

<http://ansinet.com/itj>

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSI*net*

Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Domain Specific View for Face Perception

Mary Metilda and T. Santhanam

PG and Research Department of Computer Science, DG. Vaishnav College,
Arumbakkam, Chennai-600 106, Tamil Nadu, India

Abstract: The task of detecting faces is a precursor intended for shaping the information that the face provides. A robust way to locate the faces in images, insensitive to scale, pose, style, facial expression and lighting condition is contributed as a two phase process in this study. The first phase aspects a human face against a complex background by performing a skin color analysis of the image, as it is most liable that color noises have low probability on being flesh color. During the second phase, random measurements are generated to populate the entrant face region according to the face anthropometrics using the eye location determined from the YC_bC_r color space. The combination of holistic and configural approach implemented in this study has proven to be a very speedy, effortless and exceedingly competent scheme for locating complex stimuli such as face.

Key words: Skin color segmentation, color space, anthropometrics, connected component operator

INTRODUCTION

Face processing systems are not mere mechanisms to process faces but rather mechanisms for deriving fine-grained discriminations between visually similar exemplars. Face detection and tracking are the foremost steps in face perception, since the knowledge about spatial features of a face can be used very effectively for providing a wide variety of information such as personality, social class level, age, health, gender, etc. Most of the feature extraction algorithms purposeful on a universal account will incline to produce false targets, as they are ineffectual to resolve the face locality and for this reason they are impracticable.

Faces represent complex, multidimensional, meaningful visual stimuli. Faces differ from other objects in several vital facets. By and large it is enough to identify common objects at the basic level (e.g., Chair, Car, Pen), without determining the specific exemplars of the category (my chair, joe's car, edi's pen), but for faces it is necessary to proceed beyond the general category face for determining the identity of the said individual and non-homogeneity further compounds the problem. Besides faces share the same basic configuration unlike object recognition. Relatively high discriminability between numerous faces is achieved only by small displacement of face features so that extremely reliable discrimination occurs within the class of face objects. Substantially, a domain general account for face perception is not possible.

In addition, behavioral studies also suggest that domain specific mechanism helps to reduce the complexity of the surface whose properties are decisively dependent for the accuracy of curvature estimation to accomplish a precise throughput. Nevertheless, biological studies show that regions in the fusiform gyrus is activated not only when subjects view faces, but also atleast twice as strongly for faces as for a wide variety of non face stimuli, including letters (Mario, 2005).

From these grounds, mechanisms to recognize individuals are face-specific mechanisms making fine-grained discriminations between exemplars and their accuracy at discriminating is higher, only when the entire face is presented than including the other subjects. Other findings put forward that face recognition involves special mechanisms that are more inversion-sensitive and more holistic (less part based) than those involved in object recognition.

The human face is a dynamic object making face detection a intricate apprehension in computer vision, for which ample variety of techniques have been recommended in the literature ranging from simple edge-based algorithm to composite high-level approaches utilizing advanced pattern recognition methods.

Current techniques employed in the task of face detection are classified into two distinct groups' namely feature-based approaches and image-based approaches. Feature based approaches relies on the application of face knowledge and employs known properties of facial features. Representative techniques of feature-based

approaches include color-analysis (Cahi and Nga, 1999; Michael and James, 1998), feature searching (Leung *et al.*, 1995; Liu *et al.*, 2002) and template matching (Craw *et al.*, 1992; Lamitis *et al.*, 1995). In contrast, the image-based approaches exploit pattern recognition theories and make use of training algorithms to implicitly incorporate face knowledge into the system. Typical techniques using this approach include linear subspace method, such as Eigen faces (Turk and Pentland, 1991), neural networks (Agui *et al.*, 1992; Henry *et al.*, 1998) and various statistical methods (Erik and Boon, 2001; Colmenarez and Huang, 1997). The intended face detection algorithm presented in this study uses a permutation of feature-based approaches to maximize the probability of detection.

SKIN REGION SEGMENTATION

The inspiration to use skin color analysis for initial classification of an image into probable face and non-face region stems from a number of simple but powerful characteristics of skin color.

- Color of human skin is different from the color of most other natural objects in the world.
- Under certain conditions, skin color is orientation invariant.
- It is less obscure to process skin in contrast to other facial features.

Skin color detection can be performed via computationally intensive iterative methods such as skin patch merging (Garcia and Tziritas, 1999), dynamic link architecture (Laddes *et al.*, 1993), elastic graph matching (Wiskott *et al.*, 1999) etc. However, real time applications heavily rely on the fast pixel based approaches rather than the above optimization methods to meet stringent time requirements. The proposed pixel based approach for skin tone detection is a modified and extended form of Fleck and Forsy algorithm (Mario, 2005) that performs reasonably well both in terms of skin color detection and speed.

