Face Recognition via Direct Search Optimized Gabor Filters

CĂTĂLIN-DANIEL CĂLEANU*, VASILE GUI**, FLORIN ALEXA**
*Applied Electronics, **Communications
University POLITEHNICA Timişoara
V. Pârvan Av., no. 2, 300223 Timişoara
ROMÂNIA
http://www.etc.upt.ro

Abstract: - While Gabor filter banks are one of the most successful image processing techniques for face processing tasks, finding which frequency and orientation bands are responsible for the success of the approach is still an open topic. This paper introduces the Mesh Adaptive Direct Search (MADS) algorithm for the Gabor filters frequency and orientation optimization. In this paper MADS optimization has been applied in the context of face recognition. Compared to heuristic or trial and error parameters selection, the proposed method possesses a number of advantages; among the most important of them is the higher recognition accuracy obtained on public face databases.

Key-Words: - Direct Search Optimization, Gabor Filters, Face Recognition

1 Introduction

The problem of recognizing faces has been largely investigated due to its wide range of applications. Many paradigms are available for implementing the recognition/classification phase. Some of the most important are briefly discussed in the following.

Geometric feature based matching has been employed by Brunelli, Poggio and Kanade [1], [2]. The basic idea behind their algorithm was to describe the overall configuration of the face by a vector of numerical data representing the relative position and size of the main facial features: eyes and eyebrows, nose and mouth.

Eigenfaces proposed by Turk et al. [3] are a set of orthonormal basis vectors computed from a collection of training face images. They provide a basis of low dimensional representation of the facial images and are optimal in the minimum least square error sense.

Guo et al. [4], incorporated Support Vector Machines (SVM's) with binary tree recognition for multi-class recognition. More on this topic in [5].

In 2001 Cesar et al. [6] approached facial feature recognition as a problem of matching inexact graphs where the graphs were built from regions and relationships between regions in an image.

Texture coding provides information about facial regions with little geometric structure like hair, forehead and eyebrows whereas a depth map provides us with information about regions with little texture such as chin, jaw line and cheeks. Considering this fact, BenAbdelkader et al. proved that the accuracy of face recognition systems can be improved by considering not only the texture map but also the depth map [7].

Ekenel and Sankur [8] proposed multiresolution facial recognition in 2005. They employ multiresolution analysis to decompose the image into its subbands prior to the subspace operations such as principal or independent component analysis.

Liu et al. [9] describe a novel Gabor Feature Classifier (GFC) method for face recognition. The kernels of Gabor wavelets are similar to the 2D receptive field profiles of the mammalian cortical simple cells and exhibit desirable characteristics of spatial locality and orientation selectivity.

Recently, a new trend in face recognition has become evident: the use of 3D data of the face. A 3-D model should be better for representing faces, especially to handle facial variations, such as pose, illumination etc.. Therefore, it follows logically that the best and most complete solution to this problem is to acquire/analyze/match a full 3D model of the face as represented, for example, by a 3D shape-mesh plus a 2D texture-map. See [10]-[12] for representative work in this direction. However, lots of applications still have to rely on 2D image data.

As it results from face recognition literature, the methods based on Gabor kernels give good results and are one of the most successful approaches for processing images of the human face [13]. In spite of the fact that Gabor filter banks specifications are crucial for the application success, only few papers address systematically the problem of filter frequency and orientation optimization [14] - [16]. Some of these approaches sample the parameter space uniformly. However, this strategy does not exploit the particular characteristics of the object part under test. Another possibility is to analyse the Gabor response function in the full parameter space [17] and select those parameters that best describe the particular object characteristics. However, this strategy could bias the parameter distribution to a narrow range and reduce the capability to discriminate the modelled object from others.

In order to avoid the problem of a sub-optimal and/or overly complex filter banks specifications, in [18], the face recognition problem is tackled by a *genetic algorithm* used to find an optimal basis derived from a combination of frequencies and orientation angles in the 2-D Gabor wavelet transform domain. In [19] the most useful frequencies and orientations of Gabor kernels are determined using a *floating feature selection* algorithm.

