

FACE DETECTION USING SKIN LIKELIHOOD AND HAAR FEATURES FOR DIGITAL VIDEO PROCESSING

Vishakha V. Navlakhe^{1*}, Prof. Deepak Kagate²

¹M. Tech Student, Department of C.S.E., GHRAET Nagpur, Maharashtra, India

²Professor, Department of C.S.E., GHRAET Nagpur, Maharashtra, India

¹vishakha.navlakhe@gmail.com

²deepakkagate32@gmail.com

***Corresponding Author: -**

Email ID- vishakha.navlakhe@gmail.com

Abstract: -

Face detection is an important early step in many computer vision systems. By using pixel-wise detectors, spatial analysis of skin probability and skin regions segmentation, a new method for face detection is introduced. In this project, we proposed and implemented a modified self-organizing mixture network (SOMN) which specifies the distribution of objects in image and skin and non-skin color model, skin likely-hood to exactly identify skin region of interest from image. Bayesian Decision Rule is applied to specify c as skin color or non-skin color. Finally, we are using haar like features to identify face and cascade to improve performance and efficiency. We present results of an extensive experimental study which clearly indicate high competitiveness of the proposed method and its relevance to gesture recognition.

Keywords: - Skin-color modeling, Self-Organizing Mixture Network, Haar like features.



INTRODUCTION:

In any automated system for face recognition, face tracking, and facial expression recognition, face detection is necessary. There are large number of face detection systems. For any face detection system, detection rate that is the ratio between the number of faces correctly detected by the system and the actual number of faces in the image is the important factor. For detecting and locating faces in color images, much work has been done. For this the methods like Chrominance based [1], Skin color based [2,3], Adaboost based [4], segmentation based [3], neural network-based [5], have been well studied by many researchers. There are many face detection algorithms, but the method based on skin color model, Bayesian decision and haar like rectangular features has been widely used for face detection and for its simple performance, convenient use, and high detection speed. In this context, face detection could be viewed as face segmentation problem whereby the face region is separated from background. Thus, by detecting image pixels within a range of "skin-like" colors, it is feasible to isolate face region by some thresholding stage. This method works well under different lighting conditions. The Proposed method introduces a self-organizing mixture network (SOMN) to develop accurate and robust models for image data and use the Gaussian mixture model then we applied the Bayesian decision rule to classify the pixels in the testing set and compare it with ground-true classifications. Bayesian decision rule is used to classify the image pixels according to the obtained models [6]. It helps to detect skin from different environmental variations. To design a skin pixel classifier, this project uses skin and non-skin color models. For the unconstrained nature of Web images this gives surprisingly good performance. We construct a system for detecting images containing naked people using the proposed skin classifier, which operates on the color of a single pixel. This detector is much fast since it is based on pixel-wise classification. These experiments suggest that skin color can be a more powerful cue for detecting people in unconstrained imagery than was previously suspected. We explore this point by comparing histogram and Gaussian mixture models learned from our dataset. This shows that histogram models slightly outperform mixture densities in this context. The method mentioned in this project for face detection can achieve high detection accuracy, high detection speed and reduce the false detecting rate and the missing rate.

Literature survey:

Meng Yang, Lei Zhang et. Proposed Regularized Robust Coding for Face Recognition [11] based on sparse representation-based classification (SRC). In this method the testing image is coded as a sparse linear combination of the training samples, and the representation fidelity is measured by the l_2 -norm or l_1 -norm of the coding residual. But the limitation with this algorithm is linear combination of training samples. Zuhadi Zakaria et. al. Proposed Face Detection Using Combination of Neural Network and adaboost[4]. In this paper, two well-known algorithms that is adaboost and neural network are combined. A cascade adaboost classifier is used to increase the face detection speed, neural network is used as final classifier to verify face or non-face. But the drawback with this algorithm is the complexity for combining two algorithms. Can-hui Cai et. Al. Proposed Real-Time Face Detection Using Gentle AdaBoost Algorithm and Nesting Cascade Structure [19]. In this gentle adaboost algorithm is used to train node classifiers on Haar like features set to improve generalization ability of node classifier. Nesting Cascade Structure is introduced to avoid that too many weak classifiers in a Cascade classifier will slow down the face detection speed of this Cascade classifier. Face Detection Using Fuzzy Granulation and Genetic algorithm In Color Images uses fuzzy logic and genetic factor to detect face. But the limitation of this method is that always use the combination of fuzzy logic and genetic algorithm will slow down the performance of face detection in complex background and this algorithm is less fault tolerant.

