

Improved Parameters Estimating Scheme for E-HMM with Application to Face Recognition

Bindang Xue¹, Wenfang Xue², and Zhiguo Jiang¹

¹ Image processing center, Beihang University, Beijing 100083, China
{xuebd, jiangzg}@buaa.edu.cn

² Institute of Automation, Chinese Academy of Sciences,
100088, Beijing, China
wfxue@nlpr.ia.ac.cn

Abstract. This paper presents a new scheme to initialize and re-estimate Embedded Hidden Markov Models(E-HMM) parameters for face recognition. Firstly, the current samples were assumed to be a subset of the whole training samples, after the training process, the E-HMM parameters and the necessary temporary parameters in the parameter re-estimating process were saved for the possible retraining use. When new training samples were added to the training samples, the saved E-HMM parameters were chosen as the initial model parameter. Then the E-HMM was retrained based on the new samples and the new temporary parameters were obtained. Finally, these temporary parameters were combined with saved temporary parameters to form the final E-HMM parameters for representing one person face. Experiments on ORL databases show the improved method is effective.

1 Introduction

Face recognition has been an active research topic recently and remains largely unsolved [1, 2]. Based on the recognition principle, diverse existing face recognition approaches can be briefly classified as three catalogues: geometric feature-based, principle component analysis (PCA)-like based and model based. Due to the ability to “learn” model parameters, several face recognition systems were based on E-HMM and this method appears having more promising potential [3-6]. The key problem using E-HMM for face recognition is how to train the model parameters for discovering intrinsic relations between face images and human face, and further building appropriate models based on these relations. However, the problem of choosing the initial model parameters for the training process and the problem of retraining model parameters were still left as open problems. In earlier work, Davis and Lovell had studied the problem of learning from multiple observation sequences [7] and the problem of ensemble learning [8] with multiple observation sequences being provided at one time. But how to deal with multiple observation sequences being provided at different time has not been addressed. While the retraining problem of E-HMM for face recognition is just like this problem. Under new environment, in order to improve the recognition accuracy, news training samples sets are added to the training samples sets. So it is needed to re-estimate the model parameters based on the newly formed sample sets. In this paper, a segmental scheme is presented to solve this problem.

2 E-HMM for face

A human face can be sequentially divided from top to bottom as forehead, eyes, nose, mouth and chin. Hence a human face can be viewed as a region chain. In such a way a human face can be defined as 1-D HMM. In essence, a human face image is a two dimensional object which should process as a 2-D HMM. To simplify the model processing, a specified pseudo 2-D HMM scheme is proposed. This model extends all top-down sub-regions in 1-D HMM as sub-sequences from left-hand side to right-hand side separately and uses extended sub-1-D HMM defining these sub-sequences hierarchically. This pseudo face 2-D HMM is also called E-HMM[3]. The face 2-D HMM scheme shown as fig. 1, composed of five super states (forehead, eyes, nose, mouth and chin) vertically, and the super states are extended as {3, 6, 6, 6, 3} sub states (embedded states) horizontally.

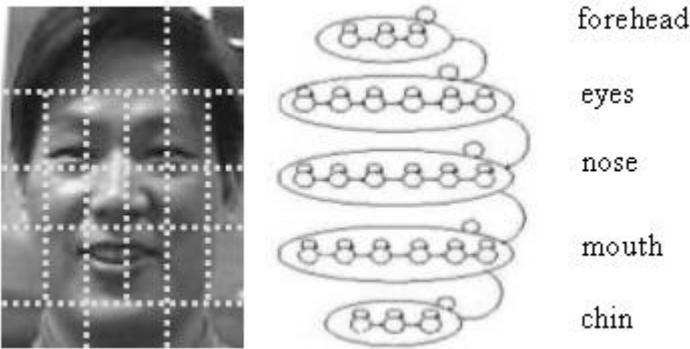


Fig. 1. E-HMM for face

An E-HMM structure can be defined by the following elements:

Super states parameters:

- N : the number of super states.
- Π : the initial super state probability distribution.
- A : the super state transition matrix. $A = \{a_{ij}, 1 \leq i, j, \leq N\}$.
- Λ : embedded 1D-HMMs, named super state. $\Lambda = \{\Lambda^i, 1 \leq i \leq N\}$.

