Proposing a high performance algorithm for face detection based on the Unscented Kalman Filter

*Bikash Lamsal and Naofumi Matsumoto (Ashikaga Institute of Technology)

Abstract - In this paper, we propose a high performance algorithm for detecting human faces in a still image. We are combining and modifying different types of algorithm in this paper. The algorithms are modified along with the low pass filter, Sobel edge detector and the Viola Jones eye detector. The low pass filter and the Unscented Kalman filter process play a vital role in removing the noises from the image. To clarify the effectiveness of our algorithm, we compare our proposed algorithms with other face detection algorithms through the benchmark tests. The ROC curve clarifies the effectiveness of our proposed algorithm.

Key words: Face detection, Unscented Kalman filter, Modified skin color detector, Modified Haar cascade classifier

1. INTRODUCTION

Face represents a physiological biometric identifier that is widely used in person recognition. Nowadays, the face detection process has become very common in computer vision systems. The process of locating the position of the face in an image is known as face detection. The process is performed either by using still or video images. In this paper, we are using still images for face detection. There seems to be a lot of applications and systems applied for face detection, but developing a system that will increase the face detection rate based on the static position of the face on an image will help to make the face detection system more advanced and crucial.

We are developing a new algorithm by using the different face detection algorithms in a different manner. We are using the different face detection process in our algorithm. The algorithm is developed in such a way that it combines the skin color detector \(^1\), the Haar cascade classifier, the eye detector \(^3\) and the Unscented Kalman filter process for detecting the faces in an image. The skin color detector is slightly modified by adding the low pass filter for removing noise from an image. The Sobel edge detector operates the edges of a face. Moreover, by selecting the skin color region of the face a facial candidate is detected. The detected facial candidate is then passed to an eye detector for checking the presence of eyes in a face. We have modified Haar cascade classifier \(^3\) by combining with the clustering algorithm. It even helps in increasing the accuracy of the face detection rate.

After combining these major face detection algorithms, we obtained a good face detection rate, but still some high frequency Gaussian noises were found in the images. So to remove the noises from a detected image, we use the Unscented Kalman Filter (UKF) process \(^9\) which remove the noise from the images at first and then passes the noise filtered images to the Modified Haar cascade classifier for detecting and verifying the face in an image under the different environmental conditions such as pose, scale, the absence of the structural component, facial expression, occlusion, Illumination variation, color region, multiple face detection and so on.

In this paper, we have positioned the different process in a proper way. The reason related to the positioning of the different process is mentioned in this paper. We have also mentioned the reason for using a low pass filter and the Unscented Kalman Filter process instead of using the Extended Kalman Filter process. This paper clearly shows the effect of noises in an image for the face detection.

Our process is applied slightly different from the other face detection algorithms implemented yet. The changes, modifications and the differences between our proposed algorithm and the others are mentioned in this paper clearly. As for evaluation metrics, we use the Receiver Operating Characteristic (ROC) curve. Finally, we clarify the effectiveness of our proposed algorithm by benchmarking using image databases of CMU-MIT \(^5\), INRIA Graz-01 \(^6\), MIT training set \(^7\), and FDDB datasets \(^8\).

In the area of face detection, many novel methods have been implemented. Yang \(^9\) provided an approach combining multiple color models for stable color based face detection. Several researches have been done regarding the success of Viola Jones algorithms \(^3\) which describes about the Haar feature calculation for face detection. The process of detecting and aligning faces by image retrieval process \(^10\) and Landmark localization \(^11\) shows the face detection process performed under different environmental conditions of occlusion, facial appearances, clutter and pose estimation. The LAB features \(^12\) have been developed from the inspiration of Haar features and Local binary pattern \(^13\) for face detection, which holds a good result for face detection but still lacks the highest face detection rate. Erdem et al. \(^14\) combined the Haar cascade classifier and Skin color detector, but were limited in these two processes only.

2. PROPOSED ALGORITHM

2.1 Structure of the proposed algorithm

The flow of the proposed algorithm is shown in Figure 1. Firstly, the skin color detector is modified by using a low pass filter, Sobel edge detector and modified Viola Jones eye detector process as shown in figure 2. Secondly, the Unscented Kalman Filter (UKF) process is applied for removing the high frequency film grain noises in the images. Finally, the noise filtered images are passed to the modified Haar cascade classifier, which is the combination of Haar cascade classifier and the Clustering algorithm.
2.2 Modified skin color detector and Modified Haar cascade classifier

The skin color detector is an algorithm that automatically recognizes skin tone and detects a face image as shown in figure 4, where the skin color region is detected but with the high frequency noises similar to the skin color. To clear this problem, we have modified the skin color detector.

