Effects of the Unscented Kalman Filter Process for High Performance Face Detector

Bikash Lamsal and Naofumi Matsumoto

***Abstract*—This paper concerns with a high performance algorithm for human face detection in still image. The algorithm has been developed for increasing the face detection rate under different environmental conditions. In this paper, the skin color detector is modified by combining with the low pass filter, Sobel edge detector and the modified Viola Jones eye detector. The Haar cascade classifier is modified by using the Clustering algorithm. The modified algorithms are then combined together with the Unscented Kalman filter (UKF) in our process. The use of the UKF in case of the face detection algorithm simply reflects the novelty of our paper. The UKF is used for removing the film grain noise from the still image. To clarify the effectiveness of our proposed algorithm, we compare our proposed algorithm with other face detection algorithms through the benchmark tests using different facial databases. The ROC curve clarifies the effectiveness of our proposed algorithm with best face detection results of 98.3%.**

***Index Terms*—Face detection, modified Haar cascade classifier, modified skin color detector, the Unscented Kalman filter.**

I. 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 [2] 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 [4] 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.

Manuscript received December 15, 2014; revised April 5, 2015.

Bikash Lamsal and Naofumi Matsumoto are with the Department of Information and Production Engineering, Ashikaga Institute of Technology, Ohmae-cho268-1, Ashikaga-shi, Tochigi prefecture, Japan (e-mail: bikashaitjp@gmail.com, matsu@ashitech.ac.jp).

DOI: 10.7763/IJIEE.2015.V5.578

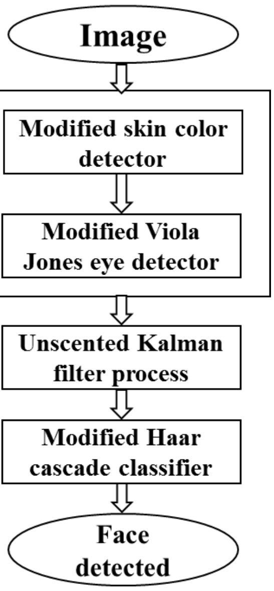


Fig. 1. Proposed algorithm.

II. Proposed Algorithm

A. Structure of the Proposed Algorithm

The structure of our proposed algorithm is reflected by the flow diagram shown in Fig. 1. Firstly, the skin color detector is modified by using a low pass filter, Sobel edge detector and color region detector of the face as shown in Fig. 2. The presence of two eyes in the facial candidates are detected by using the modified Viola Jones eye detector. Secondly, the Unscented Kalman Filter (UKF) process is used for removing the high frequency film grain noises from the images. Finally, the noise filtered images are passed to the modified Haar cascade classifier for detecting the face. The Modified Haar cascade classifier is the combination of the Haar cascade classifier and the Clustering algorithm as shown in Fig. 3.

B. Modified Skin Color Detector and Modified Haar Cascade Classifier

The skin color detector automatically recognizes skin tone and detects a face image as shown in Fig. 4, where the skin color region is detected along with the high frequency noises that are quite similar to the color of the skin. For clearing this problem, we have modified the skin color detector.

The stepwise execution process of the modified skin color detector is shown in Fig. 5. The high frequency noises are cleared by using the low pass filter, the edges of the face have been created and detected by using the Sobel edge detector, the skin color region is detected by using the color region detector and the facial candidates are detected. The modified Viola Jones eye detector is used to verify the presence of the eyes in a detected facial candidates.

In our proposed system, we are using the Haar cascade classifier for detecting the multiple faces in an image. We have modified the Haar cascade classifier by combining with the clustering algorithm [8] for clearing the false positives from the image as shown in Fig. 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 using the Haar cascade classifier and the modified Haar cascade classifier is shown in Fig. 6. The images used in Fig. 6 are taken from CMU-MIT database.