Sub-states parameters:

- N^i : the number of sub states embedded in super state Λ^i . $S^i = \{s_k^i, 1 \leq k \leq N^i\}$.
- Π^i : the initial sub state probability distribution embedded in super state Λ^i , $\Pi^i = \{\Pi_k^i, 1 \leq k \leq N^i\}$.
- A^i : sub-states transition matrix in super state Λ^i . $A^i = \{a_{kl}^i, 1 \leq k, l \leq N^i\}$.
- B^i : the sub states output probability function in super state Λ^i , $B^i = \{b_k^i(o_{xy})\}$, where o_{xy} represent the observation vector at row x and column y ($x = 1, \dots, X, y = 1, \dots, Y$), the

sub-states output probability function that is typically used is a finite mixture of Gaussian probability density function(P.D.F.)

$$b_k^i(o_{xy}) = \sum_{f=1}^F C_{kf}^i N(o_{xy} | \mu_{kf}^i, U_{kf}^i), \quad (1 \leq k \leq N^i) \tag{1}$$

Where $N(o_{xy} | \mu_{kf}^i, U_{kf}^i)$ represents f^{th} Gaussian P.D.F. with the mean vector μ_{kf}^i and the covariance matrix U_{kf}^i , C_{kf}^i is the mixture coefficient for the f^{th} mixture of the output probability function of sub state k in super state Λ^i .

So an E-HMM can be defined as $\lambda = (N, A, \Pi, \Lambda)$, where N is the number of the super states. $\Lambda = \{\Lambda^1, \dots, \Lambda^N\}$, $\Lambda^i = \{N^i, \Pi^i, A^i, B^i\}$, A^i represents the super state i , N^i is the number of embedded sub states in super state Λ^i .

3 Training of the E-HMM

Given a set face images taken from the same person, model training is estimating the corresponding model parameters and saving them in a face database. The strategy for generating observation vector sequence and the training method are similar to the methods as described in article [3].

For describing the algorithm simply, it is useful to define the following variables:

- $StartSuperstate(i)$: represents the expected number of super state Λ^i at column $y = 1$ given R observation sequences;
- $StartState(i, k)$: represents the expected number of sub state s_k^i at row $x = 1$ in super state Λ^i given R observation sequences;
- $SuperTransition(i, j)$: represents the expected transition number from super state Λ^i to super state Λ^j ;
- $StateTransition(i, k, l)$: represents the expected transition number from sub state s_k^i to sub state s_l^i in super state Λ^i ;
- $SuperTransform(i)$: represents the expected transition number from super state Λ^i ;
- $StateTransform(i, k)$: represents the expected transition number from sub state s_k^i in super state Λ^i ;
- $Component(i, k, f)$: represents the expected number of f^{th} mixture element of the output probability function of the sub state s_k^i .

Based on the above variables, part parameters of the EHMM can be further re-estimated using the following formulas:

$$\bar{\Pi}_i = \frac{StartSuperstate(i)}{\sum_{j=1}^N StartSuperstate(j)} \tag{2}$$

$$\bar{a}_{ij} = \frac{SuperTransition(i, j)}{SuperTransform(i)} \tag{3}$$

$$\bar{\Pi}_k^i = \frac{StartState(i,k)}{\sum_{k=1}^{N^i} StartState(i,k)} \quad (4)$$

$$\bar{a}_{kl}^i = \frac{StateTransition(i,k,l)}{StateTransform(i,k)} \quad (5)$$

$$\bar{C}_{kf}^i = \frac{Component(i,k,f)}{\sum_{f=1}^F Component(i,k,f)} \quad (6)$$

