Novel Approach to Enhance
Face Recognition Using Depth Maps

Yaser Saleh, Eran Edirisinghe

Department Computer Science, Loughborough University, Loughborough, LE11 3TU, UK

Y.saleh@lboro.ac.uk

Abstract - **Face recognition, although being a popular area of research and study, still has many challenges, and with the appearance of the Microsoft Kinect device, new possibilities of research were uncovered, one of which is face recognition using the Kinect. With the goal of enhancing face recognition, this paper is aiming to prove how depth maps, since not effected by illumination, can improve face recognition with a benchmark algorithm based on the Eigenface. This required some experiments to be carried out, mainly in order to check if algorithms created to recognize faces using normal images can be as effective if not more effective with depth map images. The OpenCV Eigenface algorithm implementation was used for the purpose of training and testing both normal and depth-map images. Finally, results of the experiments are presented to prove the ability of the tested algorithm to function with depth maps, also, proving the capability of depth maps face recognition's task in poor illumination.**

Keywords – **Face recognition; Kinect; Depth maps; Eigenface.**

I. INTRODUCTION

Over the years, face recognition became a popular area of research and study in computer vision, and a popular application regarding image analysis and understanding. Knowing that, in the recent years, a new device was introduced by Microsoft with the name of The Microsoft Kinect [1], the device was first introduced to be a gaming add-on, but later on, research in the area of face recognition crossed paths with the device since it had many appealing features, one feature was the depth sensor, a sensor that provided what is called "Depth Maps". Depth maps would provide a detailed pixel by pixel depth of the area in front of the Kinect, such sensor grabbed the attention of much research around different aspects of computer vision; one of the most interesting is the use of the depth sensor for face recognition, which produced many papers explaining how the sensor can be used to create new face recognition algorithms and technique [2-7].

This paper presents some tests for the capabilities of a benchmark algorithm of face recognition, using the colored images and the depth map images generated from the Kinect. Furthermore, this study examine the possibility of enhancing face recognition in poor illumination by seizing the fact that depth maps are not affected by illumination, therefore, making face recognition more robust.

This paper is divided as follows: Section I contains the introduction, Section II comments on some related work, section III explains the background of the Kinect 2 and the Eigenface based approach, Section IV explains the proposed procedure, Section V analyzes the experiments results, and finally Section VI presents the conclusion.

II. RELATED WORK

In this section, we go through some research done that relates to the concept of developing and\or using benchmark algorithms for face recognition and\or detection using the Microsoft Kinect as a capturing device. Many face databases and systems were created with the purpose of detecting and\or recognizing human faces, where a database of face images in color and as depth maps were created, with the goal of using such data to test a proposed face recognition algorithm [2, 3], however, such research that used the Microsoft Kinect 1 reported that the accuracy of using the depth maps was not as good as using the 2D images, but the accuracy using both was outstanding. While in other research a fusion between the use of normal images and depth maps was created, where although results were encouraging, some known face recognition algorithms were tested and found not able to function properly with the fused data, not taking in consideration the possibility of some face recognition algorithms functioning on depth maps alone without needing to fuse normal and depth maps together [5]. It was clear in many studies how the data that the Microsoft Kinect v1 provided was very noisy, with low resolution of depth data, also, and most importantly, research in this area didn't use any benchmark algorithms in order to check if the data provided from the depth sensor can be used without creating a custom algorithm [2, 5].

Described in [4], the use of the new Microsoft Kinect v2 with a new descriptor for face recognition. The descriptor was described as a modification of the Local Binary Patter (LBP) algorithm. The authors discussed how in the past couple of year a lot of work was done using the Microsoft Kinect v1, which had a serious problem with the fact that it was impossible to extract features such as the eyes, nose, mouth and ears, mainly because of the high noise and low resolution of the depth map. A dataset was also created, where the depth data only was recorded for 18 participants. It was also mentioned that the Pegasos Support Vector Machine (SVM) technique was used for classification, and an experiment was conducted, where the results were compared with Local Binary Patterns and Scale-Invariant Feature Transform (SIFT) algorithms' results, showing that the presented technique gave much better results than the normal LBP method, and somewhat better results than the SIFT method. This paper, although the new Microsoft Kinect v2 was used, have

presented a new algorithm handling only depth maps for face recognition, therefore leaving out the possibility of benchmark face recognition algorithms running successfully on depth maps.

