Gender classification with visual and depth images

Yannick van den Hurk

ANR: S333437

HAIT Master Thesis series nr. 12-014

THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS IN COMMUNICATION AND INFORMATION SCIENCES,
MASTER TRACK HUMAN ASPECTS OF INFORMATION TECHNOLOGY,
AT THE SCHOOL OF HUMANITIES
OF TILBURG UNIVERSITY

Thesis committee:

Prof. Dr. E.O. Postma
Dr. Ir. P.H.M. Spronck

Tilburg University
School of Humanities
Department of Communication and Information Sciences
Tilburg center for Cognition and Communication (TiCC)
Tilburg, The Netherlands

August 2012

Summary

In the last couple of years there has been an increased interest in automatic face recognition systems. New techniques of imaging for subsequent face recognition have been proposed. These techniques include the use of high-quality visual and depth imaging equipment. The costs of digital cameras that can produce high-resolution still images have dropped significantly. Recently, the Kinect device became available, a low-cost depth imaging device developed by Microsoft.

Gender determination enhances automatic face recognition by limiting the required search to the class of faces within a gender class. In automatic face recognition machine learning techniques using Principal Component Analysis (PCA) and Lineair Discriminant Analysis (LDA) have been applied with good results. The application of PCA produces so-called Eigenfaces, images that visualize the main sources of variation in the data. The goal of this study is to find out, using PCA and LDA, if it is possible to determine the gender of faces using visual and depth images, both obtained with a Kinect device. In addition, the study aims to compare the gender recognition performance obtained with visual images with that obtained with depth images. The following three research questions are addressed in the thesis.

RQ1: Is it possible to determine gender by using visual and depth imagery obtained with a Microsoft Kinect device?

RQ2: If so, what is the difference in accuracy of gender determination between the obtained visual and depth imagery?

RQ3: What principal components are most informative for determining gender using visual and depth imagery?

To answer these research questions, we created a Gender Classification by Computation (GECCO) method, which uses visual and depth facial images to provide a gender classification of faces as an output. We employed a standardized photography setup and standard machine learning algorithms for preprocessing, classification and validation. The principal components generated using PCA are either ranked according to their Eigenvalues (from large to small), i.e., Eigenvalue Ranked Principal Components (ERPCs), or ranked according to their individual recognition performance, i.e., Performance Ranked Principal Components (PRPCs).

The experimental results indicate that gender determination is possible using visual and depth imagery obtained with a Kinect device. By ranking the Eigenfaces, the best separation of gender can be found using the first 7 depth ERPCs and first 15 visual ERPCs resulting in recognition accuracies of 94,08% (visual) and 90,48 (depth), respectively. When ranking the Eigenfaces on their best individual performances, the best gender recognition is obtained using the first 49 visual PRPCs and first 57 best depth PRPCs resulting in respectable separation accuracies of 100% (visual) and 97,62% (depth). Although PRPCs outperform ERPCs, they require more computational effort.

Table of contents

1. Introduction

In the last couple of years there has been an increased interest in automatic face recognition systems. According to Zhao et al. (2003) there are at least two reasons for this trend: the first is the wide range of commercial and law enforcement applications and the second is the availability of feasible technologies after 30 years of research. In contrast to many forms of biometrical identification methods for security purposes, like fingerprint and retinal scans, automatic face recognition does not necessarily rely on the cooperation of the participants. Some typical examples of face recognition can be found in areas like entertainment, smartcards (passports, licenses), information security, law enforcement and surveillance (Zhao et al., 2003).

New techniques of imaging for subsequent face recognition have been proposed. Two main innovations are recognition using high-resolution still images and from three-dimensional (depth) scans (Bowyer, 2004, Phillips et al, 2005). In their overview of the Face Recognition Grand Challenge Phillips et al. (2005) suggest that face recognition applications require either depth information on faces, or multiple facial images, or high resolution facial images, without advocating either one. Some studies (Bowyer et al. 2004, Tsalakanidou et al. 2004) suggest that a combination of visual and depth information can achieve more accurate face recognition. Due to price drops, technological advancements and general availability of depth imaging technology this kind of research is still liable to further research and improvement (Bowyer, 2004).

Due to their broad commercial availability, costs of digital cameras that can produce high-resolution still images have dropped significantly. Until two years ago depth imaging equipment was reserved for professionals and enthusiasts who had access to the required financial resources. This situation changed when the Microsoft Corporation launched its Kinect device in 2010, utilizing technology licensed from PrimeSense (2011). The device, originally intended for entertainment purposes, contains both a RGB visual image sensor and a depth sensor. In the beginning of 2011 the Microsoft Corporation released a software development kit (SDK) for its Kinect device so that it can be used for various other purposes besides entertainment. Using the Software Development Kit SDK it is possible to acquire relatively high resolution still and depth images.

According to Zhao et al. (2003) gender determination enhances automatic face recognition by reducing the search space leading to faster and better results. A technique called Eigenface decomposition described and successfully applied in a number of studies (Kirby & Sirovich 1987,1990; Turk & Pentland 1991a,-1991b; O'Toole 1991,1993,1994; Valentin, 1994; Tsalakanidou, 2003) can be used as support for gender detection. Eigenfaces are created by applying Principal Component Analysis (Jollife, 2002) to a collection of human face images. Before the conception and application of Eigenfaces, local facial features were used for face detection by comparing the distance between ears and eyes (Kirby & Sirovich 1987,1990). These local methods turned out not to be very effective.

PCA is used for dimensionality reduction. PCA rotates the original data space (i.e., the space spanned by the values of the individual pixels in an image), so that the rotated axes form ranked principal components. Projecting the data points (i.e., images) onto the principal components (axes) yields the largest projected variance on the first principal component, the one-but-largest variance on the second principal component, and so forth. The front-ranked components often contain the most important or significant characteristics of the data (Jollife, 2002).

When applying PCA to images of faces, the components correspond to Eigenfaces. These Eigenfaces may be useful for automatic gender classification (O'Toole 1994, Buchala 2004-2005).

In a study by O'Toole et al. (1996), supported by Etemad (1997) and Buchala et al. (2004-2005) it is stated that "information related to the gender property is shared by most of the faces and hence (it) is encoded in the first few (principal) components". To see which principal components contain the most specific gender information the data should be classified based on gender.

Lineair Discriminant Analysis (LDA) is a statistical classification method in machine learning to separate two or more classes and is related to Fisher's linear discriminant (Fisher, 1936). The lineair classifier can be "used to identify which class (or group) the object belongs by making a classification decision, based on the value of a linear combination of the characteristics" (Anila & Devarajan, 2011). Zhao et al. (1998) demonstrate that when principal components rather than original images are fed to a Lineair Discriminant Analysis (LDA) classifier this significantly improves image recognition.

Some earlier research on gender recognition found that facial depth structure information yield better recognition results than recognition based on visual images (O'Toole et al, 1996). Using a local region-based feature extraction and SVM classification a gender separation performance of 94.2% was achieved by Ben-Abdelkader and Griffin (2005). In another study, utilizing a number of different techniques involving a rectangle feature vector with support vector machine classification, Shen et al. (2009) achieved an accuracy of 92,4%. In other research concerning visual images, Bekios-Calfa et al. (2011) found gender separation accuracies as high as 95,4% using lineair discriminant techniques. This study also states that if training data is scarce (300 images or less) and relatively homogeneous, the combination of PCA and LDA is the best choice .

1.1 Problem statement

The following three research questions are addressed in this thesis. Determination:

RQ1: Is it possible to determine gender by using visual and depth imagery obtained with a Microsoft Kinect device?

RQ2: If so, what is the difference in accuracy of gender determination between the obtained visual and depth imagery?

RQ3: What principal components are most informative for determining gender using visual and depth imagery?

1.2 Outline

The outline of the rest of this paper is as follows. In chapter 2 we describe our gender classification by computation (GECCO) method based on PCA and LDA. Subsequently chapter 3 describes the experimental methodology used to acquire a database of visual and depth facial images, extracting their features and separating between genders. In chapter 4 we will present the obtained results. The validity, implications and points of improvement of our findings are addressed in the discussion in section 5. We will then conclude on our findings in chapter 6.

2. Gender Classification by Computation (GECCO)

The GECCO method consists of three components as illustrated in figure 1. For input our GECCO method uses visual and depth facial images to provide a gender classification of these faces as an output. In this section we will explain these components beginning with what the sampling component of the GECCO sytem is (2.1). Next we will explain the pre-processing step of the GECCO method concerning PCA (2.2). Finally the classification component of the GECCO method will be discussed concerning the implementation of a LDA algorithm for gender separation and validation (2.3).

