Face recognition through a chaotic neural network model

International Joint Conference on Neural Networks, 2014, Beijing.
http://www.producao.usp.br/handle/BDPI/46567

Downloaded from: Biblioteca Digital da Produção Intelectual - BDPI, Universidade de São Paulo
Face Recognition through a Chaotic Neural Network Model

Luis Fernando Martins Carlos Junior and Joao Luis Garcia Rosa, Member, IEEE

Abstract—K-sets models are connectionist methods based on neuron populations, conceived through EEG analyses of the mammalian olfactory system. These models present a biologically more plausible structure and behavior than other neural networks models. K-sets have been used in many machine-learning problems, with potential application on pattern recognition while presenting novel chaotic concepts for signal processing. This paper presents the characteristics of the K-sets models and their application in a face recognition task. Our method was tested using Yale Face Database B and the results show that it outperforms popular recognition methods.

I. INTRODUCTION

Face recognition is nowadays a very attractive area of investigation for many researchers, mainly for two reasons: the large number of possible applications and the challenges found in the area. Among the challenges encountered in the field, the main ones are: image noise, variations in illumination, pose variations, facial expressions, age range, glasses, beard, mustache, etc. [1].

Many researchers have worked with the aim to develop a method robust enough to cope with the difficulties encountered in the area, giving rise to several different techniques to deal with the problem. Among the existing techniques, the main ones are: principal component analysis (PCA) [4], [5], linear discriminant analysis (LDA) [6], independent component analysis (ICA) [7], support vector machines [8], Bayesian methods [9] and neural networks [10], [11].

Despite the several years of research and the amount of different approaches for face recognition available, most methods still find it difficult to deal with the problems of the area [12]. In [13], a review of the main methods of face recognition is presented, being observed that most of the methods achieve a recognition rate greater than 90% in controlled lighting conditions. However, this performance is reduced when there are variations in illumination, pose and facial expression.

Due to the challenges of the area and the deficiencies found in current models, we propose the research and development of a method for face recognition based on Freeman K sets. The K sets are models belonging to the fourth generation of connectionism [14], which are based on populations of neurons, whose structure and behavior are considered biologically more plausible [15]. The K sets were designed as results of studies and observations of neuroscientist Walter J. Freeman on the olfactory system of animals from electroencephalograms (EEG) [16], [17].

Successfully utilized in several tasks, the K sets produce better results in complex problems, with few or noisy examples where other techniques usually do not get good results. The K sets have been used to predict time series [15], character recognition [18], recognition of complex patterns [19], clustering [20], and face recognition [21], [22], [23].

Among the features presented by K sets, we provide a few that can bring benefits to the task of face recognition, such as (i) dynamic behavior that allows better generalization and fault tolerance [15], (ii) ability to generalize in complex, noisy and with very few examples environments [14], [15], (iii) requirement of only one step to converge in learning, which in some cases makes the K sets to go much faster than other techniques, despite the higher computational cost required by each step [21].

II. OVERVIEW OF K-SETS

Freeman K sets are a family of components of increasing complexity, which have dynamic structure and behavior. They are mesoscopic models that represent an intermediate level between microscopic and macroscopic structures of the brain [24], [17]. The models were introduced by neuroscientist Walter J. Freeman in the 1970s [16] and gained prominence since the turn of the century [25], [26], [27], [28].

The family of K sets is comprised of sets K0, KI, KII, KIII, and KIV, each one designed to model a part of the brain [15]. Thus, the KI models the dentate gyrus, KII the olfactory bulb, KIII the sensory cortex while KIV models the limbic system [15], [17]. Figure 1 illustrates the hierarchy of K sets.
A. K0 set

