Quaternion Correlation Filters for Illumination Invariant Face Recognition

Dayron Rizo-Rodriguez[1], Heydi Méndez-Vázquez[1], Edel García[1], César San Martín[2], and Pablo Meza[3]

[1] Advanced Technologies Application Center. 7a # 21812 b/ 218 and 222, Rpto. Siboney, Playa, P.C. 12200, Havana, Cuba
{drizo,hmendez,egarcia}@cenatav.co.cu

[2] Center for Optics and Photonics, University of La Frontera, Casilla 54-D, Temuco, Chile
cesarsanmartin@ufro.cl

[3] Center for Optics and Photonics, University of Concepción, Casilla 160-C, Concepción, Chile
pablomeza@udec.cl

Abstract. Illumination variations is one of the factors that causes the degradation of face recognition systems performance. The representation of face image features using the structure of quaternion numbers is a novel way to alleviate the illumination effects on face images. In this paper a comparison of different quaternion representations, based on verification and identification experiments, is presented. Four different face features approaches are used to construct quaternion representations. A quaternion correlation filter is used as similarity measure, allowing to process together all the information encapsulated in quaternion components. The experiment results confirms that using quaternion algebra together with existing face recognition techniques permits to obtain more discriminative and illumination invariant methods.

1 Introduction

Variations in lighting conditions is one of the principal factors causing the deterioration of face recognition systems performance [1]. Several methods have been proposed to cope with the problem of face recognition under illumination variation [2]. Among them, in [3] taking into account the benefits of the use of quaternion algebra in image processing [4], an illumination invariant face image representation based on quaternion number structure was presented. Two representations, a complex and a quaternion one, based on image image differentiation, were constructed and compared regarding their illumination invariant properties. Both representations are transformed into frequency domain and cartesian and polar expressions are obtained. The most illumination invariant component of each representation is selected and used as face image descriptor. A simple normalized correlation is used as similarity measure. In that work [3], experimental results showed that quaternion representation is better

C. San Martin and S.-W. Kim (Eds.): CIARP 2011, LNCS 7042, pp. 467–474, 2011.

than the complex one, regarding illumination invariant and discriminative properties. However, only one component of quaternion representation is used as face image descriptor, discarding the remaining ones, where valuable discriminative information could be encapsulated. On the other hand, using the normalized correlation as similarity measure only permits to process one component at a time. More powerful tools are needed in order to use together all the components of quaternion representation to analyze face images.

Quaternion correlation filters have been specially designed for face recognition [5]. In [5], a multi-band processing analysis is performed, using discrete wavelet decomposition, to obtain the quaternion representation. The wavelet decomposition is used because directly provides a multi-resolution analysis of the image, however other face image extraction techniques can be used in order to obtain more illumination invariant and discriminative representations.

The aim of this work, is to compare the performance of quaternion correlation filters based on different face images representations, when dealing with illumination variations. Besides image differentiation and discrete wavelet decomposition, used in previous works for extracting the multi-band information of face images from which the quaternion representation is constructed, discrete cosine transform (DCT) and local binary patterns (LBP) are selected among different face images descriptors, because of the well known behavior of these methods in front of illumination variations [6,7]. The unconstrained optimal tradeoff quaternion filter (UOTQF) presented in [5], based on the traditional unconstrained optimal tradeoff filter (UOTF), is used to perform the cross-correlation based on each quaternion representation.

Verification and identification experiments were conducted in XM2VTS and Extended Yale B databases respectively. The quaternion representation constructed from LBP features showed the best performance in both face recognition experiments. The paper is organized as follows. In Section 2, face image decomposition to construct the quaternion frequency domain representation are described. In Section 3, the construction of quaternion correlation filter and its use in the recognition process is explained. The experimental results are drawn in Section 4. Finally, Section 5 gives the conclusions of the paper.

2 Face Images Quaternion Representation

Quaternion algebra was the first hypercomplex number system to be discovered, introduced by Hamilton in 1843 [8]. The cartesian representation of quaternion numbers is usually defined as follows:

$$q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \tag{1}$$

where a, b, c, d are real and $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ are orthogonal imaginary operators.

