A RELIABLE HYBRID TECHNIQUE FOR HUMAN FACE DETECTION

Ayesha Hakim, Stephen Marsland and Hans W. Guesgen
School of Engineering and Advanced Technology, Massey University, Palmerston North, New Zealand

Keywords: Face Detection, Haar-classifier, Skin-colour, Occlusion.

Abstract: The progress of computer vision technology has opened new doors for interactive and friendly computer interfaces. Human face detection is an essential step of various human-related computer applications, including face recognition, emotion recognition, lip reading, and several intelligent human computer interfaces. Since it is the basic step in such applications, it must be reliable enough to support further steps. Several approaches to detecting human faces have been proposed so far, but none of them can detect faces in all different conditions such as varying lighting conditions; frontal, profile, tilted and rotated faces; occlusions by glasses, hijab, facial hair; and noise. We propose a more reliable hybrid approach that is able to detect human faces in multiple circumstances. Moreover, a brief, but comprehensive, review of the literature is presented that may be useful to evaluate any face detection system. Our proposed approach gives up to 97% accuracy on 600 images (both simple and complicated), which is the highest accuracy rate reported to date to our knowledge.

1 INTRODUCTION

The current growth of computer technologies has paved the way to a new machinery world where human life is improved by artificial intelligence. Research efforts in human-computer interaction aim to find ways to enable computers to interact with humans in more natural ways, e.g. by recognizing their gestures, speech, hand writing, and even emotions.

Human face detection is an essential first step in almost all face-related problems. It involves localizing and extracting the face region from the rest of the image (Hjelmas and Low, 2001). The objective of face detection is to find out whether or not there are any faces present in the image and, if so, to return the location and size of each (Yang, Kreigman and Ahuja, 2002). It helps to limit the search space for facial features since the system does not have to search for features in the whole image (McDermott, 2006). It also has numerous other applications in areas such as human face recognition, emotions recognition, sign language recognition, lip reading, face focusing in cameras, and other intelligent human-computer interfaces.

Unfortunately, human face detection is not an easy task. Depending on the camera-face pose, some facial features might be partially or totally occluded, which might make it difficult to detect the face. Varying lighting conditions might result in some parts of the faces being only partially lit, which might mean that not all features are clearly visible. Also, there is a wide variety in the appearance of faces, both because of natural variation and because of additions such as facial hair and glasses. Various facial expressions and physiognomies of faces also cause problems, since they change the normal appearance of the face. Some images have a single face, others have multiple or no face at all, some may have complex backgrounds and some may be noisy due to poor quality images. All such problems are big hindrances in detecting faces from images.

We concentrate on the problem of human face detection in static images having complex background; varying lighting conditions; glasses; facial hair; various facial expressions (spontaneous and posed); considerable variations in head poses; occlusions; noise; profile, tilted, and rotated faces.

In this paper we propose a hybrid method that uses two of the most successful (and complementary) algorithms and uses them each appropriately. This provides a method that is significantly more accurate than either alone and also significantly faster, on average, than one of

Hakim A., Marsland S. and W. Guesgen H. (2010).
A RELIABLE HYBRID TECHNIQUE FOR HUMAN FACE DETECTION.
In *Proceedings of the International Conference on Computer Vision Theory and Applications*, pages 241-244

them.

See full version of paper, table comparing various published results and our experimental results on website (*http://muse.massey.ac.nz*).

2 RELATED WORK

Early efforts to develop automatic systems for human face detection began at the start of 1970s, but progress remained slow until the 1990s. The literature reveals a remarkable rise in interest in this research topic over the past decade (Hjelmas and Low, 2001). The first problem in the area of human face detection is the search for some 'standard dataset' that can be treated as a target to compare detection rates between algorithms. Unfortunately, there is no such standard comprehensive database available and the databases that most of the research groups used for evaluation consist of gray scale images. So the first step was to collect an RGB image dataset that can be used to test all of the above mentioned conditions.

