SKIN COLOUR INFORMATION AND MORPHOLOGY BASED FACE DETECTION TECHNIQUE

M. Sharmila Kumari, Akshay Kumar, Rohan Joe D'Souza, G K Manjunath and Nishan Kotian

Department of Computer Science and Engineering,
PA College of Engineering, Mangalore.
{sharmilabp@gmail.com, akshay.holmes@gmail.com, dsouza.rohanjoe@gmail.com, manjugreat.gk@gmail.com, nishank.kotian@gmail.com}

ABSTRACT

Locating and tracking human faces is a prerequisite for face recognition and/or facial expressions analysis, although it is often assumed that a normalized face image is available. In this paper, we propose a faster, yet efficient face detection approach based on mathematical morphology and skin colour information. We have devised some simple post-processing rules to eliminate non-face regions from face regions. Experimentation on our created database is conducted to reveal the performance of the proposed approach.

KEYWORDS

Skin colour, Morphology, Segmentation, Face detection.

1. INTRODUCTION

Human face detection is currently an active research area in the computer vision community. Face localization and detection is often the first step in applications such as video surveillance, human computer interface, and face recognition and image database management. Locating and tracking human faces is a prerequisite for face recognition and/or facial expressions analysis, although it is often assumed that a normalized face image is available. In order to locate a human face, the system needs to capture an image using a camera and a frame-grabber to process the image, search the image for important features and then use these features to determine the location of the face. For detecting face there are various algorithms including skin colour based algorithms. Using skin-colour as a feature for tracking a face has several advantages. Colour processing is much faster than processing other facial features. Under certain lighting conditions, colour is orientation invariant. This property makes motion estimation much easier because only a translation model is needed for motion estimation. However, colour is not a physical phenomenon; it is a perceptual phenomenon that is related to the spectral characteristics of electromagnetic radiation in the visible wavelengths striking the retina.

The accurate detection of human faces in arbitrary scenes is the most important process involved prior to face recognition. When faces could be located exactly in any scene, the recognition step

afterwards would not be so complicated. Face detection in completely unconstrained settings remains a very challenging task, particularly due to the significant pose and lighting variations. The modern face detectors are mostly appearance-based methods, which mean that they need training data to learn classifiers. Collecting a large amount of ground truth data remains a very expensive task, which certainly demands more research. In environments which have low variations, adaptation could bring very significant improvements to face detection. In this context, we proposed a simple, yet efficient algorithm for face detection which works better even in complex background. The details of the proposed methodology are presented in the following sections. We have created our own database and test the system against the ground truth data.

2. LITERATURE SURVEY

The problem of face detection goes back to early 70's, at that time the overall focus was on finding ways to detect human faces in simple constitutions where typically the face is in a passport like photo with a uniform background and uniform lighting conditions. The research on the subject was rather simple at the time and only came into attention at early 90's where more in-depth research was taking over into the problem using different algorithms, the problem began to address the issues related to detecting faces in complex backgrounds with different scales and rotation degrees, introducing statistical methods and neural networks for face detection in cluttered scenes.

There have been hundreds of reported approaches to face detection. Early Works (before year 2000) had been nicely surveyed in [21] and [4]. For instance, Yang et al. [21] grouped the various methods into four categories: knowledge-based methods, feature invariant approaches, template matching methods, and appearance-based methods. The field of face detection has made significant progress in the past decade. In particular, the seminal work by Viola and Jones [17] has made face detection practically feasible in real world applications such as digital cameras and photo organization software. Mita et al. [10] proposed joint Haar-like features, which is based on co-occurrence of multiple Haar-like features. The authors claimed that feature co-occurrence can better capture the characteristics of human faces, making it possible to construct a more powerful classifier. Another well-known feature set robust to illumination variations is the local binary patterns (LBP) [11], which have been very effective for face recognition tasks [1, 22]. In [5, 23], LBP was applied for face detection tasks under a Bayesian and a boosting framework, respectively. More recently, inspired by LBP, Yan et al. [20] proposed locally assembled binary feature, which showed great performance on standard face detection data sets.

