

# Exploring Face Space

Terence Sim

Sheng Zhang

School of Computing, National University of Singapore  
Singapore 117543, Singapore  
{tsim,zhangshe}@comp.nus.edu.sg

## Abstract

*Face recognition is a difficult problem, whether using still images or video. A robust solution is still elusive after 30 years of research. The main reason postulated for this is that two people look more alike than images of the same person under different viewing conditions, i.e. the inter-class variability is smaller than the intra-class variability. In this paper, we propose a way to investigate this, and other, phenomenon more quantitatively. This is done by exploring the space of face images. We first synthesize images under different illumination and pose, and then estimate the probability density function (pdf) for each person. The pdfs are then analyzed for their separability, and for where they overlap. Class regions, regions where the Bayes' classifier would correctly classify each person, are also determined. These class regions are subjected to k-means clustering. By examining cluster boundaries, we can determine lighting and pose conditions that make face recognition difficult. Similarly, the cluster centers tell us the viewing conditions most suited for discriminating between the persons. Our paper makes three key contributions: (1) we show how face space may be modeled and explored; (2) we show that the traditional inter-class/intra-class variability is not a good measure of the separability of two classes, and instead propose the use of the Bhattacharyya distance, and (3) we determine the viewing conditions that are best (or worst) for face recognition.*

**Keywords:** face recognition, face space, pattern recognition

## 1. Introduction

Face recognition, whether using still images or video, is a difficult problem. Even after 30 years of active research, a robust solution remains elusive. The main difficulty is that the appearance of a face changes dramatically when variations in illumination, pose, expression (to name a few) are present. When variations are absent or relatively minor, then existing face recognition systems perform very well

[10]. Changes in illumination and pose are among the most difficult to handle [10, 16], and attempts to find invariant features have largely failed [1]. Anecdotal evidence suggests that face images of two different people look more alike than images of one person under different illumination and pose [4]. In pattern recognition parlance, this means that the *inter-class* variability is smaller than the *intra-class* variability.

To date, little work has been done to study this phenomenon quantitatively. Adini et. al. [1] compared a number of popular face recognition approaches purported to be invariant to illumination and found that none of them were robust against lighting changes. Belhumeur et. al. [2] used Fisherfaces (Fisher Linear Discriminant [5]) to compensate for illumination variation. However, their focus was more on recognition accuracy than on inter- versus intra-class variability. This paper presents what we believe to be the first attempt to study *face space*, defined as the set of all images of faces under different viewing conditions. Our investigation may be considered a generalization of eigenfaces, first used by Kirby and Sirovich [8] to represent faces, and popularized by Turk and Pentland [14] for face recognition. But whereas eigenfaces are meant to give a compact representation, we are more concerned with how faces of different people overlap (or not).

More precisely, we estimate the probability density function (pdf) of each person under illumination and pose variations, analyze where these pdfs overlap, and measure their separability, as well as their inter- and intra-class distances. Our results show that the anecdotal evidence that within-class variability is larger than between-class variability to be only partially true. It depends on which "direction" of variability is being used, and hence is not reliable. We argue that the true measure of class separability must be computed using pdf distance metrics, such as the Bhattacharyya distance or the Kullback-Leibler divergence [5]. Along the way, we also discover what viewing conditions make face recognition easy, and what make it difficult. Our hope is that these explorations provide useful insights for face recognition, and paves the way for better techniques in the future.

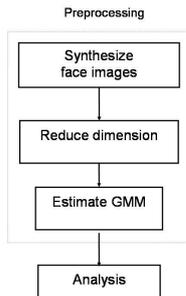


Figure 1: Schematic overview.

Although our present analysis is done using still images, the results are useful for face processing in video as well. If we regard video as a sequence of still images, then a video is simply a trajectory in face space. Where two trajectories intersect, or approach each other closely, recognition is difficult, and where they are far apart, recognition is easy. Thus a simple way to recognize faces in video is to select the frames in which the faces are far apart in face space. This obviates the need to compare every frame, and has the added flexibility that frames greatly separated in time may be compared.

The rest of this paper is organized as follows: Section 2 describes how we perform our analysis. Section 3 shows the results from analyzing the face space of five people. This is done first for frontal-pose, illumination-only images, and second for images where both illumination and pose are varying. Section 4 concludes our paper.

## 2. Method

Our approach is schematically shown in Figure 1. The idea is to model face space with a pdf for each person (the so-called “class-conditional pdf”), and then to analyze these pdfs for where they overlap, and where they do not. This follows the principle of the Bayes’ classifier, used in statistical pattern recognition [5]. The following sections provide more details.

