# Face Recognition Across Pose and Illumination

Ralph Gross, Simon Baker, Iain Matthews, and Takeo Kanade

Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213, USA
{rgross, simonb, iainm, tk}@cs.cmu.edu

The last decade has seen automatic face recognition evolve from small scale research systems to a wide range of commercial products. Driven by the FERET face database and evaluation protocol, the currently best commercial systems achieve verification accuracies comparable to those of fingerprint recognizers. In these experiments, only frontal face images taken under controlled lighting conditions were used. As the use of face recognition systems expands towards less restricted environments, the development of algorithms for view and illumination invariant face recognition becomes important. However, the performance of current algorithms degrades significantly when tested across pose and illumination as documented in a number of evaluations. In this chapter we review previously proposed algorithms for pose and illumination invariant face recognition. We then describe in detail two successful appearance-based algorithms for face recognition across pose, eigen light-fields and Bayesian face subregions. We furthermore show how both of these algorithms can be extended towards face recognition across pose and illumination.

## 1 Introduction

The most recent evaluation of commercial face recognition systems shows the level of performance for face verification of the best systems to be on par with fingerprint recognizers for frontal, uniformly illuminated faces [39]. Recognizing faces reliably across changes in pose and illumination has proved to be a much harder problem [10, 25, 39]. While the majority of research has so far focused on frontal face recognition, there is a sizable body of work on pose invariant face recognition and illumination invariant face recognition. However, face recognition across pose *and* illumination has received very little attention.

### 1.1 Multi-view Face Recognition and Face Recognition Across Pose

Approaches addressing pose variation can be classified into two categories depending on the type of gallery images they use. Multi-view face recognition is a direct

extension of frontal face recognition in which the algorithms require gallery images of every subject at every pose. In face recognition across pose we are concerned with the problem of building algorithms to recognize a face from a novel viewpoint, i.e. a viewpoint from which it has not previously been seen. In both categories we furthermore distinguish between model-based and appearance-based algorithms. Model-based algorithms use an explicit 2D [13] or 3D [11, 16] model of the face, whereas appearance-based methods directly use image pixels or features derived from image pixels [37].

One of the earliest appearance-based multi-view algorithms is described in [7]. After a pose estimation step the algorithm geometrically aligns the probe images to candidate poses of the gallery subjects using the automatically determined locations of three feature points. This alignment is then refined using optical flow. Recognition is performed by computing normalized correlation scores. Good recognition results are reported on a database of 62 subjects imaged in a number of poses ranging from $-30°$ to $+30°$ (yaw) and from $-20°$ to $+20°$ (pitch). However, the probe and gallery poses are very similar. In [38] the popular eigenface approach of Turk and Pentland [48] is extended to handle multiple views. The authors compare the performance of a parametric eigenspace (computed using all views from all subjects) with view-based eigenspaces (separate eigenspaces for each view). In experiments on a database of 21 people recorded in nine evenly spaced views from $-90°$ to $+90°$, view-based eigenspaces outperform the parametric eigenspace by a small margin.

A number of 2D model-based algorithms have been proposed for face tracking through large pose changes. In [14] separate active appearance models are trained for profile, half profile and frontal views, with models for opposing views created by simple reflection. Using a heuristic for switching between models the system is able to track faces through wide angle changes. It has been shown that linear models are able to deal with considerable pose variation as long as all the modeled features remain visible [33]. A different way of dealing with larger pose variations is then to introduce non-linearities into the model. Romdhani et al. extended active shape models [42] and active appearance models [43] using a kernel PCA to model shape and texture nonlinearities across views. In both cases models are successfully fit to face images across a full $180°$ rotation. However, no face recognition experiments were performed.

In many face recognition scenarios the pose of the probe and gallery images are different. For example, the gallery image might be a frontal "mug-shot" and the probe image might be a 3/4 view captured from a camera in the corner of a room. The number of gallery and probe images can also vary. For example, the gallery might consist of a pair of images for each subject, a frontal mug-shot and full profile view (like the images typically captured by police departments). The probe might be a similar pair of images, a single 3/4 view, or even a collection of views from random poses. In these scenarios multi-view face recognition algorithms can not be used. Early work on face recognition across pose was based on the idea of linear object classes [49]. The underlying assumption is that the 3D shape of an object (and 2D projections of 3D objects) can be represented by a linear combination of prototypical objects. It follows that a rotated view of the object

is a linear combination of the rotated views of the prototype objects. Using this idea the authors are able to synthesize rotated views of face images from a single example view. In [8] this algorithm is used to create virtual views from a single input image for use in a multi-view face recognition system. Lando and Edelman use a comparable example-based technique to generalize to new poses from a single view [32].

A completely different approach to face recognition across pose is based on the work of Murase and Nayar [37]. They showed that different views of a rigid object projected into an eigenspace fall on a 2D manifold. Using a model of the manifold they could recognize objects from arbitrary views. In a similar manner Graham and Allison observe that a densely sampled image sequence of a rotating head forms a characteristic *eigensignature* when projected into an eigenspace [20]. They use Radial Basis Function Networks to generate eigensignatures based on a single view input. Recognition is then performed by distance computation between the projection of a probe image into eigenspace and the eigensignatures created from gallery views. Good generalization is observed from half profile training views. However, recognition rates for tests across wide pose variations (e.g. frontal gallery and profile probe) are weak.

One of the early model-based approaches for face recognition is based on Elastic Bunch Graph Matching [50]. Facial landmarks are encoded with sets of complex Gabor wavelet coefficients called jets. A face is then represented with a graph where the different jets form the nodes. Based on a small number of hand-labeled examples, graphs for new images are generated automatically. The similarity between a probe graph and the gallery graphs is determined as average over the similarities between pairs of corresponding jets. Correspondences between nodes in different poses is established manually. Good recognition results are reported on frontal faces in the FERET evaluation [40]. Recognition accuracies decrease drastically though for matching half profile images with either frontal or full profile views. For the same framework a method for transforming jets across pose is introduced in [36]. In limited experiments the authors show improved recognition rates over the original representation.

## 1.2 Illumination Invariant Face Recognition

Besides face pose, illumination is the next most significant factor affecting the appearance of faces. Ambient lighting changes greatly within and between days and among indoor and outdoor environments. Due to the 3D structure of the face, a direct lighting source can cast strong shadows that accentuate or diminish certain facial features. It has been shown experimentally [2] and theoretically for systems based on Principal Component Analysis [51] that differences in appearance induced by illumination are larger than differences between individuals. Since dealing with illumination variation is a central topic in computer vision numerous approaches for illumination invariant face recognition have been proposed.

Early work in illumination invariant face recognition focused on image representations that are mostly insensitive to changes in illumination. In [2] different

image representations and distance measures are evaluated on a tightly controlled face database which varied face pose, illumination and expression. The image representations include edge maps, 2D Gabor-like filters, first and second derivatives of the gray-level image and the logarithmic transformations of the intensity image along with these representations. However, none of the different image representations is found to be sufficient by itself to overcome variations due to illumination changes. In more recent work it was shown that the ratio of two images from the same object is simpler than the ratio of images from different objects [28]. In limited experiments this method outperformed both correlation and Principal Component Analysis, but did not perform as well as the illumination cone method described below. A related line of work attempts to extract the object's surface reflectance as an illumination invariant description of the object [26, 31]. We will discuss the most recent algorithm in this area in more detail in Section 4.2. In [45] a different illumination invariant image representation, the quotient image, was proposed. Computed from a small set of example images, the quotient image can be used to re-render an object of the same class under a different illumination condition. In limited recognition experiments the method outperforms Principal Component Analysis.

