

Using Color and Geometric Models for Extracting Facial Features

Tomoyuki Ohtsuki

Sony Corporation, Tokyo, Japan

Glenn Healey

Department of Electrical and Computer Engineering, University of California, Irvine, California

We describe a method for finding the eyes and mouth in a color image of a human face. This method makes use of two new models: a skin color model and a three-dimensional (3-D) geometric model of facial features. The skin color model is based on a medical model for the reflectance of skin and applies to all kinds of skin color. The 3-D geometry model is based on medical statistics of dimensions of the human face and is used to evaluate the geometry of facial features. This model applies for any orientation of the face. The algorithm first determines candidates of facial features by color segmentation. Sets of mouth candidates and eye candidates are evaluated based on the color of the segment that surrounds them and its geometry using the skin color model and the 3-D geometry model. The features are recognized according to the candidate set with the highest likelihood. In experiments, the algorithm detects the features correctly in a large majority of face images in many orientations with a cluttered background.

Journal of Imaging Science and Technology 42: 554–561 (1998)

Introduction

Progress in hardware technology has accelerated recent research in image coding, which requires a large amount of processing power. Among image coding applications, model-based coding of face images promises to allow very low bit rate coding. In this method, encoder and decoder share the same model of face shape, expression, and the like. The encoder analyzes the input image to extract the parameters for the model, which are sent to the decoder to enable reconstruction of the image. Although significant research has been done in this field,^{1–3} most work emphasizes models and less has been done on automatic parameter extraction. Motivated by this fact, we have developed methods for facial feature detection that serve as the first step in parameter extraction.

Facial feature detection has a wide variety of applications. One of the best-known applications is face recognition where detection provides information about face position, orientation, and size. It has also been used in a mugshot system⁴ that retrieves face images from a database using the description of the shape and geometry of facial features. Some of the applications that attract the most research are in the area of man–machine interfaces. One goal is to use the position of focus as an input position instead of using the position specified by a pointing

device. Other research tries to determine human emotion from facial expression to enable a high level communication between the machine and the user. In both applications, facial feature detection can be an important step in the process.

Many facial feature detection algorithms work with frontal views of the face where almost no tilt nor turn is present.⁵ Some methods account for tilt of the face (rotation in the image plane),^{6,7} but few are designed to work for arbitrary face orientation.⁸ Many algorithms also assume that the face has a controlled background, making detection significantly easier. It is possible to assume this kind of background for applications such as teleconferencing where the environment can be controlled. However, in other situations the background is usually somewhat cluttered. Recently, some methods that use skin color have been proposed,^{9,10} while other work by Sanger and co-workers^{11,12} has combined binary image processing and color for face detection. Though these methods are related to the algorithm proposed here, the skin color models used by these approaches are not as generally applicable as the one proposed by this research. Another approach to face recognition is the use of global representations, that are often obtained by learning from sample images.^{13,14,15} However, because these approaches do not make use of physical or geometric models, performance often degrades significantly when the system is presented with images that differ from those in the training set.

The models we propose address these problems. The skin color model is based on the reflectance of skin and applies to any skin color. The facial features 3-D geometric model describes the eyes and mouth using statistics from medicine and applies to any facial orientation and scale. We believe that using these models will add robustness to systems for applications that we mentioned earlier. The

Original manuscript received December 19, 1997

ohtsuki@arc.sony.co.jp

healey@ece.uci.edu

© 1998, IS&T—The Society for Imaging Science and Technology



Figure 1. Original image.



Figure 2. Segmentation result.

system we use to demonstrate the effectiveness of these two models detects the extremities of the eyes and mouth in a cluttered outdoor background. The algorithm proceeds in several steps. First, color segmentation is performed to isolate candidate face regions as segments and candidate facial features as holes in the segments. Taking combinations of three holes within a segment as candidate facial features, the candidates are evaluated based on the color of the segment that surrounds the holes and the geometry of the extremities of the holes. This evaluation is based on the skin color model and the geometry model. The candidate set with the highest likelihood is output as the detection result. Details of the algorithm and experimental results are described in the following sections.