To date, various methods have been proposed for human face detection from images. Each method has its own benefits and limitations, but there has not been a consistent review and comparison of face detection methodologies. Since face detection is the basic step, it must be very efficient and reliable. The literature reveals that the systems that are reliable are not generally efficient, and vice versa. Moreover, a lot of costly training is usually required as a pre-processing step.

Some of the most widely used methods of human face detection are based on Haar classifiers (Viola and Jones, 2004; Lienhart and Maydt, 2002), human skin colour (Hsu, Abdel-Mottaleb and Jain, 2002; Wang and Yuan, 2001; Singh et al., 2003; Lin et al., 2008), and facial feature detection (Lee, S. Park and M. Park, 2005). These methods, although capable (to some extent) of detecting human faces alone, do not cover the wide spectrum of different conditions due to some limitations.

Viola and Jones (Viola and Jones, 2004) proposed using boosting of Haar-like features to detect the face region. Using this method, real-time detection can be achieved with the help of very simple and easily computable Haar-like features, and a cascade of boosted classifiers. AdaBoost was used to select the most representative features in a large set. On the other hand, Haar classifier detection results are highly dependent on image quality, contrast and brightness and it gives false positives or false negatives if the image is blurred or face in the image is occluded.

Most researchers report only false negatives, i.e. faces that are not detected. However, the problem of false positives (identifying a region as a face when it is not) is also a potential problem with face detection algorithms. There is also a significant difference between what is considered a correct result. Some papers report any result that includes a part of the face as correct. For frontal, non-occluded images we require that all of the eyes, nose, and lips (primary features) are included in the face 'box'.

Skin colour-based detection methods perceive skin regions over the whole image and create face candidates on the basis of several image processing techniques. These methods help to detect faces under varying poses (partially occluded or rotated), but are highly dependent on lighting conditions and are not reliable alone. Although image processing and morphological operations improve the false positives, the method still fails to differentiate between face and anything having similar colours to skin.

Feature-based detection algorithms aim to locate faces on the basis of facial features (such as eyes, eyebrows, nose, mouth, and hair-line). The problem with this method is that the image features can be highly variable due to noise, illumination, and occlusion. So this method may be used to verify the detected face, but cannot detect the face reliably.

Due to each method's limitations, they all suffer from varying conditions (mentioned in Section 1) and are not able to detect faces reliably.

3 PROPOSED APPROACH

We propose a hybrid approach for human face detection in static images based on two of the most widely used algorithms boosted Haar classifiers and skin colour, both of which are described in Section 2. First of all, the image is processed by a Haar classifier, because this method is very efficient and reliable for frontal faces. If this method fails to detect any face present in image (usually, because of its limitations as discussed in Section 2), the image is handed over to a skin colour tester. The skin colour tester consists of multiple steps including skin likelihood detection and segmentation of skin region. The resulting image passes through some morphological operations, aspect ratio test, and finally goes through a template matching test. Due to various operations and tests, skin colour-based detection is time consuming. If this method fails, the system automatically adjusts the image brightness and lighting conditions. After adjustment, the possibility of face detection by Haar classifier increases, so it is again tested by Haar.

We also experimented with an eye-detection test following face detection to avoid false positives. While it was effective for some images, the time it took and the low rate of false positives meant that we do not need to use it in the final system. Not using it also makes the system better able to deal with occlusion and camera pose.

4 EXPERIMENTAL RESULTS

Due to the unavailability of a standard accessible dataset of RGB images, there is a strong need to collect a set of images that can evaluate face detection systems under a wide range of different conditions (as mentioned in Section 1). By collecting the images from various 'accessible' sources, along with our personal and web images, we obtained a set of images that seems to fulfil all possible conditions.

We tested our system on 186 images from the Psychological Image Collection at Stirling (PICS) (http://pics.psych.stir.ac.uk, n.d.) and got a detection rate of 98% with 1 false positive. Peer and Solina (1999) reported 97.7% average detection rate on 44 images, while Wang and Sung (1999) showed almost 90% detection rate on 50 randomly selected images from PICS dataset.

