Improvised Profile Construction for Multimedia Databases in E-Learning

S. C. Premaratne, D. D. Karunaratna, G. N. Wikramanayake
K. P. Hewagamage and G. K. A. Dias.
University of Colombo School of Computing,
35, Reid Avenue, Colombo 7, Sri Lanka.

E-mail: saminda@webmail.cmb.ac.lk, {ddk ,gnw ,kph ,gkd }@ucsc.cmb.ac.lk

ABSTRACT

E-learning has acquired a prime place in many discussions recently. A number of research efforts around the world are trying to enhance education and training using e-learning. This paper briefly explains one such attempt aimed at designing a system to support video clips in e-learning and explains how profiles of the presenters in video clips can be used to improve the usefulness of e-learning systems. Then it discusses one of the main problems identified in profile construction. Finally, it presents a solution to this problem and describes a novel algorithm for improving the efficiency of the profile construction process.

1. INTRODUCTION

E-learning is one of the fastest growing areas today. The main emphasis of E-learning is the management and delivery of quality teaching material electronically with the added value of maintaining standards without the limitation of the learner access location. It includes the use of multimedia which involves more than one form of media such as text graphics, animation, audio, and video. Several approaches have been proposed to increase the acceptance and usage of existing e-learning platforms in education, but most of them are restricted in flexibility with regard to the content and adaptation to the user's skills [3]. Our primary objective is to provide an e-learning system to satisfy requirements of users with different learning objectives and learning patterns.

The paper describes a number of techniques that we have implemented to support the integration and efficient use of video clips in e-learning. In our earlier publications we have described a multimodal multimedia database system to support content-based indexing, archiving, retrieval and on-demand delivery of audiovisual content in an e-learning environment [8, 9]. In this system, a feature selection and a feature extraction sub-system have been used to construct presenter profiles. The feature extraction process transforms the video key-frame data into a multidimensional feature space as feature vectors. This process can be considered as an implicit mechanism that both summarizes and regularizes the key-frame data. Any feature extraction procedure should proceeded by a feature selection process. The main objective of the feature selection process is to identify effective and representative features of the objects involved.

In this paper, we propose a novel approach to create the profiles by introducing a profile normalization algorithm. In particular, our method places more effort on solving the profile overlapping problem by using certain parameters. This work refines our earlier approach for profile construction which averages all sample key-frame data to construct the presenter profiles [8]. From our initial work, we have observed that averaging key frames of the presenter degenerates the presenter identification process with the increase in the number of profiles and produced poor results when the illumination level is not constant. In this paper we have proposed a novel approach to overcome the illumination when constructing presenter profiles even at different illumination levels.

The remainder of this paper is organized as follows. The system architecture is shown in Section two. Section three explains the technique we have used for profile creation and profile normalization modules. Finally, sections four and five give our conclusions and address the future work possible based on this project and the experiences we have gained by using this system.

2. SYSTEM ARCHITECTURE

The main components of our architecture are: a media server, meta-data database, Ontology and object profiles, keyword extractor, keyword organizer, Feature extractor, Profile constructor and the query processor [5, 6].

The first step of the profile constructor is to extract features from the video Key-frames which containing most of the static information present in a shot, so that face recognition process can focus on key-frames only. The main inputs to the profile constructor are these key-frames stored in the multimedia database (Figure 1).

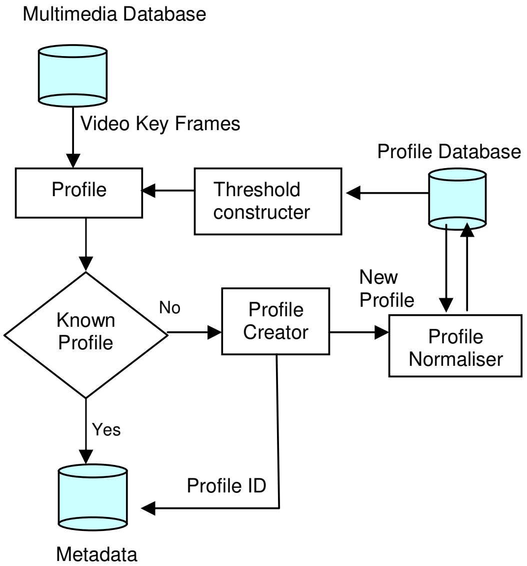

Figure 1: Profile Construction & Recognition Architecture

The profile detection and recognition process detects the faces in the key frame and try to match it with the presenter profiles available in the profile database. If the presenter in the key-frame matches with a profile then the system annotates the video shot with the profile identification and maps it with the metadata database. On the other hand, if the current presenter's key-frames do not match with the available profiles then the profile creator will create a new presenter profile and insert it in to the profile database.

2.1. Profile Construction

The profile construction is based on Principal Component Analysis (PCA) [4, 7]. The idea is to represent presenter's facial features in a transformed feature space where the individual features are uncorrelated. The feature space comprises of eigenvectors of the covariance matrix of the key-frame features. In this approach, PCA was computationally intensive when it was applied to the facespace. Through the experience we gained from our initial approach we have realized that the efficiency of our system can be improved substantially by limiting the analysis to the dominant eigenvectors of related key-frames instead of all eigenvectors of all related key-frames.

