,V} = 13, 181, 761$). Images are ranked by increasing distance from left to right and top to bottom.

fast and the Cartesian AFMT approximations were found to perform better than the direct approximation.

For pattern recognition, we used a complete and convergent set of descriptors invariant to similarity transformations. The completeness of the set gives a flexible and extensive representation of gray-level shapes. A real distance between the invariant descriptors has been used to compare objects, regardless of their pose, orientation and size, in the context of content-based image retrieval. For reducing the computing effort, the fast AFMT algorithm was preferred to the Cartesian ap-

proximation. Since the FMT is global, it suffers from local modifications, which often happens in case of complex cluttered scene or 3D objects motion. Nevertheless, the results were found to be robust even when images present small distortions and occlusions. These points make the FMT suitable in the context of the object-oriented MPEG-4 and MPEG-7 standards and argue in favor of the FMT to be incorporated in CBIR systems for gray-level shapes description.

In this context, future work will improve the efficiency for indexing and searching by means of feature quantization and hierarchical search strategies adapted to the shape space. Then, it should be interesting to test the robustness of descriptors to quantification and, especially, the influence on the reconstruction process. Other works will test the direct and inverse AFMT approximations for digital image watermarking [24], and study other Fourier transforms suited to more general groups of transformations.

ACKNOWLEDGMENT

REFERENCES

1. J.L. Mundy, A. Zisserman, and D. Forsyth, editors. *Applications of invariance in computer vision*. Lecture notes in computer science. Springer-Verlag, Berlin, Germany, 1993.

2. F. Ghorbel. Towards a unitary formulation for invariant image description ; Application to image coding. *Annals of Telecommunication*, 53(5/6):242–260, May/June 1998.

3. C.T. Zhan and R.Z. Roskies. Fourier descriptors for plane closed curves. *IEEE trans. on Computers*, 21(3):269–281, 1972.

4. F. Ghorbel. Stability of invariant Fourier descriptors and its inference in the shape classification. In *Proc. of the 11th Int. Conf. on Pattern Recognition*, pages 130–134, The Hague (Nl.), August 1992.

5. T.R. Crimmins. A complete set of Fourier descriptors for two-dimensional shapes. *IEEE trans. on Systems, Man, Cybernetics*, 12(6):848–855, November/December 1982.

6. M. Daoudi, F. Ghorbel, A. Mokadem, O. Avaro, and H. Sanson. Shape distances for contour tracking and motion estimation. *Pattern Recognition*, 2(7):1295–1304, June 1999.

7. H. Ming-Kuel. Visual pattern recognition by moment invariants. *IRE trans. on Information Theory*, 8:179–187, 1962.

8. S.S. Reddi. Radial and angular moment invariants for image identification. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 3(2):240–242, March 1981.

9. Y. Abu-Mostafa and D. Psaltis. Image normalisation by complex moments. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 7(1):46–55, January 1985.

10. I. Rothe, K. Voss, and H. Suesse. The method of normalisation to determine invariants. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 18(4):366–375, April 1996.

11. M. Pawlak. On the reconstruction aspects of moment descriptors. *IRE trans. on Information Theory*, 38(6):1698–1708, November 1992.

12. M.R. Teague. Image analysis via the general theory of moments. *J. of the Optical Society of America A*, 70(8):920–930, August 1980.

13. C.H. Teh and R.T. Chin. On image analysis by the method of moments. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 10(4):496–513, July 1988.

14. S.X. Liao and M. Pawlak. On image analysis by moments. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 18(3):254–266, March 1996.

15. S.C. Verrall and R. Kakarala. Disk-harmonic coefficients for invariant pattern recognition. *J. of the Optical Society of America A*, 15(2):289–301, February 1998.

16. D. Casasent and D. Psaltis. Scale invariant optical transform. *Optical Engineering*, 15(3):258–261, May/June 1976.

17. T. Yatagay, K. Choji, and H. Saito. Pattern classification using optical Mellin transform and circular photodiode array. *Optical Communication*, 38(3):162–165, August 1981.

18. M.R. Inggs and A.R. Robinson. Neural approaches to ship target recognition. In *Proc. of the IEEE Int. Radar Conference*, pages 386–391, Alexandria (Eg.), May 1995.

19. Y. Sheng and H.H. Arsenault. Experiments on pattern recognition using invariant Fourier-Mellin descriptors. *J. of the Optical Society of America A*, 3(6):771–776, June 1986.

20. A.E. Grace and M. Spann. A comparison between Fourier-Mellin descriptors and moment based features for invariant object recognition using neural networks. *Pattern Recognition Letters*, 12:635–643, October 1991.

21. P.E. Zwicke and Z. Kiss. A new implementation of the Mellin transform and its application to radar classification. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 5(2):191–199, March 1983.

22. F. Ghorbel. A complete invariant description for gray-level images by the harmonic analysis approach. *Pattern Recognition Letters*, 15:1043–1051, October 1994.

23. R. Milanese and M. Cherbuliez. A rotation-, translation-, and scale-invariant approach to content-based image retrieval. *J. of Visual Communication and Image Representation*, 10:186–196, doi:10.1006/jvci.1999.0411, June 1999.

24. J.J.K.. Ruanaidh and T. Pun. Rotation-, translation-, and scale-invariant digital image watermarking. In *Proc. of the IEEE Int. Conf. on Image Processing*, pages 536–539, Santa-Barbara (Ca.), October 1997.

25. J.P. Gauthier, G. Bornard, and M. Silbermann. Motion and pattern analysis: Harmonic analysis on motion groups and their homogeneous spaces. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 21(1):159–171, January 1991.

26. J. Turski. Projective fourier analysis for patterns. *Pattern Recognition*, 33:2033–2043, 2000.

27. R. Lenz. *Group theoretical methods in image processing*, volume 413 of *Lecture notes in computer science*. Springer-Verlag, Berlin, Germany, G. Goos and J. Hartmanis edition, 1990.

28. J.S. Segman, J. Rubinstein, and Y.Y. Zeevi. The canonical coordinates method for pattern deformation: Theoretical and computational considerations. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 14(12):519–523, December 1992.

29. R. Lenz. Group theoretical transforms in image processing. *Current topics in Pattern Recognition Researches*, 1:83–106, 1994.

30. R. Bracewell. *The Fourier transform and its applications*. McGraw-Hill Inc., 2nd edition, New York (USA), 1986.

31. S. Goh. *The Mellin transformation : Theory and digital filter implementation*. Ph.D. dissertation, Purdue University, West Lafayette, I.N., 1985.

32. Y. Sheng and J. Duvernoy. Circular-Fourier radial-Mellin transform descriptors for pattern recognition. *J. of the Optical Society of America A*, 3(6):885–888, June 1986.

33. H. Wechsler and G.L. Zimmerman. 2D invariant object recognition using distributive associative memory. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 10(6):811–821, November 1988.

34. Y. Sheng and C. Lejeune. Invariant pattern recognition using Fourier-Mellin transforms and neural networks. *J. of Optics*, 22(5):223–228, 1991.

35. J.L. de Bougrenet de la Tocnaye and F. Ghorbel. Scale-rotation invariant pattern recognition applied to image data compression. *Pattern Recognition Letters*, 8:55–58, July 1988.

36. S.A. Nene, S.K. Nayar, and H. Murase. Columbia Object Image Library (COIL-20). Technical Report CUCS-005-96, Columbia University, February 1996.