Another popular complex feature for face/object detection is based on regional statistics such as histograms. Levi and Weiss [8] proposed local edge orientation histograms, which compute the histogram of edges orientations in sub-regions of the test windows. These features are then selected by an AdaBoost algorithm to build the detector. Later, Dalal and Triggs [2] proposed a similar scheme called histogram of oriented gradients (HoG), which became a very popular feature for human/pedestrian detection [24, 16, 7, 3].

Training a face detector is a very time-consuming task. In early works, due to the limited computing resources, it could easily take months and lots of manual tuning to train a high quality face detector. A number of papers have been published to speed up the feature process. Wu et al. [18] proposed a cascade learning algorithm based on forward feature selection. Pham and Cham [12] presented another fast method to train and select Haar features. It treated the training examples as high dimensional random vectors, and kept the first and second order statistics to build classifiers from features.

Multiview face detection has also been explored with SVM based classifiers. Li et al. [9] proposed a multiview face detector similar to the approach in [15, 6]. They first constructed a face pose estimator using support vector regression (SVR), and then trained separate face detectors for each face pose. Yan et al. [19] instead executed multiple SVMs first, and then applied an SVR to fuse the results and generate the face pose. This method is slower, but it has lower risk of assigning. Neural networks were another popular approach to build a face detector. Early representative methods included the detectors by Rowley et al. [14] and Roth et al. [13].

It shall be observed from the above brief survey that the face detection in completely unconstrained settings remains a very challenging task, particularly due to the significant pose and lighting variations. In addition, collecting a large amount of ground truth data remains a very expensive task, which certainly demands more research. In this context, we have made an attempt to design a simple yet efficient face detector which works in real environment.

3. PROPOSED METHODOLOGY

Our algorithm is designed to overcome some of the limitations present in the existing algorithms. The existing algorithms use template matching as their primary segmentation algorithm. This process is very time consuming and hence it is proposed to detect face regions without template matching. The model that we proposed is a combination of feature-based and view based approaches. Initially the image is divided based on skin colour. This is the most vital stage as the success of the entire algorithm depends on the efficiency of the segmentation. Thus an extensive time has been given to this stage in the entire process of algorithm building. Once the segmentation of the probable skin regions are identified, then regions to be searched for potential faces have been reduced based on mathematical morphology thus improving the efficiency of the algorithm in terms of time. The very next step is removal of noisy pixels in the image. Here the noise is considered to be a part of the background or any non-face region for that matter. Thus an effective solution has been proposed in terms of morphological operations which eliminate noise and also highlight the face region. This is followed by the design of some heuristic rules based on statistical properties of an image to eliminate non-face regions from the probable face regions. The robustness can be measured from the fact that it works very well for multiple faces in complex backgrounds as it does for simple backgrounds too.

EXPERIMENTAL ILLUSTRATION:

Here, we present the proposed methodology with a sample image shown in Fig. 1.

Fig. 1. Sample image containing face and non-face regions.

Stage1: Segmentation based on Skin colour:

In this stage, probable face regions are identified based on skin color. Prior to segmentation of the skin, the given image is converted into a particular color model. In this case we convert into YCbCr color model. Then thresholds are setup to filter out non-skin pixels based on global thresholding. The probable face regions identified based on skin colour information followed by thresholding is shown in Fig. 2.

Fig. 2. Probable skin regions identified due to skin colour based thresholding.

Stage2: Morphology to eliminate noise:

It shall be observed from Fig. 2 that the skin colour segmentation rejects non-skin colours from the input image. However, the resulting image has quite a bit of noise and clutter. A series of morphological operations are performed to remove the noisy pixels in the image and a masked regions are generated which are placed on the input image to yield skin colour regions without noise and clutter.

