

Finding Facial Features Using an HLS Colour Space

Petter Ranefall, Bo Nordin, Ewert Bengtsson

Centre for Image Analysis,
Uppsala University, Sweden

Abstract. A method for finding facial features using information from an HLS colour space is presented. This method attempts to find the face with simple single band operations and in case of failure a more advanced method is used. Using this approach the nice photos which are the majority will be handled quickly, more difficult photos will be handled with more advanced methods and only a few images will be reported as failures.

1 Introduction

The general image segmentation and interpretation problem is still largely unsolved. But for constrained images, where substantial a priori information about the allowable image contents is available, rather complex segmentation and interpretation tasks can readily be solved. The recognition of facial features in images of human faces is an example where a priori information can be used to advantage. One of the first attempts to extract facial features was made by Kanade [1]. There are many possible applications for finding facial features e.g. face recognition [2,3,4], tracking the motions in the face for high compression image coding [5,6] or controlling a computer cursor simply by moving the head [7], to mention a few examples. Virtually all recent international conferences in image analysis or pattern recognition includes a few papers about face image analysis.

Our application is a rather simple one. We need to find the outline of the head as well as the position of the eyes and mouth in order to position and scale the image properly when producing driver's licenses or ID cards automatically from colour portrait images glued onto standardized application forms. The method has to be simple and robust as well as having the potential of being implemented at high speed in standard image array processing hardware.

2 Method

As a preprocessing step we convert the RGB image to HLS. Then the algorithm works in three major steps. If any of the three steps fails completely the processing reports a failure for the current image.

In the first step we first attempt to find the head contour as easily as possible by thresholding the Lightness image. If that does not work we use a Maximum Likelihood (ML) classifier on the RGB image. Secondly we search for the eyes in the Saturation image since the eyes will have very low saturation. The search area is restricted to regions

surrounded by skin. The skin is found by an ML classifier on the RGB image. Thirdly, when the eyes are found we use the natural geometry of faces and search in an area under the eyes for the lips which differs from skin in the Hue image. A problem with the Hue image is that highlighted parts of the face (e.g. the nose) may also differ from the normal skin. The mouth is found at the broadest part of this region.

In the following description of the different steps in our method the *italicised* terms will be described in the Tools section (3).

2.1 Finding the Head Contour

1. Transform the image to HLS (see figure 1 and figures 2a, b, c). There are many ways of converting RGB images to HLS proposed in the literature. In our case Hue is defined such that 0° is between blue and green and Saturation is defined as the difference between the highest and the lowest RGB-value. Lightness is defined as the average of the highest and lowest RGB-value. This definition is useful for our purposes since the white and black in the eyes gets low saturation values and the skin colour is in the middle part of the hue scale.

2. *Automatic threshold* of the Lightness image.

3. *If the head contour is not OK:*

ML train & classify the original colour image. This is done by using a pattern image (see figure 3) which marks some areas where background, clothes and head are expected.

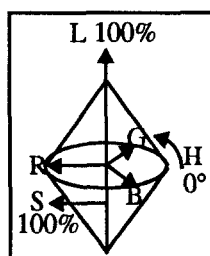


Fig. 1. HLS colour space.



Fig. 2a. Lightness.



Fig. 2b. Hue.



Fig. 2c. Saturation.

2.2 Finding the Eyes

The saturation image is used since the eyes differ from the rest of the face in this image. In some cases (e.g figure 2b) the hue image could serve equally well.

Narrow the search area by creating an *ML-classifier for skin and hair* (figure 4) and extract the skin within the head contour.

Line-by-line search within this region in the Saturation image.

2.3 Finding the Mouth

1. Extract a sub image under the eyes in the Hue image (see figure 5a). The sub image is a square with the side equal to the distance between the eyes and it is placed one half of the side length straight below the eyes.
2. *Automatic threshold* of this mouth region (see the result in figure 5b).
3. Fill holes in the marked area.
4. Find the maximum horizontal width within this marked area.

