FEATURE EXTRACTION USING ACTIVE APPEARANCE MODEL ALGORITHM WITH BAYESIAN CLASSIFICATION APPROACH

Mohammad Nuruzzaman, Azham Hussain, Hatim Mohamad Tahir, and Mohamad Amir Abu Seman

School of Computing, University Utara Malaysia, Sintok, Kedah.
nayan@egudang.com; azham.h@uum.edu.my; hatim@uum.edu.my; mohdamir@uum.edu.my

ABSTRACT. Face recognition is one of the most important and rapidly advanced active research areas of computer science. In spite of the large number of developed algorithms, real-world performance of face recognition has been disappointing. This study enhances invariant recognition of human faces and analysis to improve face verification and identification performance using Active Appearance Model (AAM) for feature extraction with Bayesian classification approach. This paper addressed some of these issues to bring face recognition more closely to being useful for real-life applications. It directed towards the illumination-invariant automatic recognition of faces and analysis to improve face verification and identification performance. To compare with other feature extraction at the end of the study, an evaluation has been done with an existing face recognition system using AAM algorithm. The experiments performed on part of the FERET color dataset. The result was satisfied with the acceptance rate more than 96%.

Keywords: Face Recognition, Face Detection, Feature Extraction, AAM Algorithm, Bayesian Classification

INTRODUCTION

Human face is an important sign to identify people. It reveals a great deal of information to a perceiver. From ancient times, hand drawing photographs used to represent human faces. Of course a person can be identified by other means than the face. Voice, body shape, fingerprint, iris pattern, hand writing or even clothing may establish identify in circumstances where facial detail may not be available. Nevertheless, a face is the most distinctive and widely used key to a person's identity. Photograph could be used to identify a person in most of circumstances. However, matching and memorizing from a large number of unfamiliar faces to recognize a specific face is not practical. This, it is crucial to have an automatic face verification system in the point of view of a wide range of commercial applications.

There are several significant research done in this area to prevent the existing algorithm to be effective. Current face recognition methods are able to identify faces from images/videos under controlled environments. However, those facial appearance changes dramatically under variations in illumination, pose, expression, aging *etc.*, are difficult to model algorithmically. In this paper, we proposed an automated and robust system to identify human faces in real-life conditions such as illumination, rotations, expression and occlusions through AAM algorithm with Bayesian classification approach. AAM is a benchmark technique that is significantly more robust and stable to accuracy comparable to human faces. We represent each images as a feature vector in some linear or feature space. Furthermore, fitting stability is quantitatively compared to another popular appearance model; AAM result demonstrates the superior stability (Toews & Arbel, 2007). In this endeavor, efforts are directed towards the automatic face verification of human faces and analysis to improve faces classification performance using AAM with Bayesian classification approach.

LITERATURE REVIEW

Feature Extraction

Facial feature extraction has aroused interest with the increasing development of modeling and digitizing techniques. Akagunduz and Ulusoy (2008) represented facial data using the scale and transform invariant features. The nose, two eye and mouth landmark located according to their position with respect to the nose. The nose point perpendicular from the center of a line connecting both eye centers. From biology (Georghiades, Belhumeur & Kriegman, 2005) the most probable range to find the nose point of a naturally formed nose along the perpendicular bisector line can be calculated from the inter eye distance and pupil radius. By normalizing this most probable nose area locally to get nose edges the nose point is found.

Active Appearance Model (AAM)

Active Appearance Model (AAM) is a generalization of the widely used Active Shape Model (ASM) approach. But uses all of the information in the image region covered by the target object, rather than just near modeled edges. It contains a statistical, photo realistic model of the shape and grey-level appearance of faces which can be generalized to almost any valid example. Matching to an image involves finding model parameters which minimize the difference between the image and a synthesized model example, projected into the image. In order to realize these benefits, the model of object appearance should be as complete as possible to synthesize a very close approximation to any image of the target image. AAM is particularly suited to the task of interpreting faces in images. Faces are highly variable, deformable objects and manifest very different appearances in images depending on pose, lighting, expression and identity of the person. Interpretation of such images requires the ability to understand this variability in order to extract useful information.