What was clear is the lack of a proper review of the standard algorithms used for face recognition being applied on depth map images, as although some research have used the new Microsoft Kinect 2, it still was not for the purpose of checking the efficiency of a standard face recognition algorithm such as the Eigenface. Also, the lack of research trying to improve on face recognition procedures when capturing scenes has poor illumination.

III.   BACKGROUND

This section presents the technologies and techniques that are required to understand the work and concepts presented within the rest of this paper. Starting with the Microsoft Kinect version 2, being the main capturing device, then, the benchmark face recognition algorithm used in the experiments was explained to provide more understanding of the experiments carried out later on.

A.   Microsoft Kinect 2

The Microsoft Kinect is a device first released by Microsoft as an add-on to the gaming console Xbox360, but, after the remarkable capabilities that was seen in the device, the development community started working and developing applications depending on the sensors embedded in the Kinect. In October 2014, the Software Development Kit for the second version of the Kinect was released by Microsoft, giving the ability for everyone to develop applications and experiment the new and improved Kinect. As a device it provides synchronized color and depth images, that can be processed as needed, also, the device has an algorithm that captures the human in 3-D, and therefor enables the human to use his\her body as the controller of the game or application. With the visual (RGB) and depth information provided by the Kinect, the possibility of finding new solutions for some classic problems in the computer vision field is higher than before. As soon as the Kinect was released, a large number of scientific papers started to appear, which shows the capabilities that can be found in the device [8].

The Kinect 2 hardware contains three main components, the RGB camera, the depth sensor, and an array of microphones installed along the Kinect as seen in Fig. 1, with the color camera providing 1920 x 1080 pixels @30 frames per second, and the depth sensor providing 512 x 424 pixels @30 frames per second data. The maximum distance the Kinect's depth sensor is capable of capturing is 4.5 meters, while the skeleton tracker of the Kinect can track up to six subjects with 25 joint points each . These components will be able to provide RGB images, depth maps, and audio signals at the same time. With these hardware components comes the software giving the developers the ability to create applications for different uses. These software tools provide the ability to handle the RGB images, capture the motion of the human, do some face recognition, and even recognize voice commands.
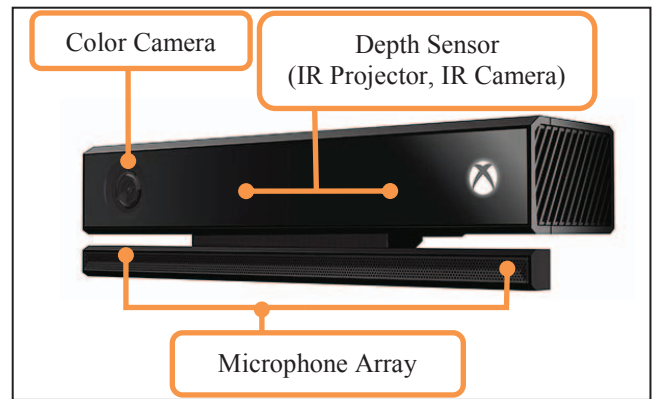

Figure 1.   Kinect 2 components

B.   Eigenface

Eigenface based face recognition was first introduced in 1991, the goal was to try and extract as much relevant information from a face image, then to encode these information as efficient as possible to then be able to compare this encoding with other models encoded and collected in a dataset.

So at the beginning, the idea was to generate eigenvectors out of the covariance matrix of the stored set of images. Those eigenvectors was explained to be thought of as "a set of features which together characterize the variation between face images" [9]. The eigenvectors can be then displayed as images which are called Eigenfaces.

