International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:2, No:8, 2008

Holistic Face Recognition using Multivariate Approximation, Genetic Algorithms and AdaBoost Classifier: Preliminary Results

C. Villegas-Quezada, and J. Climent

Abstract**—**Several works regarding facial recognition have dealt with methods which identify isolated characteristics of the face or with templates which encompass several regions of it. In this paper a new technique which approaches the problem holistically dispensing with the need to identify geometrical characteristics or regions of the face is introduced. The characterization of a face is achieved by randomly sampling selected attributes of the pixels of its image. From this information we construct a set of data, which correspond to the values of low frequencies, gradient, entropy and another several characteristics of pixel of the image. Generating a set of "p" variables. The multivariate data set with different polynomials minimizing the data fitness error in the minimax sense $(L_\infty - Norm)$ is approximated. With the use of a Genetic Algorithm (GA) it is able to circumvent the problem of dimensionality inherent to higher degree polynomial approximations. The GA yields the degree and values of a set of coefficients of the polynomials approximating of the image of a face. By finding a family of characteristic polynomials from several variables (pixel characteristics) for each face (say F_i) in the data base through a resampling process the system in use, is trained. A face (say F) is recognized by finding its characteristic polynomials and using an AdaBoost Classifier from F's polynomials to each of the F_i's polynomials. The winner is the polynomial family closer to F's corresponding to target face in data base.

Keywords**—**AdaBoost Classifier, Holistic Face Recognition, Minimax Multivariate Approximation, Genetic Algorithm.

I. Introduction

THIS paper is focused in the process of identifying an individual from the recognition of her/his face presents, interesting from the computer point of view. Although a human may solve most of the problems implied and achieve a high degree of adequate recognition since early in life, computer systems have important limitations when confronted with the same problem.

There is evidence to suggest that the human capacity for face recognition is a dedicated process, not merely an application of the general object recognition process [10]. Apparently, in human, both global and local features are used in a hierarchical manner.

Since the recognition of an object, from the point of view of a pattern classifier, may be thought of as a learning problem it may be solved by using a multivariate approximant. It is more or less common to attempt learning using, say, neural networks which try to capture the essence of the patterns under study in their architecture and their free parameters (weights). When using this approach, the network is required to learn but without overtraining, lest the model loose all generality. But an approximation function is a possible way to attain a classifier from a training data set without overtraining risks. It is wanted, in fact eager, to through this paper, accept some inherent approximation error in exchange for generality… The problem is to find the *proper form of the approximant* as well as the adequate number and value of the coefficients. Arbitrarily, a polynomial form as an approximant is selected and, thereafter, a genetic algorithm (GA) to find the coefficients and the degree of the approximation polynomials which minimize, respectively, the fitness error for the data is used. These data are used to identify the faces which form the training set. A given face which is to be identified is characterized by a polynomial which is compared with the training data set using an AdaBoost Classifier. Face recognition is then achieved by comparing it with a data base which has a single face per subject.

The rest of the paper is organized as follows. In section II some brief remarks regarding the psychophysics and neuroscience motivation of our choice of variables are made. In section III, some of the previous approaches to face recognition are reviewed. In section IV, our work and preliminary results are described. Finally, in section V conclusions are presented.

II. Psychophysics/Neuroscience Elements of Face Recognition

Psychological experiments as well as neuropsychological observations suggest that the human brain achieves face recognition from several information processing channels which are functionally independent from one another [2]. It is thought that humans are born with certain predisposition to respond to the patterns which are present in faces. Such observations suggest that faces are "special" objects in the

C. Villegas-Quezada is with Universidad Iberoamericana, Maestría en Sistemas, México, D.F. (e-mail: carlos.villegas@uia.mx).