RELATED WORKS

There are significant amount of researches on face identification. Their findings, argument, conclusions with original comments, consideration and recommendation have been taking for justification. The most fundamental problem in this area is how to identify face and extract information about the underlying emotional states and variance illumination conditions.

Georghiades, Belhumeur & Kriegman (2005) developed a model shape and local grey-level appearance using ASM to locate flexible objects in new images. To trace linear facial features, estimated corresponding parameters of face model and reproduce facial expression. However, the limitation of their research is it required facial features be highlighted with make-up for successful tracking. Having found the face shape using ASM, the face is warpped into a normalized frame, in which a model of the intensities of the shape-free face is used to interpret the image. Another study extends this work to produce a combined model of shape and grey-level appearance, but again rely on ASM to locate faces in new images. Our proposed approach can be seen as a further extension of this idea using all of the information in the combined appearance model to fit to the image.

Cootes & Taylor (2004) described an approach in which a control feedback loop between computer graphics and computer vision processes is used for a facial image coding system. Different environments in illuminations, poses, face expressions and aging are so great that the correlation between two images of the same person under different environments can be smaller than that between two images of the different persons. Image-based face recognition algorithms usually utilize whole pixel information of images so that even a small local change in illuminations, poses, expressions affects the algorithms significantly and therefore they are not usually robust to illuminations, poses, expressions and aging. On the other hand, Gabor features are known to be more robust to small variations in scaling, rotation, distortion, illumination, poses, and expressions so that they are popularly employed as features for face recognition. Their work is the most similar to ours, but both our goals and implementation differ.

Xiao, Baker, Matthews and Kanade (2004) elaborated AAM models generate by combining a model of shape variation with a model of texture variation. By texture means the pattern of intensities or colours across an image patch. To build an AAM model it requires a training set of annotated images where corresponding points have been marked on each image to build a statistical shape model. Finally, the correlations between shape and texture are learned to generate AAM model.

Bevilacqua, Cariello, Daleno and Mastronardi (2007) used AAM as the underlying basis of their model, sample mean shift and variable length Markov model to learn the relationships between trajectories of facial expressions. Others combined AAM with sound as their framework to produce sequences of a talking head. Both approaches do not deal with the expression classification directly. Chang, Bowyer & Sakar (2003) used a model based on the motion vectors of Bezier volumes. These vectors were then used in conjunction with a multi-level AMM to classify expression from image sequences. They experimented with Bayesian Classification approach.

Zhang, Potnmianos, Senior and Huang (2007) located a few high-level features, namely, eyes, nose, mouth and then 26 low-level features such as the parts of the eyes, nose, and mouth. Eyebrows are located relative to the high-level feature locations, instead of searching for all the facial features directly in the face image. The approximate locations of the high-level features are known from statistics of means and variances (relative to the nose position) gathered on a training database. The discriminate templates are used to score each potential matching image patch for a given feature. A constellation of local patches has been used as the representation. They chose the local template approach in contrast to global identity templates such as those used in Eigen faces systems. A simple Gabor jet model has been used to describe particular patches of the face corresponding to the 29 facial features found. Each patch is represented by a feature vector consisting of 40 complex elements each, representing the filter responses of Gabor filters with five different scales and eight different orientations, centered at the estimated feature location.

ALGORITHM DEVELOPMENT STRATEGY

AAM Construction & Feature Extraction

The AAM is constructed based on a training set of face images, where landmark points are marked at key positions of the face to outline the main features as shown Figure 1.

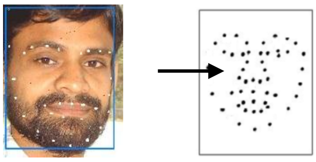


Figure 1. Landmark points of face images using AAM

The shape of a face is represented by a vector consisting of the positions of the landmarks. All shape vectors of faces are normalized into a common coordinate system. The principal component analysis is applied to this set of shape vectors to construct the face shape model as in Eq. (1).