A different approach to the problem is based on the observation that the images of a Lambertian surface, taken from a fixed viewpoint but under varying illumination lie in a 3D linear subspace of the image space [44]. A number of appearance-based methods exploit this fact to model the variability of faces under changing illumination. In [6] the eigenface algorithm of Turk and Pentland [48] is extended to Fisherfaces by employing a classifier based on Fisher's Linear Discriminant Analysis. In experiments on a face database with strong variations in illumination Fisherfaces outperform eigenfaces by a wide margin. Further work in the area by Belhumeur and Kriegman showed that the set of images of an object in fixed pose but under varying illumination forms a convex cone in the space of images [5]. The illumination cones of human faces can be approximated well by low-dimensional linear subspaces [17]. An algorithm based on this method outperforms both eigenfaces and Fisherfaces. More recently Basri and Jacobs showed that the illumination cone of a convex Lambertian surface can be approximated by a 9-dimensional linear subspace [3]. In limited experiments good recognition rates across illumination conditions are reported.

Common to all these appearance-based methods is the need for training images of database subjects under a number of different illumination conditions. An algorithm proposed by Sim and Kanade overcomes this restriction [47]. The authors use a statistical shape-from-shading model to recover the face shape from a single image and synthesize the face under a new illumination. Using this method they generate images of the gallery subjects under many different illumination conditions to serve as gallery images in a recognizer based on Principal Component Analysis. High recognition rates are reported on the illumination subset of the CMU PIE database [46].

### 1.3 Algorithms for Face Recognition across Pose and Illumination

To simultaneously address the problems of face recognition across pose and illumination a number of appearance and model-based algorithms have been proposed. In [18] a variant of photometric stereo is used to recover shape and albedo of a face based on seven images of the subject seen in a fixed pose. In combination with the illumination cone representation introduced in [5] the authors can synthesize faces in novel pose and illumination conditions. In tests on 4050 images from the Yale Face Database B the method performed almost without error. In [12] a morphable model of three dimensional faces is introduced. The model is created using a database of Cyberware laser scans of 200 subjects. Following an analysis-by-synthesis paradigm the algorithm automatically recovers face pose and illumination from a single image. For initialization, the algorithm requires the manual localization of seven facial feature points. After fitting the model to a new image the extracted model parameters describing the face shape and texture are used for recognition. The authors report excellent recognition rates on both the FERET [40] and CMU PIE [46] databases. Once fit, the model can also be used to synthesize an image of the subject under new conditions. This method was used in the most recent Face Recognition Vendor Test to create frontal view images from rotated views [39]. For 9 out of 10 face recognition systems tested, accuracies on the synthesized frontal views were significantly higher than on the original images.

## 2 Eigen Light-Fields

We propose an appearance-based algorithm for face recognition across pose. Our algorithm can use any number of gallery images captured at arbitrary poses, and any number of probe images also captured with arbitrary poses. A minimum of 1 gallery and 1 probe image are needed, but if more images are available the performance of our algorithm generally improves.

Our algorithm operates by estimating (a representation of) the light-field [35] of the subject's head. First, generic training data is used to compute an eigen-space of head light-fields, similar to the construction of eigenfaces [48]. Light-fields are simply used rather than images. Given a collection of gallery or probe images, the projection into the eigen-space is performed by setting up a least-squares problem and solving for the projection coefficients similarly to approaches used to deal with occlusions in the eigen-space approach [9, 34]. This simple linear algorithm can be applied to any number of images, captured from any poses. Finally, matching is performed by comparing the probe and gallery eigen light-fields.

### 2.1 Light-Fields Theory

**Object Light-Fields**

The *plenoptic function* [1] or *light-field* [35] is a function which specifies the radiance of light in free space. It is a 5D function of position (3D) and orientation (2D).
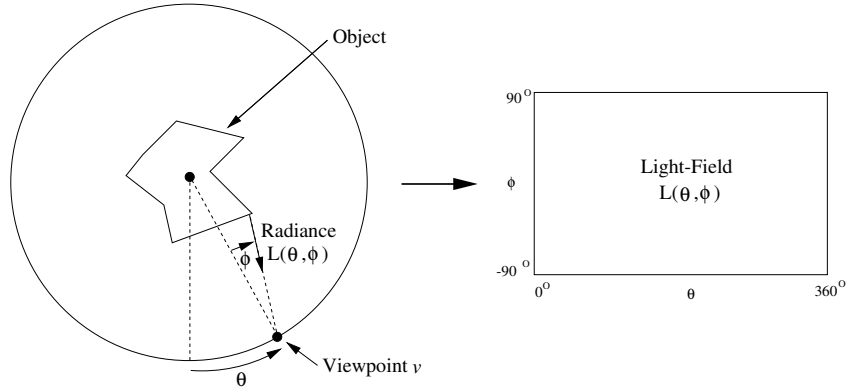
**Fig. 1.** The object is conceptually placed within a circle. The angle to the viewpoint $v$ around the circle is measured by the angle $\theta$, and the direction that the viewing ray makes with the radius of the circle is denoted $\phi$. For each pair of angles $\theta$ and $\phi$, the radiance of light reaching the viewpoint from the object is then denoted by $L(\theta, \phi)$, the *light-field*. Although the light-field of a 3D object is actually 4D, we will continue to use the 2D notation of this figure in this paper for ease of explanation.

In addition, it is also sometimes modeled as a function of time, wavelength, and polarization, depending on the application in mind. In 2D, the light-field of a 2D object is actually 2D rather than the 3D that might be expected. See Figure 1 for an illustration.

**Eigen Light-Fields**

Suppose we are given a collection of light-fields $L_i(\theta, \phi)$ of objects $O_i$ (here faces of different subjects) where $i = 1, \ldots, N$. See Figure 1 for the definition of this notation. If we perform an eigen-decomposition of these vectors using Principal Component Analysis (PCA), we obtain $d \leq N$ eigen light-fields $E_i(\theta, \phi)$ where $i = 1, \ldots, d$. Then, assuming that the eigen-space of light-fields is a good representation of the set of light-fields under consideration, we can approximate any light-field $L(\theta, \phi)$ as:

$$L(\theta, \phi) \approx \sum_{i=1}^{d} \lambda_i E_i(\theta, \phi) \tag{1}$$

where $\lambda_i = \langle L(\theta, \phi), E_i(\theta, \phi) \rangle$ is the inner (or dot) product between $L(\theta, \phi)$ and $E_i(\theta, \phi)$. This decomposition is analogous to that used in face and object recognition [37, 48]; The mean light-field could also be estimated and subtracted from all of the light-fields.

Capturing the complete light-field of an object is a difficult task, primarily because it requires a huge number of images [19, 35]. In most object recognition

scenarios it is unreasonable to expect more than a few images of the object; often just one. However, any image of the object corresponds to a curve (for 3D objects, a surface) in the light-field. One way to look at this curve is as a highly occluded light-field; only a very small part of the light-field is visible. Can the eigen coefficients $\lambda_i$ be estimated from this highly occluded view? Although this may seem hopeless, consider that light-fields are highly redundant, especially for objects with simple reflectance properties such as Lambertian. An algorithm is presented in [34] to solve for the unknown $\lambda_i$ for eigen-*images*. A similar algorithm was implicitly used in [9]. Rather than using the inner product $\lambda_i = \langle L(\theta, \phi), E_i(\theta, \phi) \rangle$, Leonardis and Bischof [34] solve for $\lambda_i$ as the least squares solution of:

$$L(\theta, \phi) - \sum_{i=1}^{d} \lambda_i E_i(\theta, \phi) \;=\; 0 \tag{2}$$

where there is one such equation for each pair of $\theta$ and $\phi$ that are un-occluded in $L(\theta, \phi)$. Assuming that $L(\theta, \phi)$ lies *completely within the eigen-space* and that enough pixels are un-occluded, then the solution of Equation (2) will be exactly the same as that obtained using the inner product [23]. Since there are $d$ unknowns $(\lambda_1 \ldots \lambda_d)$ in Equation (2), at least $d$ un-occluded light-field pixels are needed to over-constrain the problem, but more may be required due to linear dependencies between the equations. In practice, $2 - 3$ times as many equations as unknowns are typically required to get a reasonable solution [34]. Given an image $I(m, n)$, the following is then an algorithm for estimating the eigen light-field coefficients $\lambda_i$:

1. For each pixel $(m, n)$ in $I(m, n)$ compute the corresponding light-field angles $\theta_{m,n}$ and $\phi_{m,n}$. (This step assumes that the camera intrinsics are known, as well as the relative orientation of the camera to the object.)
2. Find the least-squares solution (for $\lambda_1 \ldots \lambda_d$) to the set of equations:

$$I(m, n) - \sum_{i=1}^{d} \lambda_i E_i(\theta_{m,n}, \phi_{m,n}) \;=\; 0 \tag{3}$$

   where $m$ and $n$ range over their allowed values. (In general, the eigen light-fields $E_i$ need to be interpolated to estimate $E_i(\theta_{m,n}, \phi_{m,n})$. Also, all of the equations for which the pixel $I(m, n)$ does not image the object should be excluded from the computation.)