Since morphological operations work on intensity images, the colour image is converted into a grey scale image. The intensity thresholding is performed to break up dark regions into many smaller regions so that they can be cleaned up by morphological opening. The threshold is set low enough so that it does not chip away parts of a face but only create holes in it. The morphological opening is performed to remove very small objects from the image while preserving the shape and size of larger objects in the image. A disk shaped structuring element of radius 1 is used. The hole filling is done to keep the faces as single connected regions in anticipation of a second much larger morphological opening. Otherwise, the mask image will contain many cavities and holes in the faces. The morphological opening is performed to remove small to medium objects that are well below the size of a face. A disk shaped structuring element of radius 6 is used in this case. The result of applying the mask to the grey scale version of the input image is shown in Fig. 3.

Fig.3 Mask image generated as the output of morphological operations

Step3: Removal of unwanted regions:

The output from the previous stage contains some non-face regions and hence some heuristic rules are designed to eliminate non-face regions which is achieved by converting a gray scale image to a binary image. The statistical properties of each region are used to eliminate non-face regions from face region. The area of regions in the binary image is then calculated. It shall be observed from Fig. 4 that there are there are numerous non-face regions. Few of the non-face regions occupy small areas and hence these regions can be removed by using a threshold.
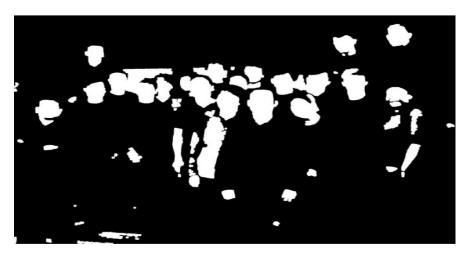

Fig.4. Face and Non-face regions after binarization

Fig.5. Some non-face regions eliminated binary image

Based on the above regions, bounding box is drawn and the results are shown in Fig. 6. It shall be observed that there are certain non-face regions which is due to the fact that the skin colour is same for both hands and face regions and hence the proposed approach produce certain false face regions. The results are shown in Fig. 6.

Fig. 6. Face and non-face regions with bounding box

Step 4: Removal of non-face regions:

It shall be observed from Fig. 6 that the output image contains many unwanted regions like fist, shoulders, jacket, exposed hands and arms etc. These are removed based on the fact that they occupy lesser area than the other regions or that the aspect ratio does not depict a face region. Thus we have set the following criteria based on which the eliminations of unwanted regions are made:

- Based on area: The area of each region is found out and the threshold is set and all the areas which are above the threshold value are accepted and others are rejected.

- Based on aspect ratio: Some regions have lesser width and greater height or vice versa, which are other than the face.

- Based on region properties: The region properties such as mean, correlation coefficient, kurtosis and moments are computed and clustering is employed to cluster face regions from non-face regions.

Upon employing the above rules, we could be able to eliminate the non-face regions and the final image with face regions identified are shown in Fig. 7.

Fig.7. Image containing only face regions with red colored bounded box.

4. EXPERIMENTAL RESULTS

In this section, we present the experimental results conducted on the dataset which is collected on our own with varied background. We have conducted experimentation using MATLAB tool on Windows platform with 4 GB RAM. Experimental results are conducted on the database which contains nearly four hundred images taken in and around our campus. Some of the sample results obtained due to proposed approach are shown in Table-1.

5. CONCLUSION

In this paper, we have proposed a methodology to detect face regions from complex background based on morphology and skin color information. In the initial stage, segmentation of probable face regions are identified based on skin color information followed by morphological mask processing to identify almost true face regions. Some simple heuristic rules have been introduced to eliminate non-face regions from probable face regions. Experimental results on our database with varied background are conducted to exhibit the performance of the proposed method.

Table-1. Experimental results on images under varied background.

REFERENCES

[1]   Ahonen, T., A. Hadid, and M. Pietik¨ainen. Face recognition with local binary patterns. In Proc. of ECCV, 2004.