### 2.1. Preprocessing

Images of five people<sup>1</sup> are first synthesized under different illumination and pose. Figures 2 and 3 show the five individuals and a sample of the various lighting conditions and viewpoints generated. We used the 3D laser-scanned face dataset provided by the University of South Florida Human

<sup>1</sup>Our technique is inherently time consuming: image rendering and statistical sampling are computationally intensive. We chose five out of the hundred people in the dataset to get results within a reasonable period of time. We have subsequently run our analyses on ten people, and the results presented here still hold.



Figure 2: The five persons used in our analysis: Persons 0 to 4, from left to right.

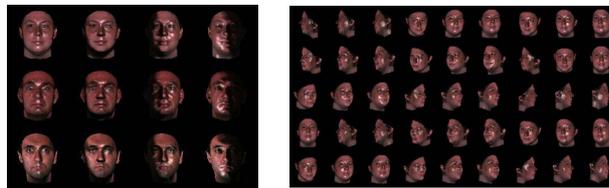


Figure 3: Images under different illumination and pose.

ID project [12]. This dataset, obtained from real human subjects, contains the 3D coordinates of each face, as well as its texture map. Rendering was done using standard computer graphics techniques: under the Phong illumination model, and employing a simple shadow buffer algorithm to compute the shadow regions [6]. For each person, a single light source (with a small amount of ambient light) was moved in front of the face, varying in azimuthal (left-right) angles of  $\pm 90^\circ$  in steps of  $20^\circ$ , and in elevation (up-down) angles of  $\pm 60^\circ$  in steps of  $20^\circ$ . This gives 70 illuminations per pose. Pose was also varied in azimuthal angles of  $\pm 90^\circ$  and elevation angles of  $\pm 60^\circ$ , both in steps of  $10^\circ$ . This gives 247 poses, making a total of  $247 \times 70 = 17290$  images per person. All images are  $120 \times 120$  in size.

It is possible to synthesize images in other ways. For instance, by the methods described in [7, 11]. Using these methods, it will be possible to render many new images from even a single input image, without first requiring a 3D mesh to be available. Our idea is thus not restricted to synthesis from a 3D model, nor constrained by traditional computer graphics techniques.

The next step is to reduce the dimension of the images, to avoid the “curse of dimensionality” [5]. This may be done using standard Principal Components Analysis (PCA) [5], and keeping 100% of the “energy”. That is, choose the smallest  $K$  such that the ratio of the sum of the  $K$  largest eigenvalues to the sum of all eigenvalues is  $> \theta$ . For our illumination-only analysis (see Section 3.1), we use  $\theta = 0.8$ , giving  $K = 8$ . The trade-off is between obtaining a PCA basis that best represents the face images on the one hand, and tractability in estimating the pdf on the other. Images of all persons are then projected into this PCA space.

In this reduced dimension space, we now estimate the pdf for each person. This is done by fitting a Gaussian

Mixture Model (GMM) using the Greedy Expectation Maximization (EM) algorithm [15]. This is a fast algorithm that estimates both the individual Gaussian components (means and covariance matrices) as well as the number of components  $M$ . The GMM pdf is defined as [5]:

$$P(\mathbf{x}) = \sum_{j=1}^M \alpha_j \text{Gauss}(\mathbf{x} | \boldsymbol{\mu}_j, \mathbf{C}_j), \quad (1)$$

$$\text{Gauss}(\mathbf{x} | \boldsymbol{\mu}_j, \mathbf{C}_j) = \frac{1}{(\det 2\pi\mathbf{C}_j)^{1/2}} \times \exp \left[ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_j)^\top \mathbf{C}_j^{-1} (\mathbf{x} - \boldsymbol{\mu}_j) \right],$$

$$\sum_{j=1}^M \alpha_j = 1 \quad (2)$$

## 2.2. Similarity measure

We begin our analysis with a similarity measure. Given two pdfs  $P_i(\mathbf{x}), P_j(\mathbf{x})$  representing two persons, how similar are they? This may be measured using the Bhattacharyya distance or the Kullback-Leibler divergence [5]. We prefer the Bhattacharyya distance because of its symmetry:

$$Dist_B(P_i, P_j) = \int \sqrt{P_i(\mathbf{x})P_j(\mathbf{x})} \, d\mathbf{x} \quad (3)$$

where a value close to 1 means that the pdfs are similar, and a value close to 0 means they are dissimilar.