Finding Facial Feature Candidates

Finding Face Candidates. The first stage of the algorithm is segmentation of the input image, which enables separation of the face region from the background. The method used is region growing and the details of the segmentation algorithm appear in the Appendix. Following segmentation we get several segments that are candidates for the face region. At this stage it is not determined which

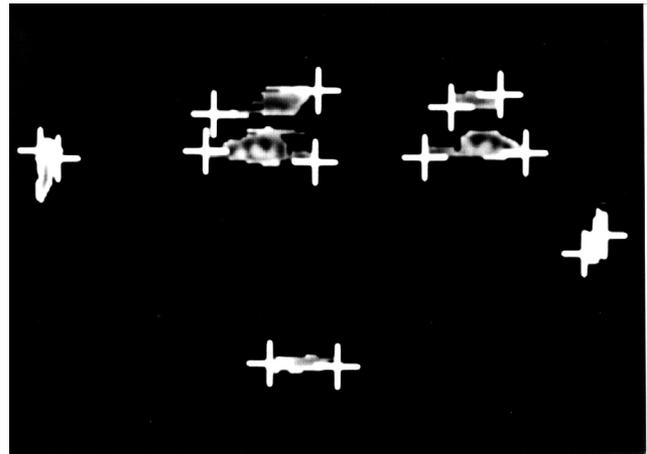


Figure 3. Detected holes and their extremities as eye candidates.

of the segments corresponds to the face. This is determined at the final stage using the two new models.

Figure 2 shows the result of the segmentation on the original image shown in Fig. 1. In Fig. 2, only the segment that corresponds to the face region, as determined at the final stage, is displayed. The background region, shown in black, is not a single segment but includes several segments.

Finding Eye Candidates. Because the eyes are of different color from the skin, they appear as two holes in the face region segment. The algorithm locates holes in each segment and considers them as eye candidates. Holes with a number of pixels less than a threshold *harea_th* are treated as small regions and are not processed in the following stages. To evaluate the candidates using the geometry model explained in a later section, the two extremities of each hole are detected as follows:

1. Calculate the center of area of the hole.
2. Calculate the axis of least second moment of the hole, the orientation of which is considered as the orientation of the hole. Express the orientation by a vector \mathbf{V} .
3. Find the most distant pixels in the hole from the center of area in the direction of \mathbf{V} and $-\mathbf{V}$. If several pixels have the same distance in one direction choose the one closest to the axis of least second moment.

The holes and their extremities found in the face region segment are shown in Fig. 3.

Finding Mouth Candidates. Due to the difference in color of the mouth from the skin, the mouth appears as a hole in the face region segment. However using the same segmentation thresholds to locate both eye and mouth regions can cause a problem as explained below.

- The hollow often present below the eyebrow causes a shadow from eyebrow to eye. Due to this shadow, using small thresholds for the segmentation produces a big hole including both eyebrow and eye. Using large thresholds, on the other hand, make it possible to separate eye from eyebrow but often divides the mouth region into several holes or makes the mouth hole very small. This is because skin color is usually more similar to mouth color than to eye color.

To avoid this problem mouth region candidates are extracted using the following method.

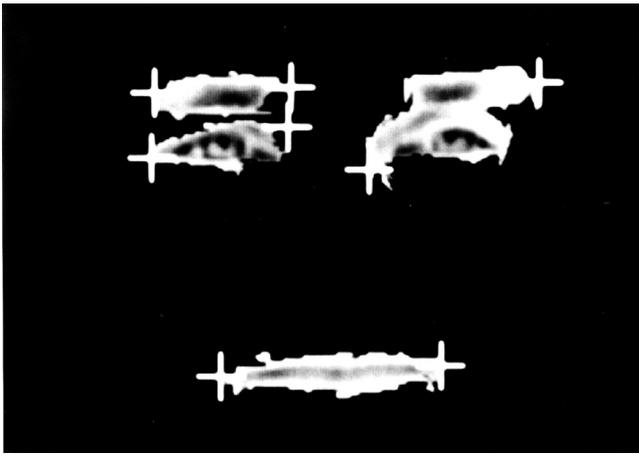


Figure 4. Detected holes and their extremities as mouth candidates.

- First, holes in each segment found using the method previously explained are filled in with the original pixel values so that segments have no holes. Then segmentation is carried out for each segment using the method in the Appendix but changing thresholds from sqr_th to $msqr_th$ and from ang_th to $mang_th$. This determines subsegment(s) for each segment. As in the case of finding face region candidates, small regions having fewer pixels than $area_th$ are not considered as subsegments. Holes in subsegments in a segment are considered as mouth region candidates for that segment.

As in finding eye region candidates, those holes with a number of pixels less than a threshold $mharea_th$ are treated as small regions and are not processed in the following stages. After finding mouth region candidates, extremities of each candidate are calculated using the same method used for finding extremities of eye region candidates. Mouth region candidates and their extremities found in the face region segment are shown in Fig. 4.

Facial Features Candidates. A facial features candidate is composed of two eye candidates and a mouth candidate. In a segment, all combinations of two eye candidates and a mouth candidate are considered as facial features candidates. The facial features candidates will be evaluated by the chromaticity of the segment surrounding the candidates and the geometry of the candidates. The evaluation method is explained in the following sections.