Using 60 images from the MMI facial expression dataset (Pantic et al., 2005), 35 images from the Indian face dataset (Jain and Mukherjee, 2002), 100 images from Libor Spacek's facial image dataset (Spacek, n.d.), and 13 images from the AR face dataset (Martinez and Benavente, 1998) we got 100% detection rate with 9, 0, 9 and 1 false positive respectively. Anisetti et al. (2006) used AR face dataset testing, which gives 72.9% average detection rate on images without black sun glasses and yellow light.

On the XM2VTS (Messer et al., 1999) sample set (54 frontal, 32 side profile and 2 dark frontal view images), we got 94% detection rate with 10 false positives. However, it gives 100% detection rate with 0 false positives on frontal faces only. Asteriadis, Nikolaidis and Pitas (2009) used only the frontal faces of XM2VTS dataset and got 99.74% average detection rate.

On 81 personal photo collections, some other test sets and the web images we got 97% detection rate with 12 false positives. Wang and Yuan (2001) gives 91.1% detection rate on web images, Hsu, Abdel-Mottaleb and Jain (2002) gives 80.35% detection rate on personal and news images, and Garcia (2004) gives 90.5% detection rate on web test set. On 75 very complicated images from yahoo news we got 86% detection rate with 12 false positives.

These images were not rescaled before testing, have different lighting conditions, multiple facial expressions, facial hair, glasses, mufflers, hijab, multiple variations of head poses and camera angle. As compared to previous researches results, on such a diverse collection of images, the average result of 97% detection rate on almost 600 images is quite satisfactory.

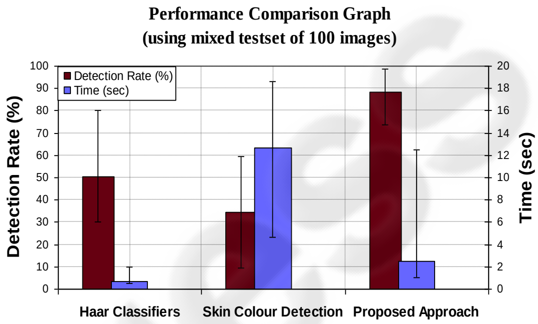

Figure 1: Performance comparison graph.

We compared the accuracy and efficiency of the proposed system with the most commonly used face detection techniques such as the Haar classifier and skin colour based detection. The methods were tested against 100 mixed images of varying conditions. The graph in Figure 1 shows that Haar classifiers gives 50.25% accuracy in 0.677 seconds per image, while the skin colour method gives 34.46% accuracy in 12.63 seconds per image, while our proposed method gives the highest accuracy of 88.50% in 2.54 seconds per image. These results are obtained by using MATLAB on a Pentium 4 CPU 3.40 GHz with 1.00 GB RAM. The high deviation of time in the proposed system is caused by the skin colour detection method being used.

Noise is the most common problem in images, caused by low quality cameras, foggy environments, dust, smoke, or motion. We tested our system against different kind of noises (Salt & pepper noise, Gaussian noise, and Speckle noise), and got satisfactory results. Two different datasets, each having 20 images, were used for testing the noise resistance. Dataset A contains all frontal face images and dataset B contains mixed complicated images.

In the case of salt & pepper noise, for dataset A the system shows 100% accuracy with 70% noise intensity, while for dataset B it shows up to 80%

accuracy with 60% noise intensity. In the case of Gaussian noise, the results fluctuate with noise variance. For dataset A, the system gives 100% accuracy up to 2% noise variance, while it fluctuates between 0 and 100 up to 6% variance, while for dataset B the accuracy fluctuates between 0 and 100 up to 4% noise variance and between 0 to 30 up to 6% variance. Similarly, in the case of Speckle noise, the results fluctuate with noise variance. For dataset A, the system gives 100% accuracy up to 2% noise variance, while it fluctuates between 90-95% up to 10% variance, while for dataset B the accuracy fluctuates between 85 and 100 up to 5% noise variance and between 75 to 100 up to 9% variance.