$$s = \bar{s} + P_s b_s \tag{1}$$

Eq. (1) where "s" is a shape vector, "$\bar{s}$" is mean shape, "P_s" is a set of orthogonal modes of shape variation and "b_s" is a set of shape parameters. In order to construct the appearance model, the image is warped to make the control points match the mean shape. Then the warped image region covered by the mean shape is sampled to extract the gray level intensity (texture) information. Similar to the shape model construction, a vector is generated as the representation as

$$g = (I_{1,} \ldots\ldots I_m)^T \tag{2}$$

Eq. (2) where *"I"* denotes the intensity of the sampled pixel in the warped image. AAM is also applied to construct a linear model $g = \bar{g} + P_g b_g$ where $\bar{g}$ is the mean appearance vector, P_g is a set of orthogonal modes of gray-level variation and b_g is a set of gray-level model parameters [11]. Thus, all shape and texture of any face can be summarized by the following formula—

$$b = \begin{bmatrix} \frac{W_s b_s}{b_g} \end{bmatrix} = \begin{bmatrix} \frac{W_s P_s (s - \bar{s})}{P_g{}^T (g - \bar{g})} \end{bmatrix} \tag{3}$$

Eq. (3) Where "W_s" is a diagonal matrix of weights for each shape parameter, allowing for difference vale unit between shape and gray scale model. AAM is applied to "*b*", where $b = Q_c$ and c is the vector for combined model.

AAM Fitting

Given a new image and constructed model, the metric used to measure the match quality between the model.

$$\Delta = |\delta I|^2 \tag{4}$$

Eq. (4) where "δI" is the vector of intensity differences between the given image and the image generated by the model that is tuned by the model parameters. AAM Fitting seeks the optimal set of model parameters that best describes the given image. Xiao, Baker, Matthews & Kanade (2004) observed that displacing each model from the correct value induces a particular pattern in the residuals. In the training phase, AAM learned a linear model that captured the relationship between parameter displacements and the induced residuals.

Searching

The searching images to match with a given image are most critical steps in this study. This step combines with other steps such as facial image capture from various illuminations and face feature extraction. Figure 2 shows that searching flows into three steps. First, Input an initial set of feature points. Second, fit and joint model to the current set of feature points to generate a set of templates. Third, use the shape constrained search method to predict a new set of feature points.

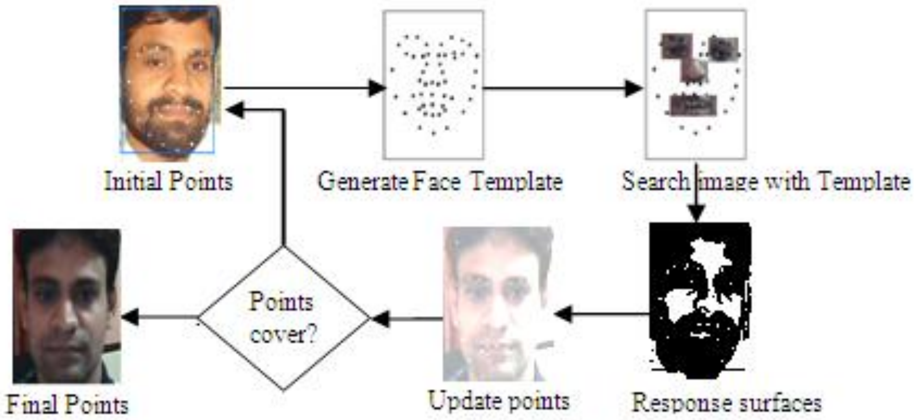


Figure 2. Searching flow using AAM

Matching

For all the training images, the corresponding model parameter vectors are used as the feature vectors. LDA is utilized to construct the subspace for face identity recognition. Given a query image, the AAM fitting is applied to extract the corresponding feature vector. Feature vector is an n-dimensional vector of numerical features that represent some object. The vectors often called feature space or feature subset selection. In order to reduce the dimensionality of the feature space, a number of dimensionality reduction techniques can be employed. This study combines Bayesian classification with feature weighting. It is based on feature vector approach. As a number of distinct features sets are exponential in the number of attributes, it is in general not feasible to perform an exhaustive search. Feature vector use feature weighting assigns a continuous weight to each feature.