Proposed methodology:**Modified SOMN:**

To estimate whether a given object of image fall under a specified range or not, a modified Self-organizing Mixture Network (SOMN) is used. SOMN is also based on self-organizing principle which improve stability, minimizing the Kullback-Leibler divergence [12] which improve the applicability and maximum likelihood approach which improve the computation performance of skin detection for face detection.

Let the number of components contained in an image are K , x is a sample from a d dimensional input space Ω which belongs to \mathbb{R}^d and prior probability or mixing parameter is P_i then the joint probability density of data sample $p(x|\theta)$ is given by:

$$p(x|\Theta) = \sum_{i=1}^K p_i(x|\theta_i) P_i \quad (1)$$

$p_i(x|\theta_i)$ is The i -th component conditional density which shows that K contain sample x . For $p_i(x|\theta_i)$, θ_i is the parameter vector. $\theta = (\theta_1, \theta_2, \dots, \theta_K)^T$. $p_i(x|\theta_i)$ has the following form for the Gaussian mixture:

$$p_i(x|\theta_i) = \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{1/2}} \exp \left[-\frac{1}{2} (x - m_i)^T \Sigma_i^{-1} (x - m_i) \right] \quad (2)$$

Where $\theta_i = \{ \Sigma_i, m_i \}$ are the covariance matrix and mean vector respectively.

To find hidden structure of image in unlabeled data, the SOMN provides a feasible solution to unsupervised learning problem. This is done by minimizing the Kullback-Leibler divergence. Let $p(x)$ is the true environmental data density function and $\hat{p}(x)$ is the estimated one then the KLD is defined as:

$$KLD = - \int \ln \frac{\hat{p}(x)}{p(x)} p(x) dx \tag{3}$$

It measures the divergence between $p(x)$ and $\hat{p}(x)$.

Using Robbins-Monro stochastic approximation method [7], minimize the KLD results in the following adaptive updating equation [8]:

$$\hat{\theta}_i(n+1) = \hat{\theta}_i(n) + \alpha(n) \left[\frac{\hat{P}_i(n) \partial \hat{p}_i(x|\hat{\theta}_i)}{\hat{p}(x|\hat{\theta}) \partial \hat{\theta}_i(n)} \right] \tag{4}$$

$$\hat{P}_i(n+1) = \hat{P}_i(n) - \alpha(n) [\hat{P}(i|x) - \hat{P}_i(n)] \tag{5}$$

Where $\alpha(n)$ is the learning rate at time steps n , $0 < \alpha(n) < 1$ and decreases monotonically and

$$\hat{P}(i|x) = \frac{\hat{P}_i \hat{p}(x|\hat{\theta}_i)}{\hat{p}(x|\hat{\theta})} \tag{6}$$

Is the estimated posterior probability of the i -th component. The learning rules for covariance matrix and mean vector for a Gaussian mixture are:

$$\begin{aligned} \Delta \hat{m}_i &= \alpha(n) \hat{P}(i|x) [x - \hat{m}_i(n)] \\ \Delta \hat{\Sigma}_i &= \alpha(n) \hat{P}(i|x) [(x - \hat{m}_i(n))(x - \hat{m}_i(n))^T - \hat{\Sigma}_i(n)] \end{aligned} \tag{7} \& \tag{8}$$

The SOMN can use inhomogeneous mixtures to derive a more general algorithm to improve the applicability of the algorithm and accuracy of density matching.

To derive a new iteration formula for \hat{P}_i we maximize the log-likelihood of the observed samples and Use the obtained conditional densities.

