Accurate and Efficient Face Recognition from Video

Ognjen Arandjelović
http://mi.eng.cam.ac.uk/~oa214

Trinity College
University of Cambridge
CB2 1TQ, UK

Abstract

As a problem of high practical appeal but outstanding challenges, computer-based face recognition remains a topic of extensive research attention. In this paper we are specifically interested in the task of identifying a person from multiple training and query images. Thus, a novel method is proposed which advances the state-of-the-art in set-based face recognition. Our method is based on a previously described invariant in the form of *generic shape-illumination* effects. The contributions include: (i) an analysis of computational demands of the original method and a demonstration of its practical limitations, (ii) a novel representation of personal appearance in the form of linked mixture models in image and pose-signature spaces, and (iii) an efficient (in terms of storage needs and matching time) manifold re-illumination algorithm based on the aforementioned representation. An evaluation and comparison of the proposed method with the original *generic shape-illumination* algorithm shows that comparably high recognition rates are achieved on a large data set (1.5% error on 700 face sets containing 100 individuals and extreme illumination variation) with a dramatic improvement in matching speed (over 700 times for sets containing 1600 faces) and storage requirements (independent of the number of training images).

1 Introduction

Computer-based face recognition continues to be a problem area of active research and outstanding practical challenges [4]. Traditionally, the focus of research has been on achieving invariance to a number of confounding factors, the most pervasive of which are variable illumination and pose, in recognition from a single image. Recently, an increasingly popular trend in the field has been to use multiple image input [5, 8]. These may be acquired as a video sequence or collected over time, for example after each successful authentication.

The appeal of using more than a single image for recognition is rooted in greater information content available, usually in the form of multiple poses [6], which allows for the creation of appearance models of higher accuracy and consequently more robust matching. However, new research difficulties are introduced as well. Of most fundamental nature is the question of how multiple appearances of a person can be used optimally to form a unified and coherent representation of the person's appearance. On the practical side, the dramatic increase in the amount of collectable data poses challenges of efficiency: it is impractical (and, in principle unnecessary) to retain all available data and burdensome to explicitly use

© 2010. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.
BMVC 2010 doi:10.5244/C.24.64

all of it every time a match is required. Consequently, the scope of research challenges is broadened to invariant recognition with additional efficiency requirements.

Generic Shape-Illumination. In this paper we introduce a novel method for face recognition from image sets, based on the concept of generic illumination-shape invariant, first proposed by Arandjelović & Cipolla [1]. Our work is motivated by the outstandingly successful recognition performance of their algorithm on the one hand and, as we show here, its efficiency shortcomings on the other. The approach we introduce inherits the high discriminability of the underlying framework, while accomplishing a large decrease in the associated storage and computational demands.

Paper Organization. In Section 2 we review the original algorithm, highlighting (i) the elements which are key to its successful recognition performance and (ii) those which create the greatest computational bottleneck. An analysis of the method's efficiency is presented in this section as well. This is followed by Section 3 with introduces the main contribution in the form of a novel recognition algorithm, and Section 4 in which the proposed algorithm is analyzed empirically. Section 5 concludes the paper with a summary.

2 Review and Limitations of the Original Algorithm

The algorithm proposed by Arandjelović & Cipolla in [1] centres around learning the effects of illumination and shape variation on the appearance of faces. This is achieved by noting that under a very general illumination model, the ratio of the spatially corresponding image intensities in two pose-matched images of a face is independent of albedo. Specifically, the image intensity i corresponding to the projection $p(\mathbf{x})$ of a point on the face $\mathbf{x}$ is assumed to be a linear function of its albedo $a(\mathbf{x})$ and ϕ, a function of illumination and face shape:

$$i(p(\mathbf{x})) = a(\mathbf{x}) \cdot \phi(\mathbf{x};\theta). \tag{1}$$

Then, then ratio of intensities in pose-matched images i_1 and i_2 is:

$$\log \frac{i_1(p(\mathbf{x}))}{i_2(p(\mathbf{x}))} = \log \frac{a(\mathbf{x}) \cdot \phi(\mathbf{x};\theta_1)}{a(\mathbf{x}) \cdot \phi(\mathbf{x};\theta_2)} = \log \frac{\phi(\mathbf{x};\theta_1)}{\phi(\mathbf{x};\theta_2)}, \tag{2}$$

which is clearly free of albedo dependency and, assuming matched pose, a function of illumination and shape only. Thus, recognition can be performed by reversing the argument and quantifying the agreement of the computed ratio of two images with the expected variation in $\rho = \log(i_1/i_2)$ across different individuals and illuminations. Arandjelović & Cipolla show that this can be achieved by learning the corresponding probability density $\mathscr{P}(\rho)$ [1].