The Bayesian classification uses “Bayes” theorem to calculate the most likely classification of an example given the attribute-value distributions of the training images. Given a test instance x_i described by the attributes A_j with values x_{ij} and the possible classifications v_c, the maximum posterior classification v_{map} is—

$$v_{map}(x_i) = \arg\max P(v_c \mid A_1 = x_{i,1 \ldots\ldots\ldots} A_n = x_{i,n}) \quad (5)$$

The attribute-value description $A_j = x_{ij}$ of an instance is usually abbreviated to x_{ij}

$$v_{map}(x_i) = \arg\max P(v_c) * P(A_1 = x_{i,1 \ldots\ldots\ldots} A_n = x_{i,n} \mid v_c)$$

The probability of an instance description, its attribute-value pairs, given its class, is hard to estimate directly from the training data because this would require a very large training set. Assuming conditional independence of the attributes, the probability can be decomposed. The decomposed conditional probabilities $P(x_{ij} \mid v_c)$ are easier to estimate as they are required less training images. To estimate the conditional probabilities of an attribute-value given a class feature vector is used. Conditional probabilities of numeric attributes are approximated by a normal distribution as shown in Eq. (6).

$$p_o(x_i, v) = P(v), p_j[x_i, c] = P(x_{ij} \mid c) \text{ for } j > 0 \qquad (6)$$
$$v_{NB}(x_i) = \arg\max \prod_{jeS} P_j(x_i, v_c)$$

Whenever the conditional independence assumption is justified, the Bayesian classification corresponds to the maximum classification. The recognition is achieved by finding the best match between the query feature vector and stored prototype feature vectors, both of which are projected onto the subspace.

COMPARISON AND ANALYSIS

This study used PCA and ICA for comparison with AAM algorithm. In the case of PCA and ICA, this study created a new GUI with the proposed system and reused the existing implemented method for PCA and ICA algorithms. Before comparison of each algorithm, training has been conducted with 1000 images from FERET database. Those images were taken from different illumination conditions. After training process is completed, each of the given images searches similar images from FERET database. This process is automated and the result is based on the training. Each of the algorithm's results for the same image is different. From the similarity percentage, it can easily be determine the performance of the face recognition algorithm as shown in Figure 3.

PCA describes shape variation allow for localized variations, and capable of a much better segmentation. Changing the weight factor of one independent component does not affect the entire shape which is one of the major advantages of PCA technique. In general, AAM has globally good segmentation results. However, AAM in some cases has great difficulties yielding locally good segmentation results because the model search in PCA-based AAM stops when the major part of an image is segmented correctly.

Additionally, ICA may give a better description of the input data set, because it does not assume that the data is drawn from a Gaussian distribution and also utilizes higher order moments of the distribution. In this way, there is no over simplification of the data set and a better representation of the shape variations occurring in the training set is obtained. In conclusion, the presented AAM showed excellent face identification accuracy as compared to the PCA and ICA. Apart from AAM, it is believed that AAM is highly promising for use in other statistical shape modeling approaches.

In our opinion, PCA-based AAM will substantially improve accuracy compared to the conventional AAM. Because shape variation allow for localized variations. This model is capable of a much better segmentation. Changing the weight factor of one independent component does not affect the entire shape, which is one of the major advantages of PCA compared to ICA. In general, PCA-based AAM yields globally good segmentation results. Table 1 shows an independent, comparative study on the three most popular appearance based face recognition algorithms (PCA, ICA and AAM) in completely equal working conditions and across all implementations. However, it was found that in terms of complexity, AAM has more features than PCA and ICA.

Table 1. Comparison of Popular Appearance-based Algorithms

PCA	ICA	AAM
Simple algorithm	Complex algorithm	Complex algorithm
Low dimensional space	Maximizes the ratio between	Statistical independence

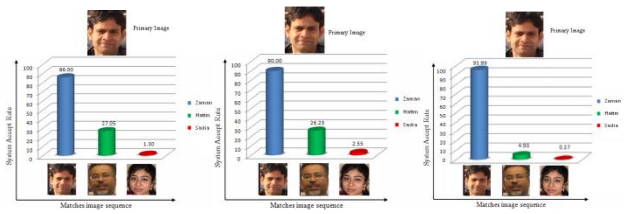
	class scatter to class scatter	
Less sample dimensional space spanned	More sample discrimination	Number of samples is comparable to or less than the dimensionality
Principal Component	Linear Discriminant	Independent Component
Best for data reduction	Best for image vector	Best in terms of time and complexity
Error rate: 1%--8%	Error rate: 5%--10%	Error rate: 1%--5%