Color space: One imperative factor that should be decided in skin color detection is the color space to operate on, in order to effectively segment color of different shades under a variety of lighting conditions, since color information can significantly simplify the task of face localization in images with complex background. The images are conventionally represented in the RGB domain for display, where the color and intensity components are correlated. This requires some preprocessing in order to

reduce the effects of unwanted variations in intensity (Michael and James, 1998). The color information expressed in the RGB color space has not provided any improvement in the detection rate. Using R, G or B independently provides the same results as using R, G and B simultaneously and exactly the same results as using the luminance (Y) only. Also the RGB components are subjected to the lighting conditions and thus face detection may fail if there are frequent changes in the lighting conditions.

Besides, a color space with a separation of the luminance and chrominance information's tend to provide better results than a color space with these information's mixed (as RGB). Consequently, the original RGB image, which incorporates luminance information, must be transformed into a color space that separates the intensity and color information of the image respectively. Various studies have shown that the skin color regions are highly correlated in their hue and saturation channels or chrominance channels (Cahi and Nga, 1999). Hence the input color image typically in the RGB format is usually converted into color components such as HSV, $YCbCr$, YIQ formats in the color space. The information of different shades is usually present only in the intensity channel (Y or V) and it captures variations in lighting conditions. Therefore C_bC_r or HS channels are used in skin color segmentation.

Skin segmentation: In this approach, the RGB image is transformed to log-opponent (IR_gB_y) values from which the texture amplitude, hue and saturation are computed (Jay, 1997). The conversion from RGB to log opponents is calculated as:

$$\begin{aligned} I &= [L(R) + L(B) + L(G)]/3 \\ R &= L(R) - L(G) \\ B &= L(B) - [L(G) + L(R)]/2 \end{aligned}$$

The $L(x)$ operator is defined as $L(X) = 105 * \log_{10}(X+1)$. A texture amplitude map is used to find regions of low texture information. Skin in images tends to have very smooth texture and so one of the constraints on detecting skin regions is to select only those regions with little texture. Hue and saturation are approximated to select those regions whose color matches that of skin and the conversions from log opponent are given as:

$$\begin{aligned} \text{Hue} &= \text{atan}^2(R_g, B_y) \\ \text{Saturation} &= \text{sqrt}(R_g^2 + B_y^2) \end{aligned}$$

Employing constraints on texture amplitude, hue and saturation marks skin regions. If a pixel falls in a particular



Fig. 1: (a) Input image (b) skin marked image

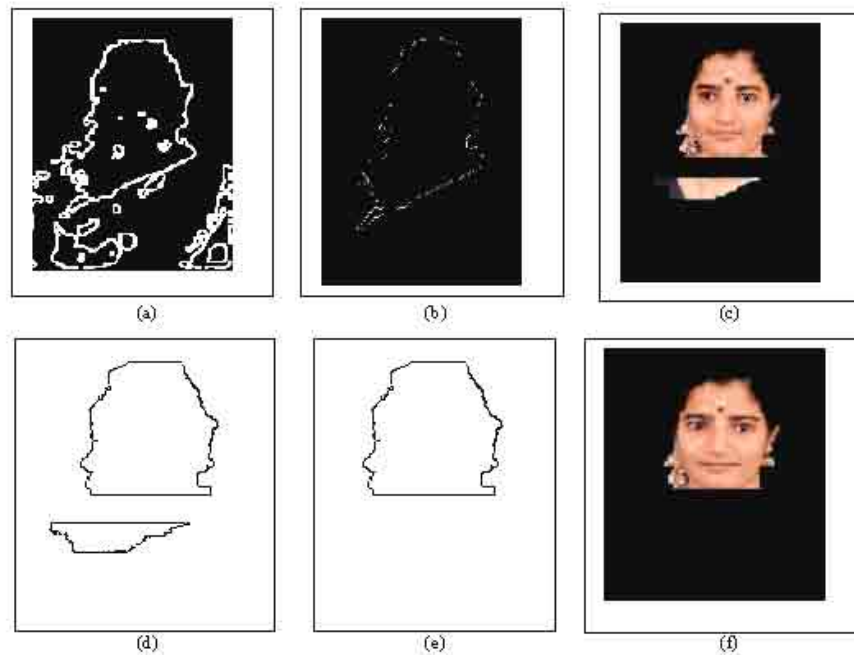


Fig. 2: (a) Skin connected component, (b) face skin connected component using compactness, (d) face skin connected component using orientation, (e) face skin connected component using solidity and (f) resultant face component

range it is marked as being skin in a binary skin array, where 1 corresponds to the coordinates being a skin pixel in the original image and 0 corresponds to a non-skin pixel. The binary skin map region is then processed using morphological operations to enlarge the skin map region to include skin and background border pixels (i.e.) regions near hair or other features (Fig. 1).