To implement the aforementioned approach, it is necessary to ensure pose alignment of face images which are being compared. Enforcing this at the time of acquisition is overly constraining in most practical applications and a synthetic approach is proposed instead. This algorithm for achieving this plays a pivotal role in the original method and is termed *image set re-illumination* by its authors.

2.1 Face Image Set Re-illumination

To facilitate automatic off-line learning of shape-illumination effects as well as the application of the learnt model in an unconstrained environment, two sets of face images need to be

aligned with respect to pose. Let the two sets be I and J:

$$I = \{i_1, \ldots, i_n\} \qquad J = \{j_1, \ldots, j_m\}. \tag{3}$$

Without loss of generality, the re-illumination algorithm described in [1] computes a set of synthetic appearances $I^* = \{i_1, \ldots, i_n\}$ which retain the original pose of I but match the illumination of some members of J (the illumination may vary within each set). Synthetic faces are computed as linear combinations of faces in J that best match those in I pose-wise.

Pose matching between members of I and J is computed by mapping both sets into the space of pseudo-illumination invariant *pose signatures* – which are distance transformed binary edge images – and formulating the solution to the correspondence as an optimization problem whereby local pose signature agreement is weighed against a global smoothness constraint. Pose signature s of $i \in I$ is then re-projected into the original image space as a linear combination of the neighbouring (pose-wise) appearances of J, as illustrated in Figure 1. For an in-depth description of the algorithm, the reader is referred to the original publication.

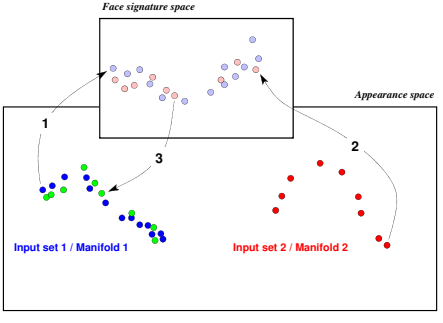


Figure 1: Conceptual illustration of the original set re-illumination algorithm.

Re-illumination Overhead. The optimization problem used to find set-to-set pose correspondences – a task of pivotal importance in the original recognition algorithm – performs matching by evaluating geodesic distances between all appearance images. This makes it both computationally expensive (as analyzed in detail in the next section) and imposes the requirement of having the original data available each time a novel sequence is matched.

2.2 Efficiency Analysis

Arandjelović & Cipolla's approach to recognition consists of two complementary algorithms. The first considers the problem of one-time off-line learning of the shape-illumination effects for human faces, while the second one concerns the application of the learnt model in assigning the identity to a novel face image set. It is the performance of the latter which is critical in practice so in this section we focus solely on its analysis.

The application of the shape-illumination invariant consists of the following steps:

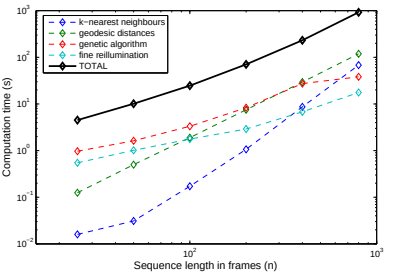
1: k-nearest neighbour computation for each face image,

2: geodesic distance estimation between all pairs of images,

3: pose correspondence optimization using a genetic algorithm,

Algorithm step	Complexity
k-neighbourhood	$\theta(n^2 \log n)$
geodesic distances	$\theta(n^3)$
genetic algorithm	$\theta(n_{gen}\, n_{chr}\, n\, k)$
set re-illumination	$O(n\, k^3)$
robust likelihood	$\theta(n \log n)$

(a)



(b)

Figure 2: (a) Asymptotic complexity of different stages of Arandjelović & Cipolla's face set matching, and (b) the actual (measured) execution time of a Matlab implementation of the algorithm as a function of the number of images in matched sets.