EXPERIMENT RESULTS AND DISCUSSION

This study performed a comprehensive comparison of classical and state-of-the-art appearance-based face recognition algorithms applied to visible images. Building on previous work, this study emphasized the role of varying the training and testing sets, as a tool to uncover strengths and weaknesses of algorithms and imaging modalities.

It becomes clear from our analysis that face images are not only a valid biometric but surely a superior one comparable to visible imagery. Face extraction and matching based on FERET database lead to a better recognition performance in face recognition algorithm as confirmed by the training set in this experiment. The results suggested that FERET database makes object recognition more robust to imaging conditions such as clarification. A face space created with a supervised testing method based on the face extraction criterion, trained on intensity and database information leads to a good face recognition performance in reduced representation conditions.

Face extraction and face matching features based on FERET database lead to a better recognition performance in face recognition algorithm as confirmed by the training set in this experiment. The results suggested that FERET database makes object recognition more robust to imaging conditions such as clarification. A face space created with a supervised testing method based on the face extraction criterion, trained on intensity and database information leads to a good face recognition performance in reduced representation conditions.



(a) PCA (86%) (b) ICA (90%) (c) AAM (96%)

Figure 3. Performance Evaluation for Face Recognition

Face extraction and face matching is used along with intensity as an input to the AAM, PCA and ICA techniques. The small sample size of database training scenarios considered major importance in this work. AAM has been successfully applied to face and facial feature localization with high rate probability of detection. In this study, face extractor and face matching using AAM compared with PCA and ICA. The experiment has been implemented for recognizing faces, which first fits an internal facial feature point of the face and then localizing the whole face. The experiments show that the enhancement of the AAM algorithm is more reliable with 96% rather than 90% and 86% for ICA and PCA respectively this means that the face extractor and face matching tested acceptance rate of 94%—96% and the error detection rate was only 1%—5%.

Traditional AAM is not robust against occlusions as shown in Figure. 4. PCA could overcome this problem as a pre-processing step to remove disturbances in the input image and to perform the AAM fitting on the obtained reconstruction. Those disturbances of facial images influence the quality of the fitting process of AAM. Thus, for the pre-processing step, PCA uses facial images which do not exhibit any disturbances. Using PCA as a pre-processing step the reconstructed image is overlaid on the original input image. It can be clearly seen that traditional AAM cannot handle occlusions directly although the fit on the reconstructed image is well defined.

Figure 4. Disturbances of Images from One Person

It is possible to reconstruct an approximation to the original images, since $x^T \approx Cn^T W^{-1} U$. PCA could enhance AAM performance by discarding small trailing Eigen values before whitening and reducing computational complexity by minimizing pair-wise dependencies. ICA set is for Image Composition Analysis. ICA is a highly complex system is designed as a filter that stops incompatible image content from being extreme by individuals that are not accepting. In its essential form, the image filter has a number of mechanisms that are designed to use fundamentals of an image to determine its type. After all the images from different illumination are collected and training process begins. Initially, manually determine a set of point location on the face at key positions for PCA and ICA. This is known as face template. The proposed AAM automatically determining faces, set of point at key positions and outline the faces. Then, AAM used extract facial features landmarks and is known to work relatively well under various illumination variations.

CONCLUSION

We have described the use of AAM with Bayesian classification approach in face recognition. It is fast and automated face verification system that is quite accurate even though pose, illumination and expression variations were present. The performance with FERET database color images was 96% accuracy with 0.01 second processing time per facial image and system has tested in real-time on live faces. This study did not replace background image. It should replace during training to prevent the detection from learning such background. In future research, background should randomly change at any time during test and training. A different aspect of training that can be examined is the way that the training examples are aligned with one another. Appropriate alignment of the features of the training examples is crucial to the performance of the system. Further research and investigations are needed to see how well the face verification performs with training data. The algorithm also needs to study the effect of deformation face detection for error rate due to the sensitivity detection techniques on the quality of AAM. We anticipate that the AAM with Bayesian classification approach will be an important method of locating deformable objects in many applications.

