Evaluating Active Shape Models for Eye-Shape Classification

Sheethal Bhat and Marios Savvides
ECE Department, Carnegie Mellon University, Pittsburgh PA, USA
{sbhat@, msavvid@ri}.cmu.edu

ABSTRACT

This paper explores the goal of applying Active Shape Models (ASMs) on the eye images to classify eye shapes and identify whether the images belong from left or right irises. In many applications, particular to data collected from single eye capture devices (such as the PIER mobile iris image acquisition device), it is of importance to be able to sort and correct mislabeled collected data. ASMs have traditionally been applied for classification or identification of a wide variety of objects ranging from faces, assembly line objects to biomedical objects such as bone structures (like the spine etc). In this paper we apply and evaluate ASM models to fit on the eye shape to determine if the image belongs to a left or right eye. The approach we employ is based on building 2 ASM models, one for the left eye and one for right eye. The best fit model is chosen as the result. Our preliminary evaluation using vanilla ASM shows that preprocessing techniques like illumination compensation, shape normalization, and accurate Iris detection are key steps required to improve the classification performance.

Index Terms— Image Processing, Image Shape Analysis, Pattern Classification, Pattern Recognition, Correlation

1. INTRODUCTION

Automatic machine recognition of people from still images is an active research area due to the increasing demand for authentication in commercial and law enforcement applications. Many iris biometric applications use single iris acquisition devices (one example is the PIER mobile iris image acquisition device) where it may be easy to mislabel the left/right iris, particularly during enrollment process. Under long enrollment periods and depending on the environment scenario in some applications it maybe easier to make mistakes due to human error in the punching in of data (fatigue, and other external parameters contribute to this). As an example, we could imagine such a system is used in the field to capture iris images for security purposes. Existing systems can be used in law enforcement or commercial situations to enroll iris images, yet need to be named correctly when saving the images in the database. We can imagine that under durations of duress, officers might easily mislabel the left and right eyes and this can lead to errors during recognition particularly if this is limited to a particular eye dataset (left vs right).

In this paper we mainly explore tackling this problem using Active Shape Models (ASMs) on the iris eye images obtained from a public database such as NIST's Iris Challenge Evaluation (ICE) dataset as well some example iris images from an LG4000 commercial system. Applying active shape models on the iris images has its own challenges unlike applying ASMs on other objects that have been previously documented (particularly in the medical field). Non-uniform illumination variations, rotation and scale changes in the testing set, that were not present in the training set, were just some of the challenges faced.

2. BACKGROUND

Active Shape Model (ASM) [1] is a linear statistical modeling method used to interpret shapes in images. It uses Principal Component Analysis (PCA) on shape features to capture and represent the shape point variations. The active shape model fitting process finds the shape which gives the least approximation error using the model parameters.

The general procedure of an ASM algorithm can be outlined as follows: an offline model is trained from a set of training images which have the shape annotated by a human expert. By analyzing the variations of the shape over the training set, a model is built which can capture these variations. To interpret a new image we must find the coefficients which best match a model instance of the image. Once the model fitting process is completed, the extracted coefficients from the fitting process can be used to classify, make measurements, or as an input to further processing. Thus an ASM is a local deformation model constrained by a global variance model. That is, it allows for variability specific to the class of objects on which the model is trained.

Suppose we have $\mathbf{s}$ sets of points $\mathbf{x}_i$ which are aligned into a common co-ordinate frame. These vectors form a distribution in the $2N$ dimensional space in which they lie where N is the number of points. If we can model this distribution, we can generate new shapes similar to those in the original training

set and then examine these new shapes to decide whether they are plausible examples. To simplify the problem, we first wish to reduce the dimensionality of the data from $2N$ to something more manageable. An effective approach is to apply Principal Component Analysis (PCA) to this data. The data forms a cloud of points in the $2N$-dimensional space, though by aligning the points they lie in a $(2N-4)$-dimensional manifold in this space. PCA computes the main axes of this point cloud, allowing one to approximate any of the original points using a model with less than $2N$ parameters. Hence by applying PCA to the data set, we can approximate any of the shapes, x using

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}\,\mathbf{b} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

where $\mathbf{P} = (\mathbf{p}_1|\mathbf{p}_2|\ldots|\mathbf{p}_t)$ is a matrix containing t eigenvectors from the data covariance matrix and $\mathbf{b}$ is a t dimensional coefficient vector given by

$$\mathbf{b} = \mathbf{P}^T(\mathbf{x} - \bar{\mathbf{x}}) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

The vector $\mathbf{b}$ defines the set of parameters of the deformable model. By varying the elements of $\mathbf{b}$ we can vary the shape $\mathbf{x}$ using Eq (1). The variance of the i^{th} parameter, b_i, across the training set is given by ρ_i. By applying limits of $\pm 3\sqrt{\rho_i}$ to the parameter b_i we ensure that the shape generated is similar to those in the original training set.