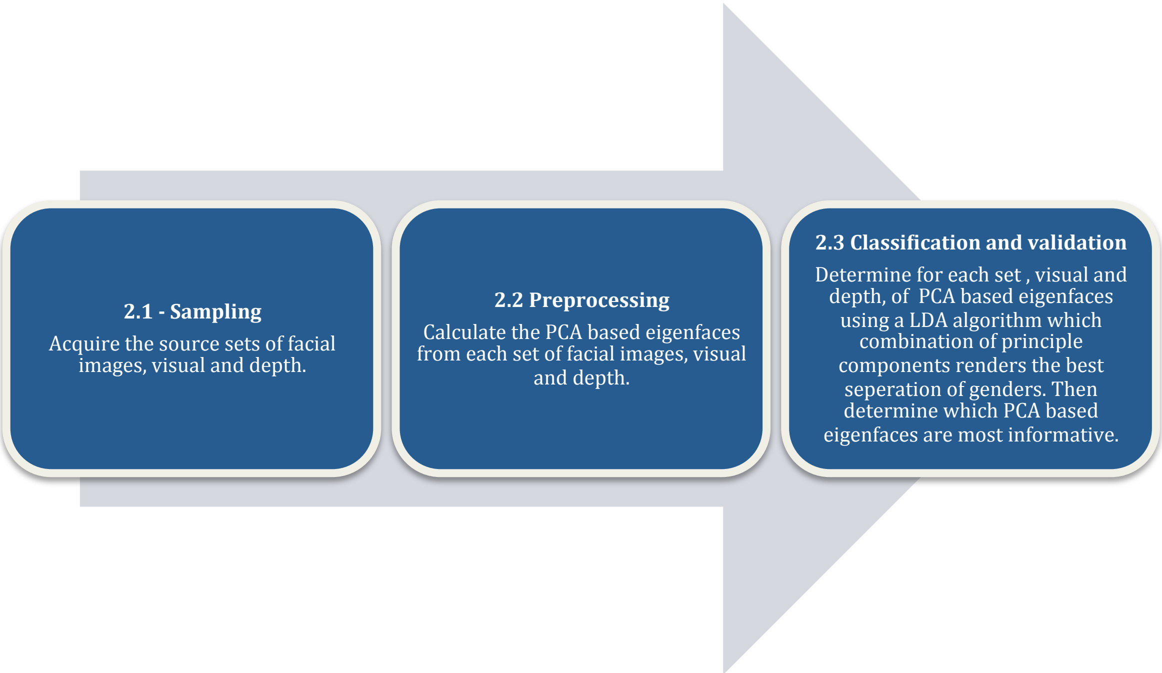


Figure 1 - Outline of GECCO method

2.1 Sampling

The sampling component of the GECCO method consists of acquiring source sets of facial images, visual and depth, for further processing. There are multiple different facial databases like the UCN, FERET and PAL datasets but these all lack depth images. The dataset for this study is being newly created due to quality and uniformity considerations and the ability to create relatively high quality depth images. For recording both the visual and depth images a Kinect device, utilizing software developed using its SDK, is used. The Kinect device is capable of capturing both visual Red Green Blue (RGB) and depth information. The visual RGB images are captured with a resolution of 1280x1024 pixels. The depth images are created by an embedded system called LightCoding. This system uses a monochrome Complementary Metal Oxide Semiconductor (CMOS) sensor to capture a laser grid that is projected by an infrared laser projector. The images of a mixed group of participants (preferably 50% male, 50% female) are captured in a lab under different lighting conditions and gaze directions, at different distances from the Kinect device. The images acquired using this method are to be stored on a computer system for further processing.

2.2 Pre-processing

The pre-processing component of the GECCO method is an Eigenface decomposition. The implementation of PCA for creating Eigenfaces mandates that the source face images are to be normalised in size and are aligned at eye and mouth levels. Our GECCO method then uses PCA to calculate Eigenfaces for both sets of facial imagery, visual and depth. Assuming that a set contains X face images, the PCA calculations generate X principal components/Eigenfaces based on that set.

The PSEUDO code for generating Eigenfaces is described by Serrano (2004) as follows:

The first step is to obtain a set S with M face images. These face images have to be normilised in size and aligned at eye and mouth levels. Each image is transformed into a vector of size N and placed into the set. $S = \{\Gamma_1, \Gamma_2, \Gamma_3 \dots \Gamma_m\}$

When this set has been obtained, the mean image has to be calculated Ψ.

$$\Psi = \frac{1}{M}\sum_{n=1}^{M}\Gamma_n$$

In this step the difference Φ between the input image and the mean image has to be found. $\Phi_i = \Gamma_i - \Psi$

Next we seek a set of M orthonormal vectors, u_n, which best describes the distribution of the data. The k^{th} vector, u_k , is chosen such that $\lambda_k = \frac{1}{M}\sum_{n=1}^{M}(u_k^T\Phi_n)^2$ is a maximum, subject to $u_l^T u_k = \delta_{lk}\begin{cases}1\ if\ l = k\\ 0\ otherwise\end{cases}$.
Note that u_k and λ_k are the eigenvectors and eigenvalues of the covariance matrix C.

1. The covariance matrix C is obtained as follows $C = \frac{1}{M}\sum_{n=1}^{M}\Phi_n\Phi_n^T$ $= AA^t$, $A = \{\Phi_1, \Phi_2, \Phi_3 \dots, \Phi_n\}$.
2. $A^T, L_{mn} = \Phi_M^T\Phi_n$
3. Now the eigenvectors v_l, u_l have been found: $u_l = \sum_{K=1}^{M} v_{lk}\,\Phi_k \qquad l = 1, \dots, M$

Using this PCA the Eigenfaces can be generated. The number of principle components or Eigenfaces is equal to the number of source facial images.

Code 1 – PSEUDO code for generating Eigenfaces

2.3 Classification and validation

The classification component of the GECCO method consists of applying a LDA classifier. Labelled data is needed to enable this supervised machine-learning algorithm. This last step determines which set(s) of the generated principal components contain the most gender specific information for visual and depth PCA based Eigenfaces, using the source images labelled by gender. This way finding the best accuracy of gender separation for the visual and depth set separately.

3. Experimental setup

In the upcoming chapter we describe the design for our dataset creation that yielded the visual and depth facial images used to implement the GECCO system (3.1). In the subsequent sections we elaborate on the pre-processing (3.2) and the classification and validation (3.3) steps of our GECCO method. Finally we will end this chapter with the criteria the obtained results should meet (3.4).

3.1 Dataset creation

To generate a database of usable visual and depth facial images for our GECCO method we constructed a fixed photography setup, which allowed us to take a large amount of uniform pictures. A white walled, practically light-tight lab with existing fluorescent ceiling lighting was used to build this setup. In this lab a Microsoft Kinect camera was fixed to a stand at a height of approximately 1,5 meters and was placed 1,0 meter from the position a participant would have to take place to be photographed. Every participant was asked to assume a frontal gaze when being photographed. The schematic layout for this setup can be seen in figure 2.