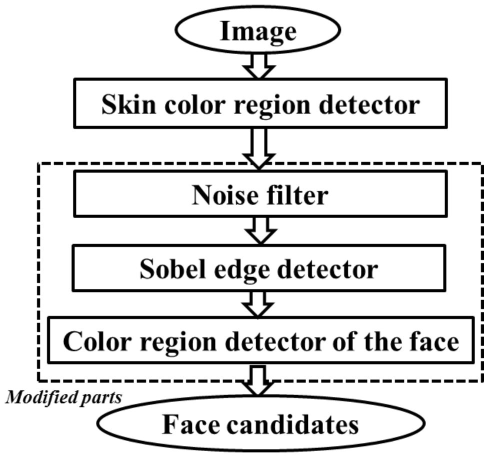


Fig. 2. Modified skin color detector.

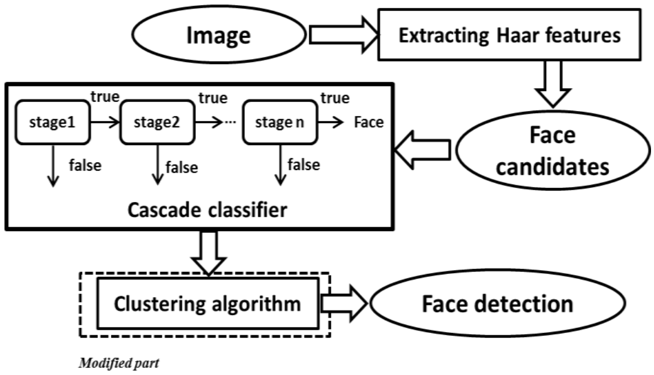
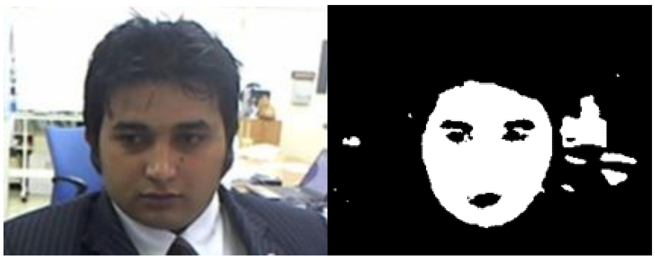
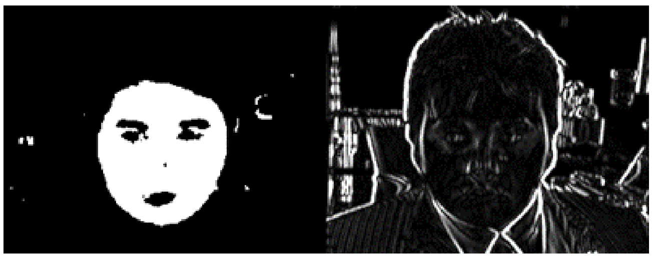


Fig. 3. Modified Haar cascade classifier algorithm.

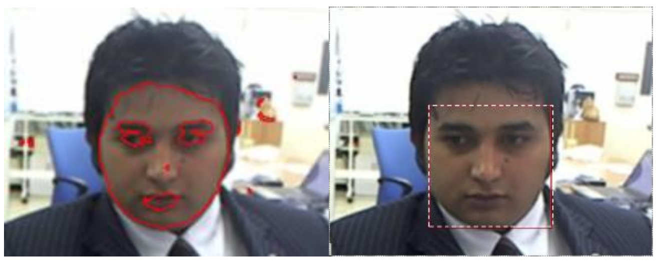


(a). RGB image. (b). Detected skin color region.

Fig. 4. Execution process of skin color detector.

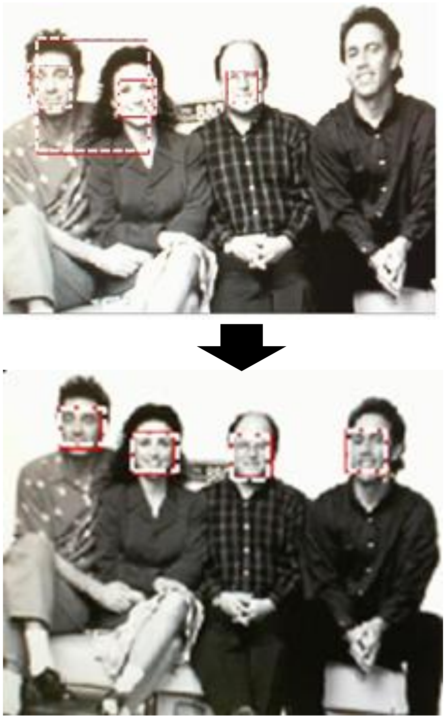


(a). Noise filtered skin region. (b). Face edge detection.