J. Climent is with Universitat Politecnica de Catalunya, Departamento de Ingeniería de Sistemas, Automática e Informática Industrial, Barcelona, España (e-mail: juan.climent@upc.es).

visual world of man. However, these possible specialization characteristics do not necessarily imply that the processes of face recognition fundamentally differ from those used to recognize other kinds of objects. The representations of the face that a human uses for its recognition, seems to be based in encoding the “low level” characteristics of a given image. Costen and others [7], mention that there exists ample evidence that face recognition depends, in good measure, of the low frequencies which a human receives when visualizing a particular face. On the other hand, several authors consider that face recognition seems to involve basically holistic recognition [3]. In the recognition of other kind of objects, it seems that such global “approach” is present to a lesser extent. Rather, the process seems to involve the search for local characteristics and analysis of the parts of the object under scrutiny. Other researchers have found that children, the younger the better, use representations based on particular signs (eyes, mouth, nose) to achieve their recognition, whereas the adults use, mainly, holistic methods.

The role of spatial frequency analysis is important. Earlier studies [9, 40] concluded that information in low spatial frequency plays a dominant role in face recognition. Recent studies have show that in identification requires the use of high-frequency components, and the low-frequency components contribute to global description.

III. Previous Works in Face Recognition

Most of the current face recognition algorithms can be categorized into two classes: image template based or geometry feature-based.

In face recognition we may think of two basic forms to try to solve the problem:

- Feature identification (such as eyes, mouth, nose, etc.)
- Templates.

In the reported works at least two images per subject are employed for the training phase. The use of 5 or 10 images per face are common and, in some cases, 20 images per subject are employed. The traditional main methods are:

- Geometrical features [4,8]
- Templates [5]
- Principal components analysis (PCA). [14,15]
- Eigenfaces. [37,38]
- Fisherfaces. [1,26]

Some other methods proposed for face recognition are: Bayesian inference [29], Elastic Bunch Graph Matching (EBGM) [24], Support Vector Machines (SVM) [30], Linear Discriminant Analysis (LDA) [13], Kernel Methods [36], Neural Networks [16, 23], Local Feature Analysis (LFA) [31].

Within the past decade, major advances have occurred in face recognition. Especially, in recent years: Elastic Graph Matching, Gabor Fisher Classifier, and applications with AdaBoost.

Regarding the utilization of Gas in face recognition there are not many reported works. Most of them use the GAs in combination with some of the methods discussed in the preceding paragraphs: PCA, templates and neural networks, among others [17,18]. Schackleton [34] uses a population of templates to recognize the face using a GA to find the best fitness of the templates to recognize within the face data base. Populations size was 100, in 200 generations with a 40 subject set. Lanitis et al [22] used a combination of PCA and a GA to achieve the recognition. A data base of 30 subjects with 690 images was used. Another application of GAs and PCA is reported by Liu and Wechsler [25]. There, GAs were used to find the best configuration (rotation) of the axis of the principal components. This was applied to the recognition of 369 subjects, using approximately 2 images per subject achieving an efficiency of 97.02%. Pinto and Sossa [32] propose a method of invariants combined with a GA for face recognition

Face recognition in almost the totality of the reviewed literature, achieves the desired goal through the analysis of the principal characteristics of the face generating a set of templates which comprise different regions of the countenance or methods which combine templates and features.

IV. Face Recognition Using Multivariate Approximation and Genetic Algorithms

A. Genetic Algorithm

This paper reports the theoretical elements and initial experiments that are being done for face recognition using a polynomial approximation (as an optimization problem) and genetic algorithms.

The GA we used has been reported in [20]. It may be described as follws:

<u>Vasconcelos Algorithm</u>

00) **Define**:
    a) *N*: the number of individuals
    b) Pc: the probability of crossover
    c) Pm: the probability of mutation
01) Randomly generate a population of size *N*.
02) Evaluate the individuals in the population
03) Sort individuals from best to worst
04) **While** *convergence criteria are not met*
05) **for** i=1 to *N*/2
06) Generate a random number *R*
07) If $R >= Pc$ cross individuals *i* and $N\text{-}i+1$
    (crossover is annular with a ring size *l/2*)
08) **endfor**
09) Randomly mutate ($l*N*Pm$) bits in the population
10) Evaluate the new *N* individuals
11) Sort all *2N* individuals from best to worst
12) Eliminate the worst *N* individuals
13) **endwhile**

This algorithm has been show to display faster convergence than others [20] for an unbiased set of arbitrary functions.

