MODELING GABOR COEFFICIENTS VIA GENERALIZED GAUSSIAN DISTRIBUTIONS FOR FACE RECOGNITION

Daniel González-Jiménez, Fernando Pérez-González, Pedro Comesaña-Alfaro, Luis Pérez-Freire, José Luis Alba-Castro

Signal Theory and Communications Department,
University of Vigo,
{danisub,fperez,pcomesan,lpfreire,jalba}@gts.tsc.uvigo.es

ABSTRACT

Gabor filters are biologically motivated convolution kernels that have been widely used in the field of computer vision and, specially, in face recognition during the last decade. This paper proposes a statistical model of Gabor coefficients extracted from face images using generalized Gaussian distributions (GGD's). By measuring the Kullback-Leibler distance (KLD) between the *pdf* of the GGD and the relative frequency of the coefficients, we conclude that GGD's provide an accurate modeling. The underlying statistics allow us to reduce the required amount of data to be stored (i.e. data compression) via Lloyd-Max quantization. Verification experiments on the XM2VTS database show that performance does not drop when, instead of the original data, we use quantized coefficients.

Index Terms— Face Recognition, Generalized Gaussian Distribution, Gabor filters, Kullback-Leibler distance, Lloyd-Max quantization, data compression, XM2VTS database

1. INTRODUCTION

Gabor filters are biologically motivated convolution kernels that have been widely used in face recognition during the last decade (see [1] for a recent survey). Basically, Gabor-based approaches fall into one of the following categories: **a)** Extraction of Gabor responses from a set of *key* points in face images and **b)** Convolution of the whole image with a set of Gabor filters. As highlighted in [1], one of the main drawbacks of these approaches (specially the ones included in category **b**) is the huge amount of memory that is needed to store a Gabor-based representation of the image. Regarding feature comparison, similarity between Gabor responses has been usually measured by means of normalized dot products (or related measures), but there is no theoretical evidence supporting this choice.

On the other hand, experiments have shown that generalized Gaussian distributions (GGD's) provide a good *pdf* approximation for the distribution of coefficients produced by several types of wavelet transforms [7, 8, 9]. To the best of our knowledge, despite the large number of papers using Gabor filters for face recognition, no statistical model has been proposed for Gabor coefficients. In this paper, we suggest that GGD's could provide a suitable modeling, and empirically validate this hypothesis using the Kullback-Leibler distance (KLD). The underlying statistics allow us to perform data compression via Lloyd-Max quantization, and open new possibilities in terms of selecting an optimal measure between Gabor reponses from a theoretical point of view.

The paper is organized as follows: Section 2 describes the baseline face recognition system used in this paper. Section 3 introduces the formulation of generalized Gaussian distributions (GGD's), as well as the modeling of Gabor coefficients using GGD's. Coefficient quantization by means of Lloyd-Max algorithm is explained in Section 4. The impact of coefficient quantization on verification performance is reported in Section 5 with experimental results on the XM2VTS database [3]. Finally, conclusions and future research lines are drawn in Section 6.

2. THE FACE RECOGNITION SYSTEM

A set of 40 Gabor filters $\{\psi_m\}_{m=1,2,\ldots,40}$ with the same configuration as in [2] (5 spatial frequencies and 8 orientations), is used to extract textural information from face images. These filters are biologically motivated convolution kernels in the shape of plane waves restricted by a Gaussian envelope [5], as it is shown next:

$$\psi_m(\vec{x}) = \frac{\left\|\vec{k}_m\right\|^2}{s^2}\exp\left(\frac{-\left\|\vec{k}_m\right\|^2\|\vec{x}\|^2}{2s^2}\right)\left[\exp\left(i\cdot\vec{k}_m^T\vec{x}\right)-\exp\left(\frac{-s^2}{2}\right)\right] \quad (1)$$

where $\vec{k}_m$ contains information about spatial frequency and orientation, and the same standard deviation $s = 2\pi$ is used in both directions for the Gaussian envelope. Figure 1 shows the real part of the 40 Gabor filters used in this paper.

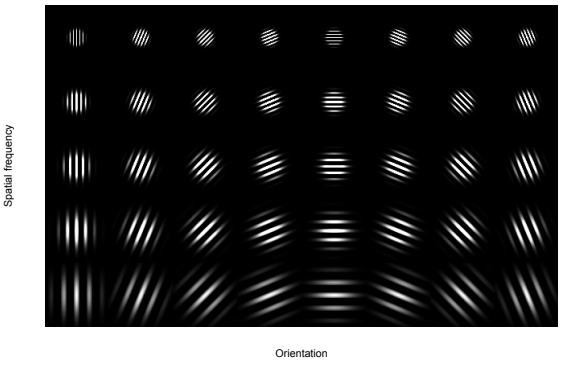

Fig. 1. Real part of the set of 40 (8 orientations $\times$ 5 scales) Gabor filters used in this paper.

