



HAL
open science

New geometric fourier descriptors for color image recognition

José Mennesson, Christophe Saint-Jean, Laurent Mascarilla

► **To cite this version:**

José Mennesson, Christophe Saint-Jean, Laurent Mascarilla. New geometric fourier descriptors for color image recognition. International Conference on Image Processing (ICIP), Sep 2010, Hong Kong, China. pp.2685-2688, 10.1109/ICIP.2010.5651939 . hal-00599963v2

HAL Id: hal-00599963

<https://hal.science/hal-00599963v2>

Submitted on 20 Jun 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NEW GEOMETRIC FOURIER DESCRIPTORS FOR COLOR IMAGE RECOGNITION

José Mennesson, Christophe Saint-Jean, Laurent Mascarilla

MIA - Laboratoire Mathématiques, Image et Applications
Avenue Michel Crépeau
17042 La Rochelle, France

ABSTRACT

This article relies on two recent developments of well known methods which are a color Fourier transform using geometric algebra [1] and Generalized Fourier descriptors defined from the group M_2 of the motion of the plane [2]. In this paper, new generalized color Fourier descriptors (*GCFD*) are proposed. They depend on the choice of a bivector B acting as an analysis plane in a colorimetric space. The relevance of proposed descriptors is discussed on several color image databases. In particular, the influence of parameter B is studied regarding the type of images. It appears that proposed descriptors are more compact with a lower complexity and better classification rate.

Index Terms— Color Fourier descriptors, geometric algebra.

1. INTRODUCTION

In the literature, there's many recent advances in terms of image recognition. The recognition process depends highly on discriminative and invariant descriptors. Among them, one can cite moment-based descriptors [3] such as Hu invariants, Legendre moments or Zernike moments. Beside this approach, SIFT (Scale-Invariant Feature Transform) descriptors are a popular choice giving very good results [4].

An alternative to these methods is to define descriptors in the frequency domain. In this framework, our paper concerns an extension of Fourier descriptors. Clearly, Fourier coefficients don't respect the classical invariances (translation, rotation and scale) and must be processed to obtain invariant descriptors. This paper propose an extension of a recent advance concerning generalized Fourier descriptors defined by F. Smach *et al.* [2].

The extension of these descriptors to the color images is generally based on a marginal processing of the three channels R, G and B. Then, descriptors extracted from each channel are concatenated to form the description vector. In order to avoid this marginal processing, our proposal is to extract descriptors from a color Clifford-Fourier transform as defined by Batard *et al.* [1].

This paper is structured as follows : In the section 2, the definition of the Clifford-Fourier transform is introduced. Then, section 3 presents the extraction of generalized descriptors from 2D Fourier transform. New geometric Fourier descriptors, corresponding to our contribution, are defined in section 4. Finally, experimentations on well-known databases are presented in the section 5.

2. A COLOR CLIFFORD-FOURIER TRANSFORM

The usual Fourier transform can be applied to $L^2(\mathbb{R}^m; \mathbb{R} \text{ or } \mathbb{C})$ functions. Defining a color image as a function from \mathbb{R}^2 to \mathbb{R}^3 , the usual definition is no more usable. Several authors have proposed to tackle this problem by an ad-hoc processing of each channel. As an example, Smach *et al.* compute three Fourier transforms on each channel while Ell and Sangwine [5] considered only two Fourier transforms: one on the luminance and the other on the chrominance plane. Intuitively, these works emphasize the notion of analysis direction in applications to color images.

Recently, Batard *et al.* [1] have defined a Fourier transform for $L^2(\mathbb{R}^m; \mathbb{R}^n)$ functions. This one is mathematically rigorous and clarifies relations between the Fourier transform and the action of the translation group through a spinor group. They show that the previously proposed generalizations for color images (i.e. $n=3$) are particular cases of their definition. In this paper, only the particular case $m = 2$ and $n = 3$ is considered and described shortly in the following.

Firstly, in the equation of the classical 2D Fourier transform, the term $e^{-i(ux+vy)}$ rotates $f(x, y)$ in the complex plane \mathbb{C} . From a mathematical point of view, it corresponds to the action of the group \mathcal{S}^1 on \mathbb{C} which can be identified as the group action of $SO(2)$ on \mathbb{R}^2 . In order to generalize this principle to color images, one have to consider the action of the matrix group of $SO(3)$ on \mathbb{R}^3 . However, as described in [1], a general and elegant expression may be written if the function corresponding to the image is embedded in the Clifford algebra $\mathbb{R}_{4,0}$.