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:2, No:8, 2008

B. Face Images for Training

The data base called ORL generated in the AT&T Laboratories at Cambridge University in the UK was used. Such base contains the faces of 40 subjects with 10 pictures for each subject. The subjects (men and women) are found with different facial expressions and some of them with beard and/or eyeglasses. All images were taken with a homogeneous background, in frontal position and with certain variations in the face's angle. The pictures are in PGM (Portable Grey Map) with 92x112 pixels and 256 levels of gray. ORL has been used in several projects involving face identification using various recognition methods [12,33,34].

For this work, only the first image of every series (frontal picture), was selected to create the training data base (40 images). One of the purposes of this project is to recognize a face using only the frontal picture of each subject. The other 9 images of each subject will be used as external pictures to be identified.

From the image of the face several characteristics from the pixels are obtained: Low frequency, Gradient, high frequency, Maximum entropy, gray level, etc. An approximate number of fifteen characteristics were obtained from the pixels. These characteristics are considered as variables, one of them is considered the dependent variable, and the others as independent variables. As part of this work, it is wished to obtained which variables are the most adequate for obtaining the face characterization.

Through the multivariate system described before, a resampling method will be done as presented in the next section.

C. Resampling for Faces

It is proposed a resampling scheme for the configuration a family of polynomials. A resampling technique is utilized to generate a number of subsets from the original image and for the original training dataset, generate a new training dataset [28].

So, the new training dataset contains a small number of sample face images. A number of subsets (Si) are generated by resampling the original training set. For each subset Si, is generated an approximation polynomial using a Genetic Algorithm. Also, the image to be recognized is resampled and approximated by the polynomial family.

D. Face Recognition as a Learning Problem

The recognition of a face starting from the digitalization of frontal pictures may be formulated as a pattern recognition problem, which in turn, may be seen as a learning problem [6,21,39]. In this work we attempt supervised learning, consisting of the acquisition of classification functions from a set of examples. Some preliminary experiments were done using a set of 40 frontal images of faces from ORL data base.

From the data base it is compared the face to be identified, in order to find holistic similarities without the need to extract geometric features or the use of templates. When, as here, samples are of the form (x_i, y_i). A learning function f such that $f(x_i)=y_i$ may be assumed. The goal is to find a function f such that the given function captures the "general patterns" present in the training data and we may apply the determined function to predict the values of y when x is given.

The function may be generalized to an *n-dimensional* space. The values used for the function that identifies the face may be attributes of the pixels (such as levels of gray or color, noise, entropy, etc.). Using this approach, it is capable to characterize a face holistically found a polynomial approximant with a GA. The approximant has the form of:

$$f(V_1,...,V_p) = \sum_{i_1=0}^{g_1} ... \sum_{i_p=0}^{g_p} C_{i_1...i_p} V_1^{i_1} ... V_p^{i_p} \qquad (1)$$

E. Genetic Algorithms in the Approximation of Functions

A family of approximating polynomials whose purpose will be to represent a face from some attributes of the pixels of its image has been selected. Polynomial approximation has been frequently performed as linear or multiple regression. In the latter case, data has to comply with certain conditions (under the L_2-norm): variables have to be normally distributed, the distributions variances must be similar for similar values of the independent variables, the dependent variable must have a mean which is on the regression lines, etc. Several "real-life" problems, (among them the data corresponding to a human face) do not comply with the aforementioned constraints. Hence, one form to solve the problem is to find an approximant seen as an optimization problem. The goal is to find the form and coefficients of a polynomial which better characterize the relations between a set of independent variable and the dependent variable according to some norm. The solution of this problem is combinatorially hard and is difficult to tackle using traditional methods [13]. To perform the optimization mentioned before, we used a method proposed before in [19, 21]. It allows us to find the form and values of the coefficients of the approximating polynomial in a way that the maximum absolute error between the data and the approximant is minimized (under the L_∞-norm). The polynomial is of the form presented in (1).

