Journal Of Harmonized Research (JOHR)

Journal Of Harmonized Research in Engineering
2(1), 2014, 147-152

ISSN 2347 – 7393

Original Research Article

FACE DETECTION USING COST SENSITIVE ADABOOST ALGORITHM AND SKIN COLORSEGMENTATION. (WITH PROBABILISTIC MULTIPOSE FACE DETECTION)

Atul Dhatrak*, Anand Karandikar, SomnathThorave, Ajit Zagade

Dept. of Comp Engg.,SCOE,Pune.

Abstract:
One can say, face detection is a mature topic in image processing as there has been a lot of work done in this area. But yet there remain some challenges that need to be addressed to use face detection for real world scenario like video surveillance. We proposed a new approach to face detection which addresses most of the challenges by using a hybrid approach which consist of combination of cost sensitive Adaboost algorithm and skin color segmentation. By applying skin color segmentation in the initial stage we eliminate the unnecessary background area and applying the Haarfeature extraction on skin color segmented area reduces the time requirement drastically. With the use of extensive training set and cost sensitive Adaboost algorithm, accuracy of the face detection is increased.

Keywords- Cost SensitiveAdaBoost algorithm, multipose, Haar feature extraction.

Introduction

Face detection is a very active research topic in thefield of computer vision and pattern recognition, which iswidely applied in the identity authentication; man-machineinterface; visual communication; virtual reality; management of public security files; content-based retrieval and many other aspects.[3] Therefore, it requires that the facedetection system with strong adaptability to all environments. It's the key for the face detection to find an effectivemethod to extract the common features of human facewhich can describe face model, namely face modelling. Thecommon features of human face include geometry,symmetry,texture etc. As the complexity of theimageincreases, it can't obtain accurate test results with onlysingle feature information, then people pay more attentionto the detection method with multi-feature information.As the number of feature

For Correspondence:
Scoe.facedetectionATgmail.com
Received on: February 2014
Accepted after revision: February 2014
Downloaded from: www.johronline.com

for face detection increase its time requirement also increase.The other factors such as noise, light, color change, posture andfacial expressions, and many other factors makes face detection more complex.Most difficult aspect of face detection is to develop a system which can be used in real time with accuracy and timeliness. Thus we proposed this hybrid approach which addresses both the issues. This method focuses on improved AdaBoostalgorithm[1] thatreduces the misclassification error of traditional AdaBoostalgorithm.The improved AdaBoost algorithm is called as Cost sensitive AdaBoostalgorithm.It also uses skin color segmentation, and combines the two kinds of methods, therebymaking full use of the advantage of high speed of skin colordetection and high detection rate and low false acceptance rateof cost sensitiveAdaBoostalgorithm.One more enhancement proposed is the use of multipose face images in the training set as a special category besides faces and non faces images. This gives us the different set of classifiers than the traditional Adaboost algorithm will give.Thus improves the capability of algorithm to detect multipose faces from image [10].

Skin Color Segmentation

The process of differentiation between skin color pixel and non skincolor pixel can be defined as skin color segmentation. However, there are some difficulties in robust detection of the skin color [4][6]. The appearance of the skin tone can be affected by the ambient of the light and shadows. Also there are various skin color tones such as Asians skin thathas big difference with Caucasians skin type. Thus to accurately identify the skin color pixel choice of color space is a crucial part.[8]RGBcolorspace is not suitable as itnon-uniformity, mixing of chrominance and luminance data and high correlation between them.Thus we use YCbCrcolor space which has the luminance and chromaticity information. The effect of irregular illumination in an image is reduced in YCbCrcolor space due to separation of brightness information from the chrominance and chromaticity.

The conversion from RGB color space to YCbCrcolor space can be done using the following equation.

$$Y=0.299R+0.587G+0.114B \quad (1)$$

$$Cb=B-Y. \quad (2)$$

$$Cr=R-Y. \quad (3)$$

There are different methods for skin colorsegmentation. In our approach, we identify the skin color using mean and covariance of chrominancecolor in a set of 100 skin sample images. The mean vector and covariance matrix is obtained from the total data of 100 skin samples by appending them together.This gives us a domain of skin color pixel which can be used to identify skin color pixel.