The stepwise execution process is shown in figure 5. The high frequency noises are cleared by using the low pass filter, the edges of the face have been created and the skin color is easy to detect which results in the proper face candidate's detection. The modified Viola Jones eye detector is used to verify the presence of the eyes in a face.

In our proposed system, we are using the Haar cascade classifier for detecting the multiple faces in an image, but sometimes the false face detections, no face detection and multiple detection in a face occurs. In order to clear this error, we have modified the Haar cascade classifier by combining it with the clustering algorithm as shown in figure 3. We choose clustering algorithm for our process, because this algorithm can be used for basic pixel level to feature level for the classification of facial images. The execution examples for correct face detection using the clustering algorithm together with some experimental results for Haar cascade classifier is shown in figure 6. The images used in Figure 6 are taken from CMU-MIT database.
2.3 The Unscented Kalman filter process

The Unscented Kalman Filter (UKF) is similar to that of the Kalman filter. Like the Kalman filter, the UKF is based on a state-space formulation, but differs in the way it represents and propagates Gaussianity system dynamics.

Applying 2D Kalman filter to image noise reduction have begun by Woods 15, here the resulting filter is a general 2D recursive filter. They suggest that it is able to use to remove noises from still images. We have used the Unscented Kalman Filter (UKF) process in our algorithm. So, the UKF is used for removing the noise from an image in our process. The Unscented Transformation (UT) is a method for calculating the statistics of a random variable which undergoes a nonlinear transformation. The two dimensional UKF for image restoration 16 is used in our process. Gaussian white noise in an image is determined by the least squares method using the state equation and the measurement equation. The image is scanned by using the non-symmetric half plane (NSHP) model. The image is scanned from the top left pixel, and scan the first column from top to the bottom. This process is applied for eight times and we can achieve a better result with the noise filtered image. This process is the 2D UKF method 16 for clearing noises in a static image. A noise filtered image helps to increase the face detection rate. The images before and after applying the UKF process is shown in Figure 7.

2.4 Why the Unscented Kalman filter process?

In our proposed algorithm, we have used the Unscented Kalman Filter process instead of the Extended Kalman filter because the UKF yield performance equivalent to the Kalman filter for linear systems, yet generalizes elegantly to nonlinear systems without the linearization steps required for the EKF. The UKF consistently achieves a better level of accuracy than the EKF in systems with severe nonlinearities.

In case of the Image processing, the film grain noises that are not removed by the other noise filters are easily removed by the use of UKF. The removal of noises in the images causes a drastic change in the face detection from an image. The noises can be filtered by using the other filters too, but the face detection accuracy is high when compared to the other process. The ROC curve in figure 8 shows the difference of accuracy using the EKF and the UKF in our proposed algorithm.

2.5 Position of the applied processes

The position or the order of different process plays a vital role in an algorithm. In our process, the UKF is positioned after the modified skin color detector and before the modified Haar cascade classifier as shown in figure 1.

In the modified skin color detector, we are using the Low pass filter at the very beginning for filtering the high Gaussian noises from the images because it is easy to detect or create an edge from the noise filtered image.

After modified skin color detector, the images are once again passed towards the UKF for clearing the film grain noises that are still not being removed by the low pass filter. Those film grain noises takes part in the false detection of a face.

We also tried to use the UKF before the modified skin color detector, but when the UKF is applied, the images are blur, which means the modified skin color detector will results to a lot of false detections.

But in case of the modified Haar cascade classifier, the images are clustered and the duplication of the clusters is ignored using the Haar classifiers, so the modified Haar cascade classifier is used at the last for the proper face detection.
3. EVALUATION OF THE PROPOSED ALGORITHM

We evaluated our proposed algorithm on four public data sets. The databases used in our process are CMU-MIT database, MIT training sets, INRIA Graz-01 dataset and FDDB database. The CMU-MIT database, including 507 grayscale frontal face images is categorized as Database 1 (DB1). Database 2 (DB2) is the combination of three different databases; CMU-MIT Database (DB1), MIT training sets and INRIA Graz-01 datasets. Database 3 (DB3) is the FDDB database. Here, we set as follows.