Evaluation by Skin Color Model

Evaluation Method. The evaluation of a facial features candidate by color is based on the angle between the representative color vector of the segment surrounding the candidate and the color vector in the skin color region, explained later, which gives the smallest angle. This angle is denoted as E_{color} in the following discussion. The representative vector \mathbf{V}_{rep} is calculated as follows:

$$\mathbf{V}_{rep} = \frac{\mathbf{V}_{ttl}}{\|\mathbf{V}_{ttl}\|} \quad (1)$$

where

$$\mathbf{V}_{ttl} = \sum_{\text{segment}} \mathbf{v}. \quad (2)$$

Summation is over all pixels in the segment and

$$\mathbf{v} = \frac{\mathbf{V}}{\|\mathbf{V}\|}, \quad (3)$$

where \mathbf{V} denotes a vector composed of the red, green, and blue (R, G, and B) values of a pixel in the segment. When a representative vector is inside the skin color region, then E_{color} is 0. The skin color region is defined in the following subsections.

Optical Properties of Skin. The reflection of light can be divided into two parts: surface (or interface) reflection, which occurs at the surface of an object due to the change of refractive index, and body reflection due to light penetrating into the body of the material that encounters scattering and absorption and later emerges from the material.¹⁶ To establish a model for the surface and body reflectance of skin, the optical properties of human skin from the medical literature were used.¹⁷⁻²¹ Human skin is composed of epidermis, a thin layer on the surface of the skin, and dermis, which lies beneath the epidermis. Surface reflection of the skin is from the epidermis surface and the reflectance is about 4 to 7% independent of the wavelength of the light. On the other hand, body reflectance is determined by the following optical properties of the two layers:

1. epidermis
 - negligible scattering
 - visible light absorption mainly by melanin
2. dermis
 - scattering mainly by collagen fibers
 - visible light absorption mainly by ingredients in blood such as hemoglobin, bilirubin, and beta-carotene.

As a result, the epidermis can be treated as an optical filter that transmits light according to wavelength, whereas the dermis has a much more complex body reflectance property due to scattering. Despite this complexity the body reflectance of dermis is mainly determined by blood content. Though there are many different human skin colors the optical properties of dermis are basically the same for all of them. The difference in skin color for different races is caused primarily by the difference in the optical properties of the epidermis, which is caused mainly by differences in melanin content.

Skin Color Model. From the optical properties of the skin, a simple model of surface and body reflectance can be derived:

1. surface reflectance: 4 to 7%
2. body reflectance:
 - $< \text{epidermis transmittance} >^2 \times < \text{dermis reflectance} >$

In the body reflectance model, epidermis transmittance is squared because the light goes through the epidermis when it enters the body and also when it leaves the body. To further simplify the model a fixed value 5% for surface reflectance R_{surf} was adopted.

The remaining factors that determine reflectance of the skin are epidermis transmittance and dermis reflectance. To make the calculation easier the transmittance of dopa-melanin is used for the epidermis transmittance in the final model. The transmittance of dopa-melanin is calculated as follows. Given the absorbance of dopa-melanin $ABM(\lambda)$ of a certain density, as shown in Fig. 5, the transmittance $TRM0(\lambda)$ at the same density is $10^{-ABM(\lambda)}$. When the density is multiplied by a factor d , the transmittance

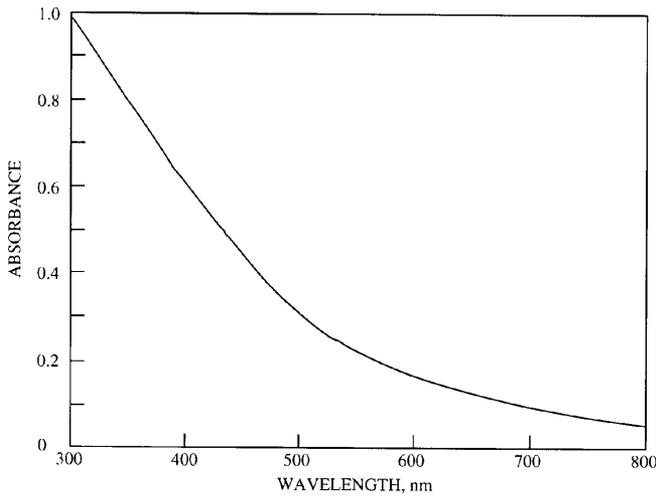


Figure 5. Absorbance of dopa-melanin (1.5 mg % in H₂O). Reproduced from Ref. 18.

$TRM(d, \lambda)$ becomes $\{TRM_0(\lambda)\}^d$. Thus, by changing d the transmittance of epidermis with any melanin density can be obtained. Using $TRM(d, \lambda)$, the total reflectance of skin can be expressed in the form

$$R_{surf} + TRM(2 \times d_0, \lambda) \times R_{derm}, \quad (4)$$

where R_{derm} denotes the dermis reflectance and d_0 corresponds to the density of melanin in the skin. Conversely, given a skin reflectance, dermis reflectance can be written in the form

$$(R_{skin} - R_{surf}) \times TRM(-2 \times d_0, \lambda), \quad (5)$$

where R_{skin} denotes the skin reflectance.