(a) (b) (c)

Fig. 1. (a) Original image (b) Image in YCbCr space(c) image using parametric single Gaussian model.

Cost Sensitive AdaBoost Algorithm

We use cost-sensitive AdaBoost algorithmthat is based on the AdaBoost algorithm, followed by short CS-Boost algorithm, the specific process is asfollows:[5][7]

1) Given the training images (x_1,y_1) ...,(x_n,y_n), x is the image of the sample, y = {-1,1}. Respectivelynon-face and face samples, WeakLearnindicates Weak learning algorithm, T is the number of iterations, Constant cost C> 1 ;[7]

2) The definition of the cost of misclassified positive and negative samples: $C_+=\sqrt{C}$, $C=1/\sqrt{C}$.

3) Initialized the weight of positive and negativesamples:

$$w_{i,j}=\begin{cases} C_+ / (C_+ N_+ + C_- N_-) & if\ y_i = +1 \\ C_- / (C_+ N_+ + C_- N_-) & if\ y_i = -1 \end{cases} \quad (4)$$

Where, N-plus or minus the number of samples;

4) Let t=1…..T

a) Normalized weight

$W_{i,j}= W_{i,j} / (\sum W_{i,j})$ for j=1 upto n.

b) Used weight Wt and training samples, called weakclassifier learning algorithm WeakLeam,get a weakclassifier ht .

c) Calculated Error rate of ht:

$$\varepsilon_t=\sum_{i=1}^{n} P_i^t | h_t(x_i)-y_i| \text{ , let } \alpha=\log (1-\varepsilon_t)/ \varepsilon_t \quad (5)$$

d) Let

$$w_{i+1,j} = \begin{cases} w_{i,j}\exp(-\alpha_t / C_\delta) & \text{if } y_i h_t(x_i)=+1 \\ w_{i,j}\exp(C_\delta \alpha_t) & \text{if } y_i h_t(x_i)=-1 \end{cases}$$

$$\text{where } C_\delta = \begin{cases} C_+ & \text{if } y_i=+1 \\ C_- & \text{if } y_i=-1 \end{cases} \quad (6)$$

Output the final strong classifier:

$$H(X)=\text{sign}\left[\sum_{t=1}^{T} \alpha_t h_t(x) \right] \quad (7)$$

Cost sensitive AdaBoost algorithm takes into account the cost ofdifferent types of classification errors, and tries to minimize thecost of wrong classification. As compared toAdaBoostalgorithm, Cost sensitive AdaBoost algorithm taken into account the cost ofthemisclassification at two levels.First, initialize theweight of the sample according to misclassification costs.Another, consider cost impact while the weight updating. The learning process will be more focused on the positivesamples with wrong classification. This could get a higherdetection rate and improve detection speed.

Cost Sensitive AdaBoostBased Face Detection Algorithm

This algorithm is based on the basic framework proposed by Viola and Jones.

A. Haar feature space

To give a face classifier speed as quickly as possible, complex features should not be used. Viola and Jonespropose a class of simple rectangular features.[1][2] The flexibilityof these features is far less thanotheradjustable classifier. It's more sensitive for edges and other simple image structure, and relatively rough, only the horizontal andvertical directions are available. However, these featuresprovide a rectangular image of the very rich characterization, which supports effective learning. Also in the calculationof high efficiency to make up their lack offlexibility.Characteristic value of the rectangular graph is black part of the black pixel values subtracting the white part of the black pixel value, one boxindicatethe rectangular area of thedetection. In such an area, these four characteristics of thesize and position can be arbitrarily chosen, assuming that therectangular area the size of 24 x 24, then all of therectangular characteristics of the data will be very large,more than 70,000.In order to perform fast computation of rectangularfeatures, the algorithm uses agraph called the integral image representation. Position (x, y) points on the map with thislocation and at the top of the left and all the pixels, expressedas