Obtaining an analytic formula for the Bhattacharyya distance of two GMMs is difficult, so we resort to a Monte Carlo approach [9]. The basic idea is to draw random samples  $\mathbf{y}_1, \dots, \mathbf{y}_R$  from a pdf  $P(\mathbf{y})$ , and approximate the expectation by a sum. That is,  $E[g(\mathbf{y})] = \int g(\mathbf{y})P(\mathbf{y}) \, d\mathbf{y} \approx \frac{1}{R} \sum_{i=1}^R g(\mathbf{y}_i)$ . We re-write Equation (3) to get:

$$Dist_B(P_i, P_j) = \int \sqrt{P_j(\mathbf{x})P_i(\mathbf{x})} \, d\mathbf{x}$$

$$= \int \left( \sqrt{\frac{P_j(\mathbf{x})}{P_i(\mathbf{x})}} \right) P_i(\mathbf{x}) \, d\mathbf{x}$$

$$\approx \frac{1}{R} \sum_{i=1}^R \sqrt{\frac{P_j(\mathbf{x}_i)}{P_i(\mathbf{x}_i)}}$$

where the samples  $\mathbf{x}_i$  are drawn from  $P_i(\mathbf{x})$ . In our experiments, 50,000 samples were drawn for each  $P_i$  to compute the Bhattacharyya distances. This was repeated 5 times, and the average was taken.

## 2.3. Inter-class vs. intra-class distance

Given two pdfs, we can also determine their inter-class and intra-class distances. Conventional methods, such as Fisher Linear Discriminant [5], use this concept to determine class

separability. The inter-class distance measures the variability between two classes (persons), while the intra-class distance measures the variability of all the images of one person. We may define these as follows:

$$Dist_{\text{Inter}}(P_i, P_j) = \|\boldsymbol{\mu}^i - \boldsymbol{\mu}^j\|,$$

$$Dist_{\text{Intra}}(P_i) = \sqrt{\lambda_{\text{max}}^i},$$

where  $\boldsymbol{\mu}^i$  = mean of pdf  $P_i$ ,  
 $\lambda_{\text{max}}^i$  = largest eigenvalue of the covariance matrix  $\mathbf{C}_i$  of  $P_i$  (4)

These definitions are similar to the MIN/MAX ratio of Brunelli et. al. [3]. Intuitively, the inter-class distance may be measured as the distance between representatives (means) of the two classes. As for intra-class variability, we note that there will be different amounts of variation in each dimension of our multidimensional space. The largest variation among all these dimensions is given by the largest eigenvalue of the covariance matrix, and serves as an upper bound on the variability within the class. We define the following ratio to measure the inter-class versus intra-class variability:

$$V_{ij} = \frac{Dist_{\text{Inter}}(P_i, P_j)}{Dist_{\text{Intra}}(P_i) + Dist_{\text{Intra}}(P_j)} \quad (5)$$

Note that if the distributions of the two classes are spherical, then their  $V_{ij}$  ratio measures the distance between the (hyper)sphere centers relative to their radii. Thus if  $V_{ij} > 1$ , then the spheres are well separated; otherwise they intersect. Hence,  $V_{ij}$  gives us another measure of similarity between the two pdfs. In most situations, however, the distributions will not be spherical, so  $V_{ij}$  could be misleading (see our experiment in Section 3.1). In such cases, one should really use the Bhattacharyya distance.

Using Equation (1), it is straightforward to obtain an analytic formula for the mean and covariance matrix in terms of the component means and covariances. From this, we may calculate the largest eigenvalue of  $\mathbf{C}$  for each class, and hence  $V_{ij}$ .

## 2.4. Class regions

In theory, the Bayes' classifier is the optimum classifier, giving the minimum expected misclassification error [5]. Classification is done as follows: let  $\omega_1, \dots, \omega_T$  be  $T$  classes, and their corresponding class-conditional pdfs be  $P(\mathbf{x} | \omega_1), \dots, P(\mathbf{x} | \omega_T)$ . Then given an unknown vector  $\mathbf{x}$ , it is assigned to the maximum a posteriori class:

$$\omega^* = \arg \max_{\omega_i} P(\omega_i | \mathbf{x})$$

$$= \arg \max_{\omega_i} P(\mathbf{x} | \omega_i) P(\omega_i) \quad (6)$$

where  $P(\omega_i)$  are the prior probabilities of the classes. In the absence of other information, the priors are usually assumed to be constant, equal to  $1/T$ . Then Equation (6) becomes the maximum likelihood estimate:  $\omega^* = \arg \max_{\omega_i} P(\mathbf{x}|\omega_i)$ .