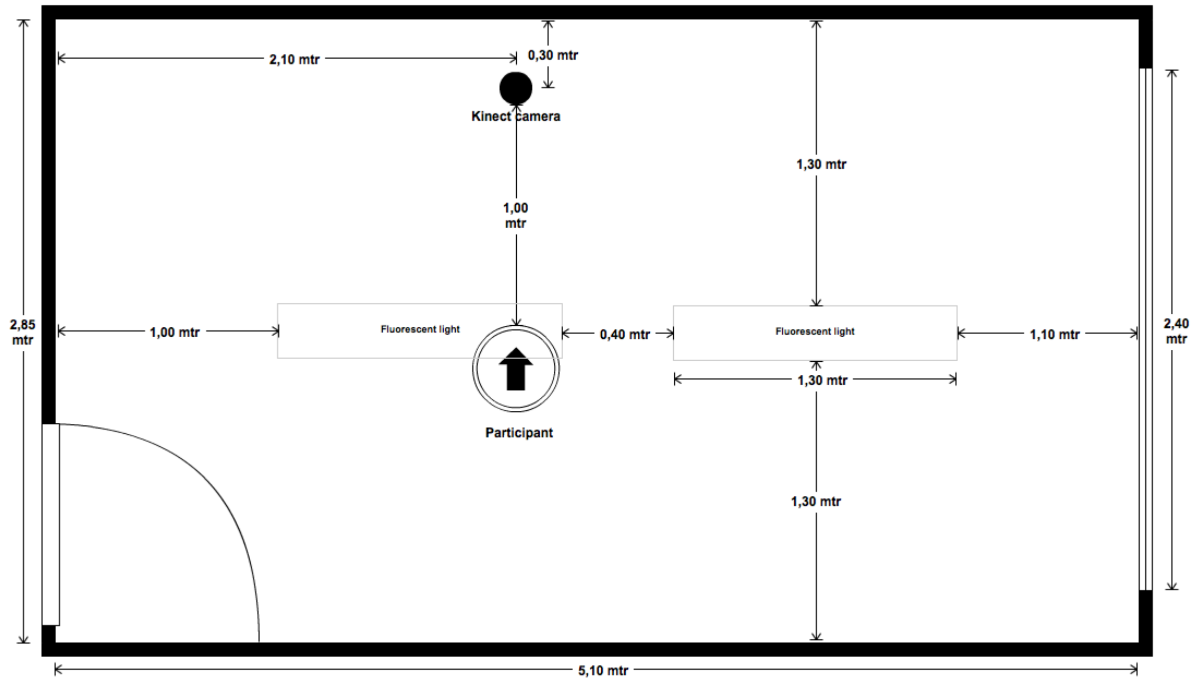


Figure 2 – Schematic layout of the experiment lab

The Kinect camera was connected to a computer workstation running custom software which allowed us at any given moment to simultaneously take a visual and depth picture of the participant, this way guaranteeing an exact match up of the visual and depth images. By using this setup we assembled a database with visual and depth images (1024x1280 pixels) of 101 participants, university students ranging in age from 18 to 60 (M = 22,15, SD=5,37). Due to quality considerations this set shrunk to a total of 84 usable face images, 39 male and 45 female. By assigning different colors to different pixel values extracted from the RAW Kinect output source, hue-saturation-value (HSV) color mapped depth images (640x480 pixels) were generated. To pixelmatch the imagery the visual images were cropped and the HSV color mapped depth images were scaled, this way achieving a resolution of 901x1201 pixels for each image.

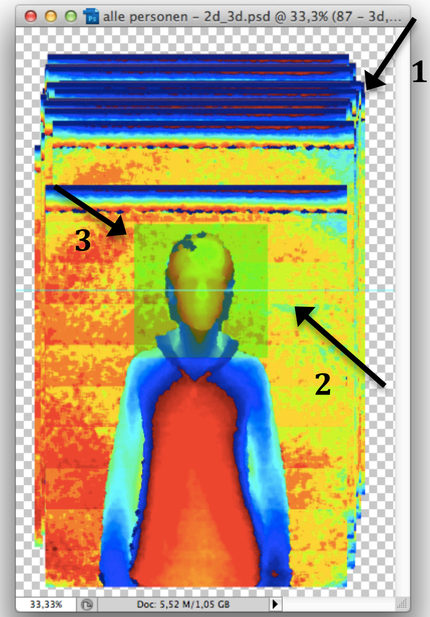


Figure 3 - Layering up images

Because the participants differ in height and the application of PCA requires a uniform image layout and orientation we used commercially available photo editing software (Photoshop CS5) to match the layered images up at the participants' eye level. The method for achieving this is shown in figure 3. Firstly the visual and depth images were imported as layers [1]. Using a ruler [2] these layers were lined up at the participants' eye level. Finally a square semi-transparant green mask [3] was used to determine which area was best to crop the layered images. Afterwards the area was cropped and every layer was exported to an individual image file. This generated two sets, visual and depth, of 84 head centred squared facial images of 400x400 pixels, resulting in the examples shown in figure 4a (visual) and figure 4b (depth).

Figure 4a – Example of visual face

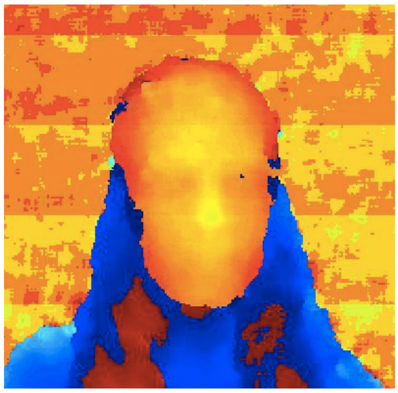


Figure 4b – Example of depth face

3.2 Pre-processing

Using a modified Matlab script originally written by Daily (2000) based on the PCA algorithm by Turk and Pentland (1991) we created two sets of PCA based Eigenfaces from the sets of 84 head centred squared facial images, visual and depth. The two datamatrices from which these Eigenfaces were drawn were saved for further processing in the classification and validation step of our GECCO method.

3.3 Classification and validation

For the classification step in our GECCO method an internal Matlab LDA algorithm was used to determine the best separation between the gender classes for each set, visual and depth. Due to the large number of principal components, equal to the number of Eigenfaces (84), two methods were developed to determine which combination of them resulted in the best accuracy of separation.

Due to software limitations the number of principal components assessed was reduced from 84 to 80, leaving out the four lowest eigenvalue ranked components. Using the first method the software starts to calculate the gender separation using the first 80 principal components ranked on their eigenvalues. In every subsequent calculation the set of to be assessed Eigenvalue Ranked Principal Components (ERPC) is then reduced by one untill all 80 sets have been processed. The schematic overview of this method is shown in figure 5a.

The second method firstly ranks every principal component from each set, visual and depth, individually for classifying gender using LDA. The goal is then to investigate which combination of these Performance Ranked Principle Components (PRPC) results in the best accuracy of gender separation. This is done by sorting the principal components from the previous calculation for each set, based on their performance. Subsequently the aim is determining the best gender separation for a set of these PRPCs. This is done by creating a set of the best first PRPC(s), calculating their gender separation capability, then adding the next best PRPC to this set and looping through this mechanism again as shown in figure 5b.

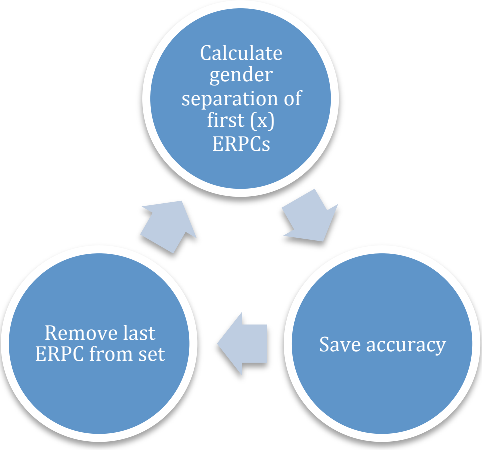


Figure 5a – Method 1

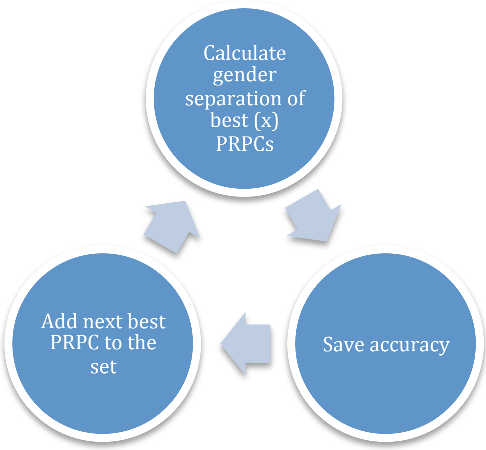


Figure 5b – Method 2

To validate these classification methods a leave-one-out cross validation (LOOCV) system is used. This system uses one observation from the data set as a validation for the training that is performed on the rest of the observations from that set. It rotates through the rest of the observations, making sure every observation in the data set has been used as a validation for training.

