Pattern Recognition Letters 40 (2014) 27–35

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Color Fourier–Mellin descriptors for image recognition ☆

J. Mennesson [a,*], C. Saint-Jean [b], L. Mascarilla [b]

[a] Laboratory LIFL, Cité scientifique, Villeneuve d'Ascq, France
[b] Laboratory MIA, Avenue Michel Crépeau, La Rochelle, France

ARTICLE INFO

Article history:
Received 26 July 2013
Available online 6 January 2014

Keywords:
Object recognition
Image retrieval
Invariant color descriptors
Frequency methods
Clifford algebra

ABSTRACT

We propose new sets of Fourier–Mellin descriptors for color images. They are constructed using the Clifford Fourier transform of Batard et al. (2010) [4] and are an extension of the classical Fourier–Mellin descriptors for grayscale images. These are invariant under direct similarity transformations (translations, rotations, scale) and marginal treatment of colors images is avoided. An implementation of these features is given and the choice of the bivector (a distinguished color plane which parameterizes the Clifford Fourier transform) is discussed. The proposed formalism extends and clarifies the notion of direction of analysis as introduced for the quaternionic Fourier–Mellin moments (Guo and Zhu, 2011). Thus, another set of descriptors invariant under this parameter is defined. Our proposals are tested with the purpose of object recognition on well-known color image databases. Their retrieval rates are favorably compared to standard feature descriptors.

1. Introduction

In the literature, there are many recent advances in terms of image recognition. The recognition process depends highly on discriminative and invariant descriptors. Two main approaches can be used: global methods which concern features calculated on the entire image [5,38,35] and local methods based on feature extraction around of keypoints (e.g. SIFT [23], GLOH [27], FAST [30]).

Among global methods, one can cite moment-based descriptors [15] such as Hu invariants [20], Legendre moments or Zernike moments [37]. Other approaches based on the computation of an histogram can be used [8]. An alternative to these methods is to define descriptors in the frequency domain. For example, the family of Fourier descriptors [38,34] is widely used because of their properties of invariance, speed of convergence, low computational time. Originally, the 1-D Fourier descriptors [9,3] are obtained through Fourier transform (FT) on a shape signature function derived from shape boundary coordinates. More recently, 2-D Fourier descriptors have been proposed. In this case, it is assumed that images contain only one object and that the background is uniform. The most common are the Generalized Fourier Descriptors (*GFD*) [34] (invariant under translation and rotation) and the Fourier–Mellin descriptors [32,10] (invariant under translation, rotation and scale) computed from the Fourier–Mellin transform of a grayscale image. This later is widely applied in the field of document processing [32,17].

Originally, the Fourier–Mellin method is based on the polar transformation of the image, followed by a Fourier transform then a Mellin transform. More recently, Derrode and Ghorbel [10] proposed a complete set of Fourier–Mellin descriptors using an analytical Fourier–Mellin transform. Three algorithms which consider the original image, its polar and log-polar forms are defined to accelerate the computation of these descriptors. This work emphasized the effect of the polar and log-polar transformation of an image which are not exact (numerical interpolation is needed). This is a well-known open issue that is currently under investigation by Fenn et al. [14,22].

Extending these approaches to color images is not straightforward because they rely on the definition of a Fourier transform on color images. More precisely, these images are no longer viewed as functions from $\mathbb{R}^2$ to $\mathbb{R}$ but from $\mathbb{R}^2$ to $\mathbb{R}^3$: the value of each pixel is not a scalar but a vector. A classical generalization to color images is the use of an *ad hoc* approach like the marginal one [34], *i.e.* a separate treatment of each red, green, blue color plane. Another method consists in encoding *RGB* color space within the space of pure quaternions. In this framework, Sangwine and Ell proposed a Quaternionic Fourier Transform (*QFT*) [31]. This one is defined by replacing the imaginary unit i in the exponential of the Fourier transform by a pure unit quaternion μ, interpreted as a direction of analysis. This latter is commonly set as the gray level axis to obtain a luminance/chrominance analysis. Based on this *QFT*, Guo and Zhu [18] derivated a quaternionic extension of the Fourier–Mellin moments [32] with application to color image registration.

☆ This paper has been recommended for acceptance by A. Fernandez-Caballero.
* Corresponding author. Tel.: +33 359632233.
E-mail addresses: jose.mennesson@univ-lille1.fr (J. Mennesson), christophe.saint-jean@univ-lr.fr (C. Saint-Jean), laurent.mascarilla@univ-lr.fr (L. Mascarilla).

http://dx.doi.org/10.1016/j.patrec.2013.12.014

Clifford algebras [19], which contains the quaternion algebras, can also be used to embed and process color images. In our previous works, the *GFD* have been extended in several ways to color images yielding the Generalized Color Fourier Descriptors [24] (*GCFD*1 and *GCFD*2) by using a Clifford Fourier transform (*CFT*) dedicated to color images [4].