Our approach proposes the design of feature extraction phase as a *single problem* of simultaneuosly estimating the best Gabor filter parameters (frequency and orientation) trough the *Direct Search* optimization technique. We show that minimizing an objective function using the Direct Search algorithm is an attractive alternative because it is both computationally less expensive then previously used optimization methods for the facial recognition and provides a higher accuracy.

The paper is organized as follows: Section 2 describes optimization technique used in the paper. Section 3 explains the feature extraction and classification scheme. Section 4 provides a brief facial database description and the comparative experimental results, and Section 5 concludes the paper.

2 Mesh Adaptive Direct Search

Pattern search methods are a particular class of direct search methods firstly analyzed by Torczon [20] for unconstrained optimization. They were also extended to bound and linear constrained optimization [21]. The research about pattern search methods is still now very flourishing: several generalizations and extensions have recently been proposed [22].

One of these extensions, called Mesh Adaptive Direct Search (MADS) [23], is a class of algorithms for nonlinear optimization that extends the Generalized Pattern Search (GPS) class by allowing local exploration, called polling, in a dense set of directions in the space of optimization variables. The operating principle of these algorithms imply searching a set of points called a pattern, which expands or shrinks depending on whether any point within the pattern has a lower objective function value than the current point.

Formalizing the problem, the aim of the algorithm is minimizing a nonsmooth function:

$$\min_{x\in X} f(x) \tag{1}$$

where $x \in \mathrm{X}$ is the vector of design parameters, $f : \mathbf{X} \rightarrow \Re$ is the cost (objective) function, and $\mathbf{X} \subset \Re^n$ is the constraint set, defined as:

$$\mathbf{X} \stackrel{\Delta}{=} \{x \in \Re^n \mid l^i \le x^i \le u^i, i \in \{1,...,n\}\} \tag{2}$$

with $-\infty \le l^i < u^i \le \infty$, for all $i \in \{1,...,n\}$. In our case, the x vector contains design parameters for the Gabor filter feature extraction module. The objective function it is represented by the error of the facial recognition system regarding the test data set, usually referred as test error.

This class of optimization algorithms computes a sequence of points that get closer and closer to the optimal point. At each step, the algorithm searches a set of points, called a *mesh*, around the current point — the point computed at the previous step of the algorithm. The mesh is formed by adding the current point to a scalar multiple of a set of vectors called a pattern. If the algorithm finds a point in the mesh that improves the objective function at the current point, the new point becomes the current point at the next step of the algorithm.

Two types of direct search algorithms have been considered: GPS and MADS. The algorithms differ in how the set of points forming the mesh is computed. The GPS algorithm uses fixed direction vectors, whereas the MADS algorithm uses a random selection of vectors to define the mesh.

Let $k \in \mathrm{N}$ denotes the iteration number, and let $x_k \in \mathbf{X}$ denote the current iterate. The pattern search algorithms have in common that after a finite number of iterations, they search for a lower cost function value than $f(x_k)$ on the points in the set:

$$L \stackrel{\Delta}{=} \{x \in \mathbf{X} \mid x = x_k \pm \Delta_k s^i d_i,\ i \in \{1, . . . , n\}\} \tag{3}$$

where $\Delta_k > 0$ is a scalar called the *mesh size factor*, and $s \in \mathrm{R}^n$ is a fixed parameter that can be used to take the different scaling of the design parameter components into account. Once the direction vectors have been defined, the GPS and MADS algorithms form the mesh by multiplying the pattern vectors by a scalar, the mesh size Δ_k, defined by the length and direction $s^i d_i$. The pattern search algorithms have a rule that selects a finite number of points in $\mathbf{X}$ on a mesh defined by:

$$M(x_0, \Delta_k) \stackrel{\Delta}{=} \{x_0 + m\Delta_k s^i d_i \mid$$

$i \in \{1, \ldots, n\}, m \in \mathrm{Z}\}$ (4)

