Journal of Engineering and Applied Sciences 12 (Special Issue 12): 9520-9527, 2017
ISSN: 1816-949X

Implementation of RGB Based Face Detection Using Threshold Values and Morphological Processing

[1]Bhavana Godvarthi, [1]Paparao Nalajala, [2]Kutikanti Chiranjeevi and [2]Ponna Madhuri
[1]Department of Electronics and Communication Engineering, Institute of Aeronautical Engineering, Dundigal, Hyderabad, India
[2]Department of Electronics and Communication Engineering, St. Peter's Engineering College, Hyderabad, India

Abstract: This present research study to detect face using threshold values and morphological processing has been used to try to replicate on a computer that which human beings are able to do effortlessly every moment of their lives, detect the presence or absence of faces in their field of vision. While it something that to a layman appears trivial to implement the necessary steps leading to the successful execution of this in an algorithm is difficult and still an unsolved problem in computer vision. In EE368, we have been given the task of using a collection of seven digital images to train and develop a system for doing just this in a competitive format. The only real limitation is that it runs under 7 min for a single file. By using this approach, we can reasonably show the fast algorithm.

Key words: Morphological processing, threshold values, EE368, digital images, field of vision, fast algorithm

INTRODUCTION

The system is a simple application of a color based segmentation scheme that takes advantages of patterns developed in the HSV, YCrCb and RGB color spaces followed in series by a matched filter/template matching system. Training data to unseen images and where prevented by the nature of the neural network to really have a grasp of the particular short comings of our system and what could be done to improve it. Assuming that a person framed in any random photograph is not an attendee at the renaissance fair or Mardi gras, it can be assumed that the face is not white, green, red or any unnatural color of that nature. While different ethnic groups have different levels of melanin and pigmentation, the range of colors that human facial skin takes on is clearly a subspace of the total color space. With the assumption of a typical photographic scenario, it would be clearly wise to take advantage of face-color correlations to limit our face search to areas of an input image that have at least the correct color components.

In pursuing this goal, we looked at three color spaces that have been reported to be useful in the literature, HSV and YCrCb spaces as well as the more commonly seen RGB space (Colmenarez and Huang, 1997). Below we will briefly describe what we found and how that knowledge was used in our system. The result of this study is the construction of hyper planes in the various color spaces that may be used to separate colors (Fig. 1).

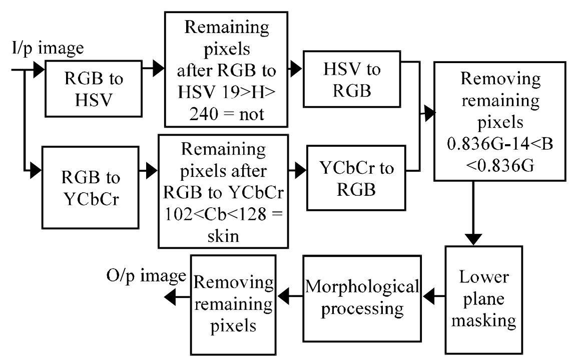

Fig. 1: Block diagram of face detection

MATERIALS AND METHODS

Implementation

RGB to HSV: While RGB may be the most commonly used basis for color descriptions, it has the negative aspect that each of the coordinates (red, green and blue) is subject to luminance effects from the lighting intensity

Corresponding Author: Bhavana Godvarthi, Department of Electronics and Communication Engineering, Institute of Aeronautical Engineering, Dundigal, Hyderabad, India

Fig. 2: RGB to HSV conversion: a) RGB and b) HSV

Fig. 3: a) RGB to YCbCr conversion and b) Remaining pixels after masking the lower image plane