Suppose $X = \{x_1, x_2, \dots, x_N\}$ is a set of N independent observations, then its log-likelihood is:

$$\begin{aligned} \ell &= \sum_{k=1}^N \ln \hat{p}(x_k|\hat{\theta}) = \sum_{k=1}^N \ln \sum_{i=1}^K \hat{p}_i(x_k|\hat{\theta}_i) \hat{P}_i \\ \ell' &= \ell + \lambda (\sum_{i=1}^K \hat{P}_i - 1) \end{aligned} \tag{9}$$

To ensure $\sum_{i=1}^K \hat{P}_i = 1$, the method of Lagrange multiplier with a constraint parameter λ is used. Calculate the partial derivative of ℓ' with respect to \hat{P}_i and set it equal to zero, after some manipulations we obtain [9]:

$$\hat{P}_i = \frac{1}{N} \sum_{k=1}^N \hat{P}(i|x_k, \hat{\theta}_i) \tag{10}$$

Equation 4 & 10 can work well together because for the true environmental data both formulas aims to achieve the estimated density function approximately that is the accurate distribution of objects in image.

The simplification of \sum_i and \hat{P}_i improves the performance of the algorithm and make it more robust [11].

Skin and Non-Skin Color Model And Bayesian Decision Rule:

After applying SOMN which specifies the distribution of objects in image then we have to apply skin and non skin color model to exactly identify skin region of interest from image.

The steps of this modeling algorithm are as follows:

- a. The mixing parameters are set to $1/K$ and the initial covariance matrices to large diagonal matrices. The initial mean vectors are set to small positive random vectors.
- b. From the training set, draw a sample x (region of interest from image) with a given probability.
- c. By using eq.(6), for each component compute the posterior probabilities and identify the skin region using the following formula:

$$i^*(x) = \arg \max_i \hat{P}(i|x),$$

- d. Using updating equations such as (7) and (8), adjust the θ_i .
- e. Repeat steps from b to d and then use the eq. (10) to update the mixing parameter to achieve better distribution in image.
- f. Repeat steps from b to e until no noticeable changes in image are observed.

For skin and non-skin color, we get the accurate probability distribution functions after training the input image. So we can be directly compute $P(c|skin)$ and $P(c|-skin)$ which denotes the probabilities that an observing color c belongs to skin and non-skin class. Applying the Bayes formula we get,

$$\frac{P(skin|c)}{P(-skin|c)} = \frac{P(c|skin)P(skin)}{P(c|-skin)P(-skin)}$$

When the above ratio is greater than a certain empirical threshold γ , namely,

$$\frac{P(c|skin)}{P(c|-skin)} > \gamma \frac{P(-skin)}{P(skin)}$$

We classify c as skin color; conversely, we classify c as non-skin color.

Haar Features:

After detection of skin color and non skin color, we use three kinds of Haar features[20]. These are two-rectangle feature, three-rectangle feature, a four-rectangle feature. The value of a two-rectangle feature is the difference between the sum of the pixels within two rectangular regions. The regions have the same size and shape and are horizontally or vertically adjacent (see Figure 1). A three-rectangle feature computes the sum within two outside rectangles subtracted from the sum in a center rectangle. Finally a four-rectangle feature computes the difference between diagonal pairs of rectangles.

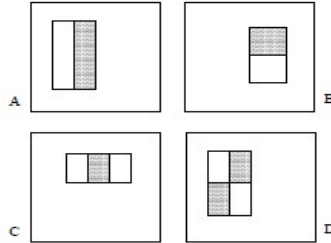


Figure 1: Example rectangle features shown relative to the enclosing detection window

The sum of the pixels which lie within the white rectangles are subtracted from the sum of pixels in the grey rectangles. Two-rectangle features are shown in (A) and (B). Figure (C) shows a three-rectangle feature, and (D) a four-rectangle feature.

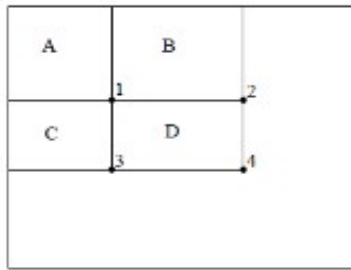


Figure 2

Figure 2 : The sum of the pixels within rectangle D can be computed with four array references. The value of the integral image at location 1 is the sum of the pixels in rectangle A . The value at location 2 is $A + B$ at location 3 is $A + C$ and at location 4 is $A + B + C + D$. The sum within D can be computed as $4 - 1 - (2 + 3)$. Using the integral image any rectangular sum can be computed in four array references (see Figure 2). Haar features are used for face detection.