This approach discusses face recognition from a two dimensional recognition point of view, trying to build a description of the face images from a set of 2D characteristics. As described by Turk and Pentland [10], the Eigenface approach steps can be summarized as follows:

1.   The training set, containing face images, is first obtained.

2.   An Eigenfaces calculation process is applied on every face image, keeping only a number of images (M) with the highest eigenvalues to span an M-dimensional subspace called the face space.

3.   Checks if the image is close enough to the face space, therefore, concluding if it is an image of a face (known or unknown).

4.   If the image was recognized as a face, the weight pattern is classified as a known or an unknown person.

In this research the OpenCV Eigenface algorithm implementation was employed to be able to make use of the Eigenface face recognizer [11].

IV.   PROPOSED PROCEDURE

The procedure used for carrying out the experiments, primarily training and testing of the Eigenface based algorithm on the normal and depth map images can be seen in Fig. 2.
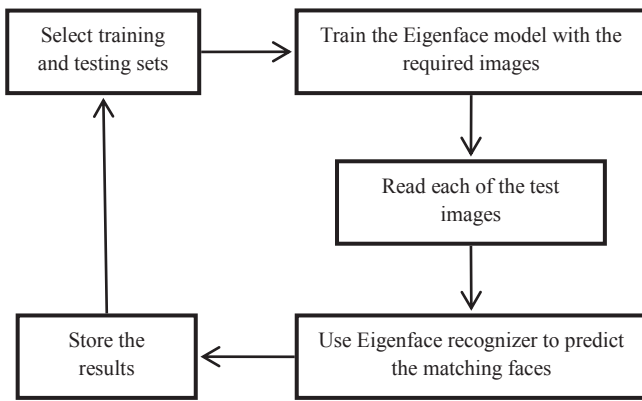
Figure 2.   Experiments procedure

The procedure starts with selecting a number of images for each person, some for training and the rest for testing, afterwards, a training process is carried out were an Eigenface face recognition model is passed the training images to be then trained. After the training is completed, the algorithm reads the test images to then be passed on to the Eigenface recognizer function, were each image is processed and a match is given. At the end, the matches produced are stored for later calculations and the whole procedure is repeated for another training\testing set.

V.   EXPERIMENTS, RESULTS AND ANALYSIS

In this section several different experiments were carried out and discussed, starting with testing the Eigenface face recognition algorithm with normal images and then depth-map images. This was performed using a dataset of 36 normal images and 36 depth-map images for 10 participants, with half of the images taken in poorly illuminated scenes, and half in normal illumination scenes, seen in Fig. 3 and Fig. 4. Moreover, the participants were displaying nine poses with two expressions at each pose. In every experiment, since each subject in the dataset has 18 normal and depth images with good illumination, and 18 normal and depth images with poor illumination, the training set for any case is set to a random 15 images, with three images left for testing.
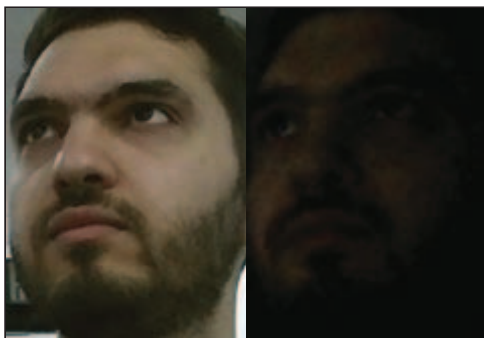

Figure 3.   A color image for a face taken in a good illumination scene (left), and a color image for the same face taken in a poor illumination scene (right)
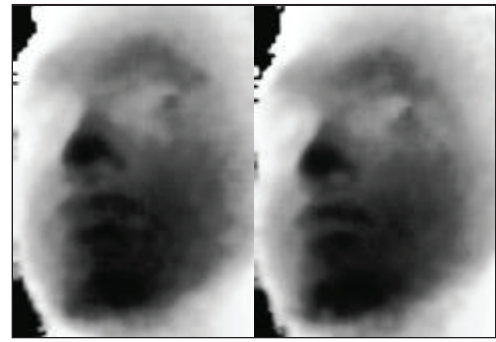