Face skin connected component operator: The skin filter described earlier tends to produce false skin regions, since there is a tendency for highly saturated reds and yellows to be detected as skin and this in turn calls for

the use of face skin connected component operators to eliminate those false regions (Fig. 2). Connected component operators are non-linear filters that eliminate parts of the image, while preserving the contour of the remaining parts. This simplification property makes them attractive for segmentation and pattern recognition applications.

Connected operators use certain decision criteria's to either retain or eradicate a connected component without affecting the other components. Three shape-based connected operators, Compactness, Orientation and Solidity are applied over the skin components to decide

whether they represent a face or not. The simple but effective decision criteria or operators make use of basic assumptions about the shape of the face to remove unwanted non-face components and are computed using the perimeter P and the size D_x and D_y of the min-max box of the connected components as configured below.

- Compactness of a connected component is defined as the perimeter of the region.

$$\text{Compactness} = P$$

The criterion is maximized in case of face objects for which a suitable threshold is fixed based on the observations of various face components. Significantly, components exhibiting high value for this operator are selected as face component.

- Orientation is but the aspect ratio of the min-max box surrounding the component

$$\text{Orientation} = \frac{D_y}{D_x}$$

Based on the observation of a number of images it is assumed that normal face components have orientation within a specific range. Using this criterion the skin region containing the face is cropped accordingly reducing the size of the candidate face region.

- Solidity of a connected component is defined as the ratio of its perimeter to the perimeter of the min-max box.

$$\text{Solidity} = \frac{P}{D_x + D_y}$$

It gives a measure of the perimeter occupied by the connected components with respect to its min-max box dimensions. The solidity operator also assumes a high value for face components. If a solidity of a component is lesser than a specified threshold value, it is eliminated otherwise retained.

Thereby the first phase deploying a skin filter together with the face skin connected component operators comprehensively isolates the potential face region from a complex background.

ENTRANT FACE SEGMENTATION

The skin filter may prove to be a generalization of the problem of face detection. To make the algorithm more robust, some additional constraints on the measurement

of the face with respect to its feature can be imposed to segment the probable face region required for face perception. Detection of the entrant face region is itself a 2-stage process which primarily extorts a feature point from the skin region, subsequently followed by exercising anthropometrical statistics of a geometric model to automatically isolate the implied face.

Eye map: Detecting the eyemap is based on the spatial arrangement of the Chrominance and Luminance components in the YC_bC_r color space (Reim-Lien *et al.*, 2002). Among the various facial features, eye and mouth are the most prominent features for segmentation in color space although the mouth map is indistinct especially in dark and light tone faces, since the red component is weaker in such cases. The Chrominance components contained in C_b and C_r and the luminance component contained in Y is employed to build two separate eyemaps (Fig. 3). These eye maps are then merged into a single map by an AND operation.

$$\text{EyeMapC} = 1/3 [C_b^2 + (\overline{C_r})^2 + (C_b/C_r)]$$

$$\text{EyeMapL} = \frac{Y(x,y) \oplus g_\sigma(x,y)}{Y(x,y) (-) g_\sigma(x,y) + 1}$$

$$\text{EyeMap} = (\text{EyeMapC}) \text{ AND } (\text{EyeMapL})$$

The resulting eyemap is then dilated, masked and normalized to brighten the eye region and suppress the other facial areas. Erosion removes small and thin isolated noise, while dilation preserves the required components. Normally the structuring element operational should not be more than that of the smallest eye which possibly will leave more connected components other than the eye. This again calls for the connected component operators to extort the contour of the eye from the isolated components while retaining the shape of the eye components (Fig. 4).

Together with the compactness and solidity operators that assume low values for eye component, a symmetric operator is also practical to retain the symmetric component of the eye. The components that fall out of symmetry are eliminated, thereby reducing the number of components and it is these retained symmetric components that enfold the eye.