The model variation corresponding to the i^{th} parameter - b_i, is known as the i^{th} *mode* of the model. The eigenvectors $\mathbf{P}$ defines a rotated co-ordinate frame transformation, aligned with the cloud of original shape vectors. Thus the vector $\mathbf{b}$ defines points in this rotated frame. Further details of the fitting are found in [1],[2] and [3].

3. TRAINING THE ASM

The performance of the ASM depends on how well we train the model by choosing the right design parameters that are: (1) the number of feature points for controlling the shape, (2) the number of training images that span the entire range of possible shapes. In this paper we used NIST's ICE dataset for the ASM and also some sample data from an LG4000 camera. The eye images are labeled in the same way as the face images were labeled in [2]. Figure 2 shows example images of different total number of labeled landmarks. The number of landmarks is selected based on the detail required and the areas of interest. Once the training data is labeled, the shapes are aligned to a reference shape so that only the shape variation of the objects is taken into consideration. We align all the shapes to the mean of the unaligned shapes. The landmarks are selected such that a larger portion of landmarks are in areas that have greater distinguishing characteristics as can be seen in Figure 3.

An example of a descriptive set of labeled points is shown in Figure 3. Different numbers of landmark vertices were selected to find the optimum number of landmarks necessary

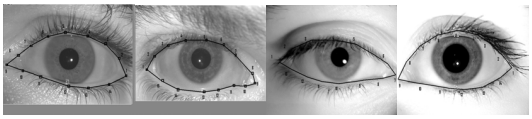

Figure 2: (a) Example Images with different total number of labeled points.

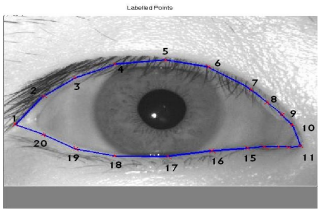

Figure 3: Shape points labeled with more points in the interest of interest.

to describe the iris shape. The optimal number found empirically, was 20 landmark points, with the side of the eye image towards the nose more densely land-marked, as the shape near the tear duct is more discriminative for determining the eye. This region is also prominently different for different people as we can observe in Fig 3. An example of right eye images and left eye images is shown in Fig 4.

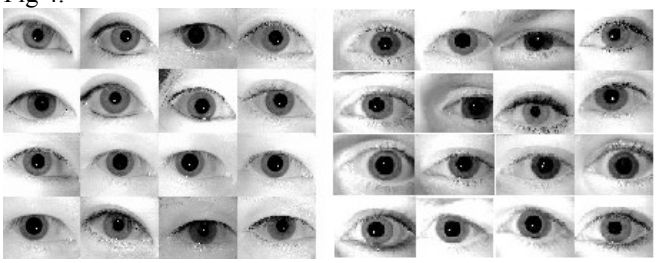

Figure 4: (a)Example Dataset of Right Eye Shapes (b) Example Dataset of Left Eye Shapes

For training, a subset of the ICE Database is taken which consists of 200 ICE images manually divided into 100 left eye images and 100 right eye images. For testing we used the remaining ICE dataset and also tested on some sample images obtained from an LG4000 camera iris acquisition device.

4. SHAPE FITTING AND RECOGNITION

Active shape models are highly sensitive to the initial position of the fitting process. If the initial position is not chosen well, then the model does not converge or fit correctly. The training dataset ideally should have all possible shape variations so that the Active Shape Model

(ASM) can learn the variations effectively. Also, for most iris images available in the public databases, the focus is on capturing clear images of the iris and not necessarily capturing the entire eye shape. Thus we are often left with images that have part of the eye shape missing and typically these should not be used for training as any corruption in the labeling will be learnt by the model.

An eye detector must be first used to localize the eye and then start the ASM fitting process. One possible eye detector that can be trained is the Adaboost Haar detector [5]. Then the ASM is applied to iteratively search for the shape around the initial estimate. The gray level statistics are used to find the desired landmarks and a multi-resolution [1] approach is used to reduce the number of iterations that are needed for a good fit.