4 Improved Parameters Estimating Scheme for E-HMM

In this paper, current training sample sets is referred as R_1 , and the model parameters can be iteratively estimated based on R_1 using formulas (2)-(6). During the estimating procedure, the variables defined above are labeled as $StartSuperstate^{R_1}(i), \dots$. When the training procedure is finished, the model parameters $\bar{\lambda}_1$ are saved, at the same time, the temporary variables $StartSuperstate^{R_1}(i), \dots, Component^{R_1}(i,k,f)$ are also saved. Once new sample sets R_2 is obtained, the whole sample sets include R_1 and R_2 . The segmental retraining scheme is that only the temporary variables $StartSuperstate^{R_2}(i)$ based on R_2 are needed to be re-estimated, then the last model parameter will be formed by combining $StartSuperstate^{R_2}(i), \dots, Component^{R_2}(i,k,f)$ with recoded $StartSuperstate^{R_1}(i), \dots, Component^{R_1}(i,k,f)$.

Another problem is how to choose a set of initial model parameters. The initial model parameters have great effect on the training procedure of the model. For example, choosing different initial model parameters will affect the convergence of the iterative training algorithm and the face recognition right rate. But there is no method to choose ideal initial model parameters now. One scheme to solve this problem is that we can divide the training sample sets into two parts R_1 and R_2 , the initial model parameters $\bar{\lambda}_1 = (\bar{\Pi}_1, \bar{A}_1, \bar{\Lambda}_1)$ are estimated based on sample sets R_1 , then we can estimate parameters $\bar{\lambda}_2 = (\bar{\Pi}_2, \bar{A}_2, \bar{\Lambda}_2)$ referring $\bar{\lambda}_1 = (\bar{\Pi}_1, \bar{A}_1, \bar{\Lambda}_1)$ as initial model parameters. In the end, it is easy to combine $\bar{\lambda}_1 = (\bar{\Pi}_1, \bar{A}_1, \bar{\Lambda}_1)$ with $\bar{\lambda}_2 = (\bar{\Pi}_2, \bar{A}_2, \bar{\Lambda}_2)$ to form the final model parameters $\bar{\lambda} = (\bar{\Pi}, \bar{A}, \bar{\Lambda})$. The initial model parameter comes from part training sample sets, so that it is better than other methods such as random initializing or choosing experiential values.

The formulas of the improved parameter estimating scheme for E-HMM are described as below:

$$\bar{\Pi}_i = \frac{StartSuperstate^{R_1}(i) + StartSuperstate^{R_2}(i)}{\sum_{j=1}^N StartSuperstate^{R_1}(j) + \sum_{j=1}^N StartSuperstate^{R_2}(j)} \quad (7)$$

$$\bar{a}_{ij} = \frac{SuperTransition^{R_1}(i,j) + SuperTransition^{R_2}(i,j)}{SuperTransform^{R_1}(i) + SuperTransform^{R_2}(i)} \quad (8)$$

$$\bar{\Pi}_k^i = \frac{StartState^{R_1}(i, k) + StartState^{R_2}(i, k)}{\sum_{k=1}^{N^i} StartState^{R_1}(i, k) + \sum_{k=1}^{N^i} StartState^{R_2}(i, k)} \tag{9}$$

$$\bar{a}_{kl}^i = \frac{StateTransition^{R_1}(i, k, l) + StateTransition^{R_2}(i, k, l)}{StateTransform^{R_1}(i, k) + StateTransform^{R_2}(i, k)} \tag{10}$$

$$\bar{C}_{kf}^i = \frac{Component^{R_1}(i, k, f) + Component^{R_2}(i, k, f)}{\sum_{f=1}^F Component^{R_1}(i, k, f) + \sum_{f=1}^F Component^{R_2}(i, k, f)} \tag{11}$$

5 Face Recognition Experiments and Results

The goal of the experiment on face recognition is just to evaluate the proposed segmental parameter re-estimating scheme, so a small database ORL face database [10] is chosen as the test datasets. ORL database contains 400 images of 40 individuals, with 10 images per individual at the resolution of 92×112 pixels. The images of the same person are taken at different times, under slightly varying lighting conditions, and with different facial expressions. Some people are captured with or without glasses. The head of the people in the images are slightly titled or rotated. Images of one person from ORL database show as Fig.2.