Anthropometrics: In human-factor analysis, a known range for human measurements can assist in deriving the probabilistic face region for face perception. Anthropometry is the biological science of human body measurements in which the rich description of human geometry developed provides qualitative information

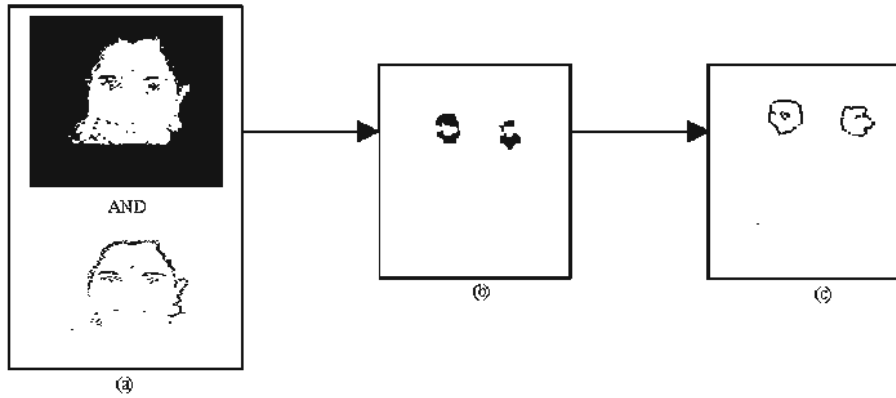


Fig. 3: (a) EyeMapC AND EyeMapL, (b) EyeMap and (c) Eye component

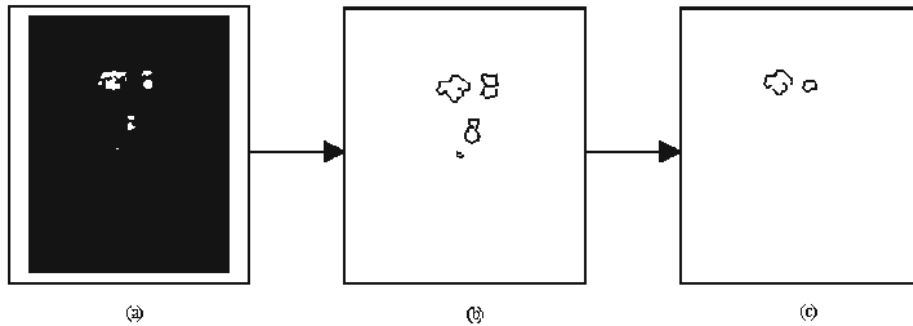


Fig. 4: (a) and (b) eye map components including noise and (c) segmented eye component

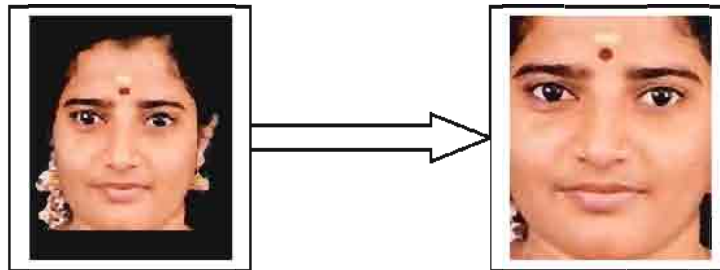


Fig. 5: Probable face region

about the shape and proposition of faces (Farkas, 1994; DeCarlo *et al.*, 1998). The theory states that all facial features are proportional to each other and the proportions of the face are as follows.

$$H_{\text{Face}} = 1.8 D_{\text{Eye}}; H_{\text{Eye}} = 1/3 H_{\text{Face}}; W_{\text{Eye}} = 0.2225 H_{\text{Face}}$$

The height of the face defined as a measure above the eyebrows to just below the face, can be determined from the dimension of the eye furnished during the

construction of the eye components. Using the above aspect ratios, the rationale behind localizing and extracting the characteristic face region from the derived skin region is accomplished (Fig. 5).

RESULTS AND DISCUSSION

The efficiency of the new-fangled face detection system discussed was analyzed using a photo gallery of 25 frontal color images, which was collected from known



Fig. 6: Extracts from the photo gallery

individuals taken from different cameras under different lighting conditions (Fig. 6). The selection of the color space and the skin segmentation operators has proven to identify the face component precisely except for an overestimation of the potential skin objects which habitually become apparent when people appear partially clothed. The resultant being a very large skin map containing the entire head arms or torso is abridged by the three skin connected component operators designed to crop the face component. This reduces a lot of computational overhead for the second phase of segmenting the characteristic face. The chrominance and luminance information easily de-embedded from the $YCbCr$ color space together with the component operators is utilized to retain the shape of the eye components while removing the other connected components. The random measurements from the eye region employed to smoothly populate the plausible face according to anthropometrical statistics extract the face region containing only the facial feature necessary for face perception.