In this paper, we define new sets of Fourier–Mellin descriptors for color images, namely *poFMD*, *CFMD* and *CFMDi*, which are different extensions of the Fourier–Mellin moments computed from the *CFT*.

In Section 2, definition and a fast implementation of the *CFT* are recalled. Then, in Section 3–5, the three differents color Fourier–Mellin descriptors are defined (their invariance under translations, rotations and scale changes are proven in (A) and (B)). Finally, in Section 6 our proposals are tested with the purpose of object recognition and retrieval on well-known color image databases. Their retrieval rates are compared to standard feature descriptors.

2. Clifford Fourier transform for color images (*CFT*)

Classical Fourier transforms [6,5,16] are usually defined for complex valued functions that suited well for gray level, and not for color images. The most immediate solution is to compute three Fourier transforms independently on each channel of the color image. This marginal method raises problems as emergence of false colors in the case of color image filtering.

To avoid this marginal treatment, Batard et al. [4] defined a Fourier transform for $L^2(\mathbb{R}^2;\mathbb{R}^4)$ functions using Clifford Algebras [19]. This one is different from other color Fourier transforms [31,12] because it clarifies relations between the Fourier transform and the action of the translation group through a spinor group. This point of view justifies the necessity of choosing a direction of analysis. It is also demonstrated in [4] that the quaternionic Fourier transform defined by Sangwine and Ell [31] is a particular case of this definition.

2.1. Definition of the CFT

The *RGB* pixels of a color image can be embedded in $\mathbb{R}^1_{4,0}$ algebra (vectors of $\mathbb{R}_{4,0}$) as follows

$$f(\mathbf{x}) = r(\mathbf{x})\mathbf{e_1} + g(\mathbf{x})\mathbf{e_2} + b(\mathbf{x})\mathbf{e_3} + 0\mathbf{e_4}. \quad (1)$$

with $\mathbf{x} = (x_1, x_2)$ and $r,\ g,\ b$ are red, green and blue channels of a color image.

The color Clifford Fourier transform *CFT* [4] of $f \in L^2(\mathbb{R}^2;\mathbb{R}^1_{4,0})$ functions (*i.e.* a color image) with respect to an unit bivector B (identifiable to an analysis plane) is the vector-valued function

$$\widehat{f}_B(\mathbf{u}) = \int_{\mathbb{R}^2} e^{\frac{1}{2}\langle \mathbf{u},\mathbf{x}\rangle B} e^{\frac{1}{2}\langle \mathbf{u},\mathbf{x}\rangle I_4 B} f(\mathbf{x}) e^{-\frac{1}{2}\langle \mathbf{u},\mathbf{x}\rangle I_4 B} e^{-\frac{1}{2}\langle \mathbf{u},\mathbf{x}\rangle B} d\mathbf{x} \quad (2)$$

where I_4 is the pseudo-scalar of $\mathbb{R}_{4,0}$ and I_4B is an unit bivector which is orthogonal to B. This color Fourier transform is invertible and the inverse of the *CFT* is given by

$$f(\mathbf{x}) = \int_{\mathbb{R}^2} e^{-\frac{1}{2}\langle \mathbf{u},\mathbf{x}\rangle B} e^{-\frac{1}{2}\langle \mathbf{u},\mathbf{x}\rangle I_4 B} \widehat{f}_B(\mathbf{u}) e^{\frac{1}{2}\langle \mathbf{u},\mathbf{x}\rangle I_4 B} e^{\frac{1}{2}\langle \mathbf{u},\mathbf{x}\rangle B} d\mathbf{u}. \quad (3)$$

A vector can be decomposed in a parallel part and an orthogonal part depending on the choice of the bivector B [19]. Being f an image and B a bivector, this decomposition is $f = fBB^{-1} = (f \cdot B + f \wedge B)B^{-1} = f_{\parallel B} + f_{\perp B}$ where $f_{\parallel B} = (f \cdot B)B^{-1}$ (resp. $f_{\perp B} = (f \wedge B)B^{-1}$) is the parallel (resp. orthogonal) projection of f on a bivector B.