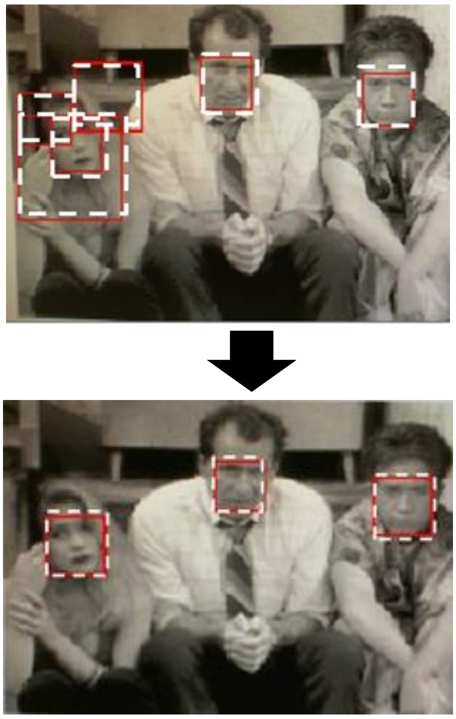


(c). Color region detection. (d). Face detected.

Fig. 5. Execution process of modified skin color detector.



(a). Effects of modified Haar cascade classifier for unrecognized face.



(b). Effect of modified Haar cascade classifier for clearing false positives.
Fig. 6. Execution examples of modified Haar cascade classifier.

C. The Unscented Kalman Filter Process

The Unscented Kalman Filter (UKF) is similar to that of the Kalman filter but works better compared 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 for removing the film grain noise from an image in our process.

Before

After

Fig. 7. Effects of the UKF process on an image from FDDB database.

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. Film grain 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 before and the after effects of the UKF process is shown in Fig. 7. The false positives are cleared after applying the UKF.

D. Why the Unscented Kalman Filter Process?

In our proposed algorithm, we have used the Unscented Kalman Filter process instead of the Extended Kalman filter (EKF) 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 can be 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 regular noises can be filtered by using the other filters too, but the film grain noises are very hard to filter because the film grain noises are hard to identify as well.

We also used the Extended Kalman filter for the experimentation and obtained the result which was comparatively low than the Unscented Kalman filter's result. The ROC curve for showing the effects of UKF, EKF and without using Kalman filter in our proposed algorithm is shown in Fig. 8. The proposed algorithm using the UKF has comparatively highest face detection rate, so we are using the Unscented Kalman filter in our process.

E. Position or Order of the Applied Processes

The position or the order of the applied process in an algorithm plays a vital role. In our process, the UKF is positioned after the modified skin color detector and before the modified Haar cascade classifier as shown in Fig. 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 through the UKF for clearing the film grain noises that are still not being removed by the low pass filter. Those film grain noises which are not removed by using the low pass filter, 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 becomes little blur, which means the modified skin color detector will results to a lot of false detections.

In case of the modified Haar cascade classifier, the Haar features are classified by using the Haar cascade classifier, which is also used for the face detection, but due to the duplication of the clusters in a group image, the clustering algorithm is used with the Haar cascade classifier, which results in the proper face detection with the true positives. The order of the different algorithms gives us a perfect result.

III. Evaluation of the Proposed Algorithm

A. Preparation for Experiments

In order to validate our proposed algorithm, we developed a C programming code for performing the benchmarking process of our proposed algorithm. Our C programming code compares our proposed algorithm of face detection with "Skin color detector," "Haar cascade classifier" and "Haar cascade classifier + skin color detector" by using the above-mentioned facial databases. We developed our system by using Visual studio 2010 and Open CV.

