

# Biometrics Systems Concepts: 2D + 3D face recognition

Tom Decroos  
r0297757

June 20, 2016

## Abstract

This report discusses the implementation of a face recognition system that makes use of 2D and 3D face representations. First, three relevant papers are discussed: a paper about using eigenfaces for facial recognition, a paper about using local binary patterns and a paper about combining 2D and 3D face recognition. Next, we describe the implementation of our face recognition system. Our implementation uses methods that are heavily based on the three papers we read. Finally, experiments are performed for 2D face recognition, 3D face recognition and the combination of 2D and 3D face recognition.

## 1 Literature study

We discuss three relevant papers about face recognition systems. The first paper provides a holistic approach for recognizing faces. The used method is eigenfaces. The second paper provides a feature based approach. The used method is local binary patterns. For each paper, a small summary is provided. Afterwards, possible advantages and disadvantages of both approaches are discussed. We focus in our summaries and discussion mostly on the parts of the papers that are relevant for this report.

### 1.1 Holistic approach: eigenfaces

The paper we read for recognizing faces with a holistic approach is ‘Use of depth and colour eigenfaces for face recognition’ by F. Tsalakanidou, D. Tzovaras and M.G. Strintzis [3]. This paper first discusses the extraction of depth maps. Most face recognition systems are based on the evaluation of 2D-intensity or colour images. These systems are subject to a variety of errors caused by differences in illumination, head rotation, face pigment or facial expressions. Biometric features from 3D measurements are expected to perform well, due to their relative independence from these differences. One way of exploiting the 3D information of the geometric characteristics of a human face is by building

a depth map. The depth map of a face is a function giving for each pixel of an image the distance to the camera. Such a depth map can be extracted from a set of 3D coordinates of points using the z-buffer algorithm.

Next, the classification method is discussed. The paper proposes a powerful global approach (PCA) in which the whole image (colour or depth map) serves as a feature vector and the variability of the face is modelled by analyzing its statistical properties using a large set of faces. PCA is a well-known, widely and successfully used method. The application of PCA for face recognition is very simple: PCA is performed in a well-defined set of images of human faces and a set of principal components is obtained. These principal components are called eigenfaces. Every face in the database can then be represented as a vector of weights. The weights are obtained by projecting the image of the face on the eigenfaces. Identification of a new test image is then done by locating the image in the database whose weights have the smallest euclidean distance from the weights of the test image. The authors of the paper observed experimentally that the accuracy of this approach quickly degrades if there are changes in scale. This can be explained by the low correlation between face images at different scales. Hence, to successfully apply this approach, all faces in the database must be uniformly scaled.

This approach can only be applied on single channel signals such as a depth map or a grayscale image. In case of face recognition from full colour images, this approach is applied to each component of the colour signal. The PCA algorithm is performed for training sets of each component. When a new image needs to be recognized, its projection to the subspace of each component is computed. The image is then assigned to the class that results in the smallest product of the euclidean distances of each component. This combination approach based on the product of the euclidean distances of each component can also be used to combine a colour and a depth map for 2D + 3D face recognition.

Finally, the eigenfaces approach was evaluated on three representations of faces: colour map, depth map and combination of colour and depth. Experiments were performed on images of 40 people. For each person, there were four colour images and two depth maps available. The training set consisted of two colour images and one depth map for each person. The test set consisted of the remaining two colour images and depth map for each person. The eigenfaces approach achieves a recognition rate of 87.5% using only colour images, 85% using depth maps and 97.5% using the combination. These results show that adding depth information leads to significant gains in recognizing faces.

## 1.2 Feature based approach: local binary patterns

The paper we read for recognizing faces with a feature based approach is ‘Face Recognition with Local Binary Patterns’ by T. Ahonen, A. Hadid and M. Pietikäinen [1]. The paper first explains the concept of local binary patterns (LBP). The original LBP operator is a powerful means of texture description. The operator labels the pixels of an image by thresholding the 3x3 neighbourhood of each pixel with the center value and considering the result as a binary

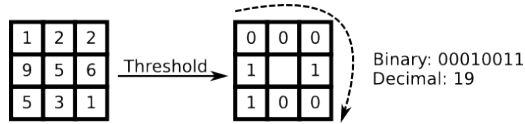


Figure 1: The basic LBP operator.

number (Figure 1). This basic operator can be extended by using neighbourhoods of different sizes and using only uniform patterns.

The histogram of the labels can then be used as a texture descriptor. This histogram contains information about the distribution of the local micropatterns, such as edges, spots and flat areas, over the whole image. For efficient face representation, one should also retain spatial information. For this purpose, the image is divided into different regions. The histograms of each separate region are then concatenated and form a spatially enhanced histogram. In this histogram, we effectively have a description of the face on three different levels of locality: the labels for the histogram contain information about the patterns on a pixel-level, the labels are summed over a small region to produce information on a regional level and the regional histograms are concatenated to build a global description of the face.

Several possible dissimilarity measures exist for histograms, such as histogram intersection, Log-likelihood statistic and Chi square statistic ( $\chi^2$ ). When the image has been divided into regions, it can be expected that some of the regions contain more useful information than others in terms of distinguishing between people. For example, eyes seem to be an important cue in human face recognition. To take advantage of this, a weight can be set for each region based on the importance of the information it contains.