Therefore, during the training phase the faces of the data base will be characterized. This approximation will be achieved using different attributes for every face, which are to be obtained by sampling the pixels of every face, as mentioned in section 4.3. Later on, a given face to be identified will also be characterized by a family of polynomials which are compared with those polynomials corresponding to the faces of the training data base.

In the GA process 50 individuals per generation and 50 generations were used. The approximating polynomials were set to 12 terms and a highest degree of 9.

F. Ensemble-Based Learning with Boosting

Traditionally, the approach used in the design of pattern recognition systems has been to experimentally compare the performance of several classifiers in order to select the best one. However, an alternative approach based on combining multiple classifiers has emerged over the last years. This approach goes under various names such as Multiple Classifier Systems (MCS) or committee or ensembles of classifiers.

Recently, a machine learning technique known as “boosting” has received considerable attention, due to its usefulness in designing ensemble-based classifiers [11, 35]. The idea behind boosting is to sequentially employ a base classifier on a weighted version of the training sample set to generalize a set of classifiers of its kind. The weights are updated at each iteration through a classification-error-driven mechanism [27].

One of the main algorithms of boosting is AdaBoost. The AdaBoost algorithm, introduced by Freund and Schapire [11], solved many of the practical difficulties of the earlier boosting algorithms [41].

In this work, it is proposed the use of AdaBoost for classified the polynomial which characterize the face to recognize with respect the other polynomial of the training data base.

V. Conclusion

The methodology under consideration does not require to calculate the geometric elements of a face or to obtain the basic position of any feature of the face; neither does it require the use of templates. In those methods, an exhaustive search to identify eyes, nose, etc, is forced to be done. Then and only it is possible to apply the methods for the recognition. The method proposed in this paper, is totally automated since it requires no knowledge of specific features. Likewise, by randomly sampling the image, only a relatively small amount of pixels to achieve polynomial approximation both during training as during identification is required.

On the other hand, the holistic approach allows the recognition of a large number of subjects which display certain facial modifications: beard, eyeglasses, etc. Other systems are hampered by these variations and some are even rendered useless. The preliminary results obtained were highly satisfactory, particularly since a single image per subject during the training phase (as opposed to other methods reported which require several images) was used.

In the following phases of this research, different (non-polynomial) multivariate approximants which combine different pixel characteristics from the faces will also be tried out. It is planned to obtain a formal analysis of the process underlying face characterization from different variables (pixel characteristics); other variables complementing or replacing the ones already tested shall be introduced. Also, the use of a parallel computer to compute the approximants using Genetic Algorithms it is planned as well.

References

[1] Belhumeur, P., Hespanha, J., Kriegman, D., “Eigenfaces vs Fisherfaces: Recognition Using Class Specific Linear Projection”, IEEE Trans. Pattern recognition and Machine Intelligence, v. 19, No. 7, 1997, pp. 711-720.

[2] Bruce, Vicki, Hancock, Peter J.B., Burton, A. Mike, “Human Face Perception and Identification”, in: Wechsler, Harry, Phillips,

[3] Bruce, Vicki, et. al. (Eds.), Face Recognition: From Theory to Applications, Springer/NATO, Germany, 1998.

[4] Brunelli, Roberto; Poggio, Tomaso; “Face Recognition through Geometrical Features”, McGraw Hill, 1995.

[5] Brunelli, Roberto; Poggio, Tomaso; “Face Recognition: Features versus Templates”; IEEE Trans. on Pattern Recognition and Machine Intelligence; v. 15; No. 10; October; 1993; pp. 1042-1052

[6] Cherkassky, Vladimir, “Inductive Principles for Learning from Data”, in: Wechsler, Harry, Phillips, P. J., Bruce, Vicki, et. al. (Eds.), Face Recognition: From Theory to Applications, Springer/NATO, Germany, 1998.