3.4 Criteria

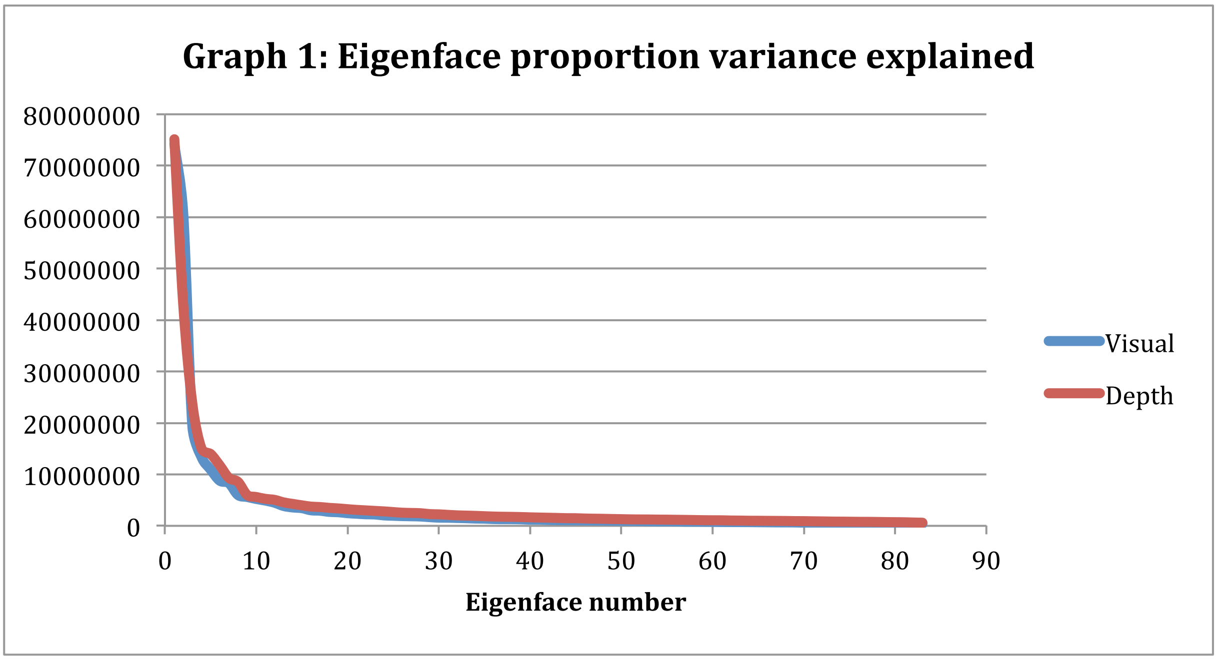
The accuracy of separation between the male and female classes has to meet some requirements. Firstly only a separation accuracy above 50% is up for evaluation because it at least defies chance. Secondly a separation accuracy closest to 100% is desirable. Furthermore a separation accuracy using the least number of principal components is also desirable due to the lower need for computational effort.

4. Results

In this chapter the results of our GECCO method will be explained beginning with the calculation of PCA based Eigenfaces. For understanding the examples of the first Eigenface from each set can be seen in figure 5a (visual) and 5b (depth). It shows that the first visual Eigenface has its highest variance in the hair and face region, being lighter and the shoulder region, being darker in hue. The first depth Eigenface shows its highest variance in the hair region, being lighter, and in the contour of the face and eye region, being darker in hue. As stated in section 2.2, PCA based Eigenfaces with a high ranking on eigenvalues depict the most variance of the data. Evidence that this is also the case for the dataset created for this study can be seen in table 1 (appendix) and graph 1.

Figure 6a - First visual Eigenface

Figure 6b – First depth Eigenface



In table 2 (appendix) and graph 2 the main results of this study are shown. These results indicate that it is possible to discriminate between genders accurately using visual as well as depth imagery.

Using the first method, by ranking the principal components on eigenvalues, the best gender discrimination for visual imagery can be found using the first 15 ERPCs, which render a separation accuracy of 94,05% based on the LDA using LOOCV. The best gender discrimination for depth imagery can be found using the first 7 ERPCs, which render a separation accuracy of 90,48% based on the LDA using LOOCV. The data shows that gender discrimination is best using visual information and drops when taking more than the first 15 ERPCs (visual) or 7 ERPCs (depth) into account.

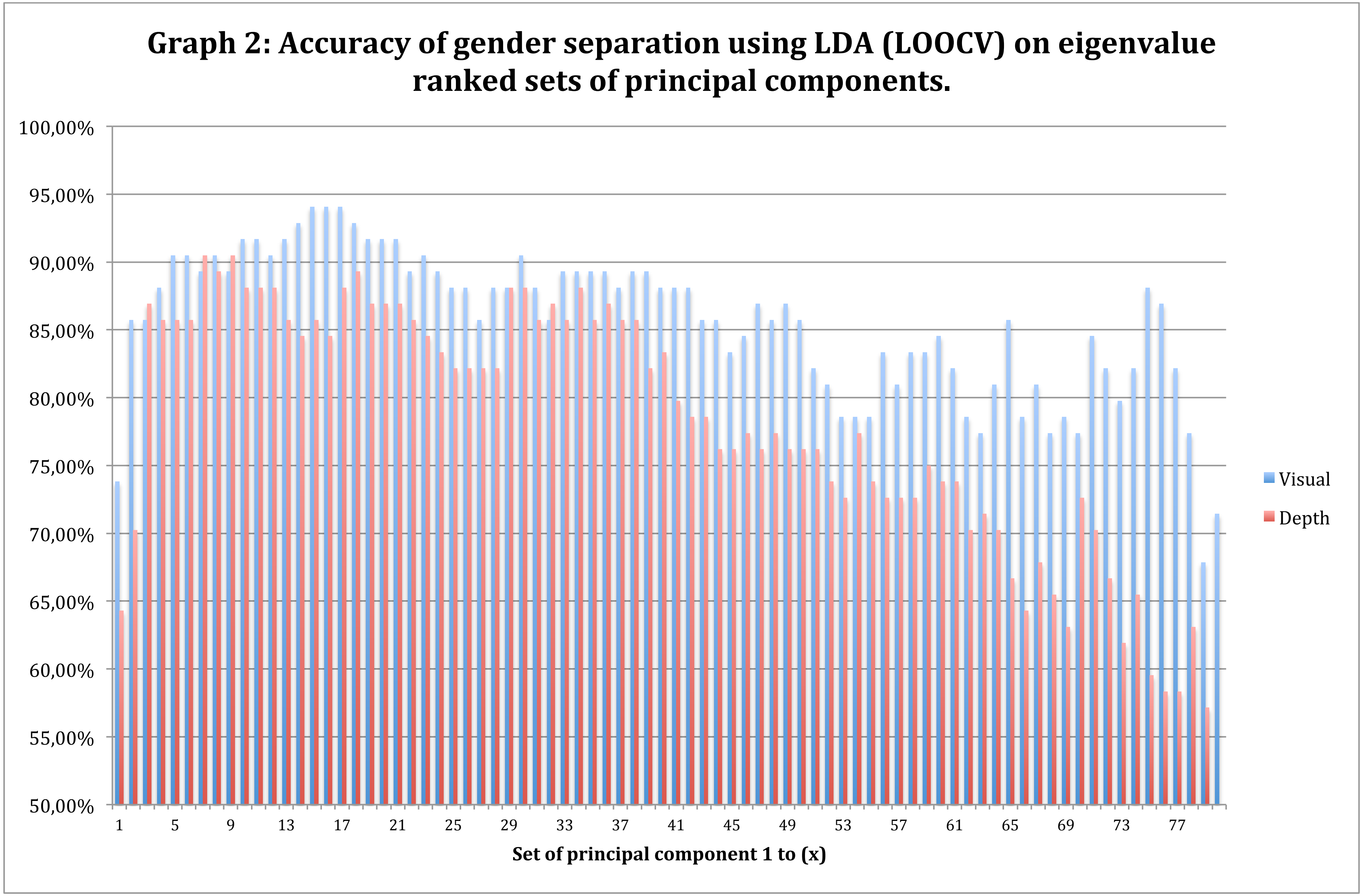
As can be seen in table 3 (appendix) and graph 3 the accuracy of gender separation for each principal component, when being considered individually, is low. This is to be expected due to the fact that the individual component's gender information is lower than that of sets of principle components.