Although we have described this algorithm for a single image $I(m, n)$, any number of images can obviously be used (so long as the camera intrinsics and relative orientation to the object are known for each image). The extra pixels from the other images are simply added in as additional constraints on the unknown coefficients $\lambda_i$ in Equation (3). The algorithm can be used to estimate a light-field from a collection of images. Once the light-field has been estimated, it can then be used to render new images of the same object under different poses. (See [49] for a related algorithm.) In [23] we show that the algorithm correctly re-renders a given object assuming a

|     Input     |   Rerendered  |    Original   |
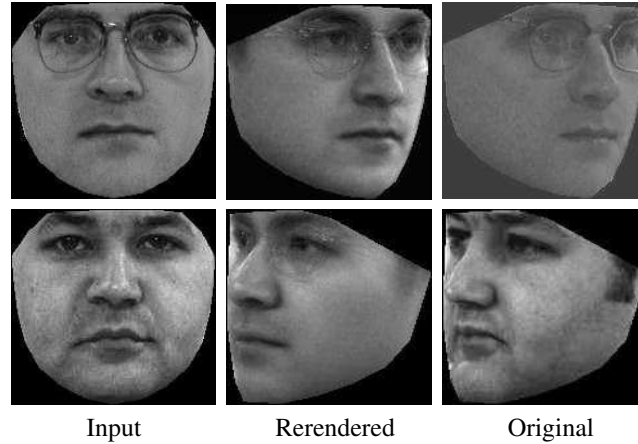
**Fig. 2.** An illustration of using our eigen light-field estimation algorithm for re-rendering a face across pose. The algorithm is given the left-most (frontal) image as input from which it estimates the eigen light-field and then creates the rotated view shown in the middle. For comparison, the original rotated view is shown in the right-most column. In the figure we show one of the better results (top) and one of the worst (bottom.) Although in both cases the output looks like a face, the identity is altered in the second case.

Lambertian reflectance model. The extent to which these assumptions are valid are illustrated in Figure 2 where we present results of using our algorithm to re-render faces across pose. In each case the algorithm received the left-most (frontal) image as input and created the rotated view in the middle. For comparison, the original rotated view is included as the right-most image. The re-rendered image for the first subject is very similar to the original. While the image created for the second subject still shows a face in the correct pose, the identity of the subject is not as accurately recreated. We conclude that overall our algorithm works fairly well, but that more training data is needed so that the eigen light-field of faces can more accurately represent any given face light-field.

## 2.2 Application to Face Recognition Across Pose

The Eigen Light-Field Estimation Algorithm described above is somewhat abstract. In order to be able to use it for face recognition across pose we need to do the following things:

**Vectorization:** The input to a face recognition algorithm consists of a collection of images (possibly just one) captured from a variety of poses. The Eigen Light-Field Estimation Algorithm operates on light-field vectors (light-fields represented as vectors). Vectorization consists of converting the input images into a light-field vector (with missing elements, as appropriate.)
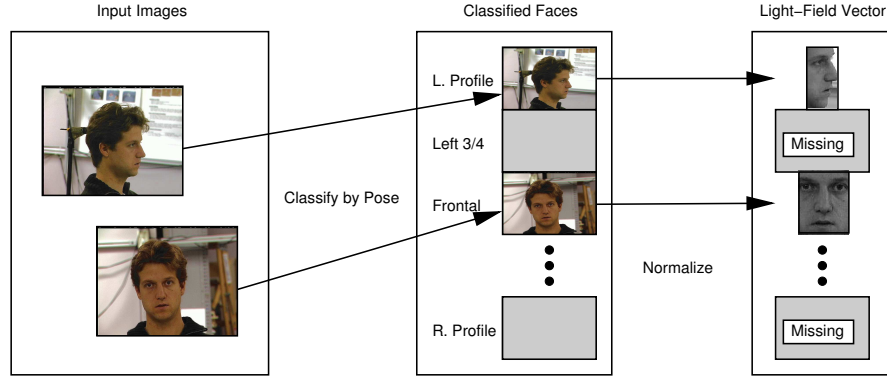
**Fig. 3.** Vectorization by normalization. Vectorization is the process of converting a set of images of a face into a light-field vector. Vectorization is performed by first classifying each input image into one of a finite number of poses. For each pose, a normalization is then applied to convert the image into a sub-vector of the light-field vector. If poses are missing, the corresponding part of the light-field vector is missing.

**Classification:** Given the eigen coefficients $a_1 \ldots a_d$ for a collection of gallery faces and for a probe face, we need to classify which gallery face is the most likely match.

**Selecting Training and Testing Sets:** To evaluate our algorithm we have to divide the database used into (disjoint) subsets for training and testing.

We now describe each of these tasks in turn.

### Vectorization by Normalization

Vectorization is the process of converting a collection of images of a face into a light-field vector. Before we can do this we first have to decide how to discretize the light-field into pixels. Perhaps the most natural way to do this is to uniformly sample the light-field angles, $\theta$ and $\phi$ in the 2D case of Figure 1. This is not the only way to discretize the light-field. Any sampling, uniform or non-uniform, could be used. All that is needed is a way of specifying what is the allowed set of light-field pixels. For each such pixel, there is a corresponding index in the light-field vector; i.e. if the light-field is sampled at $K$ pixels, the light-field vectors are $K$ dimensional vectors.

We specify the set of light-field pixels in the following manner. We assume that there are only a finite set of poses $1, 2, \ldots, P$ in which the face can occur. Each face image is first classified into the nearest pose. (Although this assumption is clearly an approximation, its validity is demonstrated by the empirical results in Section 2.3. In both the FERET [40] and PIE [46] databases, there is considerable variation in the pose of the faces. Although the subjects are asked to place their face in a fixed pose, they rarely do this perfectly. Both databases therefore contain considerable

variation away from the finite set of poses. Since our algorithm performs well on both databases, the approximation of classifying faces into a finite set of poses is validated.)

Each pose $i = 1, \ldots, P$ is then allocated a fixed number of pixels $K_i$. The total number of pixels in a light-field vector is therefore $K = \sum_{i=1}^{P} K_i$. If we have images from pose 3 and 7, for example, we know $K_3 + K_7$ of the $K$ pixels in the light-field vector. The remaining $K - K_3 - K_7$ are unknown, missing data. This vectorization process is illustrated in Figure 3.

We still need to specify how to sample the $K_i$ pixels of a face in pose $i$. This process is analogous to that needed in appearance-based object recognition and is usually performed by "normalization." In eigenfaces [48], the standard approach is to find the positions of several canonical points, typically the eyes and the nose, and to warp the input image onto a coordinate frame where these points are in fixed locations. The resulting image is then masked. To generalize eigenface normalization to eigen light-fields, we just need to define such a normalization for each pose.

We report results using two different normalizations. The first one is a simple one based on the location of the eyes and the nose. Just as in eigenfaces, we assume that the eye and nose locations are known, warp the face into a coordinate frame in which these canonical points are in a fixed location and finally crop the image with a (pose dependent) mask to yield the $K_i$ pixels. For this simple 3-point normalization, the resulting masked images vary in size between 7200 and 12600 pixels, depending on the pose.