Figure 4.   A depth map image for a face taken in a good illumination scene (left), and a depth map image for the same face taken in a poor illumination scene (right)

A.   Evaluating Eigenface with normal images

One of the common challenges of face recognition is the process of face recognition during night time or with poor illumination, where multiple enhancement procedure were presented in the image processing field to try and find a solution, but still, some dark levels in an image would make it impossible to even detect faces. Therefore, this experiment will try and observe the changes of the accuracy of the proposed face recognition algorithm.

Since each person has two colored images in each pose and expression, one in a normal illumination and one in poor illumination, in normal illumination, each person had 18 images describing nine poses with two expressions on each pose. Providing 80% of the bright images for training requires $18 \times 80\% \approx 15$ images, therefore, 15 random images in normal illumination were chosen for training, while the three set aside images not chosen and the three equivalent poor illumination images were left for testing. In order to be fair and try almost all possible train\test sets, the same process was repeated 10 times with different random images chosen for training and testing.

The results of this experiment as seen in Table 1 can explain each test process, e.g., the first test process had the images number (6, 15, and 16) as test subjects, so while other images from the 18 normal illumination images for each person are used for training, the 6th, 15th, and 16th images from the normal illumination and the 6th, 15th, and 16th images from the poor illumination are used for testing. At the end of the table an average of accuracy is calculated for each case, showing the strength of the algorithm and the effect of using images with poor illumination as test subjects.

The results of this experiment showed an average accuracy rate of 97.33% when testing the normal illumination images, where the recognition rate for poor illumination test images was only 10% showing how a large change in the light can affect the Eigenface algorithm drastically. The 10% was produced as in every case all faces were recognized as the same person; therefore, in each case one person was given the right recognition.

B. *Evaluating Eigenface with depth-map images*

The previous experiment proved that using color images alone provides unstable results, mainly because of the drastic effect of illumination, it was crucial to try and include depth map images in the recognition process to find if depth maps can replace or at least help the color images in face recognition.

Through this experiment, the depth map images taken in good illumination scenes are used, basically, this experiment repeats the previous experiment but with depth map images instead of color images. As before, three images were picked randomly for testing, while the other 15 were left to form a training set for the face recognition algorithm. The training\testing process was again repeated 10 times like the previous experiments in order to have fair sets of results.

Since the resolution of the original depth map images provided from the Kinect is much lower than the normal color images, the process of detecting faces and features from depth maps might be of a challenge compared to using colored images, nevertheless, the depth map images should be able to provide better results with face recognition in poor illumination scenes, this is mainly because of the fact that depth maps are built using infrared signals and therefore shouldn't be affected by illumination changes as much as the normal images.

Examining the results of this experiment found in Table II, at first, it was clear that with normal illumination, the average accuracy rate is 87.67%, which compared to the first experiment, presents a 10% decrease in accuracy, but when examining the Eigenface based algorithm when tested with depth images taken in poor illumination conditions, it's also clear that there is a huge increase compared to the first experiment, jumping from 10% to 82.33%, which indicates that when used with normal illumination depth images, the results of the poor illumination test subjects are very encouraging.

This experiment provided evidence that depth map images, when added to the normal images, can be able to contribute to the area of face recognition in poor illumination scenes, at least with algorithms based on the Eigenface approach, where depth map images could present acceptable results in regards to the accuracy levels when using poor illumination test images on a training set involving only depth map images taken in a normally illuminated scene.