of the environment an aspect which does not necessarily provide relevant information about whether a particular image "patch" is skin or not skin. The HSV (Jain, 1989; Lanitis *et al*., 1995) color space, however is much more intuitive and provides color information in a manner more in line how humans think of colors and how artists typically mix colors. "Hue" describes the basic pure color of the image, "saturation" gives the manner by which this pure color (hue) is diluted by white light and "Value" provides an achromatic notion of the intensity of the color. It is the first two, H and S that will provide us with useful discriminating information regarding skin. Using the reference images (truth images) provided by the teaching staff, we were able to plot the H, S and V values for face and non-face pixels and try to detect any useful trends. The results of this may be viewed in Fig. 2 from those results it is seen that the H values tend to occupy very narrow ranges towards both the bottom and top of its possible values. This is the most noticeable trend and was used by us to derive the following rule used in our face skin detection block (Moghaddam and Pentland, 1997; Treptow and Zell, 2004; Wu *et al*., 2004): ($19<H<240$) not skin and otherwise we assume that it is skin formulas:

$$H = \frac{1}{2}((R-G)+(R-B)/\surd(R-G)+(R-B) \tag{1}$$

$$S = 1-3(MIN\,(RGB)/R+G+B) \tag{2}$$

$$V = 1/3(R+G+B) \tag{3}$$

RGB to YCbCr: Similarly, we analyzed the YCbCr color space for any trends that we could take advantage of to remove areas that are likely to not be skin. Relevant plots may be viewed in Fig. 3. After experimenting with various thresholds we found that the best results were found by using the following rule ($102<Cb<128$) skin and otherwise assume that it is NOT skin and may be removed from further consideration (Moghaddam and Pentland, 1997; Treptow and Zell, 2004):

$$Y = 0.299*R+0.587*G+0.11*B \tag{4}$$

$$Cb = -0.168736*R-0.331264*G+0.5*B \tag{5}$$

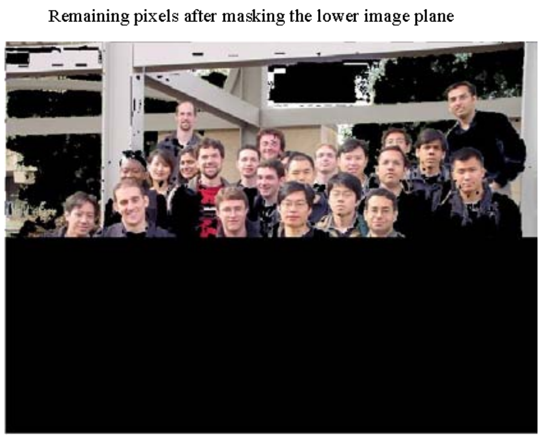

Fig. 4: Lower plane masking

$$Cr = 0.5*r-0.418688*G-0.081312*B \tag{6}$$

RGB color space: Let's not be too hard on our good friend the RGB color space, ..., she still has some useful things t o offer us to take advantage of this project. While RGB doesn't decouple the effects of luminance, a drawback that we noted earlier, it is still able to perhaps allow us to remove certain colors that are clearly out of the range of what normal skin color.

From studying and experimenting with various thresholds in RGB (Wu *et al*., 2004) space we found that the following rule worked well in removing some unnecessary pixels: 0:836G 14<B<0:836G+44) skin and 0:79G 67<B<0:78G+42) skin with other pixels being labeled as non-face and removed.

Lower plane masking: While in general it would destroy the generality of a detector in our case, we believe that its reasonable to take advantage of a priori knowledge of where faces are most likely to be and not be to remove "noise". We observed that in the training images that no faces ever appeared in the lower third of the image field. With very high probability it is likely that the scenarios where our system will be used (i.e., the testing images) that the same will be true, since, we know that the conditions in which the pictures were taken are identical. Hence, we removed the lower portion of the image from consideration to remove the possibility of false alarms originating from this region (Fig. 4).

Morphological processing

Applying the open operation: At this stage in the flow or our detector (Fig. 5), we have successfully removed the vast majority of the original pixels from consideration but we still see little specs throughout the masked image. Because we will subsequently send the image through a matched filter and the specs will be averaged out of consideration and hence could be left in and just ignored, it is preferable to remove them now in order to speed future processing (i.e., the matched filter needn't perform any wasteful calculations at these pixels). Hence, the open (erode dilate) operation was performed using a 3×3 window of all 1 sec. The result of applying this additional step is in Fig. 6. It is seen that the open operation has resulted in there being a huge reduction in the number of small "noisy" specs (Perez *et al*., 2010; Kotropoulos and Pitas, 1997).