The proposed method extracts features relative to the location of the eyes and hence it is invariant to size and tilt of the face. But in case of occlusions like spectacles the quality of the spectacles can also contribute to an

Table 1: Face detection analysis

Photo gallery	Successful detection	Over estimation	False positive	False negative
25	23	1	0	1

engorged face region, more than desired as in the last example in Fig. 6, nevertheless it adds to success. The human skin color modeling using IRgBy color mode however has serious problems when the algorithm loses track of the face, which occurs when the face does not occupy a significant area of the image resulting in a false negative. The intended system is efficient in terms of computational power and memory requirements and experimental results prove that the system gives an average deduction rate of 96% in real time under natural lighting conditions (Table 1).

CONCLUSION

The major intent of this study, to provide a rigorous cognitive process using holistic and configural perception has been achieved from experimental manipulation of faces. Face detection algorithm using holistic has the advantage of finding small faces, in addition to segmenting faces in poor quality images. Detection using

geometrical facial feature provide a good solution for detecting faces in different pose. A combination of holistic and feature based approaches have proven to be very promising from the above discussions.

Defining additional features like mouth to appreciate face dimension computations in anthropometrics can further enhance face segmentation. Besides the system can be extended to deduct multiple faces with variation in color, position, scale and orientation.

REFERENCES

- Agui, T., Y. Kokubo, H. Nagashi and T. Nagao, 1992. Extraction of face recognition from monochromatic photographs using neural networks. Proceedings 2nd International Conference Automation Robotics Computer Vision, 1, pp: 18.81-18.85.
- Cahi, D. and K.N. Nga, 1999. Face segmentation using skin color map. IEEE Trans. Circuit Syst. Video Technol., 9 (4): 551-564.
- Colmenarez, A.J. and T.S. Huang, 1997. Face detection with information-based maximum discrimination. Proceedings IEEE Conference Computer Vision Pattern Recognition, pp: 782-787.
- Craw, I., D. Tock and A. Bennett, 1992. Finding face features. Proceeding 2nd European Conference Computer Vision, pp: 92-96.
- DeCarlo, D., D. Metaxas and M. Stone, 1998. An anthropometric face model using variational techniques. Proceeding of SIGGRAPH, 1998.
- Erik, H. and K.L. Boon, 2001. Face Detection: A Survey. Comput. Vision Image Understanding, 83: 236-274.
- Farkas, L., 1994. Anthropometry of the Head and Face. Ravenpress.
- Garcia, C. and G. Tziritas, 1999. Face Detection Using Quantized Skin Color Regions Merging and Wavelet Packet Analysis. IEEE Trans. Multimedia, 1 (3): 264-277.
- Henry, A.R., B. Shumeet and K. Takeo, 1998. Neural network based face detection. IEEE Trans. PAMI, 20 (1): 23-38.
- Jay, P.K., 1997. Color Detection in Color Images. EE499 Capstone Design Project Spring.
- Laddes, M., J.C. Vorbruggen, J. Buchman, J. Lange, C.V. Malsburg, R.P. Wurtz and W. Konen, 1993. Distortion Invariant object recognition in the Dynamic Link Architecture. IEEE Trans. Comput., 42 (3): 300-311.
- Lanitis, A., C.J. Taylor and T.F. Cootes, 1995. An automatic face identification system using flexible appearance models. Image Vision Comput., 13 (5): 393-401.
- Leung, T.K., M.C. Burl and P. Perona, 1995. Finding faces in cluttered scenes using random labeled graph matching. Proc. 5th IEEE International Conference Comput. Vision, pp: 637-644.
- Liu, X., F. Xu and K. Fujimura, 2002. Real-time eye detection and tracking for driver observation under various light conditions. IEEE Intelligent Vehicle Symposium.
- Mario, F., 2005. Emerging holistic properties at face value. Assessing Characteristics of Face Perception.
- Michael, J.J. and M.R. James, 1998. Statistical color models with application to skin detection. Int. J. Comput. Vision.
- Rein-Lien, H., A.M. Mohamed and K.J. Anil, 2002. Face Detection in Color Images. IEEE Trans. on PAMI, 24 (5): .
- Turk, M. and A. Pentland, 1999. Eigenfaces for recognition. J. Cognitive Neurosci., 3 (1): 71-86.
- Wiskott, L., J. Fellous, N. Kruger and C.V. Malsburg, 1999. Face Recognition by Elastic Graph Bunch Matching. Springer-Verlag.