# STATISTICAL NON-UNIFORM SAMPLING OF GABOR WAVELET COEFFICIENTS FOR FACE RECONGNITION

Shan Du and Rabab Ward

Department of Electrical and Computer Engineering
The University of British Columbia, Vancouver, BC, Canada
{shand, rababw}@ece.ubc.ca

**ABSTRACT**

A statistics based, non-uniform sampling of the Gabor wavelet decomposition coefficients for face recognition is presented in this paper. Gabor wavelets are popularly used to decompose face images into their spatial/frequency domains. The derived Gabor coefficients generate an augmented vector, e.g., 40 times larger than the original gray-scale vector. To reduce the dimensionality, uniform sampling of the Gabor coefficients is normally used. In this paper, we propose a non-uniform sampling method of the Gabor coefficients such that the coefficients corresponding to important face features are sampled much finer than those of the other parts of the image. The non-uniform sampling is based on the local statistics of the Gabor coefficients obtained from a set of training images. This adaptation is implemented in a hierarchical fashion; a coarse-to-fine strategy results in multi-level sampling rates. After the samples are obtained, the traditional principal component analysis (PCA) is used to code the samples for the final classification. The experimental results show that the proposed non-uniform sampling of Gabor coefficients outperforms the uniform one and the popular eigenfaces method.

## 1. INTRODUCTION

Face recognition has attracted much attention in the past decade because of its wide applications in commerce and law enforcement; these include mug-shot database matching, identity authentication, access control, and surveillance. Face recognition is one of the most challenging research topics since even for the same person faces appear differently due to expression, pose, occlusion and other confounding factors in real life. A number of face recognition techniques have been proposed in recent years, among which the eigenface method [1] and the elastic graph matching (EGM) method [2] are considered to be very successful. The comprehensive FERET test [3] ranks both methods with the highest recognition accuracy.

Much research effort has demonstrated that using Gabor wavelets at the front-end of an automated face recognition system is highly effective [2][4][5][6]. The Gabor wavelets, whose kernels are similar to the 2D receptive field profiles of the mammalian cortical simple cells [7], have proven to be able to derive desirable features of spatial frequency, spatial locality, and orientation selectivity. These features are known to be robust to variations due to illumination and facial expression changes. The Gabor wavelet representation of an image is the convolution of the image with a family of Gabor kernels at different scales and different orientations. To encompass different spatial frequencies (scales), spatial localities, and orientation selectivities, all these representation results are normally concatenated and derived as an augmented feature vector. Gabor wavelets can be applied to either the entire face images or to some fiducial points on them. Lades et al. applied Gabor wavelets for face recognition via the dynamic link architecture (DLA) framework [8]. The DLA first computes the Gabor jets, and then performs a flexible template comparison among the resulting image decompositions using graph matching. Wiskott et al. further expanded DLA and developed a Gabor wavelet based elastic bunch graph matching method to recognize human faces [2].

In this work, we apply Gabor wavelets to the entire face image as done in many papers such as [6][9]. A face is represented as the convolution result of the face image with 40 Gabor wavelets (5 scales × 8 orientations). Keeping only the magnitude values in the representation, this gives a '$40 \times n \times m$' element long vector, where $n \times m$ is the length of the face vector. To reduce the dimensionality of the vector, uniform sampling is traditionally used. However, the original Gabor coefficients do not contribute equally to the face recognition task. Consequently, an analysis that would weight the coefficients according to their effectiveness in recognition should increase the system's performance. In this paper, we propose a statistics based, non-uniform sampling strategy. The experimental results show that this approach improves the recognition rate.

The remaining of this paper is organized as follows: Section 2 describes the Gabor wavelets and Gabor coefficients. In section 3, we introduce the proposed non-uniform sampling approach in detail. Sections 4 and 5 show the experimental results, our discussion and conclusions.