We performed experiments using the PC (i7-2600 CPU, 4GB main memory and Windows7 64 bit) and 1080M pixel web camera. Our system detects the faces and eyes in a red square box. The detected region either correct or false is spotted in a red square box. In this paper, the detected region is highlighted by a white square dotted box for making the detected portion clearly visible.

B. Databases Used

We evaluated our proposed algorithm on four public databases. 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. Database3 (DB3) is the FDDB database. Here, we set as follows.

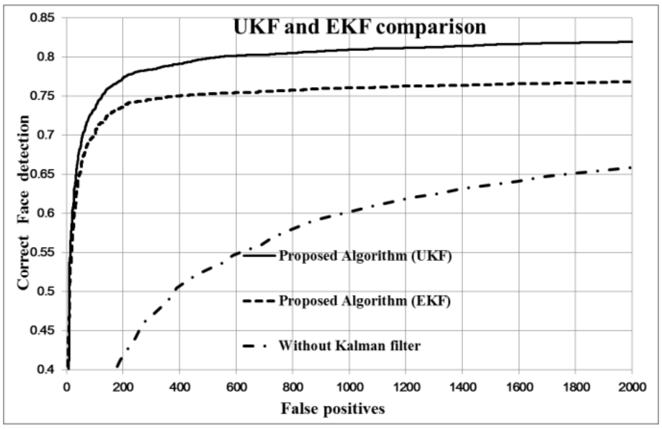


Fig. 8. Comparison of EKF and UKF using DB3.

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.

C. Benchmarking

We performed three 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 Fig. 9, Fig. 10 and Fig. 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. The curve which holds the large area in both horizontal and vertical axis is regarded to be the best result or high performance. An area of 1.0 represents a perfect test where the true positives and false positive varies but an area of 0.5 represents a worthless test where true positive and false positive rates are quite similar.

First, we have used DB1 database for comparing the face detection rate between our proposed algorithm, Viola Jones face detector [3] and LAB method [12]. We draw the ROC curve from the result of this test. The ROC curve is shown in Fig. 9. The ROC curve shows our algorithm having the correct face detection rate higher than that of Viola Jones face detector and LAB methods. The correct face detection rate for Viola Jones face detector is 94.1% [3]. The LAB method shows a very good face detection rate that started with 90% to 96% with the constant ratio of face detection. But basically, the true positive rate can reach 1 when the false positive rate is 1, our proposed algorithm shows the face detection rate of 98.3%. The ROC curve is shown in Fig. 9 which clarifies that our proposed algorithm has higher performance than Viola Jones face detector and LAB methods [12].

Second, we have used DB2 for comparing the difference of ROC curves between our proposed algorithm and the other face detectors used in our system. The ROC curves in Fig. 10 shows the results. The "Haar cascade classifier + Skin color detector" also represents our modified algorithm which shows the face detection rate without applying the Unscented Kalman filter process in our system. From the ROC curves, we clarify that our proposed algorithm has higher performance for face detection than other ones. This curve is useful for showing the participation of different algorithms in our proposed method.

Finally, we have used DB3 for comparing our proposed algorithm with XZYJ method [10] and Viola Jones face detector [10]. The ROC curve is shown in Fig. 11. The ROC curve is drawn from the first 2000 images by referring the Roc curves shown in [10]. The Proposed algorithm shows the highest face detection rate compared to the XZYJ methods and the Viola Jones face detector.

IV. Discussion

We performed various experiments and the benchmark tests under different environmental conditions using our proposed face detector.

The proposed algorithm can detect faces with different ***scales.*** The ***multi-faces*** condition is being satisfied by our algorithm as shown in Fig. 12(d). The ***occluded*** images are detected when both eyes are detected as shown in Fig. 12(a). The ***illuminated*** face with a shadow is partially detected in Fig. 12(b), but still lacks 100% detection. The multifaces ***rotation invariant*** images are not detected as shown in Fig. 12(c), where only one face in a group of 5 faces is detected. The faces with ***expression*** and ***pose*** are detected by our proposed algorithm as shown in Fig. 12(e) and Fig. 12(f).