The baseline face recognition system used in this paper relies upon extraction of Gabor responses at each of the nodes from a

$n_x \times n_y$ rectangular grid (Figure 2). All faces were geometrically normalized -so that eyes and mouth are in fixed positions-, cropped to a standard size of 150x116 pixels and photometrically corrected by means of histogram equalization and local mean removal. The region surrounding each grid-node in the image is encoded by the convolution of the image patch with these filters, and the set of responses is called a jet, $\mathcal{J}$. Therefore, a jet is a vector with 40 *complex* coefficients, and it provides information about a specific region of the image. At node $\vec{p}_i = [x_i, y_i]^T$ and for each Gabor filter ψ_m, $m = 1, 2 \ldots, 40$, we get the following Gabor coefficient:

$$g_m(\vec{p}_i) = \sum \sum I(x, y)\psi_m \left(x_i - x, y_i - y\right) \qquad (2)$$

where $I(x, y)$ represents the photometrically normalized image patch. Hence, the complete feature vector (jet) extracted at $\vec{p}_i$ is given by $\mathcal{J}(\vec{p}_i) = [g_1(\vec{p}_i), g_2(\vec{p}_i), \ldots, g_{40}(\vec{p}_i)]$. For a given a face with $n = n_x \times n_y$ grid-nodes $\{\vec{p}_1, \vec{p}_2, \ldots, \vec{p}_n\}$, we get n Gabor jets $\{\mathcal{J}(\vec{p}_1), \mathcal{J}(\vec{p}_2), \ldots, \mathcal{J}(\vec{p}_n)\}$.

3. MODELING GABOR COEFFICIENTS WITH GENERALIZED GAUSSIAN DISTRIBUTIONS (GGD'S)

Generalized Gaussian distributions have been successfully used to model coefficients produced by various types of wavelet transforms [7, 8, 9]. The *pdf* of a GGD is given by the following expression:

$$f_x(x) = Ae^{-|\beta x|^c} \qquad (3)$$

Both A and β can be expressed as a function of the so-called shaping factor c and the standard deviation σ

$$\beta = \frac{1}{\sigma}\left(\frac{\Gamma(3/c)}{\Gamma(1/c)}\right)^{1/2}$$

$$A = \frac{\beta c}{2\Gamma(1/c)}$$

where $\Gamma(.)$ is the complete gamma function. The shaping factor c is inversely proportional to the sharpness of the *pdf*. Therefore, this distribution is completely specified by two parameters, c and σ. Note that the Gaussian and Laplacian distributions are just special cases of this generalized *pdf*, given by $c = 2$ and $c = 1$ respectively.

In this paper, we attempt to model both real and imaginary parts of each Gabor coefficient using GGD's whose parameters have been obtained using the Maximum-Likelihood (ML) estimator. From a set of face images $\{\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_T\}$, we extract Gabor jets as introduced in Section 2. Regardless of the node from which they have been computed, the coefficients corresponding to a given Gabor filter ψ_m (real and imaginary parts separately) are stored together forming two sets of coefficients $\mathcal{S}_m^{real}$ and $\mathcal{S}_m^{imag}$. Now, our goal is to assess whether these distributions can be modeled using GGD's.

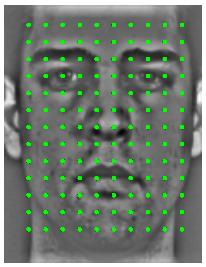

Fig. 2. Rectangular grid over the preprocessed (geometrically and photometrically normalized) face image. At each node, a Gabor jet with 40 coefficients is computed and stored.

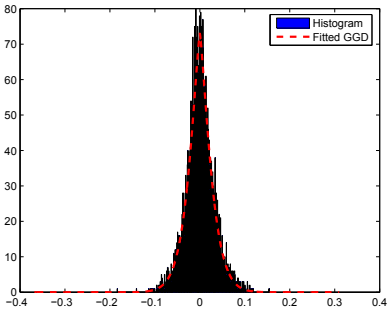

Fig. 3. Histogram for coefficient g_{34} along with the fitted GGD.