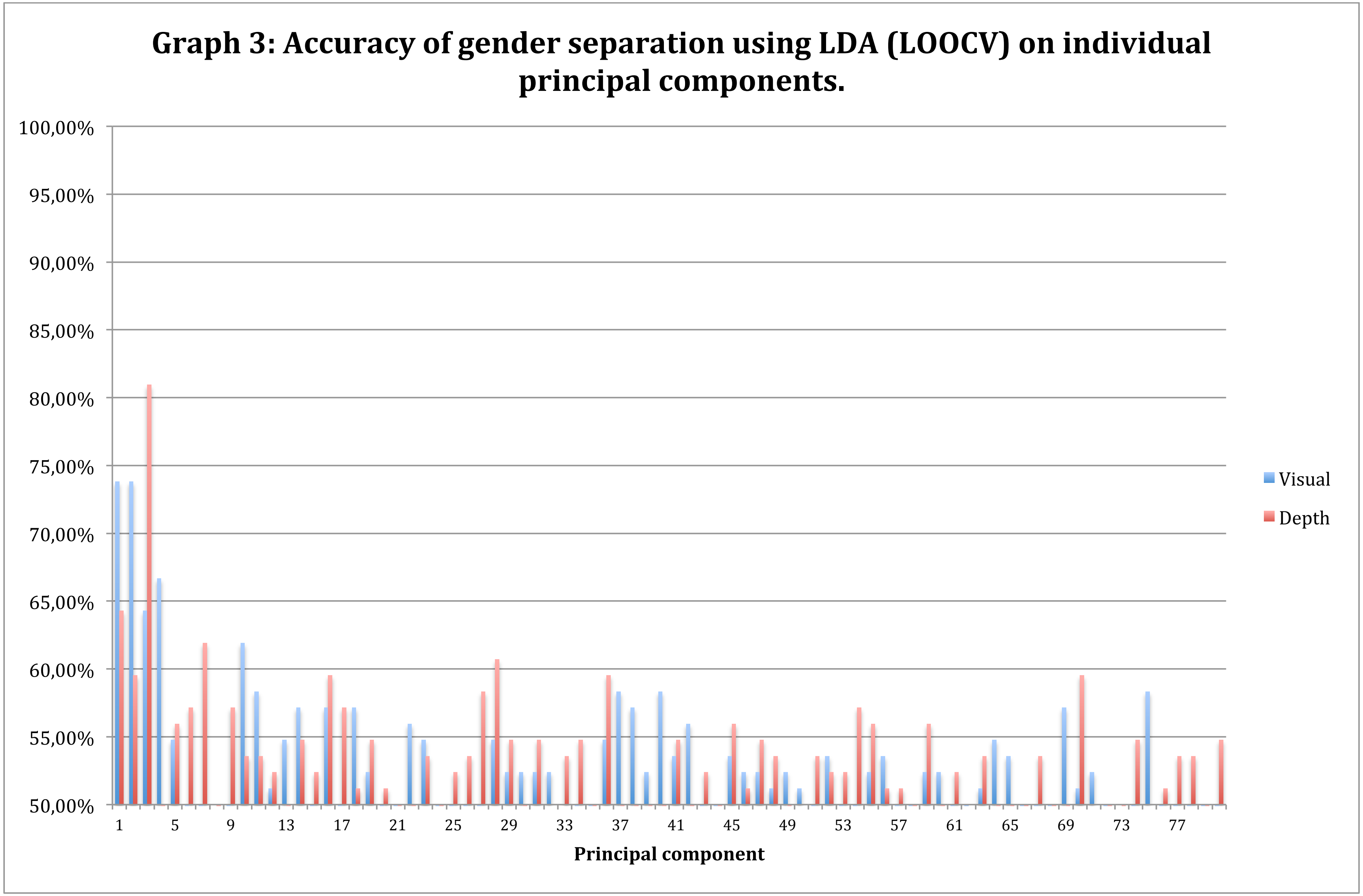
The results of the subsequent calculations concerning the sets of best performing principal components can be found in table 4 (appendix) and graph 4. These results also show that a gender separation accuracy of 100% has been achieved using the 49 best PRPCs. When taking the 57 best depth PRPCs into account it was possible to acquire a maximal gender separation accuracy of 97,62%.

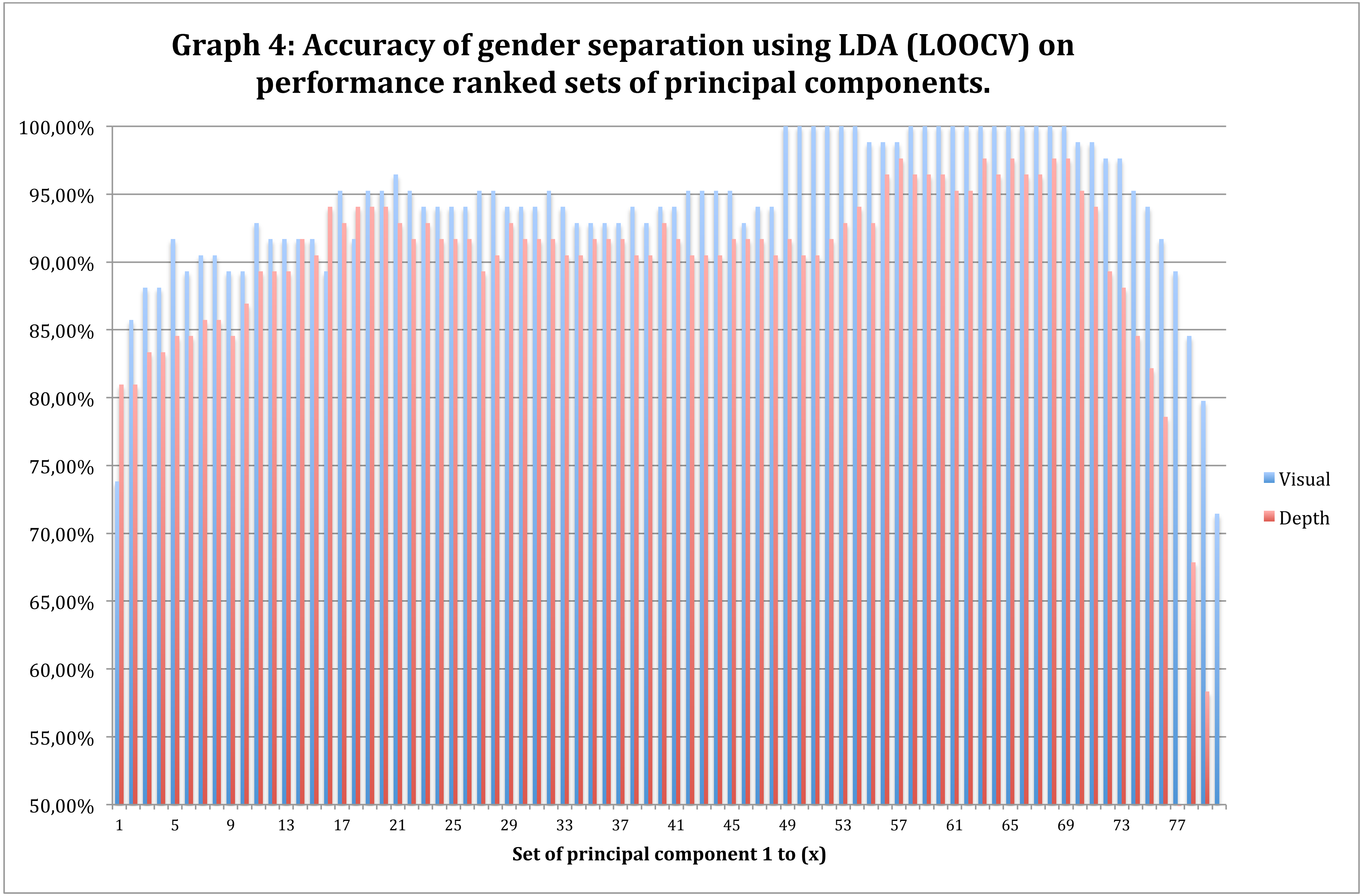
In summary, shown in table 4; gender determination is possible using visual and depth imagery obtained with a Kinect device answering RQ1. By ranking the PCA based Eigenfaces on eigenvalues, the best separation of gender can be found using the first 7 depth ERPCs and first 15 visual ERPCs resulting in respectable separation accuracies of 94,08% (visual) and 90,48 (depth). When firstly ranking the PCA based Eigenfaces on their best individual performance, the best separation of gender can be found using the first 49 best visual PRPCs and first 57 best depth PRPCs resulting in respectable separation accuracies of 100% (visual) and 97,62% (depth). This reconfirms RQ1 and answers RQ2 and RQ3. Although the results using PRPCs are higher, they can only be achieved using more computational effort. Using the sets of the best PRPCs it was not possible to transcend the 7 (depth) and 15 (visual) ERPC barriers for achieving higher gender separation accuracies. It can also be noted that both these separation accuracies, visual and depth, transcend those of other studies (Ben-Abdelkader & Griffin, 2005, Shen et al. 2009, Bekios-Calfa et al., 2011).