## 2. GABOR WAVELETS AND GABOR COEFFICIENTS

### 2.1. Gabor Wavelets

Spatial/frequency analysis has played a central role in feature extraction as it combines two fundamental domains and allows

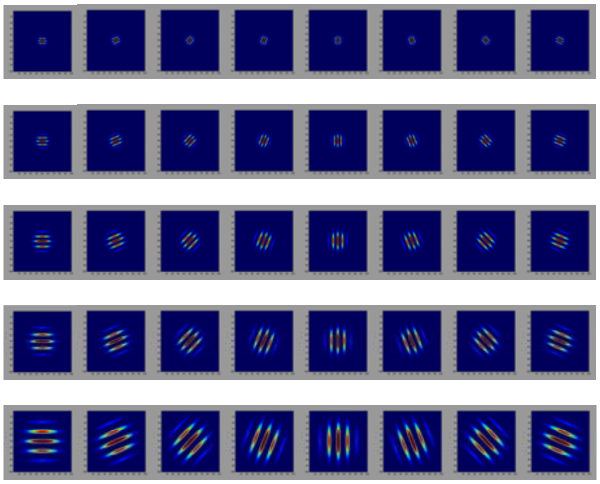

Figure 1. All 40 Gabor wavelets at different scales and orientations

the simultaneous representation of a signal in both these domains. Gabor wavelets have proven to be useful in face recognition since they extract information quanta in forms of spatial and frequency, two physically measurable quantities, combined in the most elegant way by the Heisenberg's uncertainty relation [10]. It is well known that the Gabor wavelets effectively model the receptive field profiles or cortical simple cells in the primary visual cortex [7]. The Gabor wavelet representation, therefore, captures salient visual properties such as spatial localization, orientation selectivity, and spatial frequency.

The 2-D Gabor wavelets (kernels, filters) can be defined as follows:

$$\psi_j(\vec{x}) = \frac{\|\vec{k}_j\|^2}{\sigma^2} \exp\left(-\frac{\|\vec{k}_j\|^2 \|\vec{x}\|^2}{2\sigma^2}\right)\left[\exp\left(i\vec{k}_j \vec{x}\right) - \exp\left(-\frac{\sigma^2}{2}\right)\right] \quad (1)$$

where

$$\vec{k}_j = k_m e^{i\varphi_n}$$

$$k_m = \frac{0.5\pi}{\left(\sqrt{2}\right)^m} \qquad \varphi_n = n\frac{\pi}{8}$$

The 2-D Gabor wavelet is a two-dimensional plane wave with wavelet vector $\vec{k}_j$ restricted by a Gaussian envelope function with relative width $\sigma$. The first term in the square brackets determines the oscillatory part of the wavelet. The second term makes the wavelets DC-free. The first term in (1) makes the energy of wavelets of approximately equal values. As is the case with other wavelets, the Gabor wavelet representation allows the description of spatial frequency structure of a face, as well as preserves spatial relations. The Gabor wavelets are all self-similar since they can be generated from one wavelet, the mother wavelet, by scaling and rotation. Normally Gabor wavelets at five scales ($m \in \{0,...,4\}$), and eight orientations ($n \in \{0,...,7\}$) are used. Figure 1 shows all the 40 Gabor wavelets at different scales and orientations that are used in this paper.