where $x_0 \in \mathbf{X}$ is the initial iterate. If a mesh point $x' \in \mathrm{M}(x_0, \Delta_k)$ with $f(x') < f(x_k)$ has been found, then the search continues with $x_{k+1} = x'$ and $\Delta_{k+1} = \Delta_k$ only if *complete poll* option is disabled. In this case the algorithm stops polling the mesh points as soon as it finds a point whose objective function value is less than that of the current point. If this occurs, the poll is called *successful* and the point it finds becomes the current point at the next iteration. If the algorithm fails to find a point that improves the objective function, the poll is called *unsuccessful* and the current point stays the same at the next iteration. If a complete poll is considered, the algorithm computes the objective function values at *all* mesh points. All points in L_k are tested for a decrease in $f(.)$. If $f(x') \geq f(x_k)$ for all $x_k \in L_k$, then the search continues with $x_{k+1} = x_k$ and a reduced mesh size factor. Hence, the search continues on a finer mesh. The search stops if the mesh $\mathrm{M}(x_0, .)$ has been refined a user-specified number of times.

3 Feature extraction and classification procedures

In the feature extraction problem, the task is to find an efficient way to represent the pixel data, usually trough a mapping function, from the original image space to a lower dimensional feature vector space.

Gabor filters are joint entropy minimizing frequency sensitive filters. Later extended for two dimensions [24], their use in vision systems is also biologically motivated, as the kernels of Gabor wavelets are similar to the 2D receptive field profiles of the mammalian cortical simple cells and exhibit desirable characteristics of spatial locality and orientation selectivity [9].

In this paper, the following form of a normalized 2-D Gabor filter function, in the continuous spatial domain, has been employed:

$$\psi(x, y; f, \theta) = \frac{f^2}{\pi\gamma\eta} e^{-\left(\frac{f^2}{\gamma^2}x'^2 + \frac{f^2}{\eta^2}y'^2\right)} e^{j2\pi f x'} \quad (5)$$

$$x' = x\cos(\theta) + y\sin(\theta) \quad (6)$$

$$y' = -x\sin(\theta) + y\cos(\theta) \quad (7)$$

where f is the frequency of a sinusoidal plane wave, θ is the anti-clockwise rotation of the Gaussian envelope and the sinusoid, γ is the spatial width of the filter along the plane wave, and η the spatial width perpendicular to the wave, as it was presented in [25].

In order to represent face images using Gabor filters, we have placed a square grid over the face region in the image. The responses of convolutions in an image $\xi(x, y)$, around a given pixel situated at location (x, y), are given by:

$$r_\xi(x, y; f, \theta) = \psi(x, y; f, \theta) * \xi(x, y) = \\ = \iint_{-\infty}^{\infty} \psi(x - x_\tau, y - y_\tau; f, \theta)\xi(x_\tau, y_\tau)\mathrm{d}x_\tau \mathrm{d}y_\tau \quad (8)$$

and it represents feature vectors to be further classified.

The classification problem involves designing a function to map feature vectors to the appropriate class label. In the present work a statistical classifier was employed, namely the K-Nearest Neighbour (k-NN) classifier. It is a very simple, yet powerful statistical classification method. Using a suitable distance, e.g. Euclidean (p = 1) or city-block (p = 2) distance:

$$d(\mathbf{x}, \mathbf{y}) = \left(\sum_{i=1}^{N} |x_i - y_i|^p\right)^{\frac{1}{p}} \quad (9)$$

one has to find k closest training points to the vector that should be labelled. Then, by specifying a rule, e.g. majority rule or consensus rule, it will be decided how to classify the sample. It can be shown that the performance of a k-NN classifier is always at least half of the best possible classifier for a given problem. More on this topic could be found in [26].