In our system a variant of AdaBoost is used *both* to select a small set of features *and* train the classifier. In its original form, the AdaBoost learning algorithm is used to boost the classification performance of a simple (sometimes called weak) learning algorithm. A cascade of classifiers is constructed which achieves increased detection performance while radically reducing computation time. The key insight is that smaller, and therefore more efficient, boosted classifiers can be constructed which reject many of the negative sub-windows while detecting almost all positive instances (i.e. the threshold of a boosted classifier can be adjusted so that the false negative rate is close to zero). Simpler classifiers are used to reject the majority of subwindows before more complex classifiers are called upon to achieve low false positive rates. The overall form of the detection process is that of a degenerate decision tree, what we call a “cascade” (see figure 3).

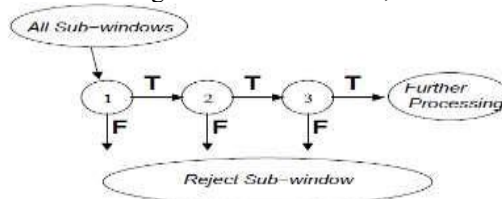


Figure 3: Schematic depiction of the detection cascade

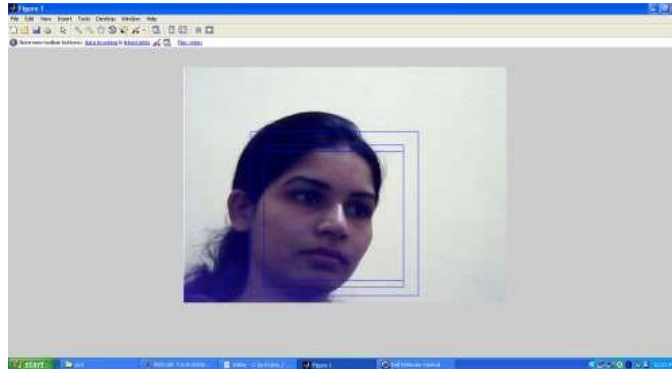
A series of classifiers are applied to every sub-window. The initial classifier eliminates a large number of negative examples with very little processing. Subsequent layers eliminate additional negatives but require additional computation. After several stages of processing the number of sub-windows have been reduced radically. The cascade structure results in fast average detection times. The speed of the cascaded detector is directly related to the number of features evaluated per scanned sub-window.

Implementation Details:

The working environment for face detection where the proposed algorithm is implemented is done using Matlab. Because in digital video face detection using skin, we need frame to select to tracking object, tracking window, tracking algorithm. These functions are easily available in Matlab.

Results Calculated:

The Proposed algorithm is implemented using Matlab. Initially a video file and the number of frames of the video are read. Then select the proper tracking algorithm and adjust the color qualification and threshold. Finally we get the tracking result that is face is detected.



Above video shows that tracking window track the face using skin likelihood and haar rectangular features.

Conclusion:

In this project, we describe a new method for face detection. The algorithm is implemented by modified SOMN to estimate whether a given object of image of video fall under a specified range or not then combine the skin on skin color model to exactly identify skin region of interest from image and then Haar like rectangular features are applied to exactly detect face. This algorithm gives good performance even with images taken from non-uniform background and different lightning and Intensity conditions. The method mentioned in this project can achieve high detection accuracy, high detection speed and reduce the false detecting rate, the missing rate by using cascade and adaboost structure. This algorithm has a lower computational cost because at each iteration, this algorithm needs only a small part of the training data. In the future work, we will improve this algorithm combined with other face detection algorithm to achieve better performance and further reduce the false detecting rate in dealing with videos with more complex background.