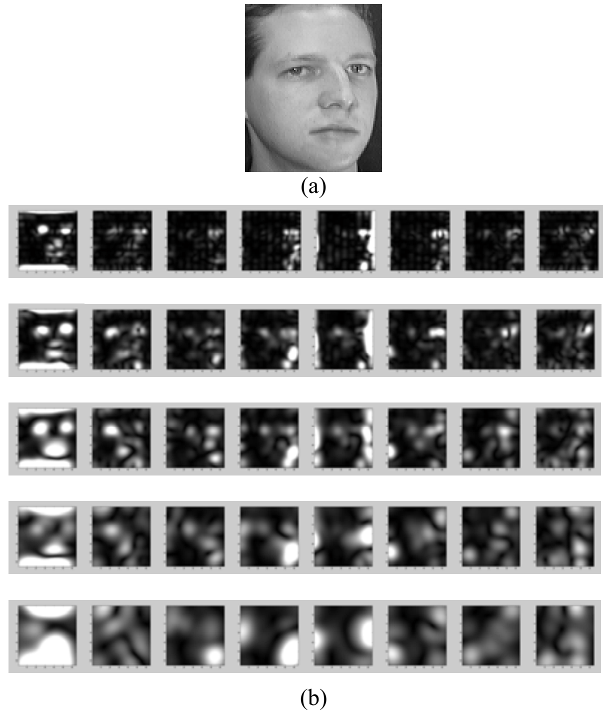

Figure 2. (a) An ORL face image, (b) Magnitude of the Gabor coefficients

### 2.2. Gabor Coefficients

The Gabor wavelet representation of an image is the convolution of the image with a family of Gabor wavelets. Let $I(x, y)$ be the gray level distribution of an image, the Gabor coefficients of image $I$ and a Gabor wavelet $\psi_j(\vec{x})$ is defined as follow:

$$O_j(\vec{x}) = I(\vec{x}) * \psi_j(\vec{x}) \quad (2)$$

where $(\vec{x}) = (x, y)$, and * denotes the convolution operator.

The Gabor coefficients of a sample image in the ORL face image database are shown in Figure 2.

## 3. STATISTICS BASED NON-UNIFORM SAMPLING OF GABOR COEFFICIENTS

Since the Gabor coefficients $O_j(\vec{x})$ consist of different local, scale, and orientation features, we concatenate all these features in order to derive a feature vector, $X$. Assuming the image size is $n \times m$, and since we are applying 40 Gabor wavelets, the size of the derived feature vector $X$ is $n \times m \times 40$. Normally, before the concatenation, one first downsamples each $O_j(\vec{x})$ by a factor $\rho$, e.g., $\rho = 8 \times 8$. This sampling is traditionally carried in a uniform fashion. Here, we should notice that if we use a fine sampling rate, the resulting vector might be too large. However, if we use a coarse sampling rate, some important features may be lost. Intuitively, the $n \times m \times 40$ coefficients do

*1. Set the initial sampling grid size (e.g., 16×16) and the minimum sampling grid size (e.g., 4×4).* (o *in the right figure is a sample point.*)

*2. For each new grid patch*
*If Local variance > Global variance and the sampling grid size > the minimum sampling grid size,*
*Then refine the sampling grid size to quarter, i.e. 8×8;*
*Otherwise, stop refining this patch.*

*3. Go to Step 2*

*4. When there is no new patch, end.*

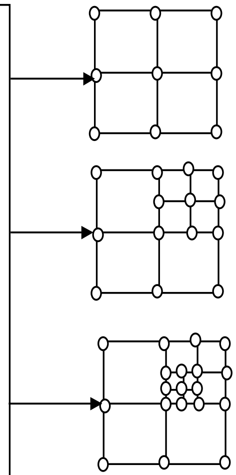

Figure 3. Non-uniform sampling algorithm

not contribute equally to the face recognition task. Therefore a uniform sampling method that treats all coefficients equally will not result in optimal performance. Thus, we here propose an approach that assigns higher importance to the more relevant parts of the augmented coefficient vector by sampling them at finer rates, and uses sparse samples at the less relevant parts.

To define the importance of the different sections of the Gabor augmented vector, we use an ensemble of $k$ training images. This process can be operated off-line to save the whole recognition processing time. First of all, we find the Gabor augmented vectors of all these $k$ training images. To find the sections in these $k$ vectors corresponding to important features, we notice that the coefficients corresponding to each such section have higher variances than the coefficients corresponding to the less relevant features. This is because important features, such as the eyes, have higher variances in their corresponding coefficients than those of less important features such as the cheeks. To identify which vector sections need to be sampled finer or less fine, we use a coarse-to-fine strategy. The Gabor coefficients are first assumed to be down sampled at a coarse rate. For each sub-section of the Gabor coefficient vector, we test the mean variance of this section among the training set; we call it the local variance. If the local variance is higher than the global variance, which is measured on the entire Gabor coefficient vectors, it means this sub-section is important, and we increase its sampling rate to its double. Otherwise, we keep it unchanged. If using higher sampling rate, the local variance is still high, we double the sampling rate again until it reaches predefined highest sampling rate. Figure 3 illustrates the process.