[2]   Dalal, N and B. Triggs. Histogram of oriented gradients for human detection. In Proc. of CVPR, 2005.

[3]   Enzweiler, M. and D. M. Gavrila. Monocular pedestrian detection: Survey and experiments. IEEE Trans. on PAMI, 31(12):2179–2195, 2009.

[4]   Hjelmas, E. and B. K. Low. Face detection: A survey. Computer Vision and Image Understanding, 83:236–274, 2001.

[5]   Jin, H., Q. Liu, H. Lu, and X. Tong. Face detection using improved lbp under bayesian framework. In Third Intl. Conf. on Image and Grahics (ICIG), 2004.

[6]   Jones, M and P. Viola. Fast multi-view face detection. Technical report, Mitsubishi Electric Research Laboratories, TR2003-96, 2003.

[7]   Laptev. Improvements of object detection using boosted histograms. In British Machine Vision Conference, 2006.

[8]   Levi, K and Y. Weiss. Learning object detection from a small number of examples: The importance of good features. In Proc. of CVPR, 2004.

[9]   Li, Y., S. Gong, and H. Liddell. Support vector regression and classification based multi-view face detection and recognition. In International Conference on Automatic Face and Gesture Recognition, 2000.

[10] Mita, T., T. Kaneko, and O. Hori. Joint Haar-like features for face detection. In Proc. of ICCV, 2005.

[11] Ojala, T., M. Pietik¨ainen, and T. M¨aenp¨a¨a. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. on PAMI, 24:971–987, 2002.

[12] Pham, M. -T. and T.-J. Cham. Fast training and selection of haar features during statistics in boosting-based face detection. In Proc. of ICCV, 2007.

[13] Roth, D., M.-H. Yang, and N. Ahuja. A SNoW-based face detector. In Proc. of NIPS, 2000.

[14] Rowley, H.A., S. Baluja, and T. Kanade. Neural network based face detection. In Proc. of CVPR, 1996.

[15] Rowley, H.A., S. Baluja, and T. Kanade. Rotation invariant neural network-based face detection. Technical report, School of Coomputer Science, Carnegie Mellow Univ., CMU-CS-97-201, 1997.

[16] Suard, F., A. Rakotomamonjy, A. Bensrhair, and A. Broggi. Pedestrian detection using infrared images and histograms of oriented gradients. In IEEE Intelligent Vehicles Symposium, 2006.

[17] Viola, P and M. Jones. Rapid object detection using a boosted cascade of simple features. In Proc. of CVPR, 2001.

[18] Wu, J., J. M. Rehg, and M. D. Mullin. Learning a rare event detection cascade by direct feature selection. In Proc. of NIPS, volume 16, 2004.

[19] Yan, J., S. Li, S. Zhu, and H. Zhang. Ensemble svm regression based multi-view face detection system. Technical report, Microsoft Research, MSR-TR-2001-09, 2001.

[20] Yan, S., S. Shan, X. Chen, and W. Gao. Locally assembled binary (LAB) feature with feature-centric cascade for fast and accurate face detection. In Proc. of CVPR, 2008.

[21] Yang, M, -H., D. J. Kriegman, and N. Ahuja. Detecting faces in images: A survey. IEEE Trans. on PAMI, 24(1):34–58, 2002.

[22] Zhang, G., X. Huang, S. Z. Li, Y. Wang, and X. Wu. Boosting local binary pattern (LBP)-based face recognition. In Proc. Advances in Biometric Person Authentication, 2004.

[23] Zhang, L., R. Chu, S. Xiang, S. Liao, and S. Z. Li. Face detection based on multi-block LBP representation. 2007.

[24] Zhu, Q., S. Avidan, M.-C. Yeh, and K.-T. Cheng. Fast human detection using a cascade of histograms of oriented gradients. In Proc. of CVPR, 2006.