$$(x,y)=\sum_{x'<x,y'<y} i(x',y') \quad (8)$$

B. Feature selection and classifier construction

In the AdaBoost learning algorithm, each training samplehas an initial weight,representative importance of currentsample. According to each candidate weak classifier on thetraining set the classification results and the weight of thesample, you can calculate the weighted classification errorrate of the current weak classifier in the training set. Eachone selected a minimum misclassification rate of weakclassifier; the minimum classification error rate ofthevalueof cut-off point is the weak classifier threshold. Thenaccording to this weak classifier obtained the results of theclassification to update the weights of training samples, reducing the weight of the sample correctly classified,increasing incorrectly classified sample weight. The purposeis to make learning algorithm to concentrate on the sample of incorrectly classified. A detailed training process is as follows:

1) Given the training images (x_1,y_1) ...,(x_n,y_n), x is the image of the sample, y = {-1,1}. Respectively non-face and face samples.

2) Initialization face and non-face images weight $w_{i,j}$=1/2m , 1/2n ,where i = 0 , 1 ; m and n respectively fornon-face and face thenumber of samples.

3)Let t = 1 , . . . , T

a)Normalized weight,

$$\text{Wi. j= Wi. j}/\sum_{j=1}^{n} Wi,j \quad (9)$$

b) For each feature f_j, trained a weak classifier h_j,calculated its weighted classification error

$$\varepsilon_j=\sum_j w_{i,j} \mid h_j(x_i)- y_i| \quad (10)$$

c) Select a weak classifier h_t of the smallest error rate ε_t, calculated

$$\beta= \varepsilon_t/(1- \varepsilon_t). \quad (11)$$

d) Update weights

$$W_{t+i,t} = W_{t,i}\beta_t^{1-e_i}, \quad (12)$$

whereε_t=1 misclassification,

ε_t= 0 correct classification

4) The final strong classifier:

$$H(x) = \text{sign}\left[\sum_{t=1}^{T} \alpha_t h_t(x)\right], \alpha_t=\log(1/\beta_t) \quad (13)$$

The cascade structure of classifier

Cascade structure could quickly filter out most ofinhuman face region using the fewestpossible terms, leavingthe region is morelikely to face sub-windows, for more complex calculations.

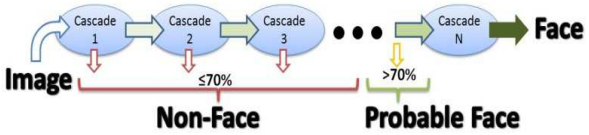

Fig.2. Cascade of classifier.

The training process of cascade structure is driven bysome detected targets, beginning set the detection targetexpected to be achieved, namely, single-layer classifier to achieve the detection rate and false positive rate, the overallfalse positive rate of classifier.[9] The specific process is asfollows:

1) User-defined values of F: the maximum acceptablesingle-layer classifier misjudgement rate, and d values:namely, single-layer classifier minimum acceptabledetection rate.

The user to select the classifier the final false positiverate Fwp;

P: Single-layer classifier for the training sample set of human face.

4) N: Single-layer classifier for the training sample set of non-face.

5)F_0=1.0, D_0=1.0, i=0;

6)If Fi>Fwp

a) 1. i=i+1, n_i=0 ; F_i= $F_{i=1}$;

b) $F_i > f * F_{i=1}$
n_i= n_i+l

Pairs of P and N, with the AdaBoost algorithm for training a classifier with n_i characteristic:
Using the current Multi-layer classifier to classify thevalidation sets, get the current multi-layer classifier Di andFi.Reduction threshold of the current first i-layer classifier, until the current multi-layer classifier detection rate atleast d * $D_{i=1}$ (This action will affect Fi of the current classifier)

c) N=Φ

d) If $F_i > F_{wp}$, Updated the non-face sets N with thecurrent classification misclassation of multi-layer classifiernon-face images.