Figure 3 shows the histogram for the real part of coefficient g_{34} along with the fitted GGD. Although it seems clear from this figure that the GGD accurately models the coefficient distribution (similar plots were obtained for the remaining coefficients), we used the Kullback-Leibler distance (KLD) [12] to assess the goodness of the fits. The Kullback-Leibler distance between two discrete distributions with probability functions P and Q, is given by $KLD(P, Q) = \sum_{i=1}^{K} P(i)\log\frac{P(i)}{Q(i)} \geq 0$, where K stands for the number of intervals in which the sample space is divided. Figure 4 (left) plots, for both real and imaginary parts of each Gabor coefficient, the KLD between the relative frequency of the coefficient and the fitted GGD. Since the obtained distances are small, it seems reasonable to conclude that generalized Gaussians are able to model Gabor coefficients accurately. Other tests, such as the χ^2 test, have been previously used to assess the quality of the fit (for instance in [11]). Applying the χ^2 test to our data leads to the same conclusion -see Figure 4 right-.

Figure 5 presents the shaping factors c and the σ parameters of the 80 GGD's modeling both real and imaginary parts of Gabor coefficients. From this figure, we can conclude:

- All the GGD's have a shaping factor c that is well below 2, and hence we can conclude that the distribution of each Gabor coefficient *is not* well modeled by a Gaussian. This fact is also confirmed with normal probability plots (statistics tool to assess whether or not a data set is approximately normally distributed), which are not displayed due to space limitations.

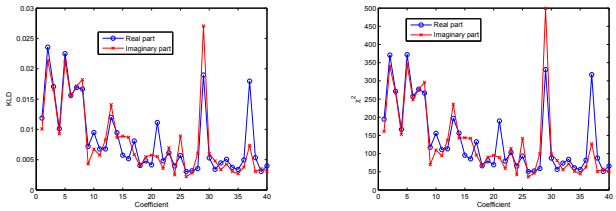

Fig. 4. KLD and χ^2 distances between the fitted GGD and the data for both real and imaginary parts of each Gabor coefficient.

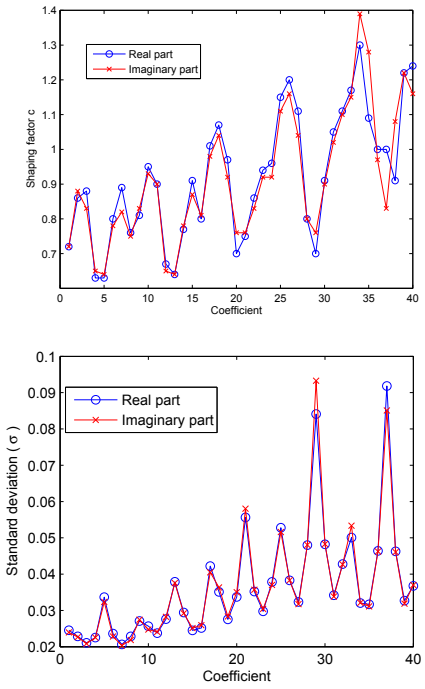

Fig. 5. Obtained c and σ GGD parameters for both real and imaginary parts of each Gabor coefficient.

- The real and imaginary parts of a given coefficient have similar c parameters. The same conclusion can be drawn for σ.
- There exists a pseudo-periodic behavior in the GGD parameters. If we examine Figure 6, which replots the shaping factors and the standard deviations for the real part of each coefficient grouped by scale subbands, it seems clear that a similar pattern emerges on each of these subbands. Further research is needed in order to provide theoretical reasons supporting this behavior.
- In an analogous way, Figure 7 replots c and σ for the real part of the coefficients grouped by orientation subbands. It can be realized that the GGD parameters increase with scale, i.e. as long as spatial frequency decreases. Taking into account the variation of Gabor filters with scale for a fixed orientation (any column from Figure 1), it is clear that a filter from the first row (1^{st} scale, highest frequency) captures texture information from a smaller neighborhood than a filter with a lower frequency does. Hence, we can assume that the information encoded in a high frequency coefficient is more correlated than the one captured by a low frequency filter and therefore, it is reasonable to conclude that the variance (and σ) should be smaller for high frequency coefficients. Moreover, the increase of c with scale means that the coefficient distributions are becoming "more Gaussian" and this fact could be explained by the same hypothesis and the central limit theorem: as long as frequency decreases, the pixels in the image patch that are taken into account for the convolution are less correlated and, applying the central limit theorem, the result of this convolution should approach a normal distribution. However, more experiments are needed to assess the validity of this hypothesis.