After some elementary calculations, Eq. (2) can be rewritten depending on this decomposition $\widehat{f}_B(\mathbf{u}) = \widehat{f}_{\parallel B}(\mathbf{u}) + \widehat{f}_{\perp B}(\mathbf{u})$ where

$$\widehat{f}_{\parallel B}(\mathbf{u}) = \int_{\mathbb{R}^2} e^{\frac{\langle \mathbf{u},\mathbf{x}\rangle}{2}B} f_{\parallel B}(\mathbf{x}) e^{-\frac{\langle \mathbf{u},\mathbf{x}\rangle}{2}B} d\mathbf{x} = \int_{\mathbb{R}^2} f_{\parallel B}(\mathbf{x}) e^{-\langle \mathbf{u},\mathbf{x}\rangle B} d\mathbf{x}, \quad (4)$$

$$\widehat{f}_{\perp B}(\mathbf{u}) = \int_{\mathbb{R}^2} f_{\perp B}(\mathbf{x}) e^{-\langle \mathbf{u},\mathbf{x}\rangle I_4 B} d\mathbf{x}. \quad (5)$$

Now, the bivectors B and I_4B can be identified to a pure imaginary number i since $B^2 = (I_4B)^2 = -1$. Eqs. (4) and (5) can be calculated using two usual fast Fourier transforms.

Depending on the application, it may be advisable to reconstruct $\widehat{f}_B$ from $\widehat{f}_{\parallel B}$ and *widehatf*$_{\perp B}$. The problem can be modelled as a system of four equations where the unknowns are the coordinates of $\widehat{f}_B$ in the basis $\{\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}, \mathbf{e_4}\}$ (see details in [25,26]).

Computational complexity of the color Clifford Fourier transform, including the reconstruction step, is $\mathcal{O}(nlog(n))$ where n is the number of pixels of the considered image. Indeed, this one requires only eight projections ($\mathcal{O}(n)$) and two fast Fourier transforms ($\mathcal{O}(nlog(n))$).

The bivector B is a required parameter of the *CFT*, hence of any derivated descriptors that are not invariant to the parameter. The choice of a given B can be left to the user, assuming some prerequired knowledge about the dataset at hand. The next subsection gives some guidelines for such choice.

2.2. Practical construction of a bivector B

An unit bivector B can be obtained with taking the geometric product of two unit vectors $\mathbf{v}_1,\ \mathbf{v}_2$, orthogonal to each other w.r.t. the quadratic form $\mathcal{Q}$. The corresponding bivector can be geometrically interpretable as the oriented plane spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$. Note that when $\mathbf{v}_1$ and $\mathbf{v}_2$ are not colinear, it is always possible to find an $\mathcal{Q}$-orthonormal basis taking the rejection of $\mathbf{v}_2$ on $\mathbf{v}_1$ and scaling to unity. If the bivector $\mathbf{B_c} = \mathbf{c} \wedge \mathbf{e}_4$ (with $\mathbf{c} = c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + c_3\mathbf{e}_3$ a normed color vector chosen by the user) is considered, the direction of analysis of the *CFT* is the same as the *QFT* by considering the unit quaternion $\mu = c_1 i + c_2 j + c_3 k$ in [13]. Let's recall that the *QFT* is used by Guo and Zhu [18] with $\mu = \frac{i+k}{\sqrt{2}}$.

In contrast of the quaternionic Fourier transform, the bivector used in the *CFT* is more general and allows, for example, to take hue planes by taking bivectors of the form $B_{\underline{\mathbf{c}}} = \mathbf{c} \wedge \mathbf{gray} = (c_1\mathbf{e_1} + c_2\mathbf{e_2} + c_3\mathbf{e_3} \wedge \frac{\mathbf{e_1}+\mathbf{e_2}+\mathbf{e_3}}{\sqrt{3}})$ with $\mathbf{c}$ a color. In this case, the two vectors must be orthogonalized using the following rejection formula:

$$\mathbf{v_3} = (\mathbf{v}_2 \wedge \mathbf{v}_1)\mathbf{v}_1^{-1}, \quad (6)$$

with $\mathbf{v}_1,\ \mathbf{v}_2$ two non orthogonal vectors, $\mathbf{v_3}$ a vector orthogonal to $\mathbf{v}_1$ in the plane generated by $\mathbf{v}_1 \wedge \mathbf{v}_2$ and $\wedge$ the outer product.

In Section 6, the sensitivity to the choice of B is tested and is emphasized by different applications.

3. The parallel-orthogonal Fourier–Mellin descriptors (*poFMD*)

In this section, we propose to compute the classical Fourier–Mellin moments (FMM) on parallel and orthogonal parts of the *CFT* as in [25] with the Generalized Fourier Descriptors. The two sets of moments can be concatenated and normalized to obtain a description vector.

3.1. Definition of the poFMM

Definition 1. The Fourier–Mellin Moments (*FMM*) are defined for an image $f \in L^2(\mathbb{R}^2, \mathbb{C})$, expressed in polar coordinates, as

$$FMM_f(m,n) = \int_{r=0}^{\infty}\int_{\theta=0}^{2\pi} r^{m-1} f(r,\theta) e^{-in\theta} d\theta dr. \quad (7)$$