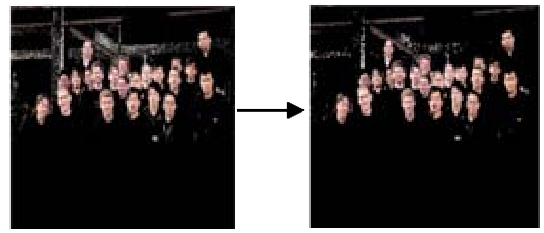

Fig. 5: Open operation

Blob removal: Removal of small blobs and grayscale transformation by "blobs", we simply mean the connected groups of pixels that remain at this stage. Here, we may apply a little additional knowledge about the way the picture was taken, ..., we know that the subjects in the photos were standing relatively closely to one another and hence, should have head sizes (measured by number of pixels) that are relatively similar.

The largest blobs should be these heads and blobs considerably smaller than the larger blobs may be safely assumed to be more "noise". In the particular sample image that we've been looking at, the sizes of the blobs from Fig. 6 were measured and ranked. The ranked sizes of the 195 remaining blobs are seen in Fig. 6. By removing blobs that are below a given threshold size we can remove even more additional noise. After experimenting with the given image studied in this report as well as the other provided images, we found that a pixel size of 200 was a good threshold value.

Hence, our blob size rule is: (blob size <200) non-face Blob and hence such blobs may be removed. Finally, we found that after this stage in our processing that all the color information that could be used within the level of sophistication feasible for this project had been and that subsequent stages could be done in grayscale without any performance degradation but with the additional benefit of a faster system that need only operate in one of

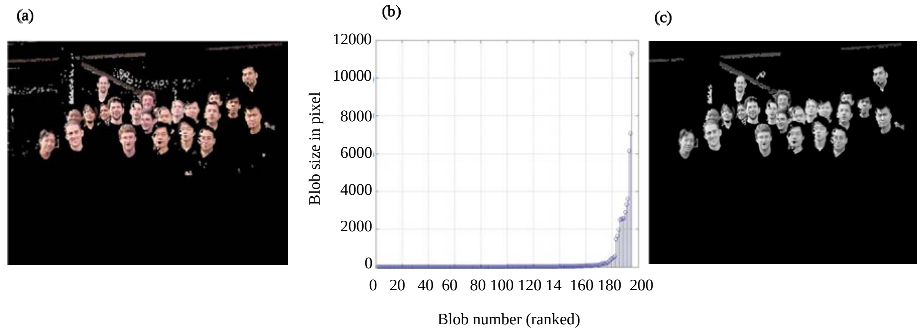

Fig. 6: a-c) Blob removal

the original three dimensions. Hence, we now transform our image to grayscale (Liu and Wechsler, 2000). This provides us with our final pre-processed image which may be seen in Fig. 6. It is important to note at this point one the main problem that we faced in this project. Note that the faces are retained at this stage in the processing but unfortunately, we have been unable to resolve them into separate blobs. Were the subjects standing with sufficient separation to do allow this we could do almost all of our necessary face detection just by working with blobs and their size statistics, etc. However, because the students in the photos are in very close clusters, multiple faces have been grouped in single blobs. This leads to complications that the template matching methodology (in our case at least) is unable to cleanly resolve in some situations.

Model the average size of head blobs in the training reference image. Remove blobs below one standard deviation.

Filtering algorithm:

1. We take pixel of n*m
2. Now we take the matched filter of k*p size
3. Now we divide the matched filter to equal no of averaging (i.e., each matrix into 1/9)
4. Now we take the center pixel value of our image and compare the values with the matched filter
5. Now we apply the same process for remaining surrounding pixels
6. Use filter (can also be called as mask/kernel/template or window)
7. The values in a filter sub image are referred to as coefficients, rather than pixel
8. Our focus will be on masks of odd sizes, e.g., 3×3, 5×5
9. Center pixel of such odd masks is well defined (Fig. 7)