Based on Eq.(1), a general expression for face images quaternion representation, at pixel (x, y), can be defined as:

$$q(x, y) = Q_1(x, y) + Q_2(x, y)i + Q_3(x, y)j + Q_4(x, y)k \tag{2}$$

where $Q_1(x,y)$, $Q_2(x,y)$, $Q_3(x,y)$ and $Q_4(x,y)$ would be four descriptions of the image at (x,y) coordinate, using some face feature extraction method.

Then, to construct the face image quaternion description, the first step is to decompose the image in four bands of information. For this purpose, different feature extraction methods are evaluated in this work: image differentiation, discrete wavelet decomposition, discrete cosine transform and local binary patterns.

Image differentiation (DIF) has shown to be a face image descriptor less sensitive to illumination effects [9]. In [9], first order derivatives of the images, in x and y directions, are used to assemble a complex representation of face images. Calculating again first order derivatives, in x and y directions, over each one of these components, results in four descriptions of the images which actually are the second order derivatives of the image: $\bigtriangledown^2_{xx}I(x,y)$, $\bigtriangledown^2_{xy}I(x,y)$, $\bigtriangledown^2_{yx}I(x,y)$ and $\bigtriangledown^2_{yy}I(x,y)$. In this way, the quaternion representation based on image differentiation, can be expressed as:

$$q_{DIF}(x,y) = \bigtriangledown^2_{xx}I(x,y) + \bigtriangledown^2_{xy}I(x,y)\mathbf{i} + \bigtriangledown^2_{yx}I(x,y)\mathbf{j} + \bigtriangledown^2_{yy}I(x,y)\mathbf{k} \quad (3)$$

In practice, image differentiation is implemented by convolving the signal with some form of linear filter that approximates derivative operator. In our case a Sobel filter, the one applied in [9], was used. In Figure 1(a), the above DIF decomposition process is graphically illustrated.

The wavelet transform can be used to decompose an image into different scales and resolutions. When applying the discrete wavelet decomposition (DWT), the face image is passed through a low pass filter and a high pass filter to get the low and high frequency components of the original image. This process is applied iteratively to the low and high frequency bands in order to obtain the representation at different scales and resolutions. The implementation of this process is made by projecting the original image to the wavelet basis function. In [5] the Daubechies family of wavelets is used to decompose the face images into four subbands: *low-low*(LL), *low-high*(LH), *high-low*(HL) and *high-high*(HH), which are encoded in the quaternion representation as in Eq.(4). The illustration of this process can be found on Figure 1(b).

$$q_{DWT}(x,y) = W_{LL}(x,y) + W_{LH}(x,y)\mathbf{i} + W_{HL}(x,y)\mathbf{j} + W_{HH}(x,y)\mathbf{k} \quad (4)$$

The discrete cosine transform (DCT) has been very used in face recognition. In [7], a method using the DCT to compensate for illumination variations was presented. The illumination variations are compensated, setting to zero the low-frequency DCT coefficients of an image in the logarithm domain and reconstructing a normalized image applying the inverse DCT. Varying the low-frequency coefficients used for the illumination compensation, a multi-resolution representation can be obtained, discarding and retaining different information of the face image each time. In Figure 1(c), four face images obtained by applying this process with different low-frequency DCT coefficients each time, are shown. The quaternion representation using this method can be expressed as:

$$q_{DCT}(x,y) = DCT_{L1}(x,y) + DCT_{L2}(x,y)\mathbf{i} + DCT_{L3}(x,y)\mathbf{j} + DCT_{L4}(x,y)\mathbf{k} \quad (5)$$