[7] Costen, N.P., Parker, D.M., Craw, I., “Effects of high-pass and low-pass spatial filtering on face identification”, Perception & Psychophysics, v. 58, 1996, pp. 602-612.

[8] Cox, Ingemar J, Ghosn, J., Joumana, Y., “Feature-Based Face Recognition Using Mixture-Distance"” NEC Research Institute, Technical Report 95-09, Princeton, NJ, October, 1995.

[9] Dailey, Matthew N., Cottrell, Garrison W., “Learning a Specialization for Face Recognition: The Effect of Spatial Frequency”, June, 1997, in Internet

[10] Ellis, H.D., “Introduction to aspects of face processing: Ten questions in need of answers”, In H. Ellis, M. Jeeves, F. Newcombe, eds., Aspects of Face Processing, pp. 3-13, Nijhoff, 1996.

[11] Freund, Y., Schapire, R.E., “A decision-theoretic generalization of on-line learning and an application to boosting”, Journal of Computer and Systems Sciences, vol. 55 (1), pp. 119-139, 1997.

[12] Gong, Shaogang, McKenna, Stephen J., Psarrou, Alexandra, Dynamic Vision: From Images to Face Recognition, Imperial College Press, London, 2000.

[13] Grotschel, Martin, Lovász, Lászlo, Combinatorial Optimization: A Survey, DIMACS Technical Report 93-29, Princeton University, May, 1993. In Internet.

[14] Hancock, P.J., Bruce, V., Burton, A.M., “Testing Principal Component Representation for faces”, Technical report, University of Stirling, UK, 1998, in Internet.

[15] Hancock, Peter J. B.; Burton, A. Mike; Bruce, Vicki; “Face processing: human perception and principal component analysis”; Memory and Cognition; vol. 24; No. 1; 1996; pp 26-40.

[16] Howell, J., Buxton, H., “Invariance in radial basis function neural networks in human face classification”. Neural Processing Letters, 2(3), pp. 26-30, 1995.

[17] Huang, Ren-Jay, Detection Strategies for face Recognition Using Learning and Evolution, Ph. D. Dissertation, George Mason University, Abstract, 1998.

[18] Isaka, Satoru, “An Empirical Study of Facial Image Feature Extraction by Genetic Programming”, Report- OMRON Advanced Systems, Inc., Santa Clara, CA, 1997, in Internet.

[19] Kuri, A., “A Methodology for the Statistical Characterization of Genetic Algorithms”, MICAI 2002: Advances in Artificial Intelligence, Lecture Notes in Artificial Intelligence, pp. 79-88, Springer-Verlag, 2002.

[20] Kuri, A., Villegas, C., "A universal Eclectic Genetic Algorithm for Constrained Optimization". Proceedings 6th European Congress on Intelligent Techniques & Soft Computing, EUFIT'98, pp. 518-522, 1998.

[21] Kuri, Angel, “Pattern Recognition via a Genetic Algorithm”, in Guzmán, A., Shulcloper, J.R., Sossa, J.H., et al. (Comp.), II Taller Iberoamericano de Reconocimiento de Patrones-La Habana, Cuba, ICIMAF-CICIPN, 1997, pp. 345-356.

[22] Lanitis, A.; Hill, A.; Cootes, T. F.; Taylor, C. J.; “Locating Facial Features Using Genetic Algorithms”; Oxford;

[23] Lawrence, S., Giles, C. L., Tsoi, A., Back, A.D., “Face recognition: A convolutional neural network approach”, IEEE Transactions on Neural Networks, 8(1), pp. 98-113, 1998.

[24] Laurenz, W., Fellous, J.M., Krüger, N., von der Malsburg, C., “Face recognition by elastic bunch graph matching”, 19 (7), pp. 775-779, 1997.