Type	Size of set	Accuracy
Visual ERPC set	15	94,08%
Depth ERPC set	7	90,48%
Visual PRPC set	49	100%
Depth PRPC set	57	97,62%

Table 4 – Summary of results


Graph 2: Accuracy of gender separation using LDA (LOOCV) on eigenvalue ranked sets of principal components.
100,00%
95,00%
90,00%
85,00%
80,00%
75,00%
70,00%
65,00%
60,00%
55,00%
50,00%
1
5
9
13
17
21
25
29
33
37
41
45
49
53
57
61
65
69
73
77
Set of principal component 1 to (x)
Visual
Depth


Graph 3: Accuracy of gender separation using LDA (LOOCV) on individual principal components.
100,00%
95,00%
90,00%
85,00%
80,00%
75,00%
70,00%
65,00%
60,00%
55,00%
50,00%
1
5
9
13
17
21
25
29
33
37
41
45
49
53
57
61
65
69
73
77
Principal component
Visual
Depth


Graph 4: Accuracy of gender separation using LDA (LOOCV) on performance ranked sets of principal components.
100,00%
95,00%
90,00%
85,00%
80,00%
75,00%
70,00%
65,00%
60,00%
55,00%
50,00%
1
5
9
13
17
21
25
29
33
37
41
45
49
53
57
61
65
69
73
77
Set of principal component 1 to (x)
Visual
Depth

5. Discussion

The obtained results using our GECCO method show that it is possible to determine gender using visual and depth imagery. Furthermore our study finds that by using visual imagery information we are able to achieve a higher gender separation accuracy compared to using depth imagery information. Below, we discuss three points: the validity of the experiment (4.1), the implications of the experiment (4.2), and the points of improvement of the GECCO method (4.3).

5.1 Validity

Concerning the validity of the experiment there are a number of marginal notes that have to be discussed. Firstly it is stated that the experiment took place in a practically light-tight lab. When reviewing the source imagery it can be seen that there are some small fluctuations in lighting. This could be the case due to daylight lunging through the panels that were blocking it. Another reason for this is that the Kinect device automatically adjusts its charge-coupled device (CCD) for the amount of light that is visible. Due to differences in body postures and height the amount of intercepted light can fluctuate, this way influencing the final image. Another remark can be made about the native resolution of the used sensors. The visual sensor has a native resolution of 1280x1024 pixels and the depth sensor has a native resolution of 640x480 pixels, four times smaller. To make sure both the visual and depth image set contain the same content, pixelmatching has been applied. This results in upscaled depth images, to achieve the same number of pixels as the visual images (901x1201). Therefore it should be noted that they originally did not have the same pixel density, thus containing less information for further processing. This could explain why this study shows a reduced gender separation capability for the depth compared to the visual data.

The low variation of participants' age due to the use of university students was insurmountable for this study but may have affected its final results. It can be argued that when people become of age some gender characteristics will fade. A study by Guo et al. (2009) shows that gender recognition accuracies can be 10% higher on adult faces than on young or senior faces, using one of the state-of-the-art methods.

Despite of clear stickering on the wall, depicting in which direction the participants had to gaze, this could not prevent the final imagery to vary slightly on this point. In other studies (Buchala et al., 2004) participants' hair was removed from the images due to its gender determining influence. In this study it was not covered to approximate a real-world scenario.

The use of PCA and LDA for obtaining, confining and classifying gender information was argumentative due to a large number of studies endorsing it. Although it has to be noted that due to the large number of machine learning steps, the classification has been subject to overfitting. By overfitting the classification model is focussing on specific random features of the training data. This leads to the used method being able to classify training examples better, but performing worse on unseen data.

5.2 Implications

The results of using the GECCO method show that PCA based Eigenfaces can be used with a LDA (LOOCV) classification algorithm to separate the source facial images based on gender accurately. Support can be found for a conjecture made by Philips et al. (2005) that one high-resolution 2D image is more powerful for face recognition than one 3D image. Also confirmation can be found for the statement by O'Toole et al. (1996), supported by Etemad (1997) and Buchala et al. (2004-2005), that "information related to the gender property is shared by most of the faces and hence (it) is encoded in the first few components".

On the other hand our results refute the following statement by Turk and Pentland (1991): "First of all the idea of Eigenfaces considers face recognition as a 2-D recognition problem, this is based on the assumption that at the time of recognition, faces will be mostly upright and frontal. Because of this, detailed 3-D information about the face is not needed. This reduces complexity by a significant bit.". Our study depicts the obtained "3-D information" visually using HSV colormapping, this way allowing the depth information to be processed using PCA and obtaining some relatively high results on gender determination. There are studies (Zhao et al., 2003, Phillips et al, 2005) which advocate to combine visual ("2-D") and depth ("3-D") data for even higher face recognition results. The HSV colormapping method could prove to be useful achieving this in the future using a revised GECCO method.

As stated in chapter four the results using ranked PCA based Eigenfaces are higher compared to the first method but they can only be achieved using more computational effort. Using the sets of best individually ranked principal components it is not possible to transcend the seven (depth) and fifteen (visual) principal component barriers for achieving higher gender separation accuracies. The question these results raise is what goal is to be pursued using the presented techniques. If the aim is pure accuracy, firstly ranking the principal components on performance would be a legitimate choice. On the other hand if speed is the aim it would be best to use the first seven (depth) and fifteen (visual) Eigenface ranked principal components.

The accuracies that our GECCO method yielded transcend those of other studies (Ben-Abdelkader and Griffin, 2005, Shen et al. 2009, Bekios-Calfa et al., 2011) but there is no speed comparison. Like stated before, some of techniques used in this study require much computational effort and therefore time.

5.3 Points of improvement

There are points that could improve the implementation of our GECCO method, some of them have already been discussed briefly in section 5.1. Firstly the lab design in future studies should be absolutely light-tight. This can be achieved with ease by using an experiment lab without any windows. Secondly it should be avoided scaling the images up or down, preserving the original pixeldensities. Recently there have been improvements in inexpensive devices equipped with depth sensors like the new Kinect for Windows device (Microsoft, 2012) and the LEAP device (LEAP, 2012), which could yield more accurate depth data.

The group of participants used for our study was small. Future studies could use larger groups, preferably varying more in age. When taking pictures, the participants should be asked to look directly into the lens(es) or some other orientation point on the device like an indicator light.

Despite of the arguments that can be made for the use PCA to obtain and confine gender information there are some studies (Belhumeur et al., 1997, He et al., 2005) advocating the use of Fisher faces or Laplacian faces for these kinds of applications. Future studies could consider using one or more of these techniques for comparison.

When considering which set of principal components to use for the final gender separation some other tactics can be used. A way to achieve this is by applying a randomization of the sets of to be assessed principal components. Using other methods it is possible to reduce chances of overfitting the data classification. Also the amount of time needed to perform all calculations and classification should be taken into account when comparing the achieved results with other studies.

6. Conclusion

The aim of this study was to create a method to examine if it is possible to determine gender by using visual and depth data obtained with a relatively cheap Kinect device. The results of our GECCO method show that this is possible. It is also evident that the visual data, acquired with the current generation of the Kinect device, yields a higher accuracy on gender separation compared to the depth data. Furthermore our GECCO method provides evidence that specific gender information is stored in the first range of Eigenfaces. The best separation of gender can be found using the first 7 depth ERPCs and first 15 visual ERPCs resulting in respectable separation accuracies of 94,08% (visual) and 90,48 (depth). The best separation of gender can be found using the first 49 best visual PRPCs and first 57 best depth PRPCs resulting in respectable separation accuracies of 100% (visual) and 97,62% (depth).

Future work should address the trade-off between speed (ERPCs) and performance (PRPCs).

7. References

Anila, S., Devarajan, N. (2011). Global and local classifiers for face recognition. *European Journal of Scientific Research 57* (4), 556 – 566.

Bekios-Calfa, J., Buenaposada, M.J. & Baumela, L. (2011). Revisiting lineair Discriminant Techniques in gender recognition. *In proceedings of the IEEE transactions on pattern analysis and machine intelligence, 33* (4), 858 – 864.