5. SHAPE CLASSIFICATION

Eye shape recognition can be regarded as an application dependent technique. If the application requires good fitting with no constraints on real-time execution then it may be possible to employ complex algorithms or combining and fusing several recognition algorithms. For the shape to converge we use the normalized mean gradients of the image around each landmarked point. We can also use scaled gradient, raw intensity, scaled intensity or un-normalized gradient instead of the normalized mean gradient. To find the closest fit, the following square error function is used

$$f(d) = (\mathbf{h}(d)\text{-}y_{imean})^T C_{yi}^{-1} (\mathbf{h}(d)\text{-}y_{imean}) \quad \ldots\ldots\ldots\ldots\ldots\ldots (3)$$

The algorithms are simpler for the purpose of eye shape classification. There are two methods that can be used: once Global ASM coefficients are obtained, a support vector machine (SVM) is used for classification. We see that using a single global SVM model gives us very good results on classification of left vs right eyes.

The second method of applying ASM for classification is to have as many ASM models as there are classes. In our case that would mean 2 different ASM models and we run both the ASMs on a test shape to verify which performs better fitting on the image. Hence for this approach, a quantitative way of measuring the fitting on the test shape must be formulated. In figure 5, an example of applying both ASM models of the Left vs Right eye on a test image of a Left eye is shown. Clearly the left ASM model performs better fitting and can be used to identify the test image as a left iris. In figure 6 we see an example of applying both ASM models on a test image of a Right Eye. We can see that in each case the fitting is better for the ASM model belonging to the correct eye shape, thus individual ASM model fitting approach provides a robust measure of classification.

5.1 Improving classification Performance

In the case of classification of Left vs Right eyes, we observe that SVM approach gives us about 100% accuracy on our dataset. However, to achieve this performance we had to manually initialize the ASM. We also tried to classify images by implementing correlation filters on the test images in small sub-windows around each landmark in order to measure the fitting process. Using this modified ASM fitting approach gave us better results in our dataset where we did not manually initialize the ASM. We built an Optimal Trade off-Synthetic Discriminant Filters (OT-SDF) filters[9] around each landmark around a small region of size 50x100 or 100x100. We found that the window size of 100 x 100 gives us the best performance in this database. An example of the windows on the fitted shape of a test image is as shown in Fig 7.

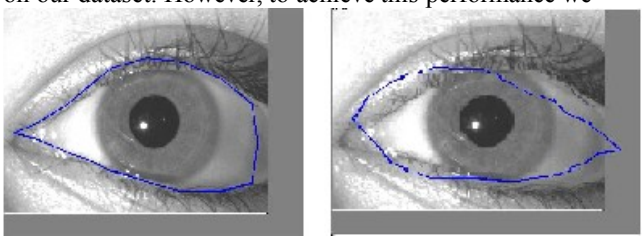

Figure 5: (a) Examples of Applying ASM of class 1 (Left Iris) on test image from class 1 (Left Iris). (b) Examples of Applying ASM of class 2 (Right Iris ASM) on test image from class 1 (Left Iris).

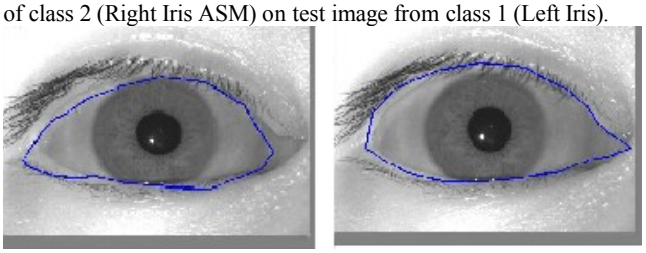

Figure 6: (a) Example of Applying ASM of class 1(Left Iris ASM) on test image from class 2 (Right iris). (b) Examples of Applying ASM of class 2(Right iris)) on test image from class 2 (Right iris).

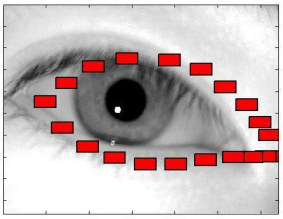

Figure 7: Correlation filter sub-windows.

Once the ASM models of each class are used to find the shape in the test image, OT-SDF correlation filters were used at each landmark iteratively and finally a score of how well that landmark actually matches the shape (based on the correlation output) is obtained. The overall score is

computed as an average of all the landmark fittings and is used for classification based on a simple comparison of which ASM model yields a higher score. In this case classification based on single threshold comparisons may not work under all conditions. Instead we compute a relative likelihood ratio using the correlation likelihood from the Left-Eye ASM model and the Right-Eye ASM model.

The below Table shows the results of applying ASM on ICE database where the model is not manually initialized and on the LG images where the test images are centered and all of the same scale and hence give better results. The results with ICE are not as good as the LG images since the ICE images are not rotation invariant and no manual initialization has been done. Also the training set does not contain all the variations present in the test set.