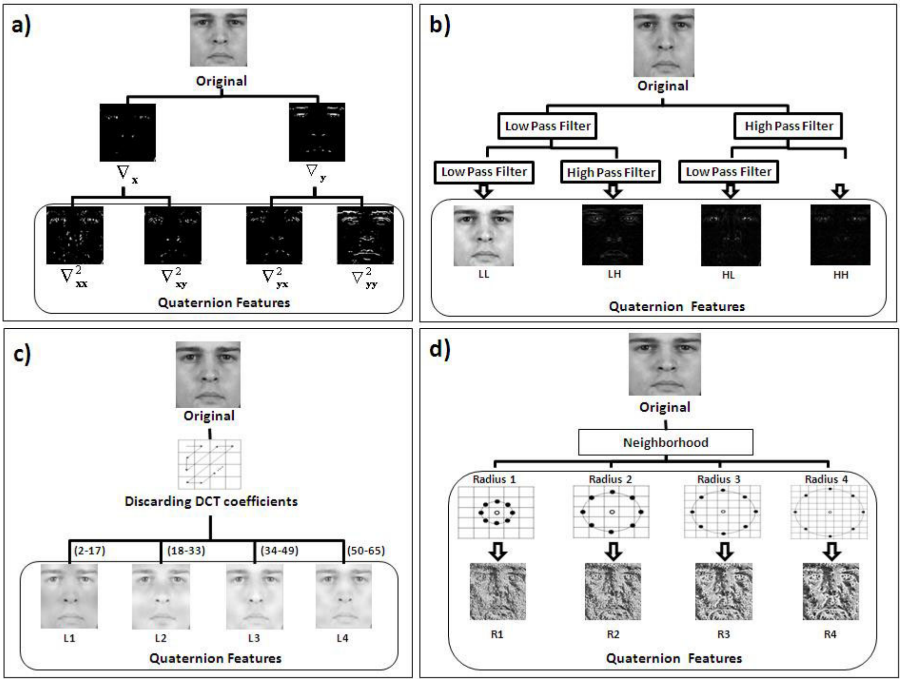

Fig. 1. Illustration of the processes of obtaining the four bands of information needed for quaternion representation, using each one of the face feature extraction method described: a) DIF, b) DWT, c) DCT and d) LBP

The local binary patterns operator (LBP) is a texture descriptor which has been very used in face analysis based on the idea that faces are composed by micro-patterns which can be well described by this operator [6]. The original LBP operator labels each pixel of an image with a value called LBP code, which corresponds to a binary number that represents its relation with the 3x3-local neighborhood. Different extensions of the original operator have been proposed. Among them, the multi-scale LBP [10], permits to codify the LBP operators at different neighborhood sizes, providing a multi-resolution analysis of face images. For the assembling of quaternion representation in Eq.(6), the LBP operator at four different radii are computed as it is shown in Figure 1(d).

$$q_{LBP}(x,y) = LBP_{R1}(x,y) + LBP_{R2}(x,y)\mathbf{i} + LBP_{R3}(x,y)\mathbf{j} + LBP_{R4}(x,y)\mathbf{k} \quad (6)$$

3 Quaternion Correlation Filters

The use of quaternion correlation filters in face recognition involves enrollment and recognition stages. During enrollment, the first step is to transform the quaternion description of training images, obtained with some of the processes described above, to the frequency domain.

The quaternion descriptions are transformed to the frequency domain, using Quaternion Discrete Fourier Transform (QDFT) [11] defined as:

$$Q(p,s) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} e^{-\mu 2\pi((pm/M)+(sn/M))} q(m,n) \tag{7}$$

where μ is any unit pure quaternion and q is the training face images quaternion representation in the form of Eq.(2).

The quaternion correlation filter is designed based on the QDFT of the training images. The unconstrained optimal tradeoff quaternion filter (UOTQF), proposed in [5], is used for this purpose. The derivation of this filter is similar to the one of the traditional unconstrained optimal tradeoff filter (UOTF) [12], and has the following closed form solution:

$$h = \gamma(\alpha D + C)^{-1} m \tag{8}$$

where h is the designed frequency domain filter represented in the vector form, m represents the frequency domain training image in the vector form, α and γ are tradeoff parameters, which can be tuned to obtain the optimal tradeoff between maximizing the discrimination ability and minimizing the output noise variance of the filter h. D is a diagonal matrix, where the main diagonal is the average power spectrum of the training images and C is also a diagonal matrix, representing the noise power spectral density (psd). Typically a white noise model is assumed, thus C takes the form of an identity matrix. In this formulation of the UOTQF filter, all vector and matrix terms are quaternion number arrays, while for the UOTF filters they are all complex number arrays.

The UOTQF are computed and stored for each subject on the training set. Then, at recognition stage, each testing image is transformed into quaternion frequency domain by applying Eq.(7) and it is cross-correlated with every UOTQF obtained at training.