Process of filtering

a	b	c
d	e	f
g	h	i

→

W1	W2	W3
W4	W5	W6
W7	W8	W9

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

a*1/9	b*1/9	c*1/9
d*1/9	e*1/9	f*1/9
g*1/9	h*1/9	i*1/9

Fig. 7: Process of filtering

Matched filtering (template matching): The first task in doing template matching is to determine what template to use. Intuitively, it seemed reasonable to us that the best template to use would be one derived by somehow averaging the some images of the students in the training images that would likely be in the testing images. We would like to find a good subset of the faces found in the training images that are clear, straight and representative of typical lighting/environmental conditions. It is also important that these images be properly aligned and scaled with respect to one another. To this end we spent considerable time manually segmenting, selecting and aligning face photos. In the end we chose 30 face images which may be seen in Fig. 8. In order to have the template reflect the shape of the faces it is trying to detect, rather than their particular coloring, etc., we applied histogram equalization to each image and removed the means. This resulted in Fig. 8. Our final template is a result of adding together the 30 face images in Fig. 8. The actual template used in the matched filtering started at 30×30 pixels by resizing this template. Its size was changed to cover different possible scaling in our test image.

When face are detected, we remove the corresponding portion of the masked input image to try to avoid multiple and false directions (Fig. 9). In addition we:

- Convert to gray scale. In our case, no more color information to extract. Apply mean removal+ histogram equalization -> flatten and bring out details