TABLE I.   EVALUATING EIGENFACE WITH NORMAL IMAGES EXPERIMENT RESULTS

Test Subjects	Normal Illumination	Poor Illumination
Test(6, 15, 16)	93.33%	10%
Test(3, 11, 15)	100%	10%
Test(1,13,17)	100%	10%
Test(6, 9, 11)	100%	10%
Test(2, 3, 5)	100%	10%
Test(7, 8, 14)	90%	10%
Test(1, 2, 7)	93.33%	10%
Test(1, 11, 14)	100%	10%
Test(2, 4, 8)	96.67%	10%
Test(5, 12, 18)	100%	10%
Average	97.33%	10%

TABLE II.   EVALUATING EIGENFACE WITH DEPTH MAP IMAGES EXPERIMENT RESULTS

Test Subjects	Normal Illumination	Poor Illumination
Test(6, 15, 16)	86.67%	73.33%
Test(3, 11, 15)	93.33%	80%
Test(1,13,17)	96.67%	93.33%
Test(6, 9, 11)	96.67%	86.67%
Test(2, 3, 5)	96.67%	96.67%
Test(7, 8, 14)	56.67%	70%
Test(1, 2, 7)	73.33%	70%
Test(1, 11, 14)	93.33%	86.67%
Test(2, 4, 8)	86.67%	83.33%
Test(5, 12, 18)	96.67%	83.33%
Average	87.67%	82.33%

VI.   CONCLUSION

This paper focused on the goal of finding if a benchmark face recognition algorithm, based on the Eigenface, can be used with the depth maps provided from the Microsoft Kinect, and to check its ability compared to its use with normal images in dark scenes. From the results of the experiments, it was established that recognition rates using depth map images show great potential, where it can be noticed that depth map images, if used for face recognition, can be very useful and accurate especially if the scenes had poor illumination.

REFERENCES

[1] Microsoft.com, "Kinect hardware". [Online]. Available: http://www.microsoft.com/en-us/kinectforwindows/meetkinect/features.aspx.[Accessed:19-Mar-2015].

[2] B. Li, A. Mian, W. Liu and A. Krishna, "Using Kinect for face recognition under varying poses, expressions, illumination and disguise", 2013 IEEE Workshop on Applications of Computer Vision (WACV), pp. 186-192, 2013.

[3] R. Hg, P. Jasek, C. Rofidal, K. Nasrollahi, T. Moeslund and G. Tranchet, "An RGB-D Database Using Microsoft's Kinect for Windows for Face Detection", 2012 Eighth International Conference on Signal Image Technology and Internet Based Systems, pp. 42-46, 2012.

[4] T. Mantecon, C. del-Bianco, F. Jaureguizar and N. Garcia, "Depth-based face recognition using local quantized patterns adapted for range data", 2014 IEEE International Conference on Image Processing (ICIP), pp. 293-297, 2014.

[5] B. Li, A. Mian, W. Liu and A. Krishna, "Face recognition based on Kinect", Pattern Analysis and Applications, pp. 1-11, 2015.

[6] G. Goswami, S. Bharadwaj, M. Vatsa and R. Singh, "On RGB-D face recognition using Kinect", 2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS), pp. 1-6, 2013.

[7] R. Min, N. Kose and J. Dugelay, "KinectFaceDB: A Kinect Database for Face Recognition", IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 44, no. 11, pp. 1534-1548, 2014.

[8] Jungong Han, Ling Shao, Dong Xu and J. Shotton, "Enhanced Computer Vision With Microsoft Kinect Sensor: A Review", IEEE Trans. Cybern., vol. 43, no. 5, pp. 1318-1334, 2013.

[9] M. Turk and A. Pentland, "Face recognition using eigenfaces", Proceedings. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 586-591.

[10] M. Turk and A. Pentland, "Eigenfaces for Recognition", Journal of Cognitive Neuroscience, vol. 3, no. 1, pp. 71-86, 1991.

[11] Docs.opencv.org, "Face Recognition with OpenCV — OpenCV 2.4.10.0 documentation", 2014. [Online]. Available: http://docs.opencv.org/2.4.10/modules/contrib/doc/facerec/facerec_tutorial.html. [Accessed: 10- Oct- 2015].