Clifford algebra $\mathbb{R}_{4,0}$ is a geometric algebra over the vector space \mathbb{R}^4 , equipped with an euclidean quadratic form, on which a new product, called geometric product [6] is defined.

Then, a color image can be viewed as a \mathbb{R}^2 into $\mathbb{R}_{4,0}$ function:

$$f(x, y) = r(x, y)e_1 + g(x, y)e_2 + b(x, y)e_3 + 0e_4$$

Within this framework, the rotation of a vector v by the angle θ , in the plane generated by a unitary bivector B can be written as

$$v \rightarrow e^{\frac{\theta}{2}B} v e^{-\frac{\theta}{2}B}$$

Bivector B is an element of the geometric algebra $\mathbb{R}_{4,0}$ which parametrizes the analysis direction. Function f can be decomposed as the sum of its parallel projection $f_{\parallel B}$ on the plane defined from B and its perpendicular projection $f_{\perp B}$ on the plane defined from $I_4 B$ (I_4 being the pseudo-scalar of $\mathbb{R}_{4,0}$). Skipping some technical details, one can prove that the values of $f(x, y)$ can be rotated by two independent rotations of the same angle, leading to the definition of a Clifford-Fourier transform:

$$\begin{aligned} \widehat{f}_B(u, v) = & \int_{\mathbb{R}^2} e^{\frac{(ux+vy)}{2}B} f_{\parallel B}(x, y) e^{-\frac{(ux+vy)}{2}B} dx dy + \\ & \int_{\mathbb{R}^2} e^{\frac{(ux+vy)}{2}I_4 B} f_{\perp B}(x, y) e^{-\frac{(ux+vy)}{2}I_4 B} dx dy \end{aligned} \quad (1)$$

Identifying the plane defined from B and $I_4 B$ to the complex plane \mathbb{C} , this color Fourier transform can be computed easily and efficiently with two complex 2D-FFT:

$$\widehat{f}_B = \widehat{f}_{\parallel B} + \widehat{f}_{\perp B} \quad (2)$$

Note that a unit bivector B can be obtained from the geometric product of two unit orthogonal vectors as $C_1 \wedge e_4$ or $C_2 \wedge C_3$ (where C_1, C_2 and C_3 are RGB colors). These two choices of bivectors are equivalent because the orthogonal bivector to the unitary bivector $B = C_1 \wedge e_4$ can be written as $C_2 \wedge C_3$ and vice versa. Obviously, an inverse transform exists and it is denoted by \widetilde{f}_B [1].

3. GENERALIZED FOURIER DESCRIPTORS

Generalized Fourier descriptors introduced by [2] are defined from the group action of M_2 . This latter group is composed of translations and rotations on the plane. Two kinds of descriptors have been defined :

- "Spectral densities"-type invariants:

$$I_1^r(f) = \int_0^{2\pi} |\widehat{f}(r, \theta)|^2 d\theta$$

- "Shift of phases"-type invariants:

$$I^{\xi_1, \xi_2}(f) = \int_0^{2\pi} \widehat{f}(R_\theta(\xi_1 + \xi_2)) \widetilde{\widehat{f}}(R_\theta(\xi_1)) \widetilde{\widehat{f}}(R_\theta(\xi_2)) d\theta$$

where f is the image, $\widehat{f}(r, \theta)$ is the Fourier transform expressed in polar coordinates in the frequency plane, ξ_1 and ξ_2 are variables of the frequency plane and R_θ is a rotation of angle θ . It must be emphasized that, by construction, I_1^r and I^{ξ_1, ξ_2} are strictly invariant in \mathbb{R}^2 with respect to the action of M_2 .

Then, the descriptor vector for the first family of invariants, namely I_1^r , is defined as follows :

$$GFD1(f) = \left\{ I_1^0(f) = |\widehat{f}(0, 0)|^2, \frac{I_1^1(f)}{I_1^0(f)}, \dots, \frac{I_1^m(f)}{I_1^0(f)} \right\}$$

where m is the number of computed descriptors. In the same way, we define the descriptor vector corresponding to the second family of invariants I^{ξ_1, ξ_2} . These ones are called *GFD2*.