Following [5], the specialized 2-D quaternion correlation (QC), is used for the cross-correlation computation. From the magnitude value of each quaternion correlation output, a similarity score is computed. A large peak value in correlation output plane is yielded in case of a genuine identity and no discernible peaks for an impostor. The fitness measure of the peak sharpness is calculated using the peak-to-sidelobe-ratio (PSR), defined in [5].

4 Experimental Evaluation

In order to compare the behavior of the quaternion representations based on the four selected face descriptors, verification and identification experiments were conducted in XM2VTS [13] and Extended Yale B [14] databases respectively.

4.1 Verification Experiment

Configuration I of the Lausanne protocol [13], designed for experiments on XM2VTS database, was used to compare the performance of the different representations on a face verification setting. Under this configuration, the 2360 face

Table 1. Verification Results in terms of TER (%)

	Eval.	Test	Dark
Q1 DIF	33.67	29.24	51.48
DWT	32.60	27.05	49.87
DIF	23.55	20.64	41.54
DCT	58.33	55.68	75.90
LBP	22.00	18.56	40.68

Table 2. Recognition Rates (%) obtained in Identification Experiment

	S1	S2	S3	S4	S5
Q1 DIF	100.0	100.0	93.14	38.60	06.05
DWT	98.22	100.0	73.71	44.73	11.92
DIF	100.0	100.0	99.24	97.15	59.07
DCT	59.11	58.33	19.42	16.66	30.60
LBP	99.56	100.0	97.71	95.83	82.92

images of 295 subject on the database, are divided into a Training, an Evaluation and a Test sets, composed of images under controlled illumination conditions used as clients and imposters. An additional set (Dark) which contains images of every subject under non regular lighting conditions is used to test the behavior of the methods in the presence of this kind of variations.

In a verification setup, the False Rejection Rate (FRR) and the False Acceptance Rate (FAR) are used as a measure of algorithms performance. The Equal Error Rate (EER), is the point where FRR = FAR. Under the selected protocol, the similarity value obtained by the classification method at this point in the Evaluation set is used as a threshold for the decision of acceptance or rejection in the Test and Dark sets.

The Total Error Rate (TER), which is the sum of FRR and FAR, is computed for each set of the database when applying the quaternion correlation filters using the four alternatives described in previous sections. The obtained results are shown in Table 1. The first row of the table corresponds to the results obtained in [3], where only one component of the quaternion representation based on DIF is selected and used as face descriptor.

It can be appreciated in the table, that quaternion correlation filters based on DIF and LBP outperform the one proposed in [5] based on DWT whether images are affected by illumination variations (Dark set) or not (Evaluation and Test sets). On the other hand, these results are also superior to use only one component of the quaternion representation. In general, the LBP method exhibits the best results in all sets of the database, while DCT achieved the worst results.

4.2 Identification Experiment

The Extended Yale B [14] database was used to conducted the identification experiments. It contains images of 28 subjects under 64 different illumination conditions. This database is usually divided into 6 subsets according to the angle of the incident illumination. Face images with frontal lighting are used as gallery and subsets S1, S2, S3, S4 and S5 grouped the images in a way that S1 contains the ones with minor variations and S5 the most affected images.

The recognition rates obtained in each subset of the database using the correlation filters based on the four representations are shown in Table 2. Also in this case, the results obtained with only one component of the quaternion

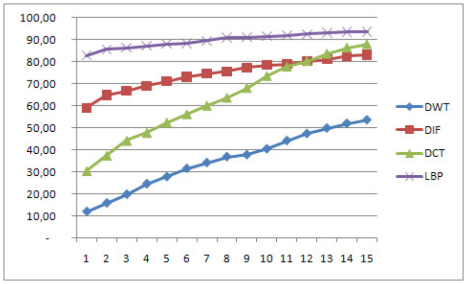

Fig. 2. Cumulative match score vs. rank curve in S5

is included in the first row of the table. Besides, in Figure 2, which represents the cumulative match score vs. rank curve, is illustrated the behavior of the four quaternion correlation filters in S5, the most difficult one respect to the illumination variations.