4: re-illumination based on pose correspondence, and

5: robust computation of likelihood of the same identity.

We use the following notation: n is the number of face images in a set, k the number of neighbourhood faces used for re-illumination, n_{gen} the maximal number of generations in the genetic algorithm iteration, n_{chr} the number of chromosomes in each generation and n_{comp} the number of Gaussian components in the mixture capturing the distribution of shape-illumination effects $\mathscr{P}(\rho)$.

Asymptotic Complexity. In *step 1*, to determine the exact k-nearest neighbourhood of each face in a set, the distance to all other faces must be computed – which requires exactly n comparisons – and the result ordered to find the smallest k, which has the average computational load proportional to $n \log n$. Thus, the entire process is $\theta(n^2 \log n)$.

In *step 2*, the estimation of geodesic distances involves the initialization of elementary within all k-neighbourhoods, which is $\theta(n\, k)$, and an application of Floyd's algorithm [2], which is $\theta(n^3)$. As $k \ll n$, the complexity of *step 2* is the same as that of Floyd's algorithm.

In a generation of the genetic algorithm applied in *step 3*, the computation of the similarity between pairs of matching pose-signatures given for every chromosome is $\theta(n)$. The look-up of geodesic distances in all k-neighbourhoods (required for imposing the smoothness constraint) is $\theta(n\, k)$. The total complexity of *step 3* is thus $\theta(n_{gen}\, n_{chr}\, n\, k)$.

Step 4 refines re-illumination results using the pose matching pairs estimated by the genetic algorithm and their k-neighbourhoods. It consists of a single $k \times k$ matrix inversion for each of n images in a set, giving the total complexity of $O(n\, k^3)$.

Finally, in *step 5*, the likelihoods corresponding to all face images are computed in $\theta(n_{comp}\, n)$, which are then ordered in further average $\theta(n \log n)$ time. In principle, n_{comp} is a constant determined by the nature of illumination-shape effects (or, in practice, by fitting an optimal mixture to the sample from the corresponding distribution), so the complexity of *step 5* simplifies to $\theta(n \log n)$.

Treating everything but n as a constant, the overall asymptotic complexity of the algorithm is $O(n^3)$. A summary is presented in Figure 2 (a).

Empirical Performance. We next profiled an implementation of the algorithm written in Matlab on an Intel Pentium 4 PC, with a 3.2GHz CPU and 2GB RAM. In all experiments

only n, the number of faces per set, was varied. Sets of sizes 25, 50, 100, 200, 400 and 800 were used. Mean computation times for different stages of the matching algorithm (estimated from 100 executions of independently drawn identity-illumination combinations for the sets matched) are shown in Figure 2 (b). In this range of n, the measured asymptote slopes were typically lower than predicted, which was especially noticeable for the most demanding computation of geodesic distances. The most likely reason for this phenomenon is the presence of large proportionality constants, associated with Matlab's *for*-loops and data allocation routines.

3 Efficient Application of Shape-Illumination Invariant

In the previous section we have seen that a major limitation of the original algorithm of Arandjelović & Cipolla stems from the need to retain and use all of the training data whenever matching is performed. This leads us to seek a representation which is compact (specifically, one that does not asymptotically grow with n, i.e. has a $O(1)$ storage requirement) yet suitable for robust pose alignment required to extract pure shape-illumination effects.

3.1 Learning a Personal Appearance Model

We show that accurate re-illumination of face sets can be achieved by maintaining two *linked* mixture models for each training image set. The focus is placed on accurately capturing the distribution of each person's pose signatures. This distribution is then related to the image space by inferring the distortion of the local, linear approximations to the appearance manifold when mapped into the pose signature space. Given a set of face appearances $I = \{i_1, \ldots, i_n\}$ (where $i_q \in \mathbb{R}^D$), the aforementioned signature and appearance space mixtures are computed through the following sequence of steps (summarized in Figure 3):