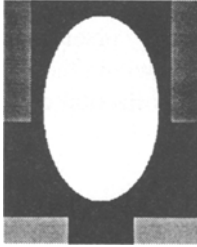


Fig. 3. Pattern.



Fig. 4. Training regions.



Fig. 5a. Contrast enhanced mouth region in the Hue image of figure 4.



Fig. 5b. Thresholded mouth region.

3 Tools

This is a more detailed description of the tools used in the previous section.

3.1 Automatic Threshold

Because of variations in skin and hair colour, lighting, background etc. we use a dynamic thresholding method which works as follows:

Split the histogram at the point S where the maximum split effect is obtained. The split effect for $S \in [x_L, x_U]$ is defined as follows [8]:

$$\sum_{k=x_L}^{x_U} (k-\bar{k})^2 - \left(\sum_{m=x_L}^S (m-\bar{m})^2 + \sum_{n=S+1}^{x_U} (n-\bar{n})^2 \right) \quad [\text{eq.1}]$$

where x_L and x_U are the lower and upper limits of the histogram, and \bar{k} , \bar{m} and \bar{n} are the mean values for the histogram and the sub histograms, respectively. If this does not separate the background from the object then the sub histogram which contains both the object and the background is split. This procedure continues until the object is separated from the background. The upper corners of the image are assumed to contain background and the centre of the image is assumed to contain object and the average in these areas defines background and object, respectively.

3.2 Test of Head Contour

Broken head contours can be detected using the classical shape measure:

$$P^2/A = \frac{\text{Length of contour}^2}{4\pi \cdot \text{Area}} \quad [\text{eq.2}]$$

where the contour in our case is measured on the upper half of the portrait and the area is the area within the same contour. We have empirically found that heads with normal contours have P^2/A values in a range from approximately 1.2 to 1.8.

3.3 ML Classifier for Skin and Hair

For each column a number (we have used 1/6 of the Y-dimension) of pixels closest to the upper contour of the head are selected as training region for hair. Only pixels in the upper half of the image are used. Around the average of the x-values within the contour a rectangle of size $X\text{-dim}/3$, $X\text{-dim}/6$ is placed standing on the horizontal centre line (figure 4). This rectangle defines the training region for skin.

3.4 Line-by-line Search for Eyes

Search for the lowest threshold value for which darker points are found on both the left and the right side of the middle within the skin area on the same horizontal line. This search also tests if conditions for the distance between the eyes and the distance to the edge of the face are fulfilled.

4 Application

ID KORT AB which produces all driver's licences and most of the ID cards in Sweden uses digital image processing based on a previous method designed by two of the authors. This has been used to produce several million licences and cards so far. The old method does however rely on rather rigid assumptions about the size and position of the photos under the camera. Since these assumptions frequently are violated many manual interventions are necessary. The company now wants to be able to process their cards with a higher degree of automation. That includes moving and resizing the photos based on the positions of the eyes and mouth.

4.1 Finding the Photo

As a preprocessing step the photo has to be found on the form. This is because many people have problems in placing the photo at the correct place in spite of rather detailed instructions. An automatic threshold applied to the colour band with the highest split effect (eq. 1) with the value for background selected as 255 and the value for object selected as 0 does the work. This often finds the contour of the head, which could make that special first operation superfluous.

The photo is then cut out from the larger digital image which has been scanned and the rest of the processing is done on this sub image. The images are subsampled to approximately 200x250 pixels to speed up the calculations.

4.2 Finding the Facial Feature Positions

Using the algorithm described above (section 2-3).

4.3 Smoothing the Background

The colour and intensity of the background is smoothed linearly from the contour of the head to the edge of the photo in order to blend the image into the ID card without disturbing effects from the background or edges of the photo.

5 Results

Figure 6 shows the result of facial feature detection of the face in figure 2. For this image the more advanced method was used to find the contour because the face was illuminated mainly from the right. Features are found in faces with darker hair and complexion or glasses (e.g figure 7) or facial hair (e.g figure 8), but glasses, facial hair as well as no hair at all could cause some problems. We will continue to study these kinds of problems.