The ***presence or absence of structural components*** in a face such as the faces with beard, masks, glasses on the eyes, etc. are also detected by our proposed algorithm. The ***image noises*** are cleared properly by our proposed algorithm, which is the main cause of the high face detection rate. The high Gaussian noises and the film grain noises in images are cleared by using our proposed algorithm.

We will be developing some new ideas for clearing the unsolved environmental conditions such as occlusion, rotation invariant, dark images and illuminated faces; to increase the face detection rate.

V. 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.

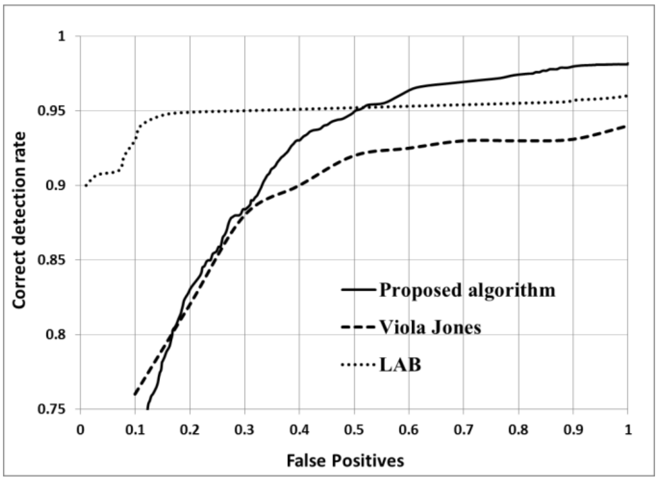


Fig. 9. Result of benchmark test for DB1.

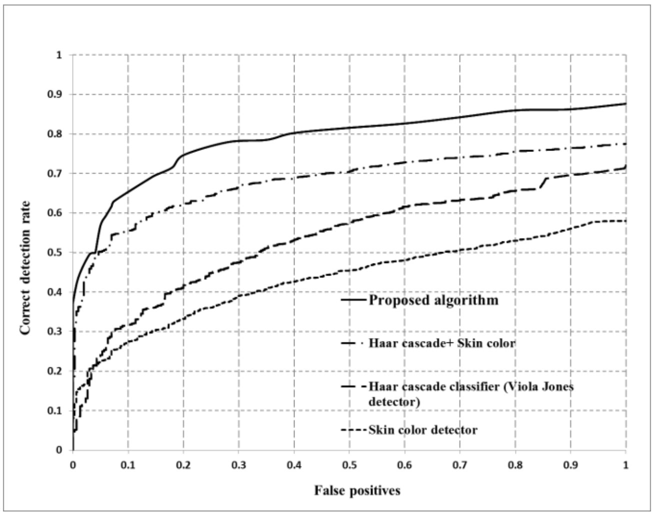


Fig. 10. Result of benchmark for DB2.

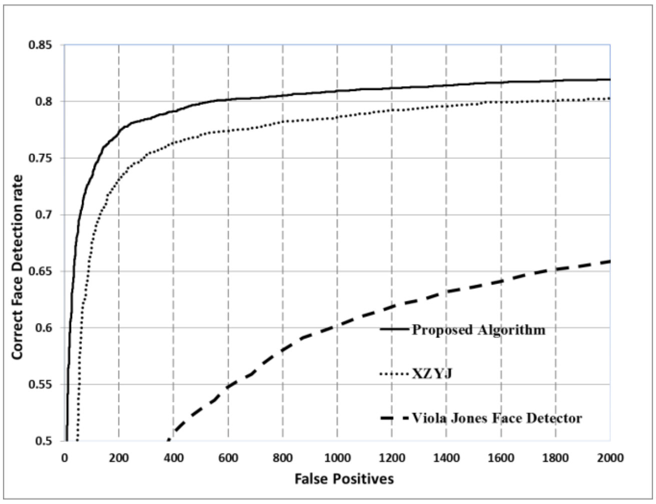
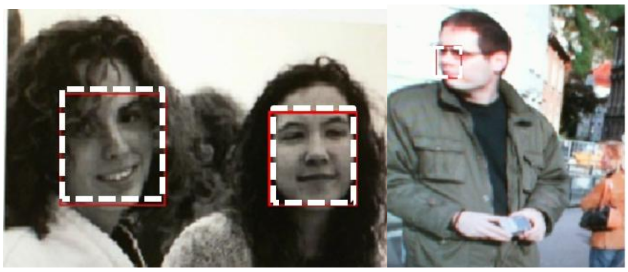


Fig. 11. Result of benchmark for DB3.