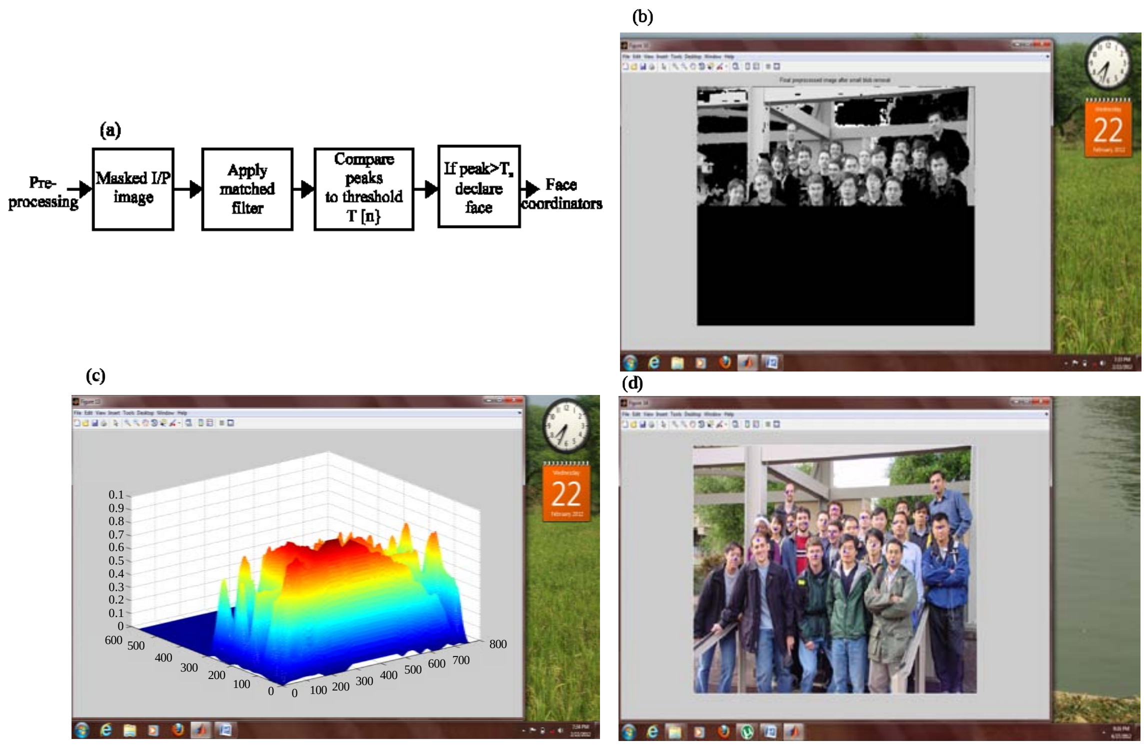

Fig. 8: a-d) Template matching

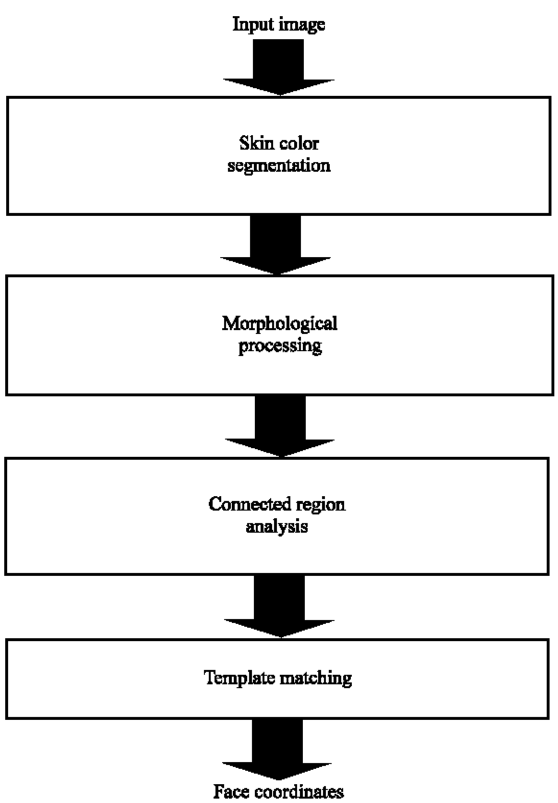

Fig. 9: Flow chart

- Manually selected a number of quality faces: centered, straight, lighting, diverse measured face dimensions and used MATLAB to uniformly scale and align them
- Efforts resulted in 26 sample faces added to produce the final template

RESULTS AND DISCUSSION

The skin template threshold values of HSV lies between 19 and 240. If the HSV threshold values lies in between the range then it is considered as skin other vice as non skin and it removes the remaining part from image (Fig. 10-16).

Same as HSV the skin template of YCbCr values lies between 102-128. If the YCbCr values lies between the given threshold it is skin other vice non skin and removes the remaining part.

After completing the skin recognization for our convience it removes the lower plane by masking technique to identify only faces then it performs open operation to mark only faces of the group.

Applications:

- Used in markets, shopping malls for face reorganization

Fig. 10: Input image RGB image

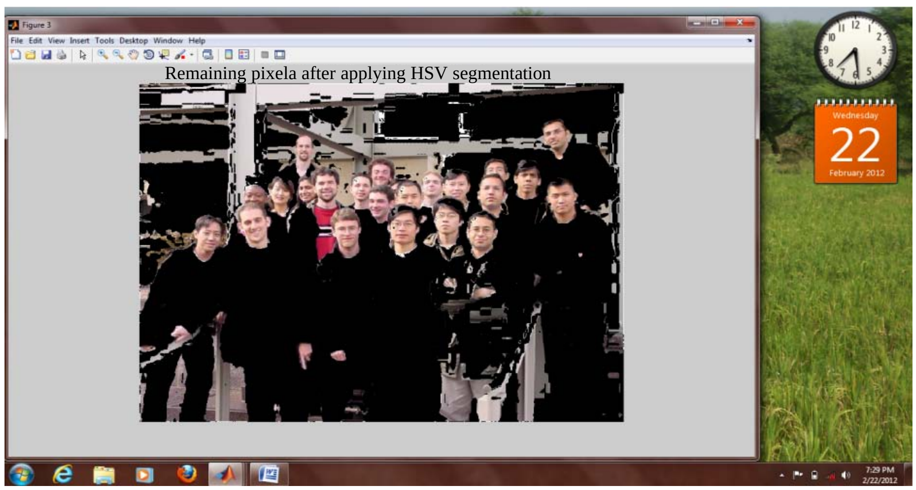

Fig. 11: HSV image

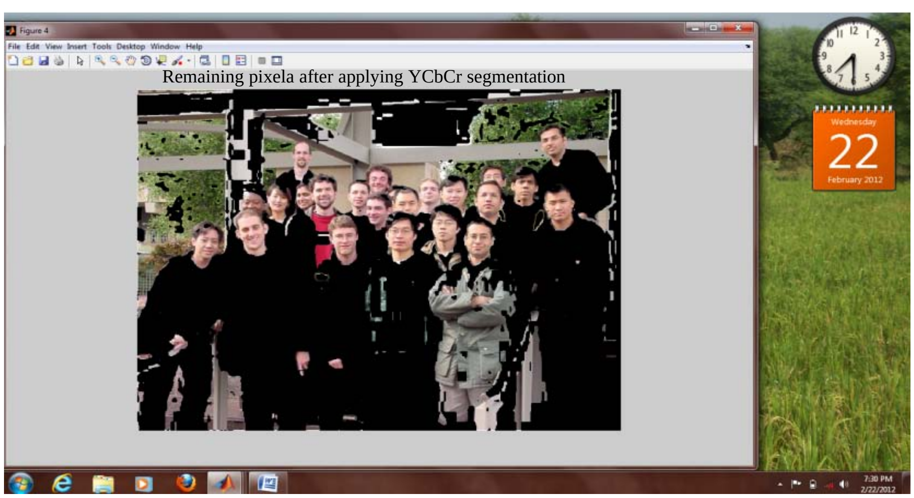

Fig. 12: YCbCr image

Fig. 13: First RGB segmentation rule image

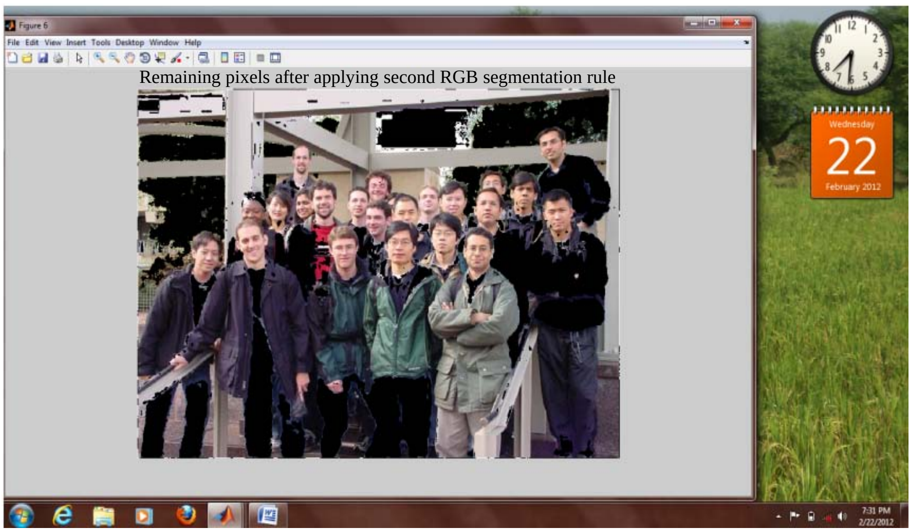

Fig. 14: Second RGB segmentation rule image

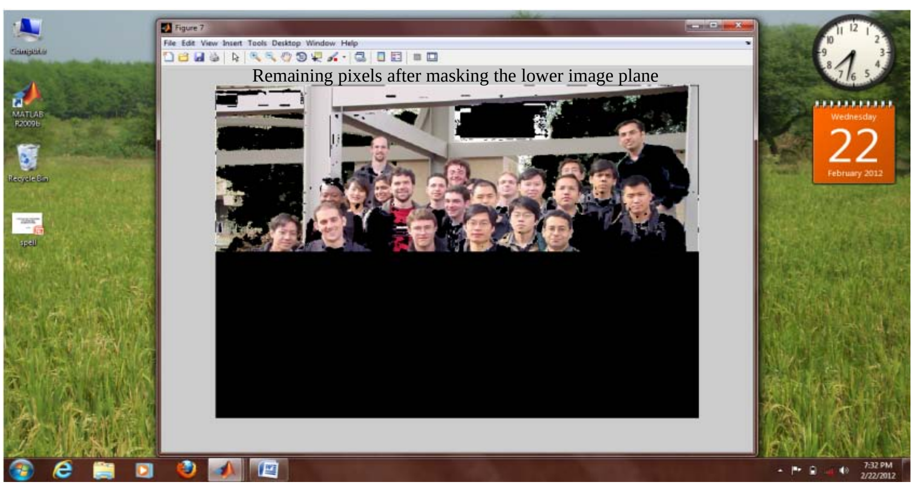

Fig. 15: Lower plane masking

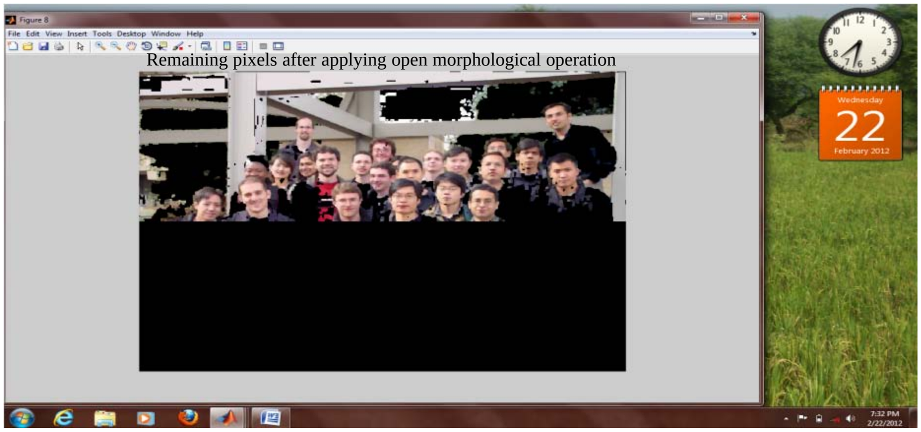

Fig. 16: Open morphological image

- Institutions, colleges, offices for attendance purpose
- In photography for face extraction

CONCLUSION

Using the algorithm described produced rather reasonable results when applied to the various training images. For the particular image that this report has been looking at were able to accurately detect the 22% faces and had no false alarms or misses. Our results when applied to the other testing images ranged from approximately 85-100%. Face detection algorithm is reasonably fast in that it performs typically in approximately 100 sec or so and is sufficiently accurate given the difficulty of the problem.

REFERENCES

Colmenarez, A.J. and T.S. Huang, 1997. Face detection with information-based maximum discrimination. Proceedings of the Conference Computer Vision Pattern Recognition, June 17-19, 1997, San Juan, Puerto Rico, pp: 782-787.

Jain, A.K., 1989. Fundamental of Digital Image Processing. 5th Edn., Prentice-Hall, Englewood Cliffs, NJ., ISBN: 9780133361650, Pages: 569.