The second normalization is more complex and is motivated by the success of Active Appearance Models [13]. This normalization is based on the location of a large number (39–54 depending on the pose) of points on the face. These canonical points are triangulated and the image warped with a piecewise affine warp onto a coordinate frame in which the canonical points are in fixed locations. The resulting masked images for this multi-point normalization vary in size between 20800 and 36000 pixels. Although currently the multi-point normalization is performed using hand-marked points, it could be performed by fitting an Active Appearance Model [13] and then using the implied canonical point locations.

**Classification using Nearest Neighbor**

The Eigen Light-Field Estimation Algorithm outputs a vector of eigen coefficients $(a_1, \ldots, a_d)$. Given a set of gallery faces, we obtain a corresponding set of vectors $(a_1^{\text{id}}, \ldots, a_d^{\text{id}})$, where id is an index over the set of gallery faces. Similarly, given a probe face, we obtain a vector $(a_1, \ldots, a_d)$ of eigen coefficients for that face. To complete the face recognition algorithm we need an algorithm which classifies $(a_1, \ldots, a_d)$ with the index id which is the most likely match. Many different classification algorithms could be used for this task. For simplicity, we use the nearest neighbor algorithm which classifies the vector $(a_1, \ldots, a_d)$ with the index:

$$\arg\min_{\text{id}} \text{dist}\left((a_1, \ldots, a_d), (a_1^{\text{id}}, \ldots, a_d^{\text{id}})\right) \;=\; \arg\min_{\text{id}} \sum_{i=1}^{d} \left(a_i - a_i^{\text{id}}\right)^2. \qquad (4)$$

All of the results reported in this paper use the Euclidean distance in Equation (4). Alternative distance functions, such as the Mahalanobis distance, could be used instead if so desired.

### Selecting the Gallery, Probe, and Generic Training Data

In each of our experiments we divided the database into three disjoint subsets:

**Generic Training Data:** Many face recognition algorithms such as eigenfaces, and including our algorithm, require "generic training data" to build a generic face model. In eigenfaces, for example, generic training data is needed to compute the eigenspace. Similarly, in our algorithm generic data is needed to construct the eigen light-field.

**Gallery:** The gallery is the set of reference images of the people to be recognized; i.e. the images given to the algorithm as examples of each person that might need to be recognized.

**Probe:** The probe set contains the "test" images; i.e. the images that will be presented to the system to be classified with the identity of the person in the image.

The division into these three subsets is performed as follows. First we randomly select half of the subjects as the generic training data. The images of the remaining subjects are used for the gallery and probe. There is therefore never any overlap between the generic training data and the gallery and probe.

After the generic training data has been removed, the remainder of the databases are divided into probe and gallery sets based on the pose of the images. For example, we might set the gallery to be the frontal images and the probe set to be the left profiles. In this case, we evaluate how well our algorithm is able to recognize people from their profiles given that the algorithm has only seen them from the front. In the experiments described below we choose the gallery and probe poses in various different ways. The gallery and probe are always disjoint unless otherwise noted.

### 2.3 Experimental Results

### Databases

We used two databases in our face recognition across pose experiments, the CMU Pose, Illumination, and Expression (PIE) database [46] and the FERET database [40]. Each of these databases contains substantial pose variation. In the pose subset of the CMU PIE database (see Figure 4), the 68 subjects are imaged simultaneously under 13 different poses totaling 884 images. In the FERET database, the subjects are imaged non-simultaneously in 9 different poses. We used 200 subjects from the FERET pose subset giving 1800 images in total. If not stated otherwise we used half of the available subjects for training of the generic eigenspace (34 subjects for PIE, 100 subjects for FERET) and the remaining subjects for testing. In all experiments (if not stated otherwise) we retain a number of eigenvectors sufficient to explain 95% of the variance in the input data.

**Fig. 4.** Pose variation in the PIE database. The pose varies from full left profile (c34) to full frontal (c27) and on to full right profile (c22). Approximate pose angles are shown below the camera numbers.

## Comparison with Other Algorithms

We compared our algorithm with eigenfaces [48] and FaceIt, the commercial face recognition system from Identix (formerly Visionics).[1]

We first performed a comparison using the PIE database. After randomly selecting the generic training data, we selected the gallery pose as one of the 13 PIE poses and the probe pose as any other of the remaining 12 PIE poses. For each disjoint pair of gallery and probe poses, we compute the average recognition rate over all subjects in the probe and gallery sets. The details of the results are included in Figure 5 and a summary is included in Table 1.

In Figure 5 we plot color-coded $13 \times 13$ "confusion matrices" of the results. The row denotes the pose of the gallery, the column the pose of the probe, and the displayed intensity the average recognition rate. A lighter color denotes a higher recognition rate. (On the diagonals the gallery and probe images are the same and so all three algorithms obtain a 100% recognition rate.)

Eigen light-fields performs far better than the other algorithms, as is witnessed by the lighter color of Figures 5(a–b) compared to Figures 5(c–d). Note how eigen light-fields is far better able to generalize across wide variations in pose, and in particular to and from near profile views.

The results in Figure 5 are summarized in Table 1. In this table we include the average recognition rate computed over all disjoint gallery-probe poses. As can be seen, eigen light-fields outperforms both the standard eigenfaces algorithm and the commercial FaceIt system.

We next performed a similar comparison using the FERET database [40]. Just as with the PIE database, we selected the gallery pose as one of the 9 FERET poses and the probe pose as any other of the remaining 8 FERET poses. For each disjoint pair of gallery and probe poses, we compute the average recognition rate over all subjects in the probe and gallery sets, and then average the results. The results are very similar to those for the PIE database and are summarized in Table 2. Again, eigen light-fields performs significantly better than both FaceIt and eigenfaces.

---

[1] In the experiments version 2.5.0.17 of the FaceIt recognition engine was used.

(a) Eigen Light-Fields - 3-Point Norm (b) Eigen Light-Fields - Multi-point Norm

(c) FaceIt                                    (d) eigenfaces
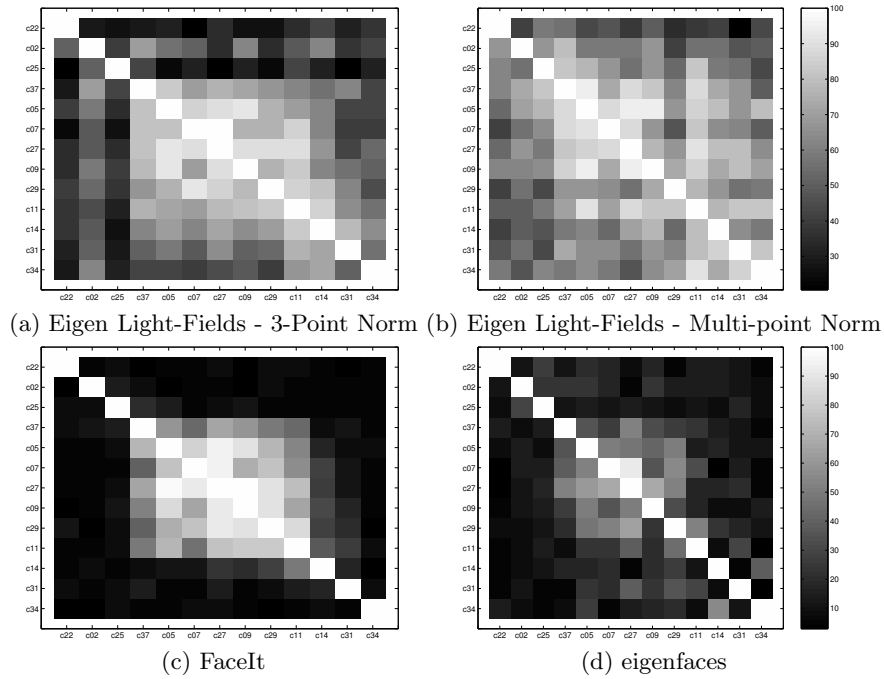
**Fig. 5.** A comparison with FaceIt and eigenfaces for face recognition across pose on the CMU PIE [46] database. For each pair of gallery and probe poses, we plot the color-coded average recognition rate. The row denotes the pose of the gallery, the column the pose of the probe. The fact that the images in (a) and (b) are lighter in color than those in (c) and (d) implies that our algorithm performs better.