[25] Liu, Chengjun, Wechsler, Harry, “Face Recognition Using Evolutionary Pursuit”, Fifth European Conference on Computer Vision, University of Freiburg, Germany, 1998, in Internet

[26] Liu, Chengjun, Wechsler, Harry, "Enhanced Fisher Linear Discriminant Models for Face Recognition", 14th International Conference on Patter Recognition , Queensland, Australia, 1998, in Internet.
[27] Lu, J., Plataniotis, K.N., Venetsanopoulos, A.N., "Ensembles-based Discriminant Learning with Boosting For Face Recognition", 2005, In Internet
[28] Lu, X., Jain, A., "Resampling for Face Recognition", In Internet.
[29] Moghaddam, B., Pentland, A., "Probabilistic Visual Learning for Object Representation", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 18, pp. 696-710, 1997.
[30] Osuna, E., Freund, R., Girosi, F., Training support vector machines: An application to face detection. 1997.
[31] Penev, P., Atick, J., Local Feature Analysis: A general statistical theory for object representation, 1996.
[32] Pinto-Elías, R., Sossa-Azuela, J.H., "Human Face Identification Using Invariant Descriptions and a Genetic Algorithm", in Coelho H. (Ed.), Progress in Artificial Intelligence-IBERAMIA 98 (6th Ibero-American Conference on AI-Lisbon, Portugal), Springer, Lecture Notes in AI-No. 1484, Germany, 1998, pp.293-302.
[33] Samaria, F.S., Harter, A.C., "Parameterization of a Stochastic Model for Human Face Identification", Proceedings of the 2nd IEEE Workshop on Application of Computer Gong, Shaogang, McKenna, Stephen J., Psarrou, Alexandra, Dynamic Vision: From Images to Face Recognition, Imperial College Press, London, 2000. Vision, Sarasota, Florida, December 1994.
[34] Schackleton, Mark, "Learned Deformable Templates for Object Recognition", IEEE GAs in Vision Colloquium, 1996, in Internet.
[35] Schapire, R.E., "The boosting approach to machine learning: An overview", MSRI Workshop on Nonlinear Estimation and Classification, Berkeley, CA, pp. 149-172, 2002.
[36] Schoelkopf, B., Smola, A., Muller, K.R., "Kernel principal components analysis", Artificial Neural Networks, ICANN97, 1997
[37] Turk, M., Pentland, A., "Eigenfaces for recognition", Journal of Cognitive Neuroscience, 3 (1), pp. 71-86, 1991
[38] Turk, M.A., Pentland, A.P., "Face Recognition Using Eigenfaces", Proceedings IEEE Computer Society Conference on Computer Vision and Pattern recognition, pp. 586-591, 1991.
[39] Vapnik, V. N., The nature of statistics learning theory. Springer Verlag. Heidelberg. 1995
[40] Wechsler, Harry, Phillips, P. J., Bruce, Vicki, et. al. (Eds.), Face Recognition: From Theory to Applications, Springer/NATO, Germany, 1998.
[41] Viola, Paul, Jones, Michael, "Rapid Object using a Boosted Cascade of Simple Features", Conference on Computer Vision and Pattern Recognition, 2001.

Carlos Villegas-Quezada received the BSc in Civil Engineering, MSc degree in Systems and Planning from Universidad Iberoamericana, México City, and Artificial Intelligence Especiality from Fundación Arturo Rosenblueth. He is currently a PhD student in the Departamento de Ingeniería de Sistemas, Automática e Informática Industrial (ESAII) en la Universitat Politecnica de Catalunya, Barcelona, España.

He is an titular professor in the Engineering Department, Maestría en Sistemas, Universidad Iberoamericana in Mexico City.

Joan Climent received the PhD in the Universitat Politécnica de Catalunya, España.

He is an Associate Researcher in the Departamento de Ingeniería de Sistemas, Automática e Informática Industrial (ESAII) in the Universitat Politecnica de Catalunya, Barcelona, España.