4 Experimental results

The following experiments results are reported against AT&T Laboratories Cambridge database of faces [27]. Five training images per person (thus 200 total training images) were randomly taken for training, and the remaining images (200 total images) are taken for testing. However, when the effect of a certain filter parameter was analyzed, in order to evaluate its usefulness in capturing relevant information, the data sets have been frozen. Otherwise, the results would have been also influenced by the random data selection process.

4.1 Sensitivity analysis based Gabor features

In this case we experimentally examine the influence of the *individual, independent* feature extraction Gabor filter parameters contribution to the recognition performance. We want to learn which frequencies and which orientations are useful for the face recognition problem.

Thus, by using the most important subset of frequency and orientation parameters, we can speed up the feature representation phase and have a more compact form of feature vectors.

Table 1 and Table 2 illustrate the independent contributions of each frequency and orientation to the recognition performance. Also different combinations of features were tested.

x (frs = 1/x)	Test error [%]
2	58.0
4	23.5
8	19.0
16	13.0
32	13.0
64	13.0
128	13.0
4 8 16	11.0
8 16 32	11.5
16 32 64	13.0
8 16 32 64	11.5
4 8 16 32 64	10.5
4 8 16 32 64 128	10.5
2 4 8 16 32 64 128	12.5

Tab. 1. The influence of Gabor filters frequency over the classification accuracy.

It could be observed that the best result, 10.5% error over the test data, is obtained for the sets Φ1 = {4, 8, 16, 32, 64} and Φ2 = {4, 8, 16, 32, 64, 128}. In order to keep the computational expense low, the Φ1 = {4, 8, 16, 32, 64} set will be used in the following experiments.

For filters orientation case, by selecting only the four most representative orientations, Θ = {pi/8, 3*pi/8, 5*pi/8, 7*pi/8}, for the total amount of eight orientation considered, the processing time has been cut in half while keeping the system accuracy low.

Orientations	Test error [%]
0	17.5
pi/8	11.5
2*pi/8	13.5
3*pi/8	13.5
4*pi/8	18.5
5*pi/8	10.5
6*pi/8	14.0
7*pi/8	11.5
pi/8 3*pi/8 5*pi/8 7*pi/8	10.0

Tab. 2. The influence of Gabor filters orientation over the classification accuracy.

4.2 Direct Search based Gabor features

The second part of the experiments aims to apply the Direct Search type techniques in order to find the optimal filter parameters vector, $x = \{\theta_1, \ldots, \theta_4, \varphi_1, \ldots, \varphi_5\}$, which minimize an objective function $f(x)$ representing the classification error over the test data set. The lower and upper bounds for orientations are set to [0, 3] respectively [0, 0,5] for frequencies. The starting point x_0 was randomly generated within specified bounds.

We found an increasing convergence speed using MADS (fig. 1); the optimization performance was found to be similar for the MADS and the GPS cases.

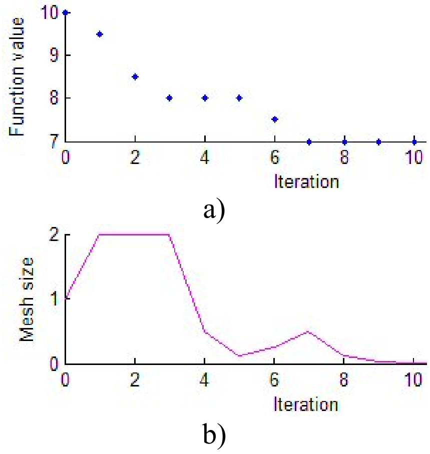

Fig. 1. a) The objective function is decreasing over iteration steps. b) The mesh size was increased after a successful pooling, otherwise it has been refined.

4.3 Sensitivity vs. Direct Search - comparative results

The final experiment compares the system accuracy using the best parameters found by the sensitivity analysis respectively MADS algorithm. Unlike previous cases, the training and testing data are generated *randomly*, in five consecutive experiments. The results are presented in Tab. 3. Over 30% improvement in accuracy was obtained by using MADS generated parameters.