4. CONTRIBUTION

In order to deal with color images, a commonly used approach consists in computing descriptors on each color channel separately. Then, they are concatenated into a unique vector (e.g. [2]). This method implies three FFT and three sets of descriptors. However, this marginal processing induces a loss of colorimetric informations that can be avoided by using color Clifford-Fourier transform.

Using the color Clifford-Fourier transform and the extraction of generalized Fourier descriptors, new color Fourier descriptors are defined as in Figure 1.

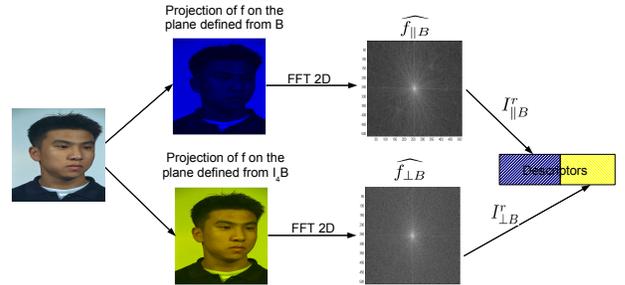


Fig. 1. Descriptors extraction from the color Clifford-Fourier transform.

Equation (2) shows that the Clifford-Fourier transform can be decomposed into two parts. So, two descriptor vectors are defined: $GCFD1_{\parallel B}$ and $GCFD1_{\perp B}$, each of them implying two complex 2D-FFT. According to the definition of $f_{\parallel B}$:

$$GCFD1_{\parallel B}(f) = \left\{ I_{\parallel B}^0(f), \frac{I_{\parallel B}^1(f)}{I_{\parallel B}^0(f)}, \dots, \frac{I_{\parallel B}^m(f)}{I_{\parallel B}^0(f)} \right\}$$

where $I_{\parallel B}^r(f) = \int_0^{2\pi} |\widehat{f}_{\parallel B}(r, \theta)|^2 d\theta$ and m is the number of computed descriptors. Similarly, $GCFD1_{\perp B}$ is defined thanks to $\widehat{f}_{\perp B}$.

Finally, the descriptor vector length is $2 \times m$:

$$GCFD1_B(f) = \{GCFD1_{\parallel B}(f), GCFD1_{\perp B}(f)\}$$

The same construction based on I^{ξ_1, ξ_2} leads to $GCFD2_B$.

5. EXPERIMENTS

In order to evaluate the discrimination capacity of our proposed descriptors, the Coil-100 [7], the color FERET [8] and the SFU [9] databases are considered.

The COIL-100 database is composed of 7200 color pictures of size 128×128 of 100 objects. Each picture has been taken with a black background and 72 different angles of view. This database, used in similar works [2], can be qualified as "easy" from an image classification point of view.

The color FERET database is composed of face images of 1408 different persons, taken from different angles of view. In our tests, a set of 2992 images containing 272 persons equally represented by 11 pictures is selected and size of images is reduced to 128×128 . This database is more difficult than the first one.

The noisy color FERET database contains the precedent color FERET database, but a gaussian noise is added to the images. The parameter σ is fixed to 0.23, which is the maximum noise used in [2].

The SFU database contains 20 objects photographed under 11 different illuminations and views. The size of images is also reduced to 128×128 . It is used to test the robustness of our descriptors under illumination variations.

5.1. Methodology

Size of descriptor vector. For the three databases, the descriptor vector length is set to 64 for each computed FFT. More precisely, $GFD1$ are built from 64 values of radius r in I_1^r and $GFD2$ are built from equally spaced values of ξ_2 in its polar domain $[0, 2\pi] \times [1, 8]$ and ξ_1 set to $(0, 1)$. Then, after concatenation, this results into 192 descriptors in the marginal approach, and in 128 descriptors for the parallel plus orthogonal part.

Classification. The classification step is performed by the SVM (Support Vector Machine) algorithm as implemented in the LIBSVM library [10]. The parameters maximizing the recognition rates are $\sigma = 0.1$ (gaussian kernel) and $C = 97$. To validate this decision step, a 10-fold cross-validation algorithm [11] is performed. Moreover, a modified SFFS algorithm (Sequential Floating Forward Selection) [11] is used to select bivectors B from their associated descriptor vectors.