Belhumeur, P.N., Hespanha, J.P. & Kriegman, D.J. (1997). Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *In proceedings of the IEEE transactions on pattern analysis and machine intelligence, 19* (7), 711 – 720.

Ben-Abdelkader, C. & Griffin, P. (2005). A local region-based approach to gender classification from face images. *In proceedings of the IEEE Computer Vision and pattern recognition,* (3), 52-57

Bowyer, K.W., Chang, K., and Flynn, P. (2004). A survey of depth and multi-modal depth + visual face recognition. *Computer Vision and Image Understanding* (101), 1 – 15.

Buchala, S., Davey, N., Frank, R.J. and Gale, T.M. (2005). Dimensionality reduction of face images for gender classification. *In proceedings of the IEEE Conference on intelligent systems.*

Buchala, S., Davey, N., Gale, T.M. and Frank, R.J. (2004). Principal Component Analysis of gender, ethnicity, age, and identity of face images. *In proceedings of the IEEE international conference on multimodal interaction.*

Dailey, M. (2012). *Matt's Matlab Tutorial Source Code Page*. Consulted on 20-06-2012, from Computer Science and Information Management program: http://www.cs.ait.ac.th/~mdailey/matlab/ .

Etemad, K. and Chellappa, R. (1997). Discriminant analysis for recognition of human face images. *Journal of Optical Society of America, 14* (8), 724 – 733.

Golomb, B. A., Lawrence, D. T. & Sejnowski, T. J. (1990). SexNet: A neural network identifies sex from human faces. *In proceedings of conference on advances in neural information processing systems series*, 572 - 577.

Guo, G., Dyer, C.R., Fu, Y. & Huang, T.S. (2009). Is gender recognition affected by age? *In proceedings of the IEEE international conference on computer vision workshop human-computer interaction,* 2032 – 2039.

He, X., Yan, S., Hu, Y., Niyogi, P.. Zhang, H-J (2005). Face recognition using Laplacianfaces. *In proceedings of the IEEE transactions on pattern analysis and machine intelligence, 27* (3), 1 – 13.

Jolliffe I.T. (2002). Principal Component Analysis, Series: Springer Series in Statistics, 2nd ed., Springer, NY, 487.

LEAP (2012). Leap Motion, 20-08-2012, from http://live.leapmotion.com/about.html.

Microsoft (2012). Product features | Microsoft Kinect for Windows, 20-08-2012, form http://www.microsoft.com/en-us/kinectforwindows/discover/features.aspx.

O'Toole, A.J., Deffenbacher, K.A., Valentin, D. and Abdi, H. (1994). Structural aspects of face recognition and the other race effect. *Memory and Cognition* (26), 146 – 160.

O'Toole, A.J., Vetter, T., Troje, N.F., Bülthoff, H.H. (1996). Sex classification is better with three-dimensional head structure than with image intensity information. *Perception* (25), 75 – 84.

Phillips, P.J., Flynn, J., Scruggs, T., Bowyer, K.W., Chang, J., Hoffman, K., Marques, J., Min, J. Worek, W. (2005). Overview of the face recognition grand challenge. *In proceedings of the IEEE computer society conference on computer vision and pattern recognition*.

PrimeSense (2011). The PrimeSensor, The PrimeSense depth sensing solution on 19-07-2012, from http://www.primesense.com/en/press-room/resources/file/4-primesense-depth-sensor-data-sheet .

Serrano, S. (2004). *Eigenface tutorial*. Consulted on 30-07-2012 from www.pages.drexel.edu/~sis26/Eigenface%20Tutorial.htm .

Shen, B.C., Chen, C.S. & Hsu, H.H. (2009). Fast gender recognition by using a shared-integral-image approach. *In proceedings of the IEEE international conference on acoustics, speech, and signal processing,* 521 – 524.

Sirovich, L. and Kirby, M. (1987). Low-dimensional procedure for the characterization of human faces. *Journal of the Optical Society of America* (4), 519 – 524.

Tsalakanidou, F., Tzocaras, D., Strintzis, M. (2003). Use of depth and colour Eigenfaces for face recognition, *Pattern Recognition Letters* (24), 1427 – 1435.

Turk, A., & Pentland, P. (1991). Eigenfaces for recognition. *The Journal of Cognitive Neuroscience , 3* (1), 71 – 86.

Turk, A., & Pentland, P. (1991). Face recognition using Eigenfaces. *In proceedings of the IEEE conference of computer vision,* 286 – 296.

Valentin, D., Abdi, H., O'Toole, A.J., Cottrell, W. (1994). Connectionist models of face processing: A survey. *Pattern Recognition 27*(9), 1209 – 1230.

Zhao, W., Krishnaswamy, A., Chellappa, R., Swets, L., Weng., J. (1998). Discriminant analysis of principal components for face recognition. *In proceedings of the international conference on automatic face and gesture recognition,* 336–341.

Zhao, W., Chellappa, R., Phillips, P.J., & Rosenfeld, A. (2003). Face recognition: a literature survey. *ACM Computing Survey's, 35*(4), 399 – 458.

8. Appendix

The following pages contain tables 1 through 4 belonging to chapter four.

Table 1 – Eigenface proportion variance explained

Principal component	Proportion variance explained (visual)	Proportion variance explained (depth)	Principal component	Proportion variance explained (visual)	Proportion variance explained (depth)
1	74231344,6587192	75165683,8648798	41	1183628,5233237	1610691,9229790
2	60752811,5521164	42960501,8275202	42	1137121,3530380	1574330,7893297
3	19318982,8097676	23743520,5032022	43	1110485,3517038	1533174,5579881
4	13370043,5028869	15041645,4695734	44	1084107,0171858	1487514,9670390
5	10937536,6313063	13967821,2922145	45	1080837,0634632	1479406,8714555
6	8775586,7725081	11753418,4690934	46	1054244,8571565	1423339,7476116
7	8309584,4371565	9314853,5590954	47	1034391,7965486	1392990,6716121
8	5993968,5178409	8543448,5053540	48	1020574,9058853	1365623,7351051
9	5609203,2411665	6001740,2422993	49	993364,8907393	1326644,0470091
10	5220270,7791626	5606980,1485323	50	974346,7679734	1297722,4106459
11	4907030,7928487	5249826,6533621	51	949269,9686940	1260745,9806312
12	4496278,5062398	5049831,9919240	52	928963,7734031	1244753,3177460
13	3849069,6624775	4559424,0729562	53	884135,4176007	1237174,9255532
14	3570407,5019749	4266353,9721020	54	867355,3413101	1213180,6343154
15	3432260,7630985	3988846,2708609	55	857134,7444323	1200868,1201174
16	3015773,0951967	3739483,6858491	56	839449,4069170	1178556,9486617
17	2935259,0634756	3655223,5173838	57	799154,5041658	1154899,5653396
18	2708410,0131680	3493377,1467408	58	794666,4044511	1128843,8068761
19	2640220,4327909	3384178,4985661	59	779221,2585627	1096818,3074389
20	2458349,0914770	3236471,8805778	60	774576,9130211	1084040,6549000
21	2351083,6745676	3101893,8541803	61	741247,5955821	1078693,4073064
22	2242548,9350442	3005682,4635455	62	730765,3639533	1037366,9222408
23	2205979,8073653	2904058,0693016	63	720462,0231880	1023721,4844456
24	2023304,5388066	2809931,4893885	64	713800,3928549	995209,8160008
25	1958368,7894096	2677032,1269499	65	696418,3475024	987886,1174751
26	1890809,3841753	2556996,8179304	66	683709,8232199	972281,0403336
27	1856780,9757879	2504944,6335325	67	665167,1611144	960659,3218750
28	1810370,1539748	2455013,4973537	68	660984,6903645	945834,4913987
29	1688242,2935324	2281810,8245221	69	653134,0184813	920834,5252969
30	1587880,6062601	2234264,9815734	70	617567,8846464	912125,7103201
31	1573735,1075183	2132872,5540959	71	616274,2796264	883944,5053110
32	1524554,4224935	2039962,2945595	72	589241,5384584	863129,3757259
33	1481005,9102748	1989269,1241833	73	578684,6344323	833146,7893557
34	1414127,6054996	1939529,5295109	74	572690,6518273	818412,7634800
35	1387634,4193158	1868783,8732548	75	551908,9869062	802911,9767578
36	1313528,5480630	1807940,9527111	76	526582,5217932	781316,4983623
37	1295706,8703430	1769608,3061204	77	516040,9265031	776237,1753248
38	1290263,6162636	1745827,5141542	78	492419,4756231	756298,4022882
39	1265854,5404081	1707232,5711202	79	475099,5764213	730830,2604136
40	1198170,9329868	1654288,2232593	80	452876,3134448	713116,6475756

Table 2 - Accuracy of gender separation using LDA (LOOCV) on eigenvalue ranked principal components.