Compared with other experimental results reported on the same database our 3.3% is one of the lowest classification error. Almost all ORL results are reported in literature using 200 training images and 200 test images (ORL 50/50) randomly selected. In spite of this fact, it is still difficult to make a fair comparison with other approaches because of different numbers of runs used to compute the test error rates. For example in [28], a mean error rate of 4% based on 20 runs is reported. In contrast, an error rate of 2.5% is reported in [29] based on ten runs, yet in [30] 0% error rate is claimed but it is not clear how many runs the result is based upon.

5 Conclusions and future work

In this work we present an automatic feature selection method that can be applied to the problem of facial recognition. The representation is based on Gabor features and our methodology selects automatically a set of parameters that are good descriptors for a particular classification task. The technique is based on the Direct Search optimization concept, which is extended, in this work, to consider optimization along *both* frequencies and orientation of the Gabor filter *simultaneously*.

Optimization technique	**Test errors in 5 experiments [%]**	**min. max. std.dev.**	**Test error mean value [%]**
Sensitivity	6.0 4.0 5.5 2.5 6.0	2.5 6.0 1.5	4.8
Direct Search	2.5 6.0 1.5 4.5 2.0	1.5 6.0 1.9	3.3

Tab. 3. The influences of the optimization technique over the classification accuracy for standard AT&T face database experimental condition.

Compared to the heuristic/sensitivity selection process, followed by using the most important subset of frequency and orientation parameters, our approach provided not only better performance but also a straightforward manner to obtain optimal parameters in a significantly reduced design time.

When the process simulation is very complex and it is not designed in a vectorised manner, Direct Search has been shown to be an attractive alternative to other optimization methods, e.g. genetic algorithms, as it is often computationally less expensive and can minimize the same types of functions.

As future work we have in view to extend the optimization process to *both* feature extraction and classification modules parameters simultaneous.

Acknowledgments

The authors acknowledge the support of CNCSIS AT41/2006 for some parts of this research.

References:

[1] T. Kanade, *Picture processing by computer complex and recognition of human faces*, Doctoral dissertation, Department of Information Science, Kyoto University, 1973.

[2] R. Brunelli, T. Poggio, Face recognition through geometrical features, in *European Conference on Computer Vision (ECCV)*, 1992, pp. 792-800.

[3] M.A. Turk, A.P. Pentland, Eigenfaces for recognition, *Journal of Cognitive Neuroscience*, Vol.1, No.3, 1991, pp. 71-86.

[4] G. Guo, S.Z. Li, K.L. Chan, Support Vector Machines for Face Recognition, *Image and Vision Computing*, No. 19, 2001, pp. 631-638.

[5] Y. Lia, S. Gong, J. Sherrah, H. Liddell, Support vector machine based multi-view face recognition and detection, *Image and Vision Computing*, No. 22, 2004, pp. 413-427.

[6] R. Cesar , E. Bengoetxea, I. Bloch, Inexact graph matching using stochastic optimization techniques for facial feature recognition, *International Conference on Pattern Recognition (ICPR)*, Quebec, Canada, 2002.

[7] C. BenAbdelkader, P.A. Griffin, Comparing and combining depth and texture cues for face recognition, *Image and Vision Computing*, No. 23, 2005, pp. 339–352.

[8] H. K. Ekenel, B. Sankur, Multiresolution Face Recognition, Image and Vision Computing, No. 23, 2005, pp. 1-9.

[9] C. Liu, H. Wechsler, A Gabor feature classifier for face recognition, in Eighth *IEEE International Conference on Computer Vision*, Vol. 2, 7-14 July 2001, pp. 270 – 275.

[10] V. Blanz, T. Vetter, Face Recognition Based on Fitting a 3D Morphable Model, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 9, September 2003, pp. 1063-1074.

[11] J. Lee, B. Moghaddam, H. Pfister, R. Machiraju, Finding Optimal Views for 3D Face Shape Modelling, *Proc. of the International Conference on Automatic Face and Gesture Recognition*, 17-19 May 2004, Seoul, Korea, pp. 31-36.