2.2 Profile Normaliser

Profile Normalizser acquires available profiles from profile database and executes the normalization algorithm and returns the profiles to the database. Since we get key-frames from different lighting conditions we have to have a proper dynamic profile normalization algorithm to maintain the efficiency of the profile matching algorithm. Therefore we concentrate on two descriptors, they are mean intensity and the standard deviation of the data set that we use to construct presenter profiles. After investigating the variation of the light and the deviation of the mean intensity and standard deviation, we propose an algorithm to normalize the profiles which provide facilities to maintain the accuracy of the system when adding new profiles to the database.

2.3 Threshold Constructor

For recognition we employee Euclidian distance algorithm to compute the distance between each profile and input face. As the minimum distance classifier, it works well when the key-frames have relatively small lighting and moderate expression variations. The weakness of this technique is that its performance deteriorates when lighting variations in the key-frames cannot be characterized as small. Therefore the threshold levels for the detection and recognition has to be changed according to the lighting variations. In our system the threshold constructor will calculate the light variation of each profile and adjust the threshold levels accordingly.

3. PROFILE CONSTRUCTION ALGORITHM

In this section of the paper we describes how we have improved the profile construction algorithm presented in [5, 6]. In our previous approach we have constructed profiles by getting the average values of the training set of faces. From the results gathered we have realized that, our system performance decreases when the video key-frames are captured in different illumination conditions [5].

The effects of illumination changes in key-frames are due to one of the two factors. The inherent amount of light reflected off the skin of the presenter, or the non-linear adjustment in internal camera control. Both of these conditions can have a major effect on facial features recognition [3]. In our initial profile construction approach lighting variations result in producing similar profile for different presenter and hence overlap of profiles in the eigenspace (Figure 2).

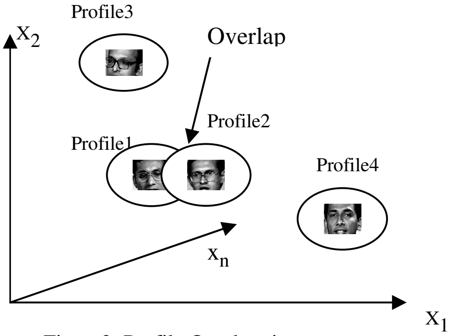

Figure2: Profile Overlapping

In Figure 2 and 5, axis X_1, X_2… X_n represents the n-dimensional eigenspace and the projection of some presenter profile to the eigenspace.

After experimenting with different parameters we have explored a strategy to overcome this problem by using standard deviation(S) and the mean intensity ( ) we developed an algorithm to implements this strategy (Equation 1).

$$\sum_{i=1}^{N}\sum_{j=1}^{n} \text{key-frame}_{j,i}(x,y) = (\text{key-frame}_{j,i}(x,y) - \overline{X}\ )\ * \quad -(1)$$

$\text{key-frame}_{j,i}(x,y) = (x, y)$ pixel value of the i^{th} key-frame of the ^{jth} presenter profile

Equation 1 describes how our method transforms the key-frames of a presenter to the eigenspace. After experimenting with different parameters we have observed that the overlapping problem of eigenfaces can be overcome by introducing a parameter . is based on the adjusted values of the standard deviation and the mean of intensity values of key-frames known to the system (Equation 2).

$$= S + \ _1 / S + \overline{X} + \ _2 - (2)$$

To obtain values for $_1$ and $_2$ we carried out experiments and analyzed results on the recognition levels of the known presenters and unknown presenters. The sample set we used to determine these two constants included presenters with different illumination variations. One such result is shown below in figure 3. The values for S and X are 100.02 and 23.24 respectively. A complete result set is obtained by varying the values of S and X. Only the best result is shown in figure 3. The recognition level can be described as the minimum value for a known face and the maximum value for a unknown face. The maximum recognition level 0.04 is obtained when S=140.02 and X=33.24. To obtain the exact values for the $_1$ and $_2$ we experimented with 20 presenters using five different datasets [1]. By evaluating this resultset we achieved constants $E_1 = 40$ and $E_2 = 10$.

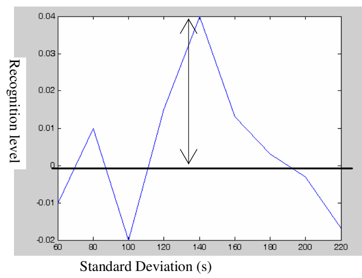

Figure 3: Determine the constants

When the key-frames are normalized by the algorithm (equation 1), we calculate the eigenvectors and eigenvalues for each set of key-frames corresponding to a particular presenter. Once the eigenfaces have been computed, each face can be viewed in the imagespace. Furthermore, good representations of the profiles can be obtained getting the largest eigenvectors available (Figure 4).