References

- [1].A.C. Rafael, D. Sidney,"Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications," *TAC*, 1(1), 18-34, 2010.
- [2].Kunqing Wu, Lianfen Huang, Hezhi Lin, Xiangping Kong, "Face Detection Based On YCbCr Gaussian Model And KL Transform," *International Symposium on Computer Science and Society*, 2011.
- [3].S. Kherchaoui et al," Face Detection Based On A Model Of The Skin Color With Constraints And Template Matching", *IEEE* 2010.
- [4].Zulhadi Zakaria, Shahrel A. Suandi," Face Detection Using Combination of Neural Network and adaboost," *IEEE* 2011.
- [5].Shahrum Shah Abdullah, Ahmad Nizam Jahari, Khairul Azha A. Aziz, Ridza Azri Ramlee,"Face Detection Using Radial Basis Function Neural Networks with Variance Spread Value," *International Conference of Soft Computing and Pattern Recognition*, 2009.
- [6]. "Skin detection using a modified Self-Organizing Mixture Network," *fg*, pp.1-6, 2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), 2013.
- [7].Marius Staring, Max A. Viergever," Adaptive Stochastic Gradient Descent Optimisation for Image Registration,"*International Journal of Computer Vision*, Volume 81, Issue 3, pp 227-239, March 2009.
- [8].H.Yin, N.Allinson,"Self-organising mixture networks for probability density estimation,"*IEEE Transactions On Neural Networks*. 12(2): 405-411, 2009.
- [9].Zhang Xuegong,"*Pattern Recognition*," Third Edition. Beijing: Tsinghua University Press, 186-191, 2010.
- [10]. J. Montenegro and W. Gómez, P. Sánchez-Orellana" A Comparative Study of Color Spaces in Skin-Based Face Segmentation,"*10th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE) Mexico City, Mexico. September 30, October 4, 2013*
- [11]. Meng Yang, Lei Zhang, Jian Yang, David Zhang, "Regularized Robust Coding for Face Recognition," *IEEE Transactions On Image Processing*, vol. 22, no. 5, May 2013.
- [12]. Wenchao Zhang, Shiguang Shan, Xilin Chen, Wen Gao," Local Gabor Binary Patterns Based on Kullback–Leibler Divergence for Partially Occluded Face Recognition," *Signal Processing Letters*, *IEEE*, Nov.2007.
- [13]. T. Kanade, J. Cohn, Y. Tian,"Comprehensive database for facial expression analysis," *CAFG*, 46–53, 2010.

- [14]. J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transaction Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.
- [15]. Ludmila I. Kuncheva, Member IEEE, "Change Detection in Streaming Multivariate Data Using Likelihood Detectors," *IEEE transactions on knowledge and data engineering*, vol. 25, No. 5, May 2013
- [16]. Gee-Sern Hsu, Tsu-Ying Chu, "A Framework for Face Detection Benchmark," *IEEE Transactions on Circuits and Systems for Video Technology*, 2013.
- [17]. J. Ruiz-del-Solar, R. Verschae, and M. Correa, "Recognition of faces in unconstrained environments: A Comparative study," *EURASIP J. Adv. Signal Process, Recent adv. Biometric Syst., A Signal Process. Perspect*, vol. 2009, pp. 1:1–1:19, Jan. 2009.
- [18]. X.Geng, D.Zhan, Z.Zhou, "Supervised Nonlinear Dim. Reduction for Visualization and Classification", *TSMCP*, 35(6), 1098–1107, 2013.
- [19]. Can-hui Cai et. Al., "Real-Time Face Detection Using Gentle AdaBoost Algorithm and Nesting Cascade Structure," 2012 IEEE International Symposium on Intelligent Signal Processing and Communication System (ISPACS 2012) November 4-7, 2012.
- [20]. Paul Viola et. Al., "Rapid Object Detection using a Boosted Cascade of SimpleFeatures," Accepted Conference On Computer Vision And Pattern Recognition.
- [21]. Mehrdad Shemshaki et al., "Face Detection Using Fuzzy Granulation and Genetic algorithm in color images," 5th International Conference on Automation, Robotics and Applications, Wellington, New Zealand, Dec 6-8, 2011,
- [22]. P. Moallem, B.S. Mousavi, S.A. Monadjemi, "A novel fuzzy rule base system for pose independent faces detection", *Applied Sof Computing*, Volume 11 Issue 2, March, 2011
- [23]. Kunqing Wu, Lianfen Huang et al. "face detection based on YCbCr Gaussian model and KL transform", *International Symposium on Computer Science and Society* 2011.