**Table 1.** A comparison of eigen light-fields with FaceIt and eigenfaces for face recognition across pose on the CMU PIE database. The table contains the average recognition rate computed across all disjoint pairs of gallery and probe poses; i.e. this table summarizes the average performance in Figure 5.

|  | Eigenfaces | FaceIt | Eigen Light-Fields 3-Point Norm | Eigen Light-Fields Multi-Point Norm |
|---|---|---|---|---|
| Average Recognition Accuracy | 16.6% | 24.3% | 52.5% | 66.3% |

Overall, the performance improvement of eigen light-fields over the other two algorithms is more significant on the PIE database than on the FERET database. This is because the PIE database contains more variation in pose than the FERET database. For more evaluation results see [22].

**Table 2.** A comparison of Eigen Light-Fields with FaceIt and eigenfaces for face recognition across pose on the FERET database. The table contains the average recognition rate computed across all disjoint pairs of gallery and probe poses. Again, eigen light-fields outperforms both eigenfaces and FaceIt.

|  | Eigenfaces | FaceIt | Eigen Light-Fields 3-Point Normalization |
|---|---|---|---|
| Average Recognition Accuracy | 39.4% | 59.3% | 75% |

## 3 Bayesian Face Subregions

Due to the complicated three dimensional nature of the face, differences exist in how the appearance of various face regions change for different face poses. If for example a head rotates from a frontal to a right profile position, the appearance of the mostly featureless cheek region will only change little (if we ignore the influence of illumination), while other regions such as the left eye will disappear, and the nose will look vastly different. Our algorithm models the appearance changes of the different face regions in a probabilistic framework [29]. Using probability distributions for similarity values of face subregions we compute the likelihood of probe and gallery images coming from the same subject. For training and testing of our algorithm we use the CMU PIE database [46].

### 3.1 Face Subregions and Feature Representation

Using the hand-marked locations of both eyes and the midpoint of the mouth we warp the input face images into a common coordinate frame in which the landmark points are in a fixed location and crop the face region to a standard 128x128 pixel size. Each image $I$ in the database is labeled with the identity $i$ and pose $\phi$ of the face in the image: $I = (i, \phi), i \in \{1, \ldots, 68\}, \phi \in \{1, \ldots, 13\}$. As shown in Figure 6 a 7-by-3 lattice is placed on the normalized faces and 9x15 pixel subregions are extracted around every lattice point. The intensity values in each of the 21 subregions are normalized to have zero mean and unit variance.

As similarity measure between subregions we use SSD (sum of squared difference) values $s_j$ between corresponding regions $j$ for all image pairs. Since we compute SSD after image normalization it effectively contains the same information as normalized correlation.

### 3.2 Modeling Local Appearance Change across Pose

For probe image $I_{i,p} = (i, \phi_p)$ with unknown identity $i$ we compute the probability that $I_{i,p}$ is coming from the same subject $k$ as gallery image $I_{k,g}$ for each face subregion $j, j \in \{1, \ldots, 21\}$. Using Bayes' rule we write:
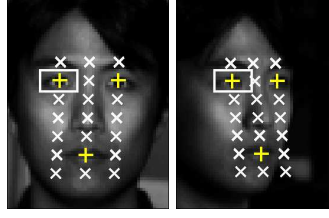
**Fig. 6.** Face subregions for two different poses of the CMU PIE database. Each face in the database is warped into a normalized coordinate frame using the hand-labeled locations of both eyes and the midpoint of the mouth. A 7x3 lattice is placed on the normalized face and 9x15 pixel subregions are extracted around every lattice point resulting in a total of 21 subregions.

$$P(i = k | s_j, \phi_p, \phi_g) = \frac{P(s_j | i = k, \phi_p, \phi_g) P(i = k)}{P(s_j | i = k, \phi_p, \phi_g) P(i = k) + P(s_j | i \neq k, \phi_p, \phi_g) P(i \neq k)} \tag{5}$$

We assume the conditional probabilities $P(s_j | i = k, \phi_p, \phi_g)$ and $P(s_j | i \neq k, \phi_p, \phi_g)$ to be Gaussian distributed and learn the parameters from data. Figure 7 shows example histograms of similarity values for the right eye region. The examples in Figure 7 show that the discriminative power of the right eye region diminishes as the probe pose changes from almost frontal (Figure 7(a)) to right profile (Figure 7(c)).

It is reasonable to assume that the pose of each gallery image is known. However, since the pose $\phi_p$ of the probe images is in general not known we marginalize over it. We can then compute the conditional densities for similarity value $s_j$ as

$$P(s_j | i = k, \phi_g) = \sum_p P(\phi_p) P(s_j | i = k, \phi_p, \phi_g)$$

and

$$P(s_j | i \neq k, \phi_g) = \sum_p P(\phi_p) P(s_j | i \neq k, \phi_p, \phi_g)$$

If no other knowledge about the probe pose is given, the pose prior $P(\phi_p)$ is assumed to be uniformly distributed. Similar to the posterior probability defined in (5) we compute the probability of the unknown probe image coming from the same subject (given similarity value $s_j$ and gallery pose $\phi_g$) as

$$P(i = k | s_j, \phi_g) = \frac{P(s_j | i = k, \phi_g) P(i = k)}{P(s_j | i = k, \phi_g) P(i = k) + P(s_j | i \neq k, \phi_g) P(i \neq k)} \tag{6}$$

In order to decide on the most likely identity of an unknown probe image $I_{i,p} = (i, \phi_p)$ we compute match probabilities between $I_{i,p}$ and all gallery images for all

(a) c27 vs. c5    (b) c27 vs. c37    (c) c27 vs. c22
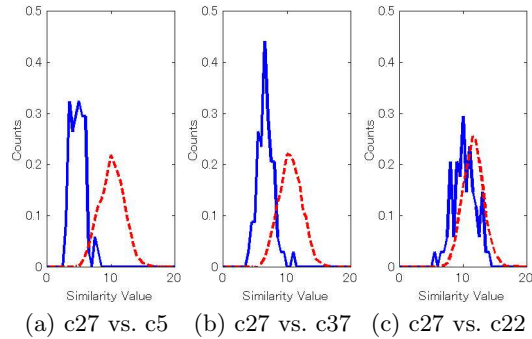
**Fig. 7.** Histograms of similarity values $s_j$ for the right eye region across multiple poses. The distribution of similarity values for identical gallery and probe subjects are shown with solid curves, the distributions for different gallery and probe subjects are shown with broken curves.

face subregions (using Equation (5) or (6)). We currently do not model dependencies between subregions, so we simply combine the different probabilities using the sum rule [30] and choose the identity of the gallery image with the highest score as recognition result.

### 3.3 Experimental Results

We used half of the 68 subjects in the CMU PIE database for training of the models described in Section 3.2. The remaining 34 subjects are used for testing. The images of all 68 subjects are used in the gallery. We compare our algorithm to eigenfaces [48] and the commercial FaceIt system.

### Experiment 1: Unknown Probe Pose

For the first experiment we assume the pose of the probe images to be unknown. We therefore must use Equation (6) to compute the posterior probability that probe and gallery images come from the same subject. We assume $P(\phi_p)$ to be uniformly distributed, i.e. $P(\phi_p) = \frac{1}{13}$. Figure 8 compares the recognition accuracies of our algorithm with eigenfaces and FaceIt for frontal gallery images. Our system clearly outperforms both eigenfaces and FaceIt. Our algorithm shows good performance up until 45° head rotation between probe and gallery image (poses 02 and 31). The performance of eigenfaces and FaceIt already drops at 15° and 30° rotation, respectively.