For each class  $\omega_i$ , the regions where  $P(\mathbf{x}|\omega_i)$  is maximum (over all other  $P(\mathbf{x}|\omega_j)$ ) is called the *class region* for  $\omega_i$ . Denote this by  $\mathcal{R}_i$ . Note that this region need not be contiguous; it may be a union of disjoint subregions. From the Bayesian perspective,  $\mathcal{R}_i$  is where class  $\omega_i$  is correctly classified, and all other classes  $\omega_j$  are misclassified.

Since we have already computed the class-conditional pdfs from Section 2.1, it is a simple matter to determine the class regions. This may be done using Monte Carlo: draw random samples  $\mathbf{x}_1, \dots, \mathbf{x}_R$  from one pdf  $P(\mathbf{x}|\omega_i)$ , and compute  $P(\mathbf{x}_k|\omega_j)$  for all other classes  $\omega_j$ . Keep the samples for which  $P(\mathbf{x}_k|\omega_i)$  is maximum. These belong to  $\mathcal{R}_i$ . Repeat this procedure for each class:  $i = 1, \dots, T$ . To better represent  $\mathcal{R}_i$ , we can perform  $k$ -means clustering [5], and calculate the cluster mean ( $\mathbf{m}_t$ ) and covariance matrix ( $\mathbf{S}_t$ ) for each cluster  $t$ .

It is possible to visualize these class regions. We can “reverse project” the cluster means back into the high dimensional image space, and display them as faces. Intuitively, these tell us the illumination and pose conditions that best distinguish person  $\omega_i$  from other people. Likewise, we can determine where  $\omega_i$  is easily confused with someone else. These occur at the boundaries of class region  $\mathcal{R}_i$ . That is, where  $P(\mathbf{x}|\omega_i) = P(\mathbf{x}|\omega_j)$  for some other class  $\omega_j$ . Near these boundaries, a small change in  $\mathbf{x}$  may put it into a different class region, resulting in a misclassification.

Although it difficult to determine the exact boundaries (whether analytically or numerically), we can get close to them using the cluster covariances. We note that for each cluster  $t$ , the covariance matrix  $\mathbf{S}_t$  tells us the spread of that cluster from its mean. We can thus compute its principal components: let them be  $\mathbf{v}_l^{(t)}$ , with corresponding eigenvalues  $\beta_l^{(t)}$ . Then the point  $\mathbf{m}_t + 2(\beta_l^{(t)})^{1/2}\mathbf{v}_l^{(t)}$  is “2 standard deviations” from the mean  $\mathbf{m}_t$ , in the direction of  $\mathbf{v}_l^{(t)}$ . This will be closer to the boundary. We can then reverse project such a point and display it as an image.

### 3. Results

#### 3.1. Illumination-only experiments

We first ran our analysis on illumination-only images. That is, for 5 people, we synthesized 555 images per person under frontal pose, with illumination varying in azimuthal angles of  $\pm 90^\circ$  in steps of  $5^\circ$ , and in elevation angles of  $\pm 70^\circ$  in steps of  $10^\circ$ . All these images were then projected using PCA down to 8 dimensional space, and the GMM was estimated for each person.

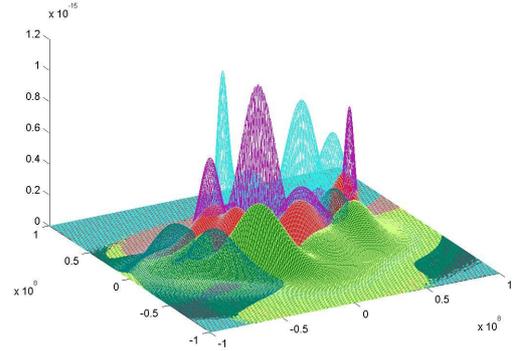


Figure 5: All the 5 GMM pdfs in different colors. This 2D slice shows the first two dimensions of the 8D space.

	P0	P1	P2	P3	P4
P0	1	1.6E-6	7.4E-6	1.3E-6	4.0E-7
P1		1	1.9E-7	6.5E-7	2.7E-5
P2			1	2.0E-5	6.3E-6
P3				1	1.9E-4
P4					1

Table 1: Pair-wise Bhattacharyya distances.