Set of principal components 1 to (x)	Accuracy of gender separation based on visual principal components	Accuracy of gender separation based on depth principal components	Set of principal components 1 to (x)	Accuracy of gender separation based on visual principal components	Accuracy of gender separation based on depth principal components
1	73,81%	64,29%	41	88,10%	79,76%
2	85,71%	70,24%	42	88,10%	78,57%
3	85,71%	86,90%	43	85,71%	78,57%
4	88,10%	85,71%	44	85,71%	76,19%
5	90,48%	85,71%	45	83,33%	76,19%
6	90,48%	85,71%	46	84,52%	77,38%
7	89,29%	**90,48%**	47	86,90%	76,19%
8	90,48%	89,29%	48	85,71%	77,38%
9	89,29%	90,48%	49	86,90%	76,19%
10	91,67%	88,10%	50	85,71%	76,19%
11	91,67%	88,10%	51	82,14%	76,19%
12	90,48%	88,10%	52	80,95%	73,81%
13	91,67%	85,71%	53	78,57%	72,62%
14	92,86%	84,52%	54	78,57%	77,38%
15	**94,05%**	85,71%	55	78,57%	73,81%
16	94,05%	84,52%	56	83,33%	72,62%
17	94,05%	88,10%	57	80,95%	72,62%
18	92,86%	89,29%	58	83,33%	72,62%
19	91,67%	86,90%	59	83,33%	75,00%
20	91,67%	86,90%	60	84,52%	73,81%
21	91,67%	86,90%	61	82,14%	73,81%
22	89,29%	85,71%	62	78,57%	70,24%
23	90,48%	84,52%	63	77,38%	71,43%
24	89,29%	83,33%	64	80,95%	70,24%
25	88,10%	82,14%	65	85,71%	66,67%
26	88,10%	82,14%	66	78,57%	64,29%
27	85,71%	82,14%	67	80,95%	67,86%
28	88,10%	82,14%	68	77,38%	65,48%
29	88,10%	88,10%	69	78,57%	63,10%
30	90,48%	88,10%	70	77,38%	72,62%
31	88,10%	85,71%	71	84,52%	70,24%
32	85,71%	86,90%	72	82,14%	66,67%
33	89,29%	85,71%	73	79,76%	61,90%
34	89,29%	88,10%	74	82,14%	65,48%
35	89,29%	85,71%	75	88,10%	59,52%
36	89,29%	86,90%	76	86,90%	58,33%
37	88,10%	85,71%	77	82,14%	58,33%
38	89,29%	85,71%	78	77,38%	63,10%
39	89,29%	82,14%	79	67,86%	57,14%
40	88,10%	83,33%	80	71,43%	50,00%

Table 3 - Accuracy of gender separation using LDA (LOOCV) on individual principal components.

Principal component	Accuracy of gender separation based on visual principal component	Accuracy of gender separation based on depth principal component	Principal component	Accuracy of gender separation based on visual principal component	Accuracy of gender separation based on depth principal component
1	73,81%	64,29%	41	53,57%	54,76%
2	73,81%	59,52%	42	55,95%	45,24%
3	64,29%	80,95%	43	48,81%	52,38%
4	66,67%	28,57%	44	14,29%	23,81%
5	54,76%	55,95%	45	53,57%	55,95%
6	48,81%	57,14%	46	52,38%	51,19%
7	44,05%	61,90%	47	52,38%	54,76%
8	50,00%	48,81%	48	51,19%	53,57%
9	27,38%	57,14%	49	52,38%	48,81%
10	61,90%	53,57%	50	51,19%	50,00%
11	58,33%	53,57%	51	36,90%	53,57%
12	51,19%	52,38%	52	53,57%	52,38%
13	54,76%	33,33%	53	0,00%	52,38%
14	57,14%	54,76%	54	48,81%	57,14%
15	48,81%	52,38%	55	52,38%	55,95%
16	57,14%	59,52%	56	53,57%	51,19%
17	36,90%	57,14%	57	2,38%	51,19%
18	57,14%	51,19%	58	47,62%	47,62%
19	52,38%	54,76%	59	52,38%	55,95%
20	48,81%	51,19%	60	52,38%	35,71%
21	33,33%	21,43%	61	45,24%	52,38%
22	55,95%	41,67%	62	48,81%	50,00%
23	54,76%	53,57%	63	51,19%	53,57%
24	32,14%	45,24%	64	54,76%	46,43%
25	48,81%	52,38%	65	53,57%	48,81%
26	47,62%	53,57%	66	38,10%	48,81%
27	45,24%	58,33%	67	44,05%	53,57%
28	54,76%	60,71%	68	7,14%	44,05%
29	52,38%	54,76%	69	57,14%	46,43%
30	52,38%	44,05%	70	51,19%	59,52%
31	52,38%	54,76%	71	52,38%	50,00%
32	52,38%	50,00%	72	20,24%	38,10%
33	50,00%	53,57%	73	50,00%	46,43%
34	42,86%	54,76%	74	47,62%	54,76%
35	44,05%	48,81%	75	58,33%	50,00%
36	54,76%	59,52%	76	14,29%	51,19%
37	58,33%	44,05%	77	40,48%	53,57%
38	57,14%	17,86%	78	28,57%	53,57%
39	52,38%	26,19%	79	45,24%	30,95%
40	58,33%	41,67%	80	48,81%	54,76%

Table 4 - Accuracy of gender separation using LDA (LOOCV) on performance ranked principal components.

Set of principal components 1 to (x)	Accuracy of gender separation based on visual principal components	Accuracy of gender separation based on depth principal components	Set of principal components 1 to (x)	Accuracy of gender separation based on visual principal components	Accuracy of gender separation based on depth principal components
1	73,81%	80,95%	41	94,05%	91,67%
2	85,71%	80,95%	42	95,24%	90,48%
3	88,10%	83,33%	43	95,24%	90,48%
4	88,10%	83,33%	44	95,24%	90,48%
5	91,67%	84,52%	45	95,24%	91,67%
6	89,29%	84,52%	46	92,86%	91,67%
7	90,48%	85,71%	47	94,05%	91,67%
8	90,48%	85,71%	48	94,05%	90,48%
9	89,29%	84,52%	**49**	**100,00%**	91,67%
10	89,29%	86,90%	50	100,00%	90,48%
11	92,86%	89,29%	51	100,00%	90,48%
12	91,67%	89,29%	52	100,00%	91,67%
13	91,67%	89,29%	53	100,00%	92,86%
14	91,67%	91,67%	54	100,00%	94,05%
15	91,67%	90,48%	55	98,81%	92,86%
16	89,29%	94,05%	56	98,81%	96,43%
17	95,24%	92,86%	**57**	98,81%	**97,62%**
18	91,67%	94,05%	58	100,00%	96,43%
19	95,24%	94,05%	59	100,00%	96,43%
20	95,24%	94,05%	60	100,00%	96,43%
21	96,43%	92,86%	61	100,00%	95,24%
22	95,24%	91,67%	62	100,00%	95,24%
23	94,05%	92,86%	63	100,00%	97,62%
24	94,05%	91,67%	64	100,00%	96,43%
25	94,05%	91,67%	65	100,00%	97,62%
26	94,05%	91,67%	66	100,00%	96,43%
27	95,24%	89,29%	67	100,00%	96,43%
28	95,24%	90,48%	68	100,00%	97,62%
29	94,05%	92,86%	69	100,00%	97,62%
30	94,05%	91,67%	70	98,81%	95,24%
31	94,05%	91,67%	71	98,81%	94,05%
32	95,24%	91,67%	72	97,62%	89,29%
33	94,05%	90,48%	73	97,62%	88,10%
34	92,86%	90,48%	74	95,24%	84,52%
35	92,86%	91,67%	75	94,05%	82,14%
36	92,86%	91,67%	76	91,67%	78,57%
37	92,86%	91,67%	77	89,29%	25,00%
38	94,05%	90,48%	78	84,52%	67,86%
39	92,86%	90,48%	79	79,76%	58,33%
40	94,05%	92,86%	80	71,43%	50,00%