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.

(a). Occlusion. (b). Illumination.

(c). Rotation invariant. (d). Multifaces.



(e). Expression . (f). Pose.

Fig. 12. Face detection under different environmental conditions.

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 by 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

[1] K. Hani, A. Mohair, J. M. Saleh, and S. A. Suandi, "Human skin color detection: A review on neural network perspective," *Intl. J. of Innovative Computing, Information and Control*, vol. 8, no. 12, pp. 8115-8131, 2012.

[2] H. R. Lien, A. M. Mohamed, and K. J. Anil, "Face detection in color images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 696-706, 2002.

[3] P. Viola and M. Jones, "Robust real-time face detection," *Intl. J. of Computer Vision*, vol. 57, no. 2, pp. 137-154, 2004.

[4] R. Wang, L. Sufei, and E. E. Kuruoglu, "A novel algorithm for image denoising based on unscented Kalman filtering," *International Journal of Information and Communication Technology*, vol. 5, pp. 343-353, 2013.

[5] H. A. Rowley, S. Baluja, and T. Kanade, "Neural network based face detection," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 20, pp. 23-38, 1998.

[6] INRIA Persons Dataset, The EU project LAVA (IST-2001-34405) and the Austrian Science Foundation (project S9103-N04).

[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

[8] V. Jain and E. L. Miller, "FDDB: A benchmark for face detection in unconstrained settings," Technical Report UM-CS-2010-009, Dept. of Computer Science, University of Massachusetts, Amherst, 2010.

[9] M. H. Yang, D. Kriegman, and N. Ahuja, "Detecting faces in images: A survey ," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 1, pp. 34–58, 2002.

[10] X. H. Shen, Z. Lin, J. Brandt, and Y. Wu, "Detecting and aligning faces by image retrieval," presented at IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013.

[11] X. Zhu and D. Ramanan, "Face detection, pose estimation and landmark localization in the wild," presented at IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012.

[12] S. Yan, S. Shan, X. Chen, and W. Gao, "Locally assembled binary (LAB) feature with feature-centric cascade for fast and accurate face detection," presented at CVPR, 2008.

[13] K. Levi and Y. Weiss, "Learning object detection from a small number of examples: The importance of good features," presented at CVPR 2004.

[14] C. E. Erdem, S. Ulukaya, A. Karaali, and A. T. Erdem, "Combining Haar feature and skin color based classifiers for face detection," *Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 36, pp. 1497-1500, 2011.

[15] J. W. Woods, "Kalman filtering in two dimensions," *IEEE Transactions on Information Theory*, vol. IT-23, no. 4, 1977.

[16] R. Wang, L. Sufei, and E. E. Kuruoglu, "A novel algorithm for mage denoising based on unscented Kalman filtering," *International Journal of Information and Communication Technology*, vol. 5, pp. 343-353, 2013.

Bikash Lamsal was born in Pokhara, Nepal in November 1987. He received his masters degree in system and information engineering from the Ashikaga Institute of Technology, Tochigi, Japan in March 2014. Currently he is a Phd scholar in information and production engineering at Ashikaga Institute of Technology, Japan since April 2014.

He is working as a research assistant in the Department of Information System Design at the Matsumoto Laboratory, Ashikaga Institute of Technology, Japan from April 2014. His research interest includes face detection, image processing, augmented reality, virtual reality, software design, robotics and algorithm design.

Naofumi Matsumoto received his doctoral degree of engineering from Sophia University, Tokyo, Japan in 1978. Currently, he is a professor in the Department of Information System Design, at Ashikaga Institute of Technology, Tochigi, Japan.

He is the member of JSME and Information Processing Society of Japan. His research interest includes system control, soft-computing and system engineering.