Architecture of face detection system using cost sensitive AdaBoost algorithm and skin color segmentation.

The system consists of two phases:
a) Training phase, and
b) Detection phase.

In training phase we train cost sensitive AdaBoost algorithm to create cascade of classifier. This cascade of classifier will be used in detection phase. The following figure shows the flow of the system in detection phase.The block diagram of the system is shown below:

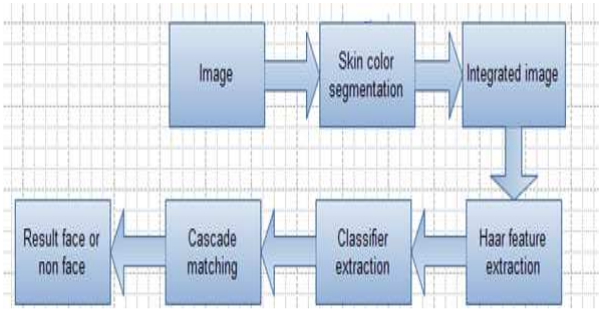

Fig. 3. Architecture diagram of face detection using Cost sensitive Adaboost algorithm and skin color segmentation.

Expected Result

This algorithm should be able to detect the faces from the images including the crucial multipose faces accurately with minimum false rate and as fast as possible.

Conclusion

We have proposed a hybrid approach for face detection which is a combination of cost sensitive Adaboost algorithm and Skin Color Segmentation. This approach aims to increase accuracy and to decrease time requirement of face detection system. It also has a new feature as probabilistic multipose face detection which

is an improvement over traditional face detection system particularly focusing on detecting faces in different poses.

Acknowledgement

This research paper cannot be considered complete without mentioning Prof.S.S.Pawar. We are also thankful to Prof. P. R. Futane,Prof. P.M. Kamde and Prof. C. A. Laulkar for their guidance and support. We wish to express true sense of gratitude toward all of them. We are grateful to them for their constant encouragement and guidance in the fulfilment of the project activity.

References

[1] Phillip Ian Wilson,Dr. John Fernandez," Facial Feature Detection Using Harr Classifiers", 2006 April.

[2] JinxinRuan ,Junxun Yin, "Multi-Pose Face Detection Using Facial Features and AdaBoost Algorithm" , 2009 Second International Workshop on Computer Science and Engineering.

[3] Liying Lang, WeiweiGu, "Study of Face Detection Algorithm for real time Face Detection System" , 2009 Second International Symposium on Electronic Commerse and Security.

[4] LI-Ying Lang, Wei-Wei Gu, "A Robustness and Real time Face Detection Algorithm in Complex Background", 2009 International Conference on Wavelet Analysis and Pattern Recognition, Baoding.

[5] Du Xingjing Zhu Dongmei, Zhao Hongyun,"Study of Fast Adaboost Face Detection Algorithm", 2010 International Conference on Computer Application and System Modeling (ICCASM 2010).

[6] Gang Li, YinpingXu，JiayingWang ,"An Improved AdaBoost Face Detection Algorithm Based on optimizing Skin Color Model", 2010 Sixth International Conference on Natural Computation (ICNC 2010).

[7] DingliSong, Bingru Yang, Fuxing Yu, "Study of The Cost-sensitive AdaBoost Face Detection Algorithm", 2010 International Conference on Computer Application and System Modeling (ICCASM 2010).

[8] Li-yingLang ,Wei-weiGu, "Study on Face Detection Algorithm based on Skin Color Segmentation and AdaBoost Algorithm", 2011.

[9] JiaMingxing, Du Junqiang, Cheng Tao, Yang Ning, Jiang Yi, Zhang Zhen, "An Improved Detection Algorithm of Face with Combining AdaBoost and SVM" ,2013 IEEE.

[10] AlirezaTofighi, S. AmirhassanMonadjemi, "Face Detection and Recognition using Skin Color and AdaBoost Algorithm Combined with Gabor Features and SVM Classifier" , 2011 International Conference on Multimedia and Signal Processing.