5.2. Results

In our experiments, descriptors are built from unit bivectors $B = c \wedge e_4$, where c is a color to be chosen and \wedge the exterior

product¹ in the Clifford algebra $\mathbb{R}_{4,0}$. Results for bivectors $B_r = red \wedge e_4$, $B_g = green \wedge e_4$, $B_b = blue \wedge e_4$ and $B_\mu = \mu \wedge e_4$ where μ is the gray level, are taken as a reference to compare with. Let us notice that the first three correspond to the marginal case as done by many scholars [5, 2].

In order to investigate the robustness against a particular bivector B choice, 100 random unit colors c_{rand} were drawn uniformly in RGB cube yielding 100 unit bivectors in the form $B_{rand} = c_{rand} \wedge e_4$. Results are shown on the figure 2.

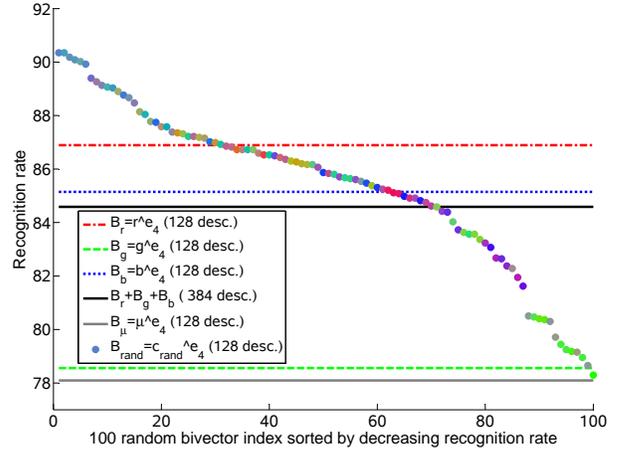


Fig. 2. color FERET: Recognition rate with $GCFD1$ for 100 random bivectors

The colors used to build the hundred random bivectors appears on each plot. Surprisingly, the rate obtained from $B_r + B_g + B_b$ is lower than that obtained from the best marginal B_r . One can suppose that there are redundant data in the three descriptor vectors which reduce the classification capacity of the SVM. The great variability of the results proves that the recognition quality depends clearly on the choice of the bivector. Our experiments show that the bivector is hard to choose for a set of heterogeneous images. However, for the color FERET database, it appears that a good choice corresponds to the background color of the images. Indeed, it yields a high energy and discriminative Fourier coefficient for most of the images.

The different recognition rates computed from the descriptors previously defined on different databases are presented in the Table 1.

Firstly, the top of the table concerning results obtained from one bivector is studied. Immediately, one can see that the recognition rates permits to define an order on the descriptors such as $GFD1 < GCFD1 < GFD2 < GCFD2$. Moreover, it can be noticed that on the noisy color FERET database, $GFD2$ descriptors give better results and are more robust against noise than $GCFD1$ descriptors with a half description length. Concerning the choice of the bivector, one

¹When 2 vectors are orthogonal their geometric product reduces to their exterior product.

	COIL-100		color FERET		Noisy color FERET				SFU	
Bivectors	<i>GFD1</i>	<i>GCFD1</i>	<i>GFD1</i>	<i>GCFD1</i>	<i>GFD1</i>	<i>GCFD1</i>	<i>GFD2</i>	<i>GCFD2</i>	<i>GFD1</i>	<i>GCFD1</i>
B_r	98.04	99.83	76.70	87.90	45.32	71.05	73.46	83.49	92.73	97.27
B_v	98.06	99.56	73.66	79.65	46.83	61.99	75.26	78.64	93.18	93.64
B_b	96.90	99.86	70.49	84.49	48.49	73.46	74.77	81.78	95.00	95.45
B_μ	98.49	99.25	73.03	78.10	55.28	62.03	77.34	80.98	95.45	93.18
$B_{rand}(\times 100)$	98.42 ± 0.3	99.54 ± 0.3	73.72 ± 1.0	85.34 ± 2.9	54.23 ± 1.7	69.64 ± 3.2	76.59 ± 0.7	82.56 ± 1.8	94.56 ± 1.5	93.00 ± 1.6
max.	98.87	99.89	76.14	90.37	57.55	77.27	78.41	87.00	97.27	96.36
$B_r + B_v + B_b$	99.9	99.92	88.03	85.53	73.16	72.16	83.25	81.12	96.36	93.63
$B_1 + B_2 + B_3(\text{SFFS})$	99.86	99.96	85.46	93.15	71.52	80.62	83.36	88.24	97.27	95.45
SIFT	100		87.55		NA				91.82	

Table 1. Recognition rates in % with *GFD1*, *GCFD1*, *GFD2*, *GCFD2* and SIFT descriptors (last line)

can remark that B_μ (luminance plus chrominance) is not a good choice for *GCFD1*. In the case of experimentations done from hundred random bivectors, the standard deviation obtained confirms that there is a relation between the recognition rates and the choice of the bivector.