1: Appearance images in I are mapped into the signature space, by detecting Canny edges and distance-transforming the resulting binary images, in the same manner as in the original algorithm:

$$I \xrightarrow{\text{pose sig.}} S(I) : \; S(I) = \{ \, s_q \mid s_q = S(i_q) \, \}. \tag{4}$$

2: The computed set of pose signatures is then used to estimate an approximation to the underlying probability density function, in the form of a Q-component mixture of Probabilistic PCA [7]:

$$\hat{p}(s) = \sum_{q=1}^{Q} \left[\alpha_q \cdot \hat{G}(s \, ; \, \hat{\mu}_q, \, \hat{\Sigma}_q) \right], \tag{5}$$

where $\hat{G}(s \, ; \, \hat{\mu}_q, \, \hat{\Sigma}_q)$ is a multivariate Gaussian function in $\mathbb{R}^D$, with the mean $\hat{\mu}_q$ and the covariance matrix $\hat{\Sigma}_q$:

$$\hat{G}(s \, ; \, \hat{\mu}_q, \, \hat{\Sigma}_q) = \exp\left\{ -\frac{1}{2} (s - \hat{\mu}_q)^T \, \hat{\Sigma}_q \, (s - \hat{\mu}_q) \right\}, \tag{6}$$

By construction, the covariance matrix can be written as a sum of a full covariance in at most d directions and an isotropic complementary covariance, uniform across

different components in the mixture:

$$\hat{\Sigma}_q = \overbrace{\hat{\mathbf{P}}_q\,\hat{\Lambda}_q\,\hat{\mathbf{P}}_q^T}^{\text{principal subspace}} + \overbrace{\hat{\rho}\,\hat{\mathbf{C}}_q\,\hat{C}_q^T}^{\text{complementary subspace}} \tag{7}$$

$$\hat{\mathbf{P}}_q^T\,\hat{\mathbf{C}}_q = \mathbf{0} \qquad \hat{\mathbf{P}}_q^T\,\hat{\mathbf{P}}_q = \mathbf{1}^{(d)} \qquad \hat{\mathbf{C}}_q^T\,\hat{\mathbf{C}}_q = \mathbf{1}^{(D-d)}, \tag{8}$$

where $\mathbf{1}^{(d)}$ is the $d \times d$ identity matrix. The maximal dimensionality of the principal subspace spanned by the columns of $\hat{\mathbf{P}}_q$ captures the inherently low-dimensional structure of the personal face manifold and is a free parameter of the algorithm.

Model selection, that is, the inference of the parameters of the mixture – component centres, principal covariances and the complementary noise variance – is performed by minimizing the description length of the model and data [3], where the number of free parameters is:

$$p_{\text{free}} = \overbrace{Q\,D}^{\text{component means}} + \overbrace{Q\,D\,d}^{\text{principal bases}} + \overbrace{Q\,d}^{\text{principal covariance}} + \overbrace{1}^{\text{noise variance}}. \tag{9}$$

3: The signature space mixture is then implicitly re-projected into the image space, forming also a Q-component mixture:

$$\hat{p}(s) \xrightarrow{\text{image space}} p(i):\ p(i) = \sum_{q=1}^{Q} \Big[\alpha_q \cdot G(s\,;\,\mu_q,\,\Sigma_q)\Big]. \tag{10}$$

The unknown parameters of mixture components are computed from the known correspondences between the image and signature space, that is all i_r and s_r, and the component-wise likelihoods in the signature space $\hat{G}(s_r\,;\,\hat{\mu}_q,\,\hat{\Sigma}_q)$. Thus, the means are given by:

$$\mu_q = \sum_{r=1}^{n} i_r\,\hat{G}(s_r\,;\,\hat{\mu}_q,\,\hat{\Sigma}_q), \tag{11}$$

while the covariance matrices Σ_q are computed by fitting a Probabilistic PCA model $\Sigma_q = \mathbf{P}_q\,\Lambda_q\,\mathbf{P}_q{}^T + \rho\,\mathbf{C}_q\,\mathbf{C}_q{}^T$ to the set of intermediate estimates Σ'_q:

$$\Sigma'_q = \sum_{r=1}^{n} (i_r - \mu_q)\,(i_r - \mu_q)^T\,\hat{G}(s_r\,;\,\hat{\mu}_q,\,\hat{\Sigma}_q). \tag{12}$$