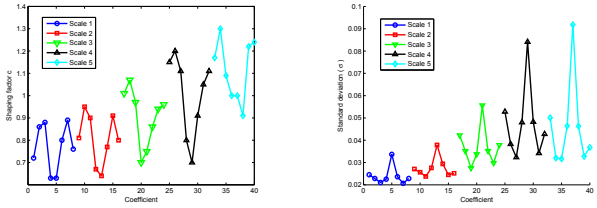

Fig. 6. Shaping factors c and standard deviations σ for the real part of Gabor coefficients grouped by scale subbands.

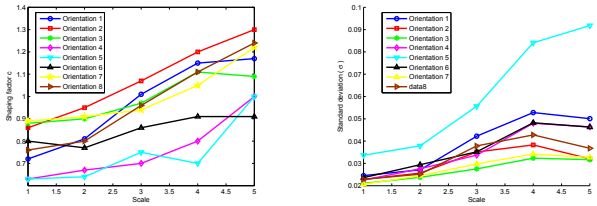

Fig. 7. Shaping factors c and standard deviations σ for the real part of Gabor coefficients grouped by orientation subbands.

4. COEFFICIENT QUANTIZATION BY MEANS OF LLOYD-MAX ALGORITHM

Now that we have a way to statistically model Gabor coefficients, a wide range of applications arises. As highlighted in [1], one of the drawbacks of Gabor-based approaches is the large amount of data that must be stored. Hence, we can think of reducing storage via coefficient quantization. To achieve this goal, we used the Lloyd-Max quantizer (the one with minimum mean squared error (MSE) for a given number N_L of representative levels).

In our case, a face is represented by n jets, each one comprising 40 complex coefficients. Assuming that each coefficient is represented by 16 bytes (8 for the real part + 8 for the imaginary part), a total of $16 \times 40 \times n$ bytes are needed per face. After GGD modelling and data quantization, instead of storing the original coefficient, we only need to keep two indices (one for the real part and another for the imaginary part) per coefficient ($2 \times 40 \times n$ indices per face). Hence, using N_L quantization levels, we can represent a coefficient with $2 \times \lceil \log_2(N_L) \rceil$ bits. In [6], it was shown that a given image can be reconstructed using the Gabor responses extracted from a sparse graph (like the rectangular grid shown in Figure 2). Figure 8 presents the reconstruction of the face in Figure 2 using different quantization levels, along with the reconstruction using the original coefficients. As can be seen, the reconstructed face with only just 4 quantization levels is already quite accurate, and the differences in quality between $N_L = 8, \ldots, 512$ levels and the original coefficients are not easily noticeable from a perceptual point of view.

5. FACE VERIFICATION ON THE XM2VTS DATABASE

In order to assess the impact of data compression on system performance, we conducted verification experiments using the XM2VTS database on configuration I of the Lausanne protocol [4]. The XM2VTS database contains synchronized image and speech data recorded on

295 subjects (200 clients, 25 evaluation impostors, and 70 test impostors) during four sessions taken at one month intervals. The database was divided into three sets: a training set, an evaluation set, and a test set. The training set was used to build client models, while the evaluation set was used to estimate thresholds that discriminate between client and impostor attempts. These thresholds are chosen so that False Acceptance Rate (FAR) equals False Rejection Rate (FRR) on the evaluation set. Finally, using the obtained thresholds, we measure the FAR and FRR on the separate test set. Table 1 presents FAR, FRR and Total Error Rate (TER=FAR+FRR) over the test set varying the number of quantization levels, along with the performance using the original coefficients. In [10], the authors adapt statistical tests to compute confidence intervals around Half Total Error Rates (HTER $=$ TER$/2$) measures, and to assess whether there exist statistically significant differences between two approaches or not. Using this analysis, we confirmed that performance was significantly worse only for $N_L = 2$ and $N_L = 4$ quantization levels. For the remaining ones, the performance was even better than that with original coefficients, although we can not conclude that significant improvements were achieved. In any case, these results suggest that noise reduction may be achieved via coefficient quantization.

6. CONCLUSIONS AND FURTHER RESEARCH

This paper has shown that Gabor coefficients extracted from face images can be accurately modeled using generalized Gaussian distributions. This finding opens a wide range of possibilities. As a first attempt, we took advantage of the underlying statistics to reduce data storage via Lloyd-Max quantization. No degradation was observed even with severe compression using 8 quantization levels. Further research is needed to investigate the pseudo-periodic behavior of GGD parameters described in Section 3. Moreover, we have demonstrated that the distributions of Gabor coefficients are far from being Gaussian, as long as the obtained shaping factors c are well below 2 for all coefficients.