The bottom of the table presents results achieved from a choice of a triplet of bivectors. The order previously defined is no more valid. However, the selection of a triplet of bivectors with an SFFS algorithm does not improve significantly results for *GFD1* and *GFD2* while results obtained from *GCFD1* and *GCFD2* are really better.

Using the SIFT implementation given by [12, 13], we can compare our descriptors with SIFT descriptors. One can see that the results are lower than results achieved from *GCFD1* or *GCFD2* (Table 1). For the SFU database, it is due to the lack of interest points that can be extracted from these images.

6. CONCLUSION

We can conclude that we defined new color Fourier descriptors from a Clifford-Fourier transform [1] and from generalized descriptors [2]. We obtain much better recognition rates with *GCFD1* and *GCFD2* than those with the marginal approach for many choices of bivector B . Their computational costs are still low and the descriptor vector size is only twice larger. Some robustness against gaussian noise is obtained with *GCFD2* descriptors. On chosen image databases, proposed descriptors gave higher recognition rates than the original SIFT descriptors. Unfortunately, selecting a good bivector remains still an open problem. However, \widehat{f}_B does not depend on bivector B . A deep work should be to defined *GFD* of Smach *et al.* directly for multi-vector valued functions. Finally, these first promising results encourage us to go farther with the general definition of the Clifford-Fourier transform [1]. Indeed, this one is able to process 4D signals like for example color+infrared images.

7. REFERENCES

- [1] T. Batard, M. Berthier, and C. Saint-Jean, "Clifford fourier transform for color image processing," in *Geometric Algebra Computing in Engineering and Computer Science*, E. Bayro-Corrochano and G. Scheuer-
- [2] F. Smach, C. Lemaître, J. P. Gauthier, J. Miteran, and M. Atri, "Generalized fourier descriptors with applications to objects recognition in svm context," *Journal of Mathematical Imaging and Vision*, vol. 30, no. 1, pp. 43–71, 2008.
- [3] J. Flusser, T. Suk, and B. Zitova, *Moments and Moment Invariants in Pattern Recognition*, Wiley, Chichester, 2009.
- [4] D. Lowe, "Object recognition from local scale-invariant features," 1999, pp. 1150–1157.
- [5] T.A. Ell and S.J. Sangwine, "Hypercomplex fourier transforms of color images," *IEEE Transactions on Image Processing*, vol. 16, no. 1, pp. 22–35, 01 2007.
- [6] D. Hestenes, H. Li, and A. Rockwood, "New algebraic tools for classical geometry," pp. 3–26, 2001.
- [7] S. Nene, S. K. Nayar, and H. Murase, "Columbia object image library (coil-100)," 1996, Tech. Rep. CUCS-006-96.
- [8] P.J. Phillips, H. Wechsler, J. Huang, and P. Rauss, "The feret database and evaluation procedure for face recognition algorithms," *Image and Vision Computing*, vol. 16, no. 5, pp. 295–306, 1998.
- [9] K. Barnard, L. Martin, B. Funt, and A. Coath, "A data set for colour research," vol. 27, pp. 147–151, 2002.
- [10] Chih C. Chang and Chih J. Lin, *LIBSVM: a library for support vector machines*, 2001.
- [11] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: a review," *ACM Comput. Surv.*, vol. 31, no. 3, pp. 264–323, 09 1999.
- [12] A. Vedaldi and B. Fulkerson, "VLFeat: An open and portable library of computer vision algorithms," 2008.
- [13] M. Muja and D. Lowe, "Fast approximate nearest neighbors with automatic algorithm configuration," in *VISS-APP (1)*, 2009, pp. 331–340.

mann Eds, Eds., chapter 8, pp. 135–162. Springer Verlag, 2010.