Kotropoulos, C. and I. Pitas, 1997. Rule-based face detection in frontal views. Proceedings of the 1997 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP-97) Vol. 4, April 21-24, 1997, IEEE, Munich, Germany, ISBN:0-8186-7919-0, pp: 2537-2540.

Lanitis, A., C.J. Taylor and T.F. Cootes, 1995. An automatic face identification system using flexible appearance models. Image Vision Comput., 13: 393-401.

Liu, C. and H. Wechsler, 2000. Evolutionary pursuit and its application to face recognition. IEEE. Trans. Pattern Anal. Mach. Intell., 22: 570-582.

Moghaddam, B. and A. Pentland, 1997. Probabilistic visual learning for object representation. IEEE. Trans. Pattern Anal. Mach. Intell., 19: 696-710.

Perez, C.A., C.M. Aravena, J.I. Vallejos, P.A. Estevez and C.M. Held, 2010. Face and iris localization using templates designed by particle swarm optimization. Pattern Recognit. Lett., 31: 857-868.

Treptow, A. and A. Zell, 2004. Combining adaboost learning and evolutionary search to select features for real-time object detection. Proceedings of the Congress on Evolutionary Computation (CEC2004) Vol. 2, June 19-23, 2004, IEEE, Portland, Oregon, USA., ISBN:0-7803-8515-2, pp: 2107-2113.

Wu, B., H. Ai, C. Huang and S. Lao, 2004. Fast rotation invariant multi-view face detection based on real adaboost. Proceedings of the 6th IEEE International Conference on Automatic Face and Gesture Recognition, May 19, 2004, IEEE, Seoul, South Korea, ISBN:0-7695-2122-3, pp: 79-84.