4: Note that in general, the principal directions of Σ_q and $\hat{\Sigma}_q$ do not correspond. To infer the correspondence between the directions in the regions of appearance and pose-signature spaces dominated by respectively Σ_q and $\hat{\Sigma}_q$, we find the optimal linear transformation $\mathbf{R}_q$ relating the projections of pose-signatures of training faces in the pose-signature space and the projections of original images in the appearance space:

$$\mathbf{R}_q = \arg\min_{\mathbf{R}} \left\{ w^T \left[\mathbf{R}\,\hat{\mathbf{P}}_q\,(\Delta_q\mathbf{S}) - \mathbf{P}_q\,(\Delta_q\mathbf{I})\right]^T \left[\mathbf{R}\,\hat{\mathbf{P}}_q\,(\Delta_q\mathbf{S}) - \mathbf{P}_q\,(\Delta_q\mathbf{I})\right] w \right\}, \tag{13}$$

where

$$\Delta_q S = \left[s_1 - \hat{\mu}_q \mid \ldots \mid s_n - \hat{\mu}_q \right] \qquad \Delta_q I = \left[i_1 - \mu_q \mid \ldots \mid i_n - \mu_q \right], \tag{14}$$

and w is a vector of weights, scaling the contribution of each face according to its Mahalanobis proximity to $\hat{\mu}_q$:

$$w(r) = (s_r - \hat{\mu}_q)\, \hat{\Sigma}_q^{-1}\, (s_r - \hat{\mu}_q),\ r = 1, \ldots, n. \tag{15}$$

Minimization in (13) can performed by a straightforward but cluttered differentiation of the quadratic form.

3.2 Matching Image Sets and Personal Appearance Models

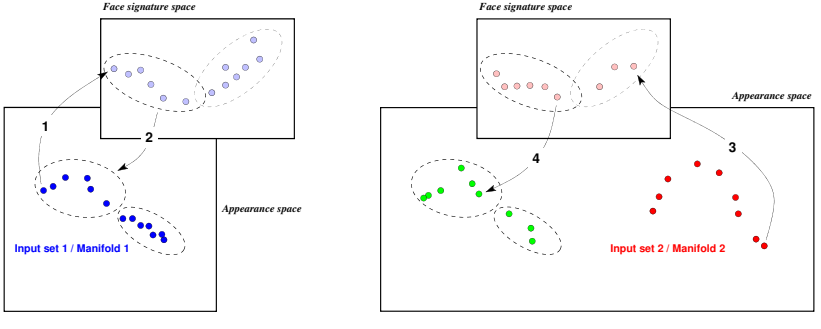
The appearance model described in the previous section was constructed so as to allow efficient re-illumination and matching of a novel set of face images. Specifically, a novel face i_r is used to compute a synthetically re-illuminated face i_r^* by independently re-illuminating it with each pair of mixture components $G(s_r\ ;\ \hat{\mu}_q,\ \hat{\Sigma}_q)$ and $\hat{G}(s_r\ ;\ \hat{\mu}_q,\ \hat{\Sigma}_q)$, and choosing the best pair under the learnt shape-illumination invariant. Thus, $i_r^* = i_{r,q^*}$ where:

$$q^* = \arg\max_q \mathscr{P}(\log i_{r,q} - \log i_r), \tag{16}$$

and

$$i_{r,q} = \mathbf{P}_q\, \mathbf{R}_q\, \hat{\mathbf{P}}_q^T\, (s_r - \hat{\mu}_q) + \mu_q. \tag{17}$$

The postulated shape-illumination effects are then extracted as in the original algorithm and used to compute the overall likelihood of the model [1].



(a) (b)

Figure 3: Conceptual illustration of the proposed method (compare with Figure 1).