**DB1:** CMU + MIT test sets; grayscale images, including 507 front faces.

**DB2:** CMU-MIT + MIT + INRIA Graz-01: MIT database includes training set images with 2429 colors and grayscale images, INRIA Graz-01 database includes 1316 color images with and without faces. The total number of images used DB2 is 507 + 2429 + 1316 = 4252 images. The images are color and grayscale images taken under different environmental conditions.

**DB3:** FDDB database; including 5171 faces in a set of 2845 images.

In order to validate our proposed algorithm, we developed a C programming code for performing the benchmarking process of our proposed algorithm. We performed different benchmarking tests based on the above-mentioned databases, i.e., DB1, DB2 and DB3 to validate our proposed algorithm. For the tests, we have also developed a system of benchmark testing support.

We consider a two-class prediction problem (binary classification), in which the outcomes are labeled either as positive or negative. We use ROC curves to evaluate the position of the proposed algorithm. A ROC curve shows the performance of our algorithm on DB1, DB2 and DB3 comparing with other face detectors on figure 9, 10 and 11 respectively. We use the correct detection rate as opposed to the true positive for the y-axis of the ROC curve to facilitate comparison with other detectors. The x-axis is the false positive. Performance of a face detector is measured by the area under a ROC curve. An area of 1.0 represents a perfect test; an area of 0.5 represents a worthless test.

![Fig.8. Comparison of EKF and UKF using DB3](image)

![Fig.9. Result of benchmark test for DB1](image)

![Fig.10. Result of benchmark for DB2](image)

![Fig.11. Result of benchmark for DB3](image)
4. DISCUSSION

Using our proposed face detector, we performed lots of experiments and the benchmark tests under various environmental conditions.

The proposed algorithm can detect faces with different scales. The multi-faces condition is being satisfied by our algorithm as shown in Figure 12 d. The occluded images are detected when both eyes are detected as shown in Figure 12 a.. The illuminated face with a shadow is partially detected in Figure 12 b, but still lacks 100% detection. The multifaces rotation invariant images are not detected as shown in figure 12 c.. The faces with expression and pose are detected by our proposed algorithm as shown in figure 12 e and 12 f.

The image noises are cleared properly by our proposed algorithm, which is the main cause of the high face detection rate.

The face detection rate is higher compared to the other face detection process. The face detection rate varies according to the Database used because of the combination of the different types of images. Color images, Grayscale images, noisy images and the images with different environmental conditions.

5. CONCLUSION

In this paper, we presented a high performance algorithm for face detection. We combined the different types of algorithms related to the face detection and image processing to work under different conditions. The used algorithms were modified by adding different features related to increase the face detection rate. In our face detector, the faces were detected individually by the two face detectors and verified by two types of eye detectors. The difficult issues on face detection were cleared stepwise with the flow of the proposed algorithm. Our algorithm was able to detect both the color and grayscale images taken under different environmental conditions, using different databases, i.e., DB1 and DB 2 and DB3. We performed our experiments using the PC (i7-2600 CPU, 4GB main memory and Windows7 64 bit) and 1080 pixel web camera.

The proposed algorithm calculates the face detection in 30 fps (frame per second), which means 0.033 second time is required for calculating the image per frame. This is the average calculated time of our proposed algorithm.

Our proposed algorithm while compared with other face detectors algorithms gives higher correct face detection rate. The individual comparison of our proposed algorithm with Viola Jones algorithm and LAB using DB1 also shows higher face detection rate by our side.

The overall comparison of the used algorithm and the proposed algorithm using DB2 also shows the percentage ratio of the participation of different algorithms for our proposed process.

The comparison of the algorithm with XZYJ and Viola Jones face detector using DB3 also showed the result from our side.

The Unscented Kalman filter was very helpful for filtering noises in images. The noise-filtered images and the images with noises had a great difference in case of face detection. The face detection rate for our proposed algorithm is 98.3%. This is the highest face detection rate compared to the other process related to face detection.

References
6). INRIA persons dataset, The EU project LAVA (IST-2001-34405) and the Austrian Science Foundation (project S9103-
(04)

7) K.K. Sung: “Learning and example selection for object and pattern recognition”, MIT, Artificial Intelligence Laboratory and Center for Biological and Computational Learning, (1996)