K0 is the simplest structure in the hierarchy of K sets and the basic element needed for the construction of all other higher levels [29]. K0 sets represent non-interactive collections of neurons that share common inputs and outputs, where each K0 models a population of approximately 10^4 neurons [16], [17]. The K0 set is governed by a point attractor with zero output and stays at equilibrium except when perturbed. It is described by a dependent state, according to the following second order ordinary differential equation (ODE):

\[ (a + b) \frac{d^2 P(t)}{dt^2} + (a + b) \frac{dP(t)}{dt} + P(t) = F(t) \]  \hspace{1cm} (1)

where \(a = 0.22\) and \(b = 0.72\) are biologically determined time constants. \(P(t)\) indicates the node activation as a function of time. \(F(t)\) is the weighted sum of the activation of neighboring nodes. K0 has a weighted input and an asymptotic sigmoid function for output \(Q(x, t)\), given by:

\[ Q(x, t) = \begin{cases} 
q(1 - e^{-(x-x(T))/x_0}), & x(T) > x_0 \\
-1, & x(T) < x_0 
\end{cases} \]  \hspace{1cm} (2)

where \(q = 5\) is the parameter that specifies the slope and the maximum asymptote of the curve. This sigmoid function was modeled from experiments on biological neural activation [16], [15].

B. KI sets

The KI set is formed by K0 units interconnected through lateral feedback of same signs, so we can have excitatory and inhibitory KI sets [29]. The dynamics of KI is described as convergence to non-zero fixed point. If KI has sufficient functional connection density then it is able to maintain a non-zero state of background activity by mutual excitation (or inhibition). KI typically operates far from thermodynamic equilibrium output. Neural interaction by stable mutual excitation (or mutual inhibition) is fundamental to understanding brain dynamics [24].

C. KII sets

The KII is formed by at least two KI sets (or four K0 sets) with dense functional interconnections [29]. Examples include the olfactory bulb, hippocampal regions, and the prepyriform cortex [20]. KII has four types of interactions: excitatory-excitatory, inhibitory-inhibitory, excitatory-inhibitory, and inhibitory-excitatory [14]. KII sets are responsible for the emergence of limit cycle oscillation due to the negative feedback between neural populations. Transitions from point attractor to limit cycle attractor can be achieved through a suitable level of feedback gain or by input stimuli. Systematic analysis has identified regions of stability and transitions between limit cycle and fixed point dynamics [20].

D. KIII sets

KIII sets consist of several interconnected KII sets, and describe a given sensory system in brains, e.g., olfactory, visual, auditory and somatosensory modality [14]. They generate broadband, aperiodic/chaotic oscillations as background activity by combining negative and positive feedback among several KII populations with incommensurate frequencies. The increase in nonlinear feedback gain is driven by the input results in a destabilization of background activity and leads to the emergence of a amplitude modulation (AM) pattern in KIII. KIII sets are responsible for the embodiment of meaning in AM patterns of neural activity shaped by synaptic interactions that have been modified through learning in their layers [20].

It has been shown that the KIII sets can be used as an associative memory that encodes the input data into nonconvergent spatio-temporal oscillations [30], [27]. Comparing with convergent recurrent networks, the KIII chaotic memories present several advantages:

1) They produce robust memories based on relatively few learning examples even in noisy environments;
2) The encoding capacity of a network with a given number of nodes is exponentially larger than their convergent counterparts;
3) They can recall the stored data very quickly, just as humans and animals can recognize a learned pattern within a fraction of a second [14].

III. USED DATA SET

In this paper we used the “Yale Face Database B” [31]. This database has 5760 images of 10 individuals over 64 lighting conditions in nine poses (frontal pose, five poses with 12° and three poses with 24° from the axis of the camera). In order to capture the images of this database, a geodesic lighting equipment with 64 Xenon strobes was constructed. The images of each pose were divided into four subsets (12°, 25°, 50° and 77°) according to the angle of the light source with the axis of the camera. Subsets 1, 2, 3 and 4 have respectively 7, 12, 12 and 14 variations of light for each pose of each individual. A sample of images from this database can be seen in Figure 2.

The original size of the images is 640 x 480 pixels. In our experiments, all images were automatically cropped using...
Fig. 2. Image examples from Yale Face Database B. a. One image for each one of nine different poses. b. One image representing each one of the lighting subsets - Subset 1 (12°), Subset 2 (25°), Subset 3 (50°), and Subset 4 (77°).

the framework of object detection Viola-Jones [32], where the face region is extracted from original image and resized to 56 x 56 pixels.