The paper used the CSU Face Identification Evaluation System for testing the performance of its proposed algorithm. The results show that  $\chi^2$  was the best distance measure between histograms for recognizing faces. They also show that the LBP algorithm is quite robust with respect to its parameters, such as neighbourhood size. Changes in the parameters may cause big differences in the length of the feature vector, but the overall performance is not necessarily affected significantly. When compared to other algorithms such as Eigenfaces, Fisherfaces or Elastic Bunch Graph Matching, the LBP algorithm significantly outperforms them all.

### 1.3 Discussion

Experiments in [2] suggest that the Eigenfaces approach works best when used on a face space that was generated from a lot of images with diverse conditions (eg. lighting or facial expression). This might be a problem, because our training and test set only consist of five persons each. This means that our face space would be based on only five images. We fear that this face space would incorporate too many specific characteristics of these five images, which would

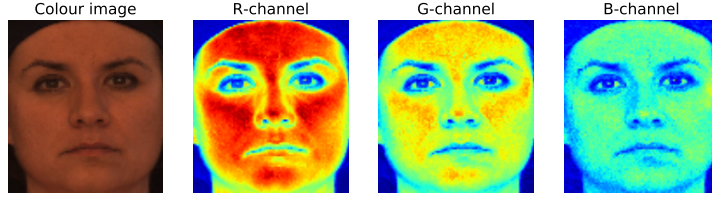


Figure 2: Example of a colour image being split into its three different colour channels

lead to a less accurate result on new images. The LBP algorithm generates features from each image in an independent way, the size of the dataset does not affect its performance. This is a useful characteristic when dealing with small datasets.

It has also been observed experimentally that the performance of the Eigenfaces approach degrades quickly as the scale changes. Intuitively, this is explained by the low correlation between faces and images at different scales. The LBP algorithm considers not only shape, but also texture information to represent face images. That is why we expect the LBP algorithm to be a bit more robust with respect to scale than the Eigenfaces approach.

## 2 Methods

In this section, we describe our methods for face recognition on 2D colour images and 3D point clouds. We describe the conversion of a colour image in three different channels and the conversion of a point cloud into a depth map. Next, we state our used similarity measure between channels and motivate this choice. We list possible ways of combining similarity measures of different channels. Finally, we explain a preprocessing procedure for 2D images.

### 2.1 Channels

Algorithms such as Eigenfaces or the LBP algorithm require images to be represented as a matrix  $X$ . The element  $x_{ij}$  then represents the value of the pixel at position  $(i, j)$ . Colour images do not have a single value per pixel, they have three. A colour image can easily be converted into a grayscale image, but then some information about the image is lost. To prevent information loss, all colour images are split in their three colour channels (Figure 2).

Converting a 3D point cloud into a 2D matrix that still represents a face is not as trivial however. We convert the point cloud into a depth map. This depth map should show a front view of the face. The value of each pixel in this depth map should represent how close the face was to the camera in the location of that pixel. We construct a depth map in the following way.

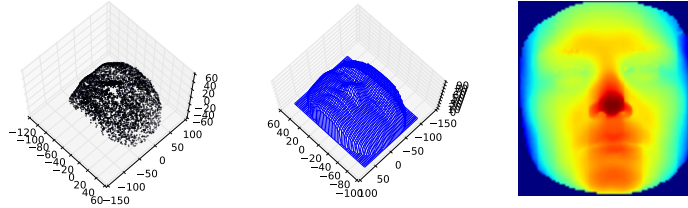


Figure 3: Conversion of a 3D point cloud of a face into a depth map by constructing a grid and interpolating the Z-coordinates of that grid.

The first step is to remove outlier points that definitely do not belong to the face. These outlier points can be detected using statistical measures such as the robust Mahalanobis distance or we can also set arbitrary thresholds based on some exploratory data analysis. All observed faces of the dataset used in this report showed no  $x, y$  or  $z$ -coordinates with an absolute value greater than 200, while all outliers always had the same specific coordinate: (999999, 999999, 999999). Since the 3D point clouds used in this report only contain one specific type of outlier, a threshold that filters these points is deemed good enough to remove all outliers. All data points should satisfy the constraint:  $|x| < 500 \wedge |y| < 500 \wedge |z| < 500$ .

### 3 Experiments

We have conducted experiments to attempt to answer the following research questions.

- Q1 Is it better to preprocess the face images or do the raw images perform better?
- Q2 Does fusion significantly raise the recognition rate?
- Q3 What is the contribution of depth to our face recognizer?
- Q4 What is the best way to fuse match scores?

### 4 Conclusion

### References

- [1] Timo Ahonen, Abdenour Hadid, and Matti Pietikäinen. Face recognition with local binary patterns. In *Computer vision-eccv 2004*, pages 469–481. Springer, 2004.

- [2] Kyong I Chang, Kevin W Bowyer, and Patrick J Flynn. An evaluation of multimodal 2d+ 3d face biometrics. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(4):619–624, 2005.
- [3] Filareti Tsalakanidou, Dimitrios Tzovaras, and Michael G Strintzis. Use of depth and colour eigenfaces for face recognition. *Pattern Recognition Letters*, 24(9):1427–1435, 2003.