The average runtime were about 15 seconds on a DECstation 5000/120. On a more advanced workstation this would decrease to a few seconds which would be acceptable for production.



Fig. 6. Face with difficult illumination.

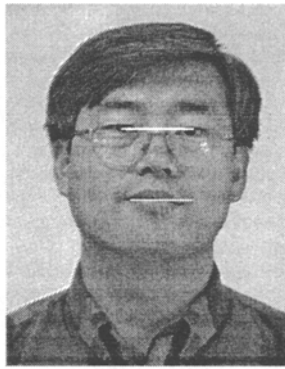


Fig. 7. Face with glasses.

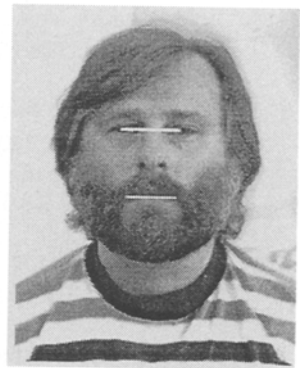


Fig. 8. Face with facial hair.

6 Discussion

Computer interpretation of facial images has recently been discussed in many papers as pointed out in the introduction. Depending on the application the images can be assumed to be more or less constrained. In our case we have a substantial amount of a priori information about the images. If they do not fulfil these conditions they are not accepted for making ID-cards. Still there is significant variation in the way people may look and a few images are relatively hard to process. This has been utilized in our algorithms which attempts to use simple and quick methods when possible but reverts to more demanding processing to deal with the more difficult cases. This increases the robustness of the method without a corresponding penalty in processing time.

The algorithm has so far been tested on a limited number of cases. It will be installed in the production line of ID KORT AB and used to process about a million images per year. This will provide reliable statistics about its performance in the near future.

Another aspect of colour photo processing is to normalize the colour, so that very dark/light images will be improved. This must be done without removing the variance due to skin and hair colour. The lightness compensation must also be done without changing the perceived hue. Perhaps this should be done in an (approximately) perceptually uniform colour space (e.g CIELAB or CIELUV). We will investigate these problems in the near future.

Acknowledgement

We are very grateful to ID KORT AB for their financial support and cooperation.

References

1. T. Kanade: Picture processing system by computer complex recognition. PhD thesis, Department of Information Science, Kyoto University, Japan, 1973.
2. Chung-Lin Huang, Ching-Wen Chen: Human facial feature extraction for face interpretation and recognition. Proc. of the 11th IAPR Int. Conf. on Pattern Recognition, The Hague, The Netherlands, August 1992, vol. 2, pp. 204-207.
3. Alex Pentland, Baback Moghaddam, Thad Starner: View-based and modular eigenspaces for face recognition. Proc. of the 1994 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, Seattle, Washington, USA, June 1994, pp. 84-91.
4. Guangzheng Yang, Thomas S. Huang: Human face detection in a complex background. Pattern Recognition, Vol. 27, No 1, pp.53-63, 1994
5. K. Aizawa, H. Harashima, T. Saito: Model-based analysis synthesis image coding (MBASIC) system for a person's face. Proc. of GLOOBECOM-87, November 1987, pp. 45-49.
6. Thomas S. Huang, Subhash C. Reddy: Human face motion analysis. Proc. of the Int. Workshop on Visual Form, Capri, Italy, May 1991, pp 287-292.
7. Thomas C. Chang, Thomas S. Huang, Carol Novak: Facial feature extraction from color images. Proc. of the 12th IAPR Int. Conf. on Pattern Recognition, Jerusalem, Israel, October 1994, vol. 2, pp 39-43.
8. Otto Milvang: An adaptive algorithm for color image quantization. Proc. of the 5th Scandinavian Conf. on Image Analysis, Stockholm, Sweden, June 1987, vol. 1, pp. 43-47.