Figure 4 shows various 2D views of the pdfs for Persons 0 and 1. These provide a glimpse into the 8D space. The 2 GMM pdfs do not seem to overlap in regions of high probability, and their Bhattacharyya distance of  $1.6 \times 10^{-6}$  confirms this. The third row of Table 4 shows the number of GMM components discovered by the Greedy EM algorithm for each class. Table 1 shows the pair-wise Bhattacharyya distances for the 5 people. Since the table is symmetric, we display only half the entries. It is clear that all the pdfs are well-separated: the distances are all close to 0, with the largest entry being less than  $2 \times 10^{-4}$ . This means that a Bayes’ classifier should perform very well.

Figure 5 displays the pdfs of all 5 people (in different colors). This is a 2D slice of the first two dimensions. More overlapping regions can be seen.

Table 2 shows the symmetric  $V_{ij}$  ratios for all pairs of

	P0	P1	P2	P3	P4
P0	0	0.288	0.332	0.361	0.580
P1		0	0.324	0.159	0.336
P2			0	0.419	0.610
P3				0	0.237
P4					0

Table 2: Pair-wise symmetric  $V_{ij}$  ratios.

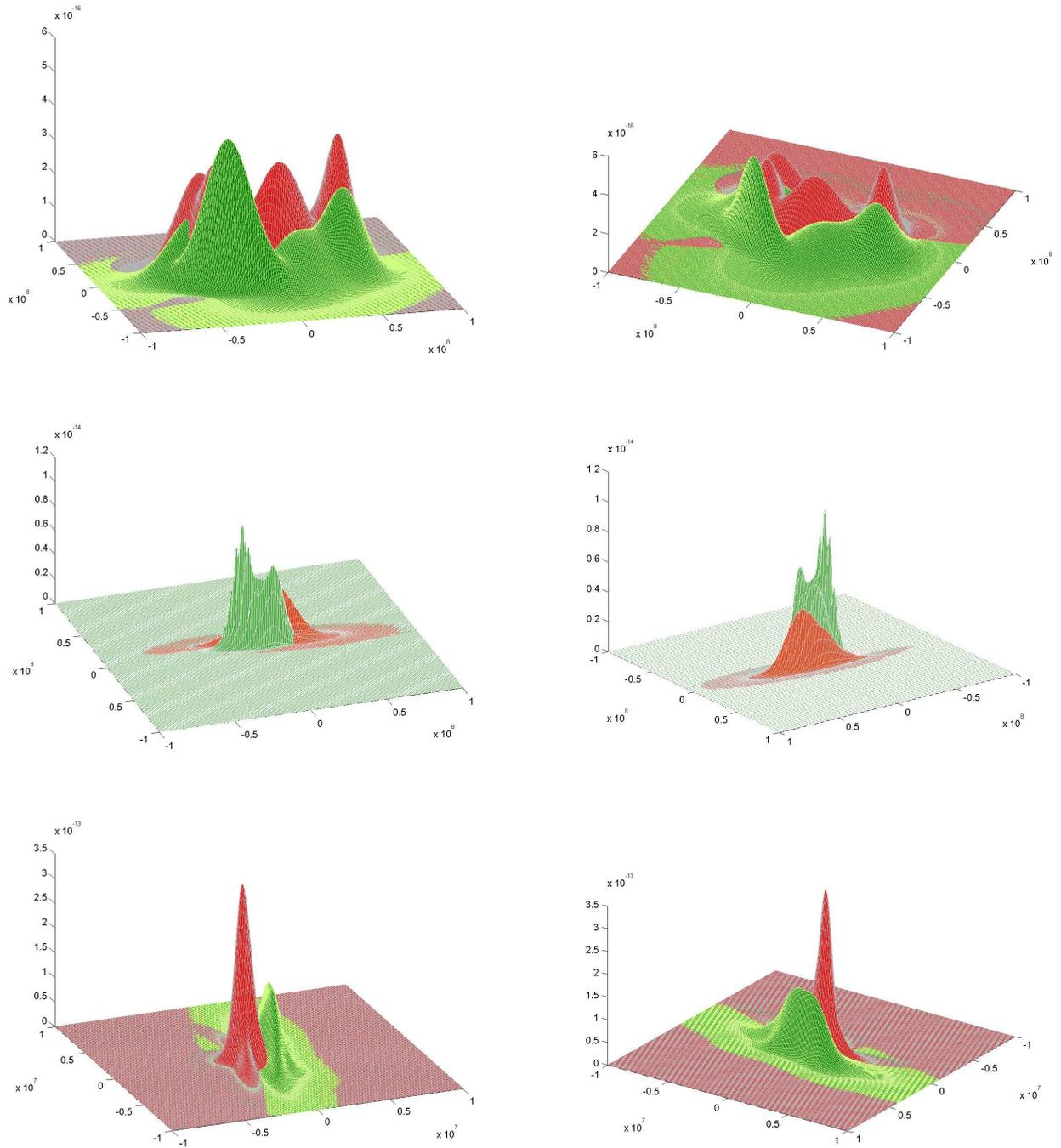


Figure 4: Different 2D views of the pdfs of Persons 0 (red) and 1 (green). (Top row) The first two dimensions of the 8D space. (Middle row) The third and fourth dimensions. (Bottom row) The last two dimensions.

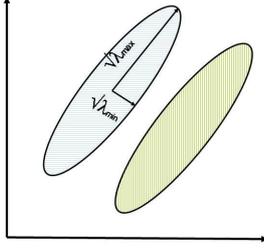


Figure 6: Skinny pdfs in 2D are well-separated, but their  $V_{ij}$  ratio is less than 1.

pdfs. Since all values are less than 1, it means that the inter-class distances are indeed smaller than the intra-class distances. This seems to confirm the anecdotal evidence that other researchers have observed. However, this table needs to be interpreted in light of the Bhattacharyya distances of Table 1. The  $V_{ij}$  ratios suggest that the pdfs overlap, while the Bhattacharyya distances indicate that they are well separated. How do we reconcile these contradictory results?

One possibility is that the pdfs are rather “skinny” in the 8D space. A simple example will serve to illustrate this. Consider Figure 6, which shows the contours of 2 skinny pdfs in 2D space. It is clear that they are well-separated, but their  $V_{ij}$  ratio is less than 1 because of their elongated shapes. This is because the  $V_{ij}$  ratio uses the largest variance in the denominator (Equations (5) and (4)), which is an over-estimate of the true shape of the pdfs. One can consider a slightly different ratio, the  $Z_{ij}$  ratio, that uses the smallest variance instead:

$$Z_{ij} = \frac{\|\mu^i - \mu^j\|}{\sqrt{\lambda_{\min}^i} + \sqrt{\lambda_{\min}^j}} \quad (7)$$

Table 3 shows these ratios for the 5 pdfs. It is clear that the  $Z_{ij}$  ratios are all greater than 1, indicating that the pdfs are separated. Thus inter-class and intra-class distances have to take into account the “direction” of the variability. In some directions the variability is small, in others, it is large. Hence, the Bhattacharyya distance is a better measure of separability, and should be used in place of inter- and intra-class distances. Previous attempts at measuring class separability for face recognition could not use the Bhattacharyya distance because the class-conditional pdfs were unknown. This is due to the small number of training images, which are insufficient to estimate the pdfs. Hence, researchers tend resort to inter-class/intra-class distances, which we have shown to be unreliable.

Let’s now look at the class regions. Recall from Section 2.4 that a class region  $\mathcal{R}_i$  is where class  $\omega_i$  is correctly classified by the Bayes’ classifier. Recall also that we performed  $k$ -means clustering on each class region. Our ex-

	P0	P1	P2	P3	P4
P0	0	41.59	47.47	47.93	72.21
P1		0	45.98	20.90	41.44
P2			0	55.14	75.98
P3				0	26.53
P4					0

Table 3: Pair-wise symmetric  $Z_{ij}$  ratios.

Region	$\mathcal{R}_0$	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$	$\mathcal{R}_4$
# clusters	10	10	8	15	12
# GMM components	5	5	4	8	6

Table 4: Class regions, their clusters, and number of GMM components.

periments discovered different numbers of clusters for each class region, as shown in Table 4. This indicates that disjoint clusters (subregions) are indeed present. Note also that the number of clusters is roughly twice the number of GMM components.

Figure 7 shows a number of faces “reverse projected” and rendered from the means and “2 std. dev.” points of some of these clusters. It is obvious that the means (top row) are more recognizable than the “2 std. dev.” points (bottom row), which are closer to the boundaries of the class region. The illumination angles of these images are also shown. These are estimated by interpolating the known illuminations of nearby points. The angles are  $(\alpha^\circ, \beta^\circ)$ , denoting the azimuth and elevation respectively. Typically, the boundary images are under extreme illumination conditions, where most of the face is shadowed and very little detail is visible. This confirms the intuition that a well-lit face is easier to recognize than a poorly-lit one.