REFERENCES

Toews, M., & Arbel, T. (2007). *A Probabilistic Model to Learn Detect, Localize and Classify Patterns in Arbitrary Images*. McGill University, Montreal, Canada.

Cootes, T. F., Edwards, G., & Taylor, C. J. (2001). Active appearance models, *IEEE Trans*. PAAMIJ, 2001, Vol. 23, No. 6, pp. 681–685.

Edwards, G. J., Cootes, T. F., & Taylor, C. J. (2005). Face Recognition using Active Appearance Models in *Proceeding of European Conference on Computer Vision*, Springer, Vol. 2, pp. 581-695.

Cootes, T. F., & Taylor, C. J. (2004). Statistical models of appearance for computer vision. *Technical report from Imaging Science and Biomedical Engineering*, University of Manchester.

Xiao, J., Baker, S., Matthews, I., & Kanade, T. (2004). Real-time combined 2D+3D active appearance models, *IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 2, pp. 535–542.

Akagunduz, E., & Ulusoy, I. (2008). 3D face representation using scale and transform invariant features, *16th IEEE Signal Processing, Communication and Applications Conference*.

Bevilacqua, V., Cariello, L., Daleno, D., & Mastronardi, G. (2007). Pseudo 2D Hidden Markov Models for Face Recognition Using Neural Network Coefficients, *IEEE*, pp. 107-111.

Boehnen, C., & Russ, T. (2005). A fast multi-modal approach to facial feature detection, Vol.1, pp. 135–142.

Chang, K., Bowyer, K. W., & Sarkar, S. (2003). Comparison and combination of ear and face images in appearance-based biometrics, *IEEE Trans on Pattern analysis and machine intelligence*, Vol. 25, No. 9, pp. 19-20.

Georghiades, A. S., Belhumeur, P. N., & Kriegman, D. J. (2005). Illumination cone models for face recognition under variable lighting and pose, *IEEE Trans. on Pattern Analysis & Machine Intelligence*, 23(6):643–660.

Lanitis, C. J., Taylor, C. J., & Cootes, T. F. (2002). Towards automatic simulation of ageing effects on face images, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 24, No. 4, pp. 442–455.

Lu, X., & Jain, A. (2008). Deformation Modeling for Robust 3D Face Matching, *Transactions on Pattern Analysis and Machine Intelligence*, 2008, Vol. 30, No. 8, pp. 1347-1356.

Nastar, C., Moghaddam, B., & Pentland, A. (2007). Generalized Image Matching: Statistical Learning of Physically-Based Deformations, *Computer Vision and Image Understanding*, 2007, Vol. 65, No. 2, pp. 179-191.

Wang, Y., Chua, C., & Ho, Y. (2002). Facial feature detection and face recognition from 2D and 3D images, *Pattern Recognition Letters*, 2002, Vol. 23, No. 10, pp. 1191–1202.

Zhang, Z., Potnmianos, G., Senior, A. W., & Huang, T. S. (2007). Joint Face and Head Tracking inside Multi-Camera Smart Rooms. *Signal, Image and Video Processing Special Issue on Multi Sensor Object Detection and Tracking*, Vol. 1, No. 2, pp. 163-178, Springer, London.

Nuruzzaman, M., Azham, H., & Hatim, M. T. (2013). Towards Increasing Web Application Development Productivity through Object-Oriented Framework, *International Journal of Future Computer and Communication*, Vol. 2, No. 3, pp. 220-225.

BIOGRAPHY

Mohammad Nuruzzaman obtained his Bachelor degree from University Utara Malaysia in 2010. His experience sparked a lifetime passion for learning, and intends to carry on cutting-edge research in software industry. Currently, he is working as a system analyst at Accenture. Mr. Nuruzzaman is pursuing Master of Science at School of Computing, University Utara Malaysia. His research interests include requirements engineering, software visualization, modeling, and network security. He has several publication and award related to these areas.