### Experiment 2: Known Probe Pose

In the case of known probe pose we can use Equation (5) to compute the probability that probe and gallery images come from the same subject. Figure 9 compares the
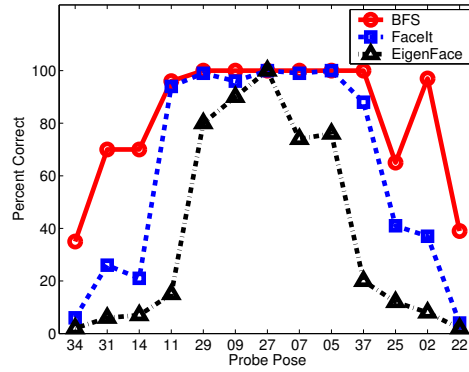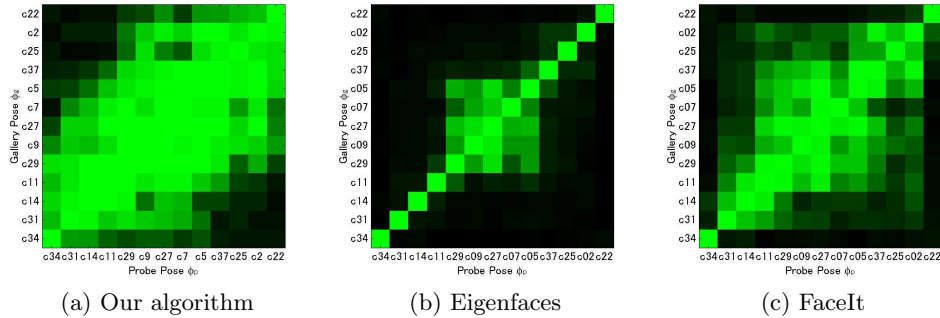
**Fig. 8.** Recognition accuracies for our algorithm (labeled 'BFS'), eigenfaces and FaceIt for frontal gallery images and unknown probe poses. Our algorithm clearly outperforms both eigenfaces and FaceIt.
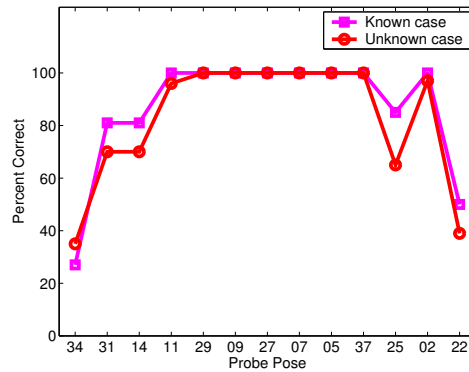


**Fig. 9.** Comparison of recognition accuracies of our algorithm for frontal gallery images for known and unknown probe poses. Only small differences are visible.

recognition accuracies of our algorithm for frontal gallery images for known and unknown probe poses. Only small differences in performances are visible. Figure 10 shows recognition accuracies for all three algorithms for all possible combinations of gallery and probe poses. The area around the diagonal in which good performance is achieved is much wider for our algorithm than for either eigenfaces or FaceIt. We therefore conclude that our algorithm generalizes much better across pose than either eigenfaces or FaceIt.

(a) Our algorithm          (b) Eigenfaces          (c) FaceIt

**Fig. 10.** Recognition accuracies for our algorithm, eigenfaces and FaceIt for all possible combinations of gallery and probe poses. Here lighter pixel values correspond to higher recognition accuracies. The area around the diagonal in which good performance is achieved is much wider for our algorithm than for either eigenfaces or FaceIt.

## 4 Face Recognition Across Pose and Illumination

Since appearance-based methods use image intensities directly they are inherently sensitive to variations in illumination. Drastic changes in illumination such as between indoor and outdoor scenes therefore cause significant problems for appearance-based face recognition algorithms [25, 39]. In this section we describe two different ways of handling illumination variations in facial imagery. The first algorithm extracts illumination invariant subspaces by extending the previously introduced eigen light-fields to Fisher light-fields [24], mirroring the step from eigenfaces [48] to Fisherfaces [4]. The second approach combines Bayesian face subregions with an image preprocessing algorithm, that removes illumination variation prior to recognition [21]. In both cases we demonstrate results for face recognition across pose *and* illumination.

### 4.1 Fisher Light-Fields

Suppose we are given a set of light-fields $L_{i,j}(\theta, \phi)$, $i = 1, \ldots, N, j = 1, \ldots, M$ where each of $N$ objects $O_i$ is imaged under $M$ different illumination conditions. We could proceed as described in Section 2.1 and perform Principal Component Analysis on the whole set of $N \times M$ light-fields. An alternative approach is Fisher's Linear Discriminant (FLD) [15], also known as Linear Discriminant Analysis (LDA) [52], which uses the available class information to compute a projection better suited for discrimination tasks. Analogous to the algorithm described in Section 2.1, we now find the least squares solution to the set of equations:

$$L(\theta, \phi) - \sum_{i=1}^{m} \lambda_i W_i(\theta, \phi) = 0 \tag{7}$$

where $W_i, i = 1, \ldots, m$ are the generalized eigenvectors computed by LDA.

**Table 3.** A comparison of the performance of eigen light-fields and Fisher light-fields with FaceIt on three different face recognition across pose and illumination scenarios. In all three cases, eigen light-fields and Fisher light-fields outperform FaceIt by a large margin.

|  | Eigen Light-Fields | Fisher Light-Fields | FaceIt |
|---|---|---|---|
| Same pose, Different illumination | - | 81.1 | 41.6 |
| Different pose, Same illumination | 72.9 | - | 25.8 |
| Different pose, Different illumination | - | 36.0 | 18.1 |

## Experimental Results

For our face recognition across pose and illumination experiments we used the pose and illumination subset of the CMU PIE database [46]. In this subset, 68 subjects are imaged under 13 different poses and 21 illumination conditions. Many of the illumination directions introduce fairly subtle variations in appearance, so we selected 12 of the 21 illumination conditions which span the set widely. In total we used $68 \times 13 \times 12 = 10,6084$ images in the experiments.

We randomly select 34 subjects of the PIE database for the generic training data and then remove this data from the experiments (see Section 2.2). There are then a variety of ways of selecting the gallery and probe images from the remaining data:
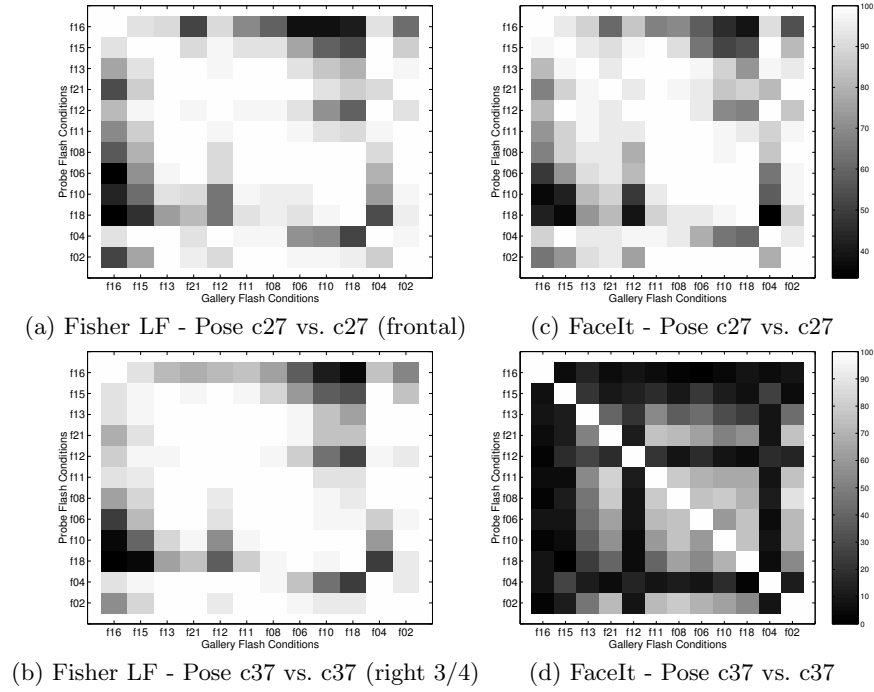
Same Pose, Different Illumination: The gallery and probe poses are the same. The gallery and probe illuminations are different. This scenario is like traditional face recognition across illumination, but is performed separately for each pose.