4 Empirical Analysis

We now turn our attention to an empirical analysis of theoretical arguments put forward in the preceding sections of the paper. To make the reported experimental results directly comparable with those of the original work [1], the same database of face motion sequences was

	L-1	L-2	L-3	L-4	L-5	L-6	L-7	Average
L-1	–	100	99	100	100	99	99	99.5 (0.5)
L-2	100	–	100	100	100	100	100	100.0 (0.0)
L-3	100	100	–	99	99	99	97	99.0 (1.0)
L-4	100	100	99	–	99	99	97	99.0 (1.0)
L-5	96	100	94	97	–	100	100	97.8 (2.3)
L-6	96	97	97	95	100	–	96	96.8 (1.6)
L-7	96	99	95	97	99	97	–	97.2 (1.5)
Average	98.0 (2.0)	99.3 (1.1)	97.3 (2.2)	98.0 (1.8)	99.5 (0.5)	99.0 (1.0)	98.2 (1.6)	98.5 (1.7)

Figure 4: Average correct recognition rate and its standard deviation (in brackets) for different training and query illuminations.

used for evaluation – please consult the original publication for a thorough description [1]. By matching sets of 70×70 pixel images of face appearance of 100 different people across different 7 illumination conditions (we will refer to these as *L-1*, ..., *L-7*), we investigated the following:

- the average recognition rate for each combination of training/query illuminations
- the separation of obtained inter-class and intra-class distances, and
- the improvement in the average time taken to compute a single match.

4.1 Results and Discussion

The average recognition rates across training and query illuminations are tabulated in Figure 4. Firstly, we can observe that a very high accuracy is achieved, with the average error of 1.5%. What is more, our algorithm exhibits low dependence of performance on the exact training or query illuminations, as witnessed by the small standard deviation of the correct recognition rate. Good separation of inter-class and intra-class distances is further corroborated by the corresponding Receiver-Operator Characteristics – a few corresponding to a representative and an extreme illumination condition are shown in Figure 5, with a typical distance matrix in Figure 6 (a).

It is interesting to observe the asymmetry of the matrix in Figure 4 – an illumination which is found favourable for training the algorithm (in terms of the average recognition rate across different query illuminations) was not necessarily a good choice for the acquisition of novel data. What is more, the choice of the illumination used for training was found to be of greater importance than that used to acquire novel data, as witnessed by the greater variation in the average recognition across the former set. While this is certainly good news from a practical standpoint, it is comforting to provide a plausible explanation why this may be the case, especially since the same phenomenon was not observed in the evaluation of the original method in [1]

Firstly, note that this asymmetry is not necessarily surprising considering that the manner in which face sets are matched in our algorithm is indeed less symmetric than theirs. While Arandjelović & Cipolla directly match two sets of images, here a set of novel images is matched with a *semi-parametric distribution*. Specifically, note that the highest average recognition rate was obtained when illumination *L-2* was used for training, which corresponds to a dominant mild frontal light. On the other hand, the most extreme illumination

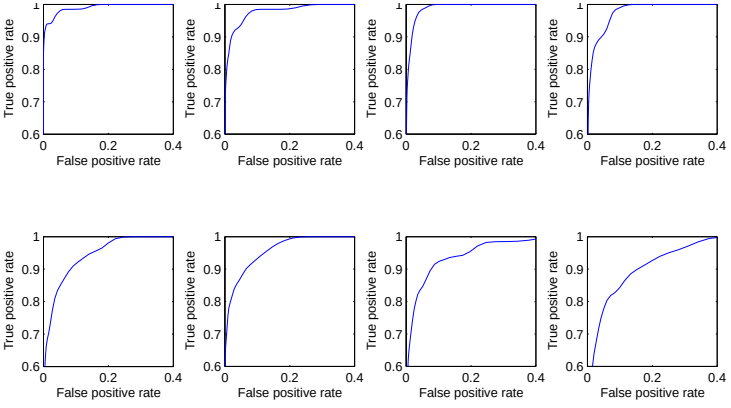
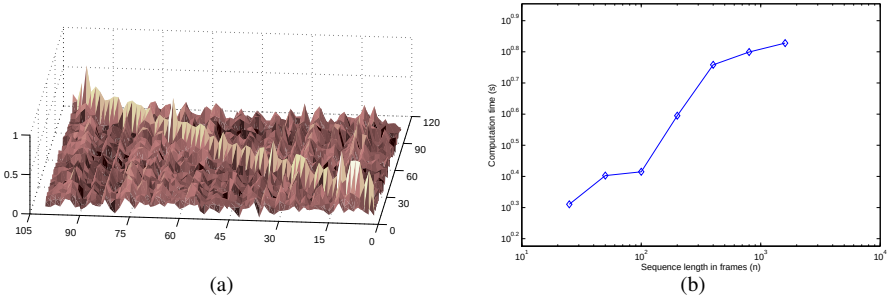


Figure 5: Receiver-Operator Characteristics corresponding to training data acquired in *L-2* (top row) and *L-6* (bottom row) and queried with *L-3* to *L-6*, and *L-2* to *L-5* respectively.