5 CONCLUSIONS

This paper has presented a hybrid system for human face detection from static images, by combining two methods and some pre-processing steps that is more efficient than either and not too much less computationally efficient than the better of the two. We have shown that it is fairly robust to common image problems such as noise and occlusions.

A brief review of the literature is presented along with the most comprehensive set of RGB images which fulfils all possible conditions for evaluation of any face related application. The proposed system is not complex and covers a wide range of human skin colours.

We intend to follow two fronts of research from here. The first is to use these results and extend them to video images, which will combine tracking with face detection. We will investigate further algorithmic speed-ups for this to work in real time. The second is to use the segmented face for human emotion recognition which is the main focus of our research.

REFERENCES

Anisetti, M., Bellandi, V., Damiani, E., Beverina, F., Arnone, L. & Rat, B. 2006, ‘A3fd: Accurate 3d face detection’, *Proc. of IEEE Int'l. Conf. on Signal-Image Technology and Internet Based Systems.*

Asteriadis, S., Nikolaidis, N., Pitas, I. 2009, ‘Facial feature detection using distance vector fields’, *Pattern Recognition*,1388-1398.

Garcia, C., Delakis, M. 2004, ‘Convolutional face finder: A neural architecture for fast and robust face detection’, *IEEE Trans on Pattern Analysis and Machine Intelligence*, 1408–1423.

Hjelmas, E., Low, B. K. 2001, ‘Face Detection: A Survey’, *Computer Vision and Image Understanding,* 236-274.

Hsu, R., Abdel-Mottaleb, M., Jain, A. K. May 2002, ‘Face Detection in Colour Images’, *IEEE Trans. on Pattern Analysis and Machine Intelligence.*

Jain, V., Mukherjee, A. 2002, ‘The Indian Face Database’.

Lee, T., Park, S., Park, M. 2005, ‘A New Facial Features and Face Detection Method for Human-Robot Interaction’, *Proc. of the 2005 IEEE Int'l. Conf. on Robotics and Automation*, 2063-2068.

Lienhart, R., Maydt, J. 2002, ‘An Extended Set of Haar-like Features for Rapid Object Detection’, *IEEE ICIP*.

Lin, H., Yen, S., Yeh, J., Lin, M. 2008, ‘Face Detection Based on Skin Colour Segmentation and SVM Classification’, *Second Int'l. Conf. on Secure System Integration and Reliability Improvement*, 230-231.

Martinez, A. M., Benavente, R. 1998, The AR Face Database. *CVC Technical Report #24*.

McDermott, J. 2006, ‘Facial Feature Extraction for Face Characterization’.

Messer, K., Matas, J., Kittler, J., Luettin, J., Maitre, G. 1999, ‘XM2VTSDB: the extended M2VTS database’.

Pantic, M., Valstar, M. F., Rademaker, R., Maat, L. 2005, ‘Web-based Database for Facial Expression Analysis’, *Proc. IEEE Int'l Conf. Multmedia and Expo*.

Peer, P., Solina, F. 1999, ‘An Automatic Human Face Detection Method’, *Proc. of the 4th Computer Vision Winter Workshop (CVWW’99)*, 122–130.

Singh, S. Kr., Chauhan, D. S., Vatsa, M., Singh, R. 2003, ‘A Robust Skin Colour Based Face Detection Algorithm’, *Tamkang Journal of Science and Engineering*, 227-234.

Spacek, L., Facial images dataset.

Viola, P., Jones, M. J. 2004, ‘Robust real-time object detection’, *Int'l Journal of Computer Vision*, 137-154.

Wang, J. G., Sung, E. 1999, ‘Frontal view face detection and facial feature extraction using color and morphological operations’, *Pattern Recognition*, 1053-1068.

Wang, Y. Yuan, B. 2001, ‘A novel approach for human face detection from colour images under complex background’, *Pattern Recognition*, 1983-1992.

Yang, M., Kreigman, J., Ahuja, N. 2002, ‘Detecting Faces in Images: A Survey’, *IEEE Trans. On Pattern Analysis and Machine Intelligence.*