### 3.2. Illumination + pose experiments

We also performed our analysis when both illumination and pose were varied. The first problem we encountered was that PCA could no longer be used to reduce the dimension of face space. Keeping 90% of the “energy” required more than 200 dimensions, far too high for estimating a GMM. On the other hand, using only 8 dimensions meant that only 56% of the energy was retained. This resulted in a PCA space that poorly represented all the variations.

We thus resorted to a non-linear dimension reduction technique called Isomap [13]. This was able to aggressively reduce the dimension of face space, while keeping the “residual variance” (which may be regarded as the opposite of “energy”) small. Indeed, Isomap gave us a 6D space at a cost of 0.1 in residual variance. The large number of images (86450 for 5 persons) made computational time prohibitive. We therefore analyzed only two persons: Person 0 and 1. Their separability measures are shown in

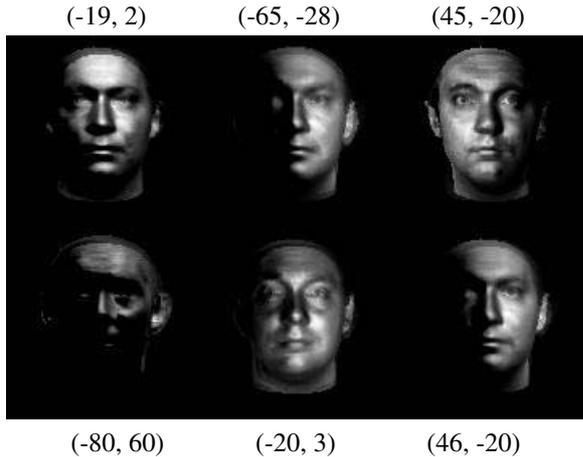


Figure 7: (Top row) Cluster means from some clusters of Persons 0, 1, 2. (Bottom row) “2 std. dev.” images of the same clusters, which are closer to the boundaries. These images tend to be under extreme illumination, making recognition difficult. For Person 1 (middle column), the boundary image is in fact a different person (Person 0), indicating that we have crossed the cluster boundary. The illumination conditions are also shown (see text for explanation).

$Dist_B(0, 1)$	$V_{01}$	$Z_{01}$
0.69	0.14	1.59

Table 5: Separability measures under pose and illumination variation.

Table 5, and a visualization of the first two dimensions of their pdfs is shown in Figure 8.

However, visualization of the cluster means and boundary points is difficult in this case. This is because unlike PCA, Isomap does not provide a way to reconstruct a high dimensional point from its lower dimensional projection. Isomap only finds a lower dimensional structure that is equivalent (in the sense of preserving geodesic distances) to its high dimensional counterpart.

Nevertheless, we may draw the following conclusions:

1. There is significant overlap in the class-conditional pdfs, as measured by a Bhattacharyya distance of 0.69. This is much worse than when pose was fixed and illumination was allowed to vary (c.f. Table 1, which gives a Bhattacharyya distance of  $1.6 \times 10^{-6}$ ).<sup>2</sup>
2. The  $V_{ij}$  and  $Z_{ij}$  ratios are smaller and larger than 1, respectively. This indicates again that such inter-class/intra-class distances are not reliable measures of

<sup>2</sup>Strictly speaking, this comparison is unfair, because the dimensions were reduced using different reduction methods: Isomap versus PCA.

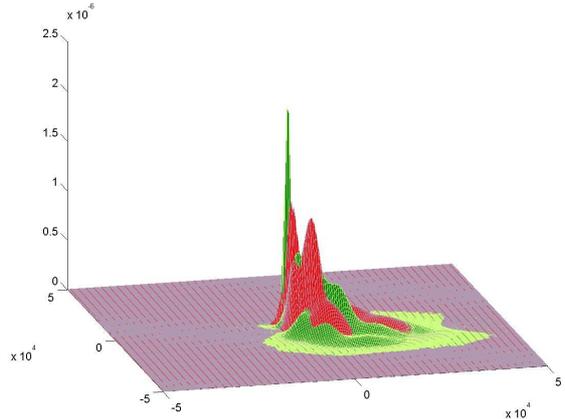


Figure 8: A 2D slice of the 2 pdfs (red and green) under pose and illumination variation. There is significant overlap, which is confirmed by a large Bhattacharyya distance.

separability.