Gabor-based face recognition systems have used distances for jet comparison that are not supported by theoretical evidence (cosine distance, as in [2], is one of the most accepted). We think that, based on the GGD modeling of Gabor coefficients, optimal ways to compare jets (from a theoretical point of view) could be obtained, and we will focus our efforts in this direction.

Acknowledgements

This project has been partially supported by Xunta de Galicia in the Competitive research units program Ref. 150/2006; Spanish MEC under the projects PRESA (TEC2005-07212) and DIPSTICK (TEC2004-02551/TCM); and the European Networks of Excellence BIOSECURE and ECRYPT.

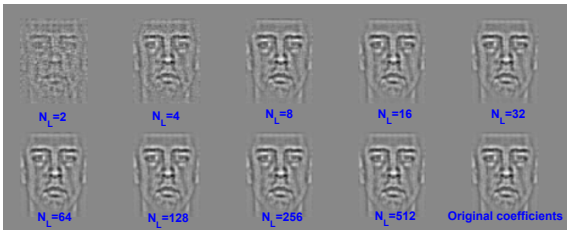

Fig. 8. Face reconstruction [6] using original and quantized coefficients.

Table 1. Face Verification on the XM2VTS database. False Acceptance Rate (FAR), False Rejection Rate (FRR) and Total Error Rate (TER) over the test set using both raw and compressed data.

	Test Set		
	FAR(%)	FRR(%)	**TER(%)**
$N_L = 2$	12.15	18.25	**30.40**
$N_L = 4$	4.19	8.00	**12.19**
$N_L = 8$	3.49	5.50	**8.99**
$N_L = 16$	3.85	5.50	**9.35**
$N_L = 32$	3.71	5.00	**8.71**
$N_L = 64$	3.53	5.50	**9.03**
$N_L = 128$	3.57	5.00	**8.57**
$N_L = 256$	2.74	4.75	**7.49**
$N_L = 512$	3.66	4.75	**8.41**
Raw data	3.79	5.25	**9.04**

7. REFERENCES

[1] Linlin Shen, Li Bai, "A Review on Gabor Wavelets for Face Recognition," in *Pattern Analysis and Applications*, Vol. 9, No. 2, pp. 273–292, 2006.

[2] Wiskott, L., Fellous, J.M., Kruger, N., von der Malsburg, C., "Face Recognition by Elastic Bunch Graph Matching", in *IEEE TPAMI*, 19(7), 775-779, 1997.

[3] Messer, K., Matas, J., Kittler, J., Luettin, J., and Maitre, G. XM2VTSDB: "The extended M2VTS database," in *Proc.AVBPA 1999*, pp. 72-77.

[4] Luttin, J. and Maître, G. Evaluation protocol for the extended M2VTS database (XM2VTSDB). *Tech. report RR-21*, IDIAP, 1998.

[5] Daugman, J.G., "Complete Discrete 2D Gabor Transforms by Neural Networks for Image Analysis and Compression", in *IEEE Trans. on Acoustics, Speech and Signal Processing* vol. 36, no. 7, pp. 1169-1179, July 1988.

[6] Potzsch, M., Maurer, T., Wiskott, L., von der Malsburg, C. "Reconstruction from Graphs Labeled with Responses of Gabor Filters," in *ICANN 1996* 845-850.

[7] Van de Wouwer, G., Scheunders, P., Van Dyck, D., "Statistical Texture Characterization from Discrete Wavelet Representations", in *IEEE TIP*, Vol. 8, No. 4, pp. 592–598, 1999.

[8] Hernández, J.R., Amado, M., Pérez-González, F., "DCT-domain watermarking techniques for still images: Detector performance analysis and a new structure," in *IEEE TIP* Vol. 9, No.1, pp. 55–68, January 2000.

[9] Minh N. Do, Vetterli, M., "Wavelet-Based Texture Retrieval Using Generalized Gaussian Density and Kullback-Leibler Distance", in *IEEE TIP*, Vol. 11, No. 2, pp. 146–158, 2002.

[10] Bengio, S., Mariethoz, J. "A statistical significance test for person authentication", in *Proceedings Odyssey 2004*, 237–244.

[11] Muller, F.,"Distribution shape of two-dimensional DCT coefficients of natural images," in *Electronics Letters*, vol.29, no.22, pp.1935-1936, 28 Oct 1993.

[12] T. M. Cover and J. A. Thomas, "Elements of Information Theory", John Wiley & Sons, New York, 1991.