Figure 4: A Profile Developed in Face Space

In Figure 4, a presenter profile is developed into the facespace. Then for the threshold construction process, threshold value (V_t) is determined to select the eigenvectors which is most suitable to construct a presenter profile. Algorithm has been developed to calculate V_t by analyzing the behavior of each profile projection to the face space using different number of eigenvectors (equation 4).

$$V_t = ((V_{max} + V_{min}) * n)^{1/5} \quad - (3)$$

V_{max} = Maximum Eigenvalue
V_{min} = Minimum Eigenvalue
V_t = Threshold Value for Eigenvector Selection
n = Number of Eigenvector for a presenter

Using the above algorithm we eliminate the eigenvectors less than V_t when constructing a presenter profile. This algorithm helps to reduce the processing speed of the system and the efficiency of the recognition process. As shown in Figure 5, our algorithms (Equation 1, 2 and 3) can successfully rearrange profiles and overcome the profile overlapping.

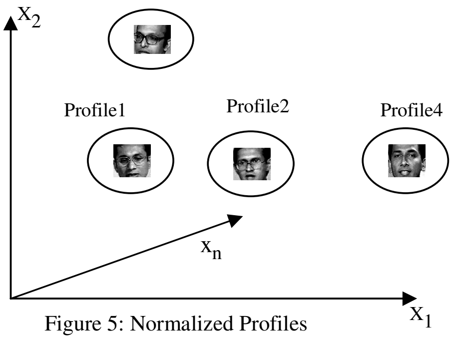

Figure 5: Normalized Profiles

4 CONCLUTION

We conducted experiments on key-frame sets that contained different illumination variations. For all these variations, our algorithm achieved much better performance than the previous method [5, 6]. We tested the algorithm using two different counts of key frames of the same presenter to construct his profile. For the initial testing we have used 4 frames per presenter and for the second testing we have increased the key frames per presenter from 4 to 8. We were able to maintain an 80% recognition rate even when the profile database expanded to 20 (Figure 4). The recognition rate with the previous algorithm was 70%.

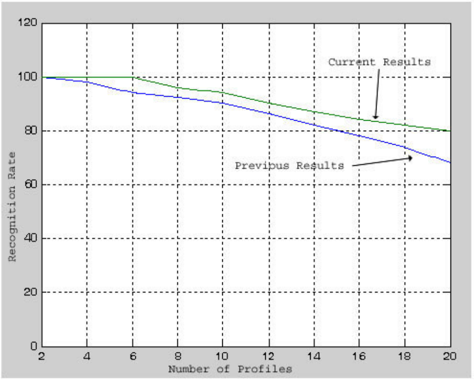

Figure 4: Results

We also have observed that effective normalization of the video key-frames greatly increases the performance of the profile matching system. From the final results obtained we can conclude that the new algorithm proposed in this paper works well under controlled environments and the recognition algorithm took advantage of the environmental constraints to obtain high recognition accuracy.

5. FUTURE WORK

The work that had been done can be expanding in several directions. The algorithm can be improved in order to recognize more complicated video key-frames such as identify presenter in different poses, although our system works well under small variations in orientation. All current person recognition algorithms fail under the vastly varying conditions under which humans need to and are able to identify other people. Next generation person recognition systems will need to recognize people in real-time and in much less constrained situations.

6. ACKNOWLEDGEMENT.

This research is supported by the National Science Foundation (NSF) under the grant No: NSF/2005/SCH/07 and Japan International Cooperation Agency (JICA). The authors also acknowledge each individual appearing in our face database.

References

[1] P. Jonathon and H. Moon and S. A. Rizvi and P. J. Rauss, *The FERET Evaluation Methodology for Face Recognition Algorithms*, IEEE Trans. Pattern Analysis and Machine Intelligence, Volume 22, October 2000, pp. 1090-1104.

[2] Vikrant Kobla, David Doermann, King-Ip (David) Lin and Christos Faloutsos, "*Compressed domain video indexing techniques using DCT and motion vector information in MPEG video*", Storage and Retrieval for Image and Video Databases (SPIE), 1997, pp. 200-211.

[3] L. Lorente and L. Torres, "*Face Recognition of Video Sequences in a MPEG-7 Context Using a Global Eigan Approach*", International Workshop on Very Low Bit-rate Video Coding, Urbana, Illinois, 1998.

[4] A. Pentland, B. Moghaddam, and T. Starner, "*View-based and modular eigenspaces for face recognition*", MIT Media Laboratory Perceptual Computing Section, Technical Report 245, 1994.

[5] S. C. Premaratne, D. D. Karunaratna, G. N. Wikramanayake K. P. Hewagamage and G. K. A. Dias. *Implementation of a Profile Based Video Segmentation System.* International Conference on Information Management in a Knowledge Society, Grand Hyatt Mumbai, Maharashtra, India, 2005, pp 89 – 100.

[6] S. C. Premaratne, D. D. Karunaratna, G. N. Wikramanayake K. P. Hewagamage and G. K. A. Dias. *Profile Based Video Segmentation System to Support E-Learning.* International Information Technology Conference, Proceedings of 6th International Conference, 2004, pp. 74-81.

[7] M. A. Turk and A. P. Pentland, "*Face recognition using eigenfaces*," in *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, Maui, Hawaii, 1991, pp. 586–591.