Different Pose, Same Illumination: The gallery and probe poses are different. The gallery and probe illuminations are the same. This scenario is like traditional face recognition across pose, but is performed separately for each possible illumination.

Different Pose, Different Illumination: Both the pose and illumination of the probe and gallery are different. This is the hardest and most general scenario.

We compare our algorithms with FaceIt under these three scenarios. In all cases we generate every possible test scenario and then average the results. For "same pose, different illumination", for example, we consider every possible pose. We generate every pair of disjoint probe and gallery illumination conditions. We then compute the average recognition rate for each such case. We average over every pose and every pair of distinct illumination conditions.

The results are included in Table 3. For "same-pose, different illumination," the task is essentially face recognition across illumination separately for each pose. In this case, it makes little sense to try eigen light-fields since we know how poorly eigenfaces performs with illumination variation. Fisher light-fields becomes Fisher faces for each pose which empirically we find outperforms FaceIt. Example illumination "confusion matrices" are included for two poses in Figure 11.

(a) Fisher LF - Pose c27 vs. c27 (frontal)    (c) FaceIt - Pose c27 vs. c27

(b) Fisher LF - Pose c37 vs. c37 (right 3/4)    (d) FaceIt - Pose c37 vs. c37

**Fig. 11.** Example "confusion matrices" for the "same-pose, different illumination" task. For a given pose, and a pair of distinct probe and gallery illumination conditions, we color-code the average recognition rate. The superior performance of Fisher light-fields is witnessed by the lighter color of (a–b) over (c–d).

For "different pose, same illumination," the task reduces to face recognition across pose but for a variety of different illumination conditions. In this case there is no intra-class variation and so it makes little sense to apply Fisher light-fields. This experiment is the same as Experiment 1 in Section 2.3 but the results are averaged over every possible illumination condition. As we found for Experiment 1, eigen light-fields outperforms FaceIt by a large amount.

Finally, in the "different pose, different illumination" task both algorithms perform fairly poorly. However, the task is very difficult. If the pose and illumination are both extreme, almost none of the face is visible. Since this case might occur in either the probe or the gallery, the chances that such a difficult case occurs is quite large. Although more work is needed on this task, note that Fisher light-fields still outperforms FaceIt by a large amount.

### 4.2 Illumination Invariant Bayesian Face Subregions

In general, an image $I(x, y)$ is regarded as product $I(x, y) = R(x, y)L(x, y)$ where $R(x, y)$ is the reflectance and $L(x, y)$ is the illuminance at each point $(x, y)$ [27].

Original PIE images



Processed PIE images
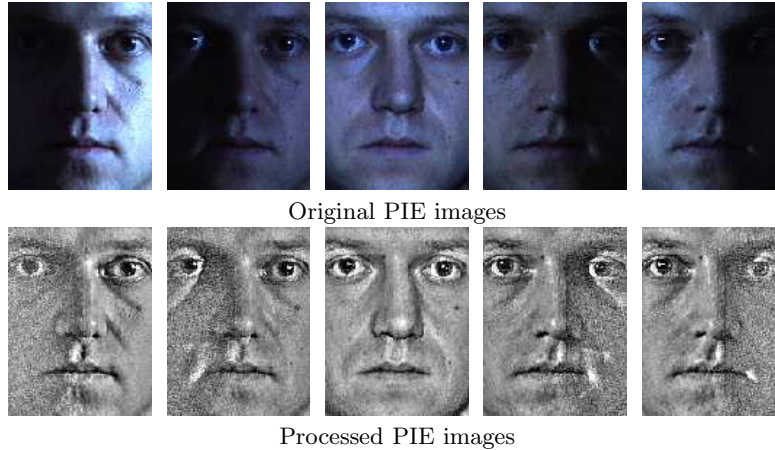
**Fig. 12.** Result of removing illumination variations with our algorithm for a set of images from the PIE database.

Computing the reflectance and the illuminance fields from real images is, in general, an ill-posed problem. Our approach uses two widely accepted assumptions about human vision to solve the problem: 1) human vision is *mostly* sensitive to scene reflectance and *mostly* insensitive to the illumination conditions, and 2) human vision responds to local changes in contrast rather than to global brightness levels. Our algorithm computes an estimate of $L(x, y)$ such that when it divides $I(x, y)$ it produces $R(x, y)$ in which the local contrast is appropriately enhanced. We find a solution for $L(x, y)$ by minimizing

$$J(L) = \iint_\Omega \rho(x, y)(L - I)^2 dx dy + \lambda \iint_\Omega (L_x^2 + L_y^2) dx dy \qquad (8)$$

Here $\Omega$ refers to the image. The parameter $\lambda$ controls the relative importance of the two terms. The space varying permeability weight $\rho(x, y)$ controls the anisotropic nature of the smoothing constraint. See [21] for details. Figure 12 shows examples from the CMU PIE database before and after processing with our algorithm. We use this algorithm to normalize the images of the combined pose and illumination subset of the PIE database. Figure 13 compares the recognition accuracies of the Bayesian Face Subregions algorithm for original and normalized images using gallery images with frontal pose and illumination. The algorithm achieves better performance on normalized images across all probe poses. Overall the average recognition accuracy improves from 37.3% to 44%.

## 5 Conclusion

One of the most successful and well-studied approaches to object recognition is the *appearance-based* approach. The defining characteristic of appearance-based algo-
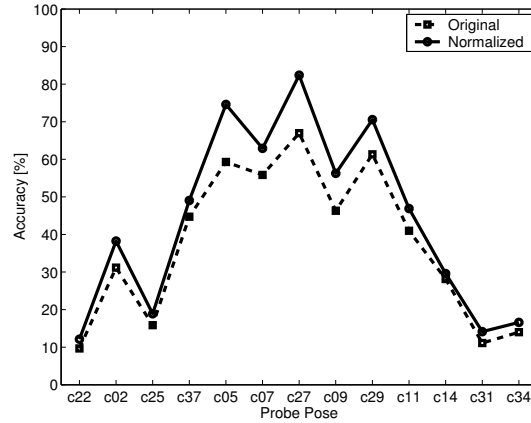
**Fig. 13.** Recognition accuracies of the Bayesian Face Subregions algorithm on original and normalized images using gallery images with frontal pose and illumination. For each probe pose the accuracy is determined by averaging the results for all 21 different illumination conditions. The algorithm achieves better performance on normalized images across all probe poses. The probe pose is assumed to be known.

rithms is that they directly use the pixel intensity values in an image of the object as the features on which to base the recognition decision. In this chapter we described an appearance-based method for face recognition across pose based on an algorithm to estimate the eigen light-field from a collection of images. Unlike previous appearance-based methods our algorithm can use any number of gallery images captured from arbitrary poses and any number of probe images also captured from arbitrary poses. The gallery and probe poses do not need to overlap. We showed that our algorithm can reliably recognize faces across pose and also take advantage of the additional information contained in widely separated views to improve recognition performance if more than one gallery or probe image is available.

In eigen light-fields all face pixels are treated equally. However, differences exist in how the appearance of various face regions change across face poses. We described a second algorithm, Bayesian face subregions, which derives a model for these differences and successfully employs it for face recognition across pose. Finally, we demonstrated how to extend both algorithms towards face recognition across both pose and illumination. Note, however, that for this task recognition accuracies are significantly lower, suggesting that there still is room for improvement. For example the model-based approach of Romdhani [41] recently achieved better results across pose on the PIE database then the appearance-based algorithms described here.