Iris Images	Test Images	Correctly classified	% classified
Left - ICE	1528	955	62.5%
Right - ICE	1425	964	67.7%
Left - LG	24	22	91%
Right - LG	28	25	89%

Table 1: automatic fitting of vanilla ASM test results on ICE images and LG iris Images

Figures 8 and Figures 9 below show some examples of good convergence and bad convergence respectively which is key to a good classification result. Here class 1 denotes Left Eyes and class 2 denotes Right Eye.

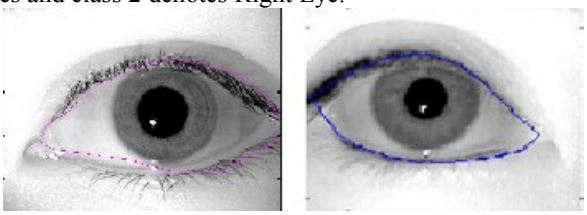

Figure 8. (a) Example of good convergence on class 1. (b) Example of good convergence on class 2.

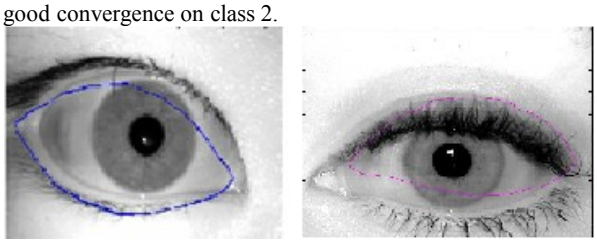

Figure 9. (a) Example of bad convergence on class 2 due to poor initialization. (b) Example of bad convergence over class 1 due to poor initialization.

6. CONCLUSIONS

We have shown application of vanilla Active Shape Models(ASM)[2] for eye-shape analysis to determine left vs right eye/iris, the initial preliminary results reported are encouraging ~65% but still require significant more work to achieve high-identification accuracy. The biggest challenge is initialization which greatly affects the fitting performance. When we tested the vanilla ASM model on classifying left vs right eye shapes, we see that manually giving correct initialization of the ASM model lead to 100% correct classification on a subset of the images. Thus model shape parameters contain the discriminative power for high identification accuracy however that depends on the initialization parameters and the search space capabilities of the ASM to perform a good fit on the eye shape. Using correlation methods, when tested on centered and rotated eye shapes of constant scale also we get better than 89% correct classification so that helped the fitting process by using local correlation to enhance the shape fitting process of the ASM. Our initial results are promising, leaving great room for improvement. Future work will focus in building more robust ASM search space & fitting to improve the resulting model accuracy, as doing so will definitely improve the identification performance accuracy since we have shown that the manually initialized shapes had the discrimination power using SVMs for successful identification of left/right irises on the subset.

7. REFERENCES

[1] T. F. Cootes, C. J. Taylor, D.H. Cooper, and J. Graham, "Active Shape models – Their training and application," In *Proceedings of Computer Vision and Image Understanding*, Vol. 61, pp. 38–59, 1995.

[2] G. Hamarneh, R. Abu-Gharbieh and T. Gustavsson, Active Shape Models Part I: Modeling Shape and Gray Level Variations , Proceedings of the Swedish Symposium on Image Analysis, SSAB 1998

[3] R. Abu-Gharbieh, G. Hamarneh and T. Gustavsson, Active Shape Models Part II: Image Search and Classification , Proceedings of the Swedish Symposium on Image Analysis, SSAB 1998

[4] C. W. Hsu and C. J. Lin, A comparison of methods for multiclass support vector machines, IEEE Transactions on Neural Networks,vol. 13, pp. 415 425, Feburary 2002.

[5] P. Viola and M. Jones, Robust Real Time Object Detection, *IEEE ICCV Workshop Statistical and Computational Theories of Vision*, July 2001.

[6] M. Turk and A. Pentland, Eigenfaces for Recognition, Journal of *Cognitive Neuroscience*, Vol. 3, pp.72-86, 1991.

[7] P.Belhumeur, J. Hespanha, and D. Kriegman, Eigenfaces vs Fisherfaces: Recognition Using Class Specific Linear Projection, *IEEE Trans. PAMI*, Vol.19. No.7, pp.711- 720, 1997.

[8] R. Gross and V. Brajovic, An Image Preprocessing Algorithm for Illumination Invariant Face Recognition, *4th International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA)*, Springer, June, 2003

[9] A. Mahalanobis, B.V.K. Vijaya Kumar, and D.Casasent, Minimum average correlation energy filters, Applied Oprics, Vol. 26, pp. 3633-3630. 1987.