L-6 (strong lateral light source), was the least forgiving in training. Considering the nature of our model, a likely explanation stems from the higher non-linearity of appearance manifolds acquired in extreme lighting conditions, producing large contrasts between illuminated and shadowed areas of the moving face. In such cases, the distortion of principal directions between the image space mixture components and those in the signature space – discussed in Section 3 – is likely to be non-uniform across a region dominated by a single component invalidating the implicit assumption to the contrary. This suggest that an improvement may be achievable by using mixtures of higher complexity (greater number of components) than that which minimizes the model and data description length.



(a) (b)

Figure 6: (a) A typical distance matrix, showing good separation of distances obtained in comparisons of image sets which correspond in identity and those which do not. Different illuminations were used for training and querying the algorithm. (b) The average time taken for a single match (compare with Figure 2).

Finally, the average time required to match a query set with a single personal appearance model is shown as a function of n, the number of faces in the set, in Figure 6 (b)[1]. Drastic improvement is demonstrated, a 1600-image set requiring only 6.7 s, an over 700-fold speed-up[2] over the original algorithm of Arandjelović & Cipolla (see the corresponding plot in

[1]The same computer as that in Section 2.2 was used: Intel P4 PC, with a 3.2GHz CPU and 2GB RAM.

[2]The time required to match 1600-image sets using the original algorithm was extrapolated from Figure 2 in

Section 2.2).

5 Summary and Conclusions

We proposed a novel method for face recognition from image sets, based on the generic shape-illumination invariant previously described in the literature. The key contribution is a personal appearance representation suitable for exploiting the invariant, which consists of two mixture models – in image space and the corresponding pose-signature space – linked through a set of linear transformations. On a large data set, our algorithm achieved a recognition accuracy comparable to that of the original method based on the aforementioned invariant, with a dramatic improvement in storage requirements and matching speed.

References

[1] O. Arandjelović and R. Cipolla. Face recognition from video using the generic shape-illumination manifold. *In Proc. European Conference on Computer Vision (ECCV)*, 4: 27–40, May 2006.

[2] T. H. Cormen, C. E. Leiserson, and R. L. Rivest. *Introduction to Algorithms*. MIT Press, June 1990. ISBN 978-0-262-03293-3.

[3] P. Grünwald. *The Minimum Description Length principle.* MIT Press, June 2007. ISBN 978-0-262-07281-6.

[4] Bowyer K. W. Flynn P. J. O'Toole A. J. Scruggs W. T. Schott C. L. Sharpe M. Phillips, P. J. *Face Recognition Vendor Test 2006 and Iris Challenge Evaluation 2006 Large-Scale Results.* DIANE Publishing Company, March 2007. ISBN 978-1-422-31289-6.

[5] C. Shan. *Video Search and Mining*, chapter Face Recognition and Retrieval in Video, pages 235–260. Springer, May 2010. ISBN 978-3-642-12899-8.

[6] J. Stallkamp, H. K. Ekenel, and R. Stiefelhagen. Video-based face recognition on real-world data. *In Proc. IEEE International Conference on Computer Vision (ICCV)*, pages 1–8, October 2007.

[7] M. E. Tipping and C. M. Bishop. Probabilistic principal component analysis. *Journal of the Royal Statistical Society*, 61(3):611–622, July 1999.

[8] Y. Yan and Y-J. Zhang. *Encyclopedia of Artificial Intelligence*, chapter State-of-the-Art on Video-Based Face Recognition, pages 1455–1461. IGI Global, August 2009. ISBN 978-1-599-04850-5.

Section 2.2, the corresponding estimate being $\sim$ 4800 s or 80 minutes.