From the table, the quaternion filter based on DIF presents the best results in subsets S1, S2 S3 and S4, following by the one based on LBP for a little margin. However in S5, where the illumination variations are the greatest, the LBP method perform much better, which is corroborated in Figure 2, being the only method achieving more than 90% of correct classification in this subset. In this case, both LBP and DIF representation, and even the one based on DCT, perform significantly better than the one based on DWT and the use of the normalize cross-correlation of only one component of the quaternion.

5 Discussion and Conclusions

This paper presents a comparison of quaternion correlation filters based on different face images representations. Quaternion representations based on DIF, DWT, DCT and LBP are constructed and cross-correlated using the unconstrained optimal tradeoff quaternion filter (UOTQF).

The obtained results in verification and identification, confirm the hypothesis that using jointly the multi-band information encoded in a quaternion representation permits to retain more discriminative information of face images than only one component even though this is the most invariant one. On the other hand, selecting adequate face image descriptors, it is possible to improve the face recognition results on an specific problem.

In this case, the LBP method was selected because of its well known behavior dealing with the illumination problem on face recognition [6]. As was expected, the quaternion representation based on this face descriptor achieves the better results in the conducted experiments. Surprisingly, the DCT method, which has been also very used for illumination invariant face recognition [7], achieved the worst result. We think that the multi-resolution analysis based on this method and the way in which the DCT coefficients are discarded to form the quaternion need to be improved.

It can be concluded that quaternion algebra is a powerful mathematical tool that can lead to excellent results in face recognition problems, specially because of the possibility of encapsulating and processing together the multi-band information. For the continuity of this work, it is necessary to obtain better face descriptors and to analyze other correlation filters on quaternion domain, in order to improve the recognition results.

References

1. Adini, Y., Moses, Y., Ullman, S.: Face recognition: The problem of compensating for changes in illumination direction. IEEE Transactions on Pattern Analysis and Machine Intelligence 19(7) (1997)
2. Zou, X., Kittler, J., Messer, K.: Illumination invariant face recognition: A survey. In: BTAS 2007 (September 2007)
3. Rizo-Rodríguez, D., Méndez-Vázquez, H., García-Reyes, E.: Illumination Invariant Face Image Representation Using Quaternions. In: Bloch, I., Cesar Jr., R.M. (eds.) CIARP 2010. LNCS, vol. 6419, pp. 434–441. Springer, Heidelberg (2010)
4. Bayro, E.: Geometric Computing for Perception Action Systems, pp. 115–136. Springer, New York (2001)
5. Chunyan, X., Savvides, M., Vijayakumar, B.: Quaternion correlation filters for face recognition in wavelet domain. In: Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (2005)
6. Marcel, S., Rodriguez, Y., Heusch, G.: On the recent use of local binary patterns for face authentication. International Journal on Image and Video Processing Special Issue on Facial Image Processing (2007) IDIAP-RR 06-34
7. Chen, W., Er, M.J., Wu, S.: Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain. IEEE Transactions on Systems, Man and Cybernetics, Part B 36(2), 458–466 (2006)
8. Hamilton, W.: Elements of Quaternions. Longmans, Green, London, UK (1866)
9. Garea, E., Kittler, J., Messer, K., Mendez, H.: An illumination insensitive representation for face verification in the frequency domain. In: IEEE Computer Society, ICPR 2006, pp. 215–218 (2006)
10. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(7), 971–987 (2002)
11. Moxey, C., Sangwine, S., Ell, T.: Hypercomplex correlation techniques for vector images. IEEE Transaction on Signal Processing 51, 1941–1953 (2003)
12. Vijayakumar, B., Savvides, K., Venkataramani, K., Xie, C.: Spatial frequency domain image processing for biometric recognition. In: IEEE International Conference on Image Processing, pp. 53–56 (2002)
13. Messer, K., Matas, J., Kittler, J., Jonsson, K.: Xm2vtsdb: The extended m2vts database. In: Second International Conference on Audio and Video-based Biometric Person Authentication, pp. 72–77 (1999)
14. Lee, K.C., Ho, J., Kriegman, D.J.: Acquiring linear subspaces for face recognition under variable lighting. IEEE Transactions on Pattern Analysis and Machine Intelligence 27(5), 684–698 (2005)