After the non-uniform sampling of the Gabor coefficients, we apply principal component analysis (PCA) to code the derived samples and then classify the face images using the simplest nearest neighbor method.

## 4. EXPERIMENTAL RESULTS

Figure 4. Samples of face images in the ORL database

The ORL face database (developed at Olivetti Research Laboratory, Cambridge, U. K.) is used to examine the proposed approach. The ORL database is composed of 400 images with ten different images for each of the 40 distinct subjects. The images vary across pose, size, time, and facial expression. All the images are taken against a dark homogeneous background with the subjects in an upright, frontal position, with tolerance for some titling and rotation of up to about $20^o$. There is some variation in scale of up to about 10%. The spatial and gray-level resolutions of the images are 92×112 and 256, respectively. Figure 4 shows some images in the ORL face image database. They are resized to 64×64. Here we use 5 images per subject for training, and the other 5 for testing.

To recognize a particular input face, the system compares the feature vectors of this face with all those of the database faces using the Euclidean distance nearest-neighbor classifier. Denoting the feature vector of the probe face image as $p$, and that of a database face as $f$, then the Euclidean distance between the two is

$$d = \sqrt{(f_0 - p_0)^2 + (f_1 - p_1)^2 + \ldots + (f_{M-1} - p_{M-1})^2} \qquad (3)$$

where

$$p = [p_0 \quad p_1 \quad \cdots \quad p_{M-1}]^T$$
$$f = [f_0 \quad f_1 \quad \cdots \quad f_{M-1}]^T$$

And $M$ is the number of features. A match is found when the minimal value of $d$ is obtained.

Figure 5 shows the results corresponding to the uniform sampling of the Gabor coefficients and Figure 6 shows those corresponding to the non-uniform sampling. We can see that the sections of coefficients with high variances correspond to salient face features, thus those features get sampled at a finer rate while the other face parts corresponding to less relevant coefficients get sampled at a coarser rate. Figure 7 shows the correct recognition rates of the eigenface, uniform sampling and our non-uniform sampling methods. It shows that our new method can significantly improve the recognition rate.

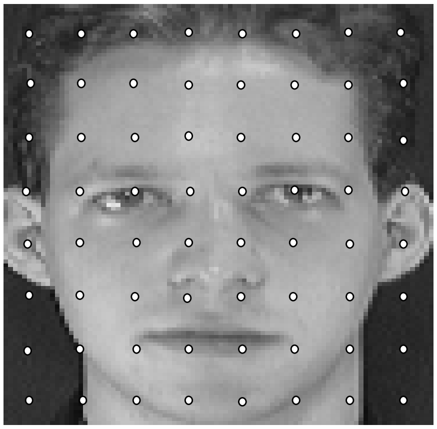

Figure 5. Uniform sampling (64 samples)

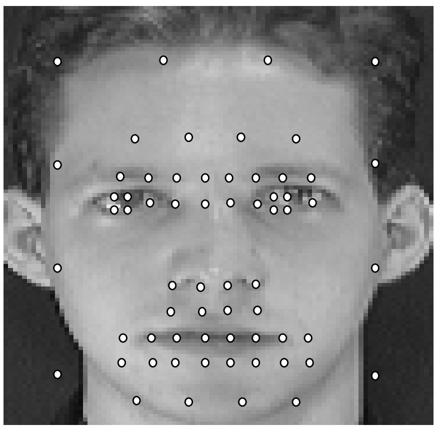

Figure 6. Non-uniform sampling (64 samples)