Firstly, the first six face images of one person are used to train the E-HMM, and the remaining four images are used to test the system. In order to evaluate the improved parameter estimating scheme, we divide the first six training images into two equal parts R_1 and R_2 . At first step, R_1 is used to train the model to get the initial model parameters $\bar{\lambda}_1 = (\bar{\Pi}_1, \bar{A}_1, \bar{\Lambda}_1)$, then R_2 is used to train the model parameters $\bar{\lambda}_2 = (\bar{\Pi}_2, \bar{A}_2, \bar{\Lambda}_2)$. At last, the final model parameters $\bar{\lambda} = (\bar{\Pi}, \bar{A}, \bar{\Lambda})$ are obtained quickly based on the improved parameter estimating scheme presented in this paper.



Fig. 2. Images of one person from ORL database

Given a test face image, recognition is to find the best matching E-HMM model within a given face model database and predicting the matching probability. Usually the model corresponding to the maximum likelihood is assumed to be the right choice

revealing the identity among the given face model database. Let there are P individuals in the database, given a face image t , the matching maximum likelihood probability rule is prescribed as:

$$P(O^t | \lambda_k) = \text{Max}(P(O^t | \lambda_p)) \quad (1 \leq k, p \leq P) \quad (12)$$

So the recognition result is that the face image t is corresponding to k^{th} person in the database.

Table 1 simply compares the recognition results of HMM trained using different parameter estimating method. The improved scheme achieves 99.5% correct recognition rate on ORL face database.

Table 1. Recognition results of different methods

Methods	Right recognition rate (%)
Pseudo-HMM[11]	90-95
E-HMM[3]	98.5
Segmental scheme	99.5

6 Conclusions

This paper describes an improved segmental scheme to initialize and re-estimate E-HMM parameter. The advantage of the improved parameter estimating scheme is that the E-HMM parameters re-estimating process has good ability of adaptation: when new sample set was added to the training sample, the information of the new sample set could be conveniently combined into the E-HMM, and the calculation complex was reduced. Besides, the improved parameter estimating scheme provides an answer for the problem of choice initial E-HMM parameters.

Future work will focus on sequential learning algorithm for E-HMM with application to face recognition.

References

1. Chellappa R., Wilson C.L., Sirohey S. Human and machine recognition of face: A survey. Proc. IEEE, 1995,83(5):705-740.
2. Zhao W., Face recognition: A Literature Survey. CS-TR-4167, University of Maryland, Oct. 2000
3. A.V. Nefian, M.H. Hayes, Maximum likelihood training of the embedded HMM for face detection and recognition, Proc. of the IEEE International Conference on Image Processing, ICIP 2000, Vol. 1, 10-13 September 2000, Vancouver, BC, Canada, pp. 33-36.
4. S. Eickeler, S. Muller, etc. Recognition of JPEG compressed face images based on statistical methods. Image and Vision Computing, 2000 (18):279-287.
5. F. Wallhoff, S. Eickeler, etc. A comparison of discrete and continuous output modeling techniques for a pseudo-2D hidden Markov model face recognition system.. Proceedings of International Conference on Image Processing, 2001(2):685-688.

6. H. Othman ,T. Aboulnasr. A simplified second-order HMM with application to face recognition, in the IEEE International Symposium on Circuits and Systems, 2001(2): 161 – 164.
7. Davis, Richard I. A. and Lovell, Brian C. and Caelli, Terry. Improved Estimation of Hidden Markov Model Parameters from Multiple Observation Sequences. In *International Conference on Pattern Recognition*, Quebec City, Canada, August 11-14 II, 2002,;168-171.
8. Davis, Richard I. A. and Lovell, Brian C. Comparing and Evaluating HMM Ensemble Training Algorithms Using Train and Test and Condition Number Criteria. *Pattern Analysis and Applications* .2003 (6):327-336.
9. Rabiner L., A tutorial on HMM and selected applications in speech recognition , *Proc. IEEE* ,1989, 77(2):257-286.
10. ORL Face database , Cambridge ,AT&T Laboratories Cambridge. (<http://www.uk.research.att.com/facedatabase.html>)
11. Samaria F., *Face Recognition Using Hidden Markov Models*, PhD thesis, University of Cambridge,1994.