## 4. Concluding Remarks

We conclude with a summary of the main ideas:

1. To better understand the challenges in face recognition, we proposed to investigate *face space*: the set of all appearance variation of faces. In this paper, we generated variations in both pose and illumination.
2. We showed how face space may be explored using Gaussian Mixture Models to represent the class-conditional pdfs.
3. We measured separability of the pdfs in two ways: using Bhattacharyya distance, and the inter-class/intra-class ratio. From our results, it is clear that the Bhattacharyya distance gives a more dependable measure of separability. The inter-class/intra-class ratio depends on the direction of variability, and is therefore not as reliable.
4. We showed how the class regions may be determined from the pdfs. These regions are where one person is correctly classified (by the Bayes’ classifier), while other people are misclassified. Each region is in turn a collection of disjoint clusters (subregions).
5. The class regions also reveal the viewing conditions that best discriminates each person from the rest. We “reverse project” a few cluster means to visualize this.

6. By considering points that are “2 standard deviations” away from the cluster means, we get an approximation to the cluster boundaries. These images tend to be under extreme viewing conditions, making recognition difficult. In some cases, we would have crossed the cluster boundary, into another person’s cluster.

Our exploration of face space is just beginning. There are a number of interesting directions we intend to pursue in the near future:

1. We intend to investigate other observations that researchers have reported. For instance, in [10], it was noted that males are easier to recognize than females, and that face recognition performance decreases as the logarithm of the number of classes. By constructing face space pdfs for males vs. females, and by measuring the size of the class regions as more people are added, we can confirm if these observations are true.
2. Besides illumination and pose, we also intend to synthesize different facial expressions. This will add a new kind of variability to face space. We can also synthesize beards, moustaches, and eye-glasses, and investigate these effects.
3. We can further exploit the realism of computer animation to synthesize facial motion. This will allow us to model the temporal aspect of video, treating video as more than a collection of static images.
4. Face space is a high dimensional space. Is there a lower dimensional (feature) space that can more compactly represent face space? What is the structure of the class-conditional pdf in this space? Ideally, the feature space should still maintain the separability of the pdfs. We intend to explore various techniques to discover such a feature space.

## Acknowledgments

Funding for this work was provided by Research Grant #R-252-000-146-112, at the National University of Singapore.

## References

- [1] Y. Adini, Y. Moses, and S. Ullman. Face Recognition: The Problem of Compensating for Changes in Illumination Direction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):721–732, 1997.
- [2] P. Belhumeur, J. Hespanha, and D. Kriegman. Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7), 1997.
- [3] R. Brunelli and T. Poggio. Face recognition: Features versus templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(10), 1993.
- [4] J. Daugman. Face and Gesture Recognition: Overview. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):675–676, 1997.
- [5] R. Duda, P. Hart, and D. Stork. *Pattern Classification, 2nd edition*. John Wiley and Sons, 2000.
- [6] J. Foley, A. van Dam, S. Feiner, and J. Hughes. *Computer Graphics: Principles and Practice*. Addison Wesley, 1993.
- [7] A.S. Georghiadis, P.N. Belhumeur, and D.J. Kriegman. From Few to Many: Generative Models for Recognition Under Variable Pose and Illumination. In *IEEE International Conference on Automatic Face and Gesture Recognition*, 2000.
- [8] M. Kirby and L. Sirovich. Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12(1):103–108, 1990.
- [9] D. MacKay. *Introduction to Monte Carlo Methods, in Learning in Graphical Models*. MIT Press, 1999.
- [10] P. Phillips, P. Grother, R. Micheals, and D. Blackburn. Facial Recognition Vendor Test 2002. <http://www.frvt.org>.
- [11] T. Sim and T. Kanade. Combining Models and Exemplars for Face Recognition: An Illuminating Example. In *Workshop on Models versus Exemplars in Computer Vision, Computer Vision Pattern Recognition Conference*, 2001.
- [12] S. Sudeep. Human ID Project at University of South Florida, 2001. <http://marathon.csee.usf.edu/HumanID/index.html>.
- [13] J. B. Tenenbaum, V. de Silva, and J. C. Langford. A Global Geometric Framework for Nonlinear Dimensionality Reduction. *Science*, 290:2319–2323, 2000.
- [14] M. Turk and A. Pentland. Eigenfaces for Recognition. *Journal of Cognitive Neuroscience*, 3(1), 1991.
- [15] N. Vlassis and A. Likas. A Greedy EM Algorithm for Gaussian Mixture Learning. *Neural Processing Letters*, 15(1):77–87, February 2002.
- [16] W. Zhao and R. Chellappa and A. Rosenfeld and P. Phillips. Face recognition: A literature survey. Technical Report CAR-TR-948, Center for Automation Research, University of Maryland, 2000.