## 5. DISCUSSION AND CONCLUSIONS

In this paper, we propose a non-uniform sampling algorithm to sample the augmented Gabor coefficients. Our strategy is based on sampling the more relevant sections of the Gabor coefficients in a finer fashion than the less relevant ones. The variances of the coefficients corresponding to the same sections of the training vectors are used to determine the relevance of the coefficients. This is because of our observation that the salient features of a face, e.g. the eyes, have larger variances among different people than other features, e.g., the cheeks. Therefore our algorithm obtains more samples on the more relevant features of a face and less samples on the less relevant ones. As the sampling is performed in a hierarchical way, we generate multi-level sampling rates. Because we save some bits on the less relevant coefficients, we can use finer sampling rate for the more relevant ones (The corresponding results are shown in Figure 5 and Figure 6, where 8×8 sampling size for the uniform coefficient sampling. For the non-uniform one, we get 16×16 sampling size on the forehead and cheeks while 2×2 sampling size on the eyes. Both have 64 samples). Thus with vectors of the same length, we can reach higher recognition accuracy. The experimental results have shown that our non-uniform sampling of Gabor coefficients has better recognition accuracy than uniform sampling of Gabor

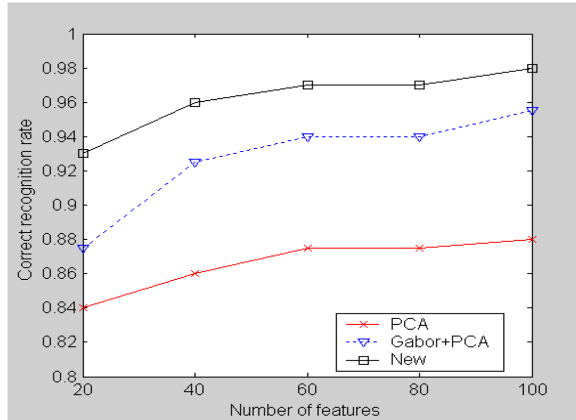


Figure 7. Comparison of PCA (Eigenface), Gabor plus PCA (uniform sampling), and the new method with the statistical selection (non-uniform sampling) of Gabor coefficients

coefficients method and eigenface method.


## 6. ACKNOWLEDGEMENT

The authors would like to thank the Natural Sciences and Engineering Council of Canada (NSERC) for their support.


## 7. REFERENCES


[1] M. Turk and A. Pentland, "Eigenfaces for recognition," J. of Cognitive Neuroscience, vol. 3, no. 1, 1991.

[2] L. Wiskott, J. M. Fellous, and C. Von Der Malsburg, "Face recognition by elastic bunch graph matching," IEEE Trans. Patt. Anal. Mach. Intell., vol. 19, pp. 775-779, 1997.

[3] P. J. Phillips, H. Wechsler, J. Huang, and P. Rauss, "The FERET database and evaluation procedure for face recognition algorithms," Image and Vision Computing Journal, vol. 16, no. 5, pp. 295-306, 1998.

[4] J. G. Daugman, "Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression," IEEE Trans. Patt. Anal. Mach. Intell., vol. 36, pp. 1169-1179, 1988.

[5] T. S. Lee, "Image representation using 2D Gabor wavelets," IEEE Trans. Patt. Anal. Mach. Intell. vol. 18, pp. 959-971, 1996.

[6] C. Liu and H. Wechsler, "Independent component analysis of Gabor features for face recognition," IEEE Trans. Neur. Net., vol. 14, no. 4, pp. 919-928, 2003.

[7] D. Field, "Relations between the statistics of natural images and the response properties of cortical cell," Journal of Opt. Soc. Amer. A, vol. 4, no. 12, pp. 2379-2394, 1987.

[8] M. Lades, J. C. Vorbruggen, J. Buhmann, J. Lange, C. van der Malsburg, R. P. Wurtz, and W. Konen, "Distortion invariant object recognition in the dynamic link architecture," IEEE Trans. Comput., vol. 42, pp. 300-311, 1993.

[9] C. Liu, "Gabor-based kernel PCA with fractional polynomial models for face recognition," IEEE Trans. Patt. Anal. Mach. Intell., vol. 26, no. 5, pp. 572-581, 2004.

[10] J. G. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," Journal of the Optical Society of America, vol. A2, no. 7, pp. 1160-1169, 1985.