Consequently the reflectance of skin with a given dermis reflectance as a function of melanin density in the epidermis can be described as

$$R_{surf} + (R_{skin} - R_{surf}) \times TRM(e, \lambda) \quad (6)$$

where e denotes the degree of the effect of melanin added to the original skin reflectance.

This is the model we use to calculate the skin color region. By changing the original skin reflectance and e to proper values, any skin reflectance is obtainable. The reason to use R_{skin} instead of R_{derm} is that neither a simple model for dermis reflectance nor proper data for dermis reflectance is available.

The reflectance curves shown in Fig. 6 are used to determine the skin color region. These are representative of fair Caucasian skin. One curve is at a normal state and the other at an erythematous state. The erythematous state is a congested state right after sunburn, which gives the skin a reddish appearance. Here, normal state means that the blood content in dermis is low and erythematous state means that it is high. These two states define the lower extremity and upper extremity of blood content that defines the skin color region. By converting reflectance to R, G, and B values, the skin color obtained by using these two curves as original skin reflectance and by changing e generates two curves in RGB space. The skin color region

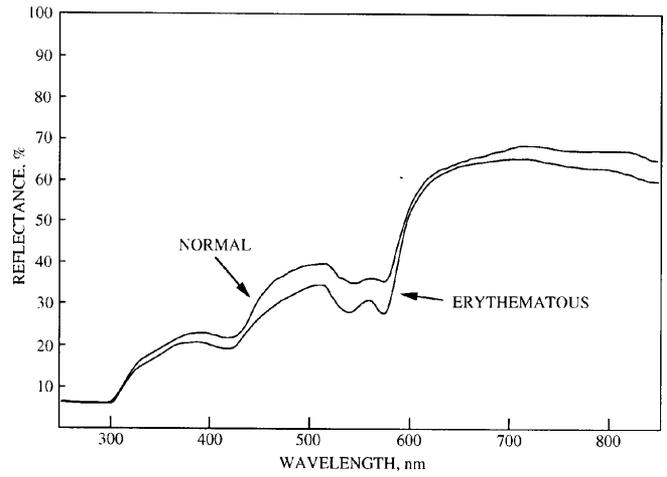


Figure 6. Reflectance spectra of normal and erythematous skin. Reproduced from Ref. 18.

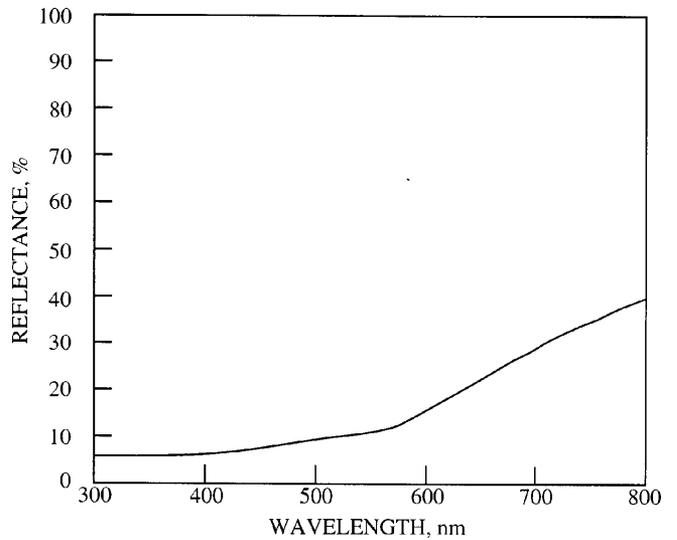


Figure 7. Reflectance spectra of dark Negro skin. Reproduced from Ref. [20].

in RGB space is the area between these two curves. The largest and smallest e , which determines the ends of the two curves, is determined as follows. The largest e , which gives the darkest appearance, is chosen so that the synthesized skin reflectance curve using the normal skin reflectance and the e value lies below the skin reflectance curve of dark Negro skin shown in Fig. 7. The smallest e is decided so that the largest value in the synthesized skin reflectance curve using the normal skin reflectance and the e value is 100%.

Because the luminance of the spectral radiance depends on light source intensity, normalized colors r , g , and b are used, where

$$r = \frac{R}{N}, g = \frac{G}{N}, b = \frac{B}{N}, N = \sqrt{R^2 + G^2 + B^2}. \quad (7)$$

Figure 8 shows r and g values of the skin color region. As is clear from the synthesis process described above, any skin color should fall within this skin color region.

The conversion from reflectance to R, G, and B values is computed using an illumination model and a camera