## Acknowledgments

## References

[1] E. Adelson and J. Bergen. The plenoptic function and elements of early vision. In Landy and Movshon, editors, *Computational Models of Visual Processing*. MIT Press, 1991.

[2] 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.

[3] R. Basri and D. Jacobs. Lambertian reflectance and linear subspaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(2):218–233, 2003.

[4] 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):711–720, 1997.

[5] P. Belhumeur and D. Kriegman. What is the set of images of an object under all possible lighting conditions. *International Journal of Computer Vision*, 28(3):245–260, 1998.

[6] P. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):711–720, July 1997.

[7] D. Beymer. Face recognition under varying pose. Technical Report 1461, MIT AI Laboratory, 1993.

[8] D. Beymer and T. Poggio. Face recognition from one example view. A.I. Memo No. 1536, MIT AI Lab, 1995.

[9] M. Black and A. Jepson. Eigen-tracking: Robust matching and tracking of articulated objects using a view-based representation. *International Journal of Computer Vision*, 36(2):101–130, 1998.

[10] D. Blackburn, M. Bone, and P. J. Philips. Facial recognition vendor test 2000: evaluation report, 2000.

[11] V. Blanz, S. Romdhani, and T. Vetter. Face identification across different poses and illumination with a 3D morphable model. In *Proceedings of the Fifth International Conference on Face and Gesture Recognition*, pages 202–207, 2002.

[12] V. Blanz and T. Vetter. Face recognition based on fitting a 3D morphable model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(9):1063–1074, 2003.

[13] T. Cootes, G. Edwards, and C. Taylor. Active appearance models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6):681–685, 2001.

[14] T. Cootes, G. Wheeler, K. Walker, and C. Taylor. View-based active appearance models. *Image and Vision Computing*, 20:657–664, 2002.

[15] K. Fukunaga. *Introduction to statistical pattern recognition*. Academic Press, 1990.

[16] A. Georghiades, P. Belhumeur, and D. Kriegman. From few to many: Generative models for recognition under variable pose and illumination. In *Proceedings of the Fourth International Conference on Face and Gesture Recognition*, pages 277–284, 2000.

[17] A. Georghiades, D. Kriegman, and P. Belhumeur. Illumination cones for recognition under variable lighting: Faces. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1998.

[18] A. Georghiades, D. Kriegman, and P. Belhumeur. From few to many: Generative models for recognition under variable pose and illumination. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6):643–660, 2001.

[19] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen. The lumigraph. In *Computer Graphics Proceedings, Annual Conference Series (SIGGRAPH)*, pages 43–54, 1996.

[20] D. Graham and N. Allinson. Face recognition from unfamiliar views: subspace methods and pose dependency. In *3rd International Conference on Automatic Face and Gesture Recognition*, pages 348–353, 1998.

[21] R. Gross and V. Brajovic. An image pre-processing algorithm for illumination invariant face recognition. In *4th International Conference on Audio-and Video Based Biometric Person Authentication (AVBPA)*, pages 10–18, June 2003.

[22] R. Gross, I. Matthews, and S. Baker. Appearance-based face recognition and light-fields. Technical Report CMU-RI-TR-02-20, Robotics Institute, Carnegie Mellon University, 2002.

[23] R. Gross, I. Matthews, and S. Baker. Eigen light-fields and face recognition across pose. In *Proceedings of the Fifth International Conference on Face and Gesture Recognition*, pages 1–7, 2002.

[24] R. Gross, I. Matthews, and S. Baker. Fisher light-fields for face recognition across pose and illumination. In *Proceedings of the German Symposium on Pattern Recognition (DAGM)*, pages 481–489, 2002.

[25] R. Gross, J. Shi, and J. Cohn. Quo vadis face recognition? In *Third Workshop on Empirical Evaluation Methods in Computer Vision*, 2001.

[26] B. Horn. Determining lightness from an image. *Computer Graphics and Image Processing*, 3(1):277–299, 1974.

[27] B. Horn. *Robot Vision*. MIT Press, 1986.

[28] D. Jacobs, P. Belhumeur, and R. Basri. Comparing images under variable illumination. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 610–617, 1998.

[29] T. Kanade and A. Yamada. Multi-subregion based probabilistic approach toward pose-invariant face recognition. In *IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA2003)*, pages 954–959, Kobe, Japan, 2003.

[30] J. Kittler, M. Hatef, R. Duin, and J. Matas. On combining classifiers. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 20(3):226–239, 1998.

[31] E. Land and J. McCann. Lightness and retinex theory. *Journal of the Optical Society of America*, 61(1), 1971.

[32] M. Lando and S. Edelman. Generalization from a single view in face recognition. In *International Workshop on Automatic Face-and Gesture-Recognition*, 1995.

[33] A. Lanitis, C. Taylor, and T. Cootes. Automatic interpretation and coding of face images using flexible models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):743–756, 1997.

[34] A. Leonardis and H. Bischof. Robust recognition using eigenimages. *Computer Vision and Image Understanding*, 78(1):99–118, 2000.

[35] M. Levoy and M. Hanrahan. Light field rendering. In *Computer Graphics Proceedings, Annual Conference Series (SIGGRAPH)*, pages 31–41, 1996.

[36] T. Maurer and C. von der Malsburg. Single-view based recognition of faces rotated in depth. In *International Workshop on Automatic Face and Gesture Recogition*, pages 248–253, Zurich, Switzerland, 1995.

[37] H. Murase and S. Nayar. Visual learning and recognition of 3-D objects from appearance. *International Journal of Computer Vision*, 14(1):5–24, 1995.

[38] A. Pentland, B. Moghaddam, and T. Starner. View-based and modular eigenspaces for face recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 84–91, 1994.

[39] P. J. Phillips, P. Grother, J. Ross, D. Blackburn, E. Tabassi, and M. Bone. Face recognition vendor test 2002: Evaluation report, March 2003.

[40] P. J. Phillips, H. Moon, S. Rizvi, and P. Rauss. The FERET evaluation methodology for face-recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1090–1104, 2000.

[41] S. Romdhani, V. Blanz, and T. Vetter. Face identification by matching a 3D morphable model using linear shape and texture error functions. In *Proceedings of the European Conference on Computer Vision*, pages 3–19, 2002.

[42] S. Romdhani, S. Gong, and A. Psarrou. Multi-view nonlinear active shape model using kernel PCA. In *10th British Machine Vision Conference*, volume 2, pages 483–492, 1999.

[43] S. Romdhani, A. Psarrou, and S. Gong. On utilising template and feature-based correspondence in multi-view appearance models. In *6th European Conference on Computer Vision*, volume 1, pages 799–813, 2000.

[44] A. Shashua. *Geometry and Photometry in 3D visual recognition*. PhD thesis, MIT, 1994.

[45] A. Shashua and T. Riklin-Raviv. The Quotient image: class-based re-rendering and recognition with varying illumination conditions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(2):129–139, 2001.

[46] T. Sim, S. Baker, and M. Bsat. The CMU pose, illumination, and expression database. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(12):1615–1618, 2003.

[47] T. Sim and T. Kanade. Combining models and exemplars for face recognition: An illuminating example. In *Workshop on Models versus Exemplars in Computer Vision*, 2001.

[48] M. Turk and A. Pentland. Face recognition using eigenfaces. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1991.

[49] T. Vetter and T. Poggio. Linear object classes and image synthesis from a single example image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):733–741, 1997.

[50] L. Wiskott, J. Fellous, N. Kruger, and C. von der Malsburg. Face recognition by elastic bunch graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):775–779, 1997.

[51] W. Zhao and R. Chellappa. Robust face recognition using symmetric shape-from-shading. Technical report, Center for Automation Research, University of Maryland, 1999.

[52] W. Zhao, A. Krishnaswamy, R. Chellappa, D. Swets, and J. Weng. Discriminant analysis of principal components for face recognition. In H. Wechsler, P. J. Phillips, V. Bruce, and T. Huang, editors, *Face Recognition: From Theory to Applications*. Springer Verlag, 1998.

# Index