[12] A. M. Bronstein, M. M. Bronstein, R. Kimmel, Three-dimensional face recognition, *Intl. Journal of Computer Vision (IJCV)*, Vol. 64/1, August 2005, pp. 5-30.

[13] Y. Wang, C.-S. Chua, Face recognition from 2D and 3D images using 3D Gabor filters Image

and Vision Computing, Vol. 23, No. 11, 1 October 2005, pp. 1018-1028.

[14] L. Wiskott, J.M Fellous, N. Kruger, C. Von der Malsburg, Face Recognition by Elastic Bunch Graph Matching, *Intelligent Biometric Techniques in Fingerprint and Face Recognition*, Springer-Verlag, 1999.

[15] S. Arca, P. Campadelli, R. Lanzarotti, A face recognition system based on automatically determined facial fiducial points, *Pattern Recognition*, Vol. 39, No. 3 , March 2006, pp. 432-443.

[16] M. Xiaoyan, M.H. Hassoun, P. Watta, Combining Gabor features: summing vs. voting in human face recognition, *Systems, Man and Cybernetics*, Vol. 1, pp. 737- 743.

[17] B. Gokberk, L. Akarun, E. Alpaydin, Feature selection for pose invariant face recognition, *Proceedings of the 16th International Conference on Pattern Recognition*, Vol. 4, 2002., pp. 306- 309.

[18] X. Wang, H. Qi, Face recognition using optimal non-orthogonal wavelet basis evaluated by information complexity, *Proceedings of the 16th International Conference on Pattern Recognition*, Vol. 1, 11-15 Aug. 2002, pp. 164 – 167.

[19] B. Gökberk, M. O. İrfanoğlu, L. Akarun, E. Alpaydın, *Selection of Location, Frequency, and Orientation Parameters of 2D Gabor Wavelets for Face Recognition*, Lecture Notes in Computer Science, Springer, Vol.3161, p. 138-146, 2005.

[20] V. Torczon, *On the convergence of Pattern Search algorithms*, SIAM Journal on Optimization, No. 7, 1997, pp. 1-25.

[21] R.M. Lewis, V. Torczon, *Pattern Search algorithms for Bound Constrained Minimization*, SIAM Journal on Optimization, No. 9, 1999, pp. 1082-1099.

[22] C. Audet, J.E. Dennis Jr., *Analysis of Generalized Pattern Searches*, SIAM Journal on Optimization, No. 13, 2003, pp. 889-903.

[23] C. Audet, J. E. Dennis, Jr., *Mesh Adaptive Direct Search Algorithms for Constrained Optimization*, SIAM Journal on Optimization archive, Vol. 17, No. 1, 2006, pp. 188 – 217.

[24] Daugman, J.G.,. 1985. Uncertainty relation for resolution in space, spatial frequency and orientation optimized by 2D visual cortical filters. Journal of the Optical Society ofAmerica, 2(7), 1160–1169.

[25] V. Kyrki, *Local and global feature extraction for invariant object recognition*, Ph.D. thesis, Lappeenranta University of Technology, 2002.

[26] T. Mitchell, *Machine Learning*, McGraw Hill, 1997.

[27] http://www.cl.cam.ac.uk/research/dtg/ attarchive/facedatabase.html

[28] M. Wang and S. Chen, Enhanced fmam based on empirical kernel map, *IEEE Trans. Neural Networks*, vol. 16, no. 3, 2005, pp. 557-564.

[29] M. E. Er, W. Chen, and S. Wu, High speed recognition based on discrete cosine transform and RBF neural networks,"*IEEE Trans. Neural Networks*, vol. 16, no. 3, 2005, pp. 679-691.

[30] H. Zhang, B. Zhang,W. Huang, and Q. Tian, Gabor wavelet associative memory for face recognition, *IEEE Trans. Neural Networks*, vol. 16, no. 1, 2005, pp. 275-278.