IV. KIII SETS FOR FACE RECOGNITION

Our task is to identify individuals from the Yale B dataset. 450 images from the frontal pose (45 per person) were used. For training, the subset 1 with 70 images (7 per person) was used and for test, the subsets 2, 3 and 4 with 380 images (38 per person) were used.

The proposed method consists of two modules: feature extraction and classification, as shown in figure 3.

The feature extraction is performed in order to extract the most relevant features of the face image through an analysis applied directly on the image pixels, performing a dimensionality reduction on the data that will be processed. For that, the method Subspace Linear Discriminant Analysis (LDA) [33] was employed. This method consists of two steps: first, the face image is projected onto a sub-space through Principal Component Analysis (PCA) and then a Linear Discriminant Analysis (LDA) is used to obtain a linear classification on the sub-space.

For application of the PCA step, consider \( G \) a matrix \( M \times p \), where \( M \) is the number of images in training set, and \( p \) is the linearization of the image pixels. The number of principal components must be choice between 1 and \( M - 1 \) [34]. In our case (between 1 and 69), we choose 50 principal components, because it represents 99% of total variance. Next LDA is applied to reduce the dimension to \( c - 1 \) [35], where \( c \) is the number of classes, in this case \( c = 10 \), thus the LDA output is 9.

For classification, the extracted features are submitted to the KIII model. The parameters used in KIII model were optimized to solve the classification task [27].

A. KIII training

Data were normalized between -1 and 1 and then presented to the KIII during training phase. The network consists of three KII layers of length \( N \), where \( N \) is the input length (9, the resulting dimension from LDA), and it was trained using Hebbian learning in third layer, with learning rate \((\alpha) 0.5 \). Each one of 70 training samples was presented during 300 cycles (milliseconds) during active phase, followed by 200 cycles without stimuli (resting phase).

B. KIII test

After the training phase, the 380 test samples were also presented to the network for 300 active cycles and 200 resting cycles. The output obtained for test samples is compared with the output obtained from training samples. The Mahalanobis distance [36] is used to calculate the distances from unknown images to each class and the k-nearest neighbor (k-NN) is used to classify them.

C. Experiments and Results

For a comparison of the proposed method, the same experiment was performed with traditional techniques Eigenfaces [5] and Fisherfaces [35]. In addition to the traditional methods, we employed a method which uses KIII model and Discrete Cosine Transform (DCT) [23], for being the closest method to the proposed method.

Results of our experiments can be seen on Table 4 and in Figures 5, 6 and 7. The results are presented in two standard ways: (1) a table showing performance at rank 1 (recognition rate within the top one match), and Cumulative Match Score (CMS) curve [37], showing the cumulative results for ranks 1 and higher.

V. DISCUSSION

The results show that the proposed method (PCA+LDA+KIII) obtained best results both in rank 1 and in CMS curve.

In the experiments, the method was trained using the subset with less variation in lighting Subset 1 (12°) and
was tested using the subsets with more variation in lighting, that is, Subset 2 (25°), Subset 3 (50°), and Subset 4 (77°). From the Subset 1, a representation for recognizing the face images from other subsets is created. The results show that the method achieves best results for Subset 2. However, the results are worse for Subset 3 and worst for Subset 4. This is expected, because the recognition complexity is increased. Nevertheless, the difference between the proposed method and the second best result (Fisherfaces) is increased. This reinforces the observation presented in [19], that states as an advantage of KIII-sets for complex pattern recognition tasks and not for regular pattern recognition.

VI. CONCLUSION

This paper presents a method for face recognition based on a KIII neural network model, inspired by the biological neural system. It has been demonstrated that the proposed method achieves best recognition rates and is superior than some popular methods of face recognition, specially under conditions of higher complexity, like wide variations in illumination. For future work, more studies seeking an improvement in the proposed method should be investigated. We believe that we are on the right track, and mimicking the biological brain is an effective way to solve complex patterns recognition problems such as face recognition.

ACKNOWLEDGMENTS

Luis Fernando Martins Carlos Junior is grateful to Brazilian Agency CAPES (process DS-7904230/M). Joao Luis G. Rosa is grateful to the Brazilian Agency FAPESP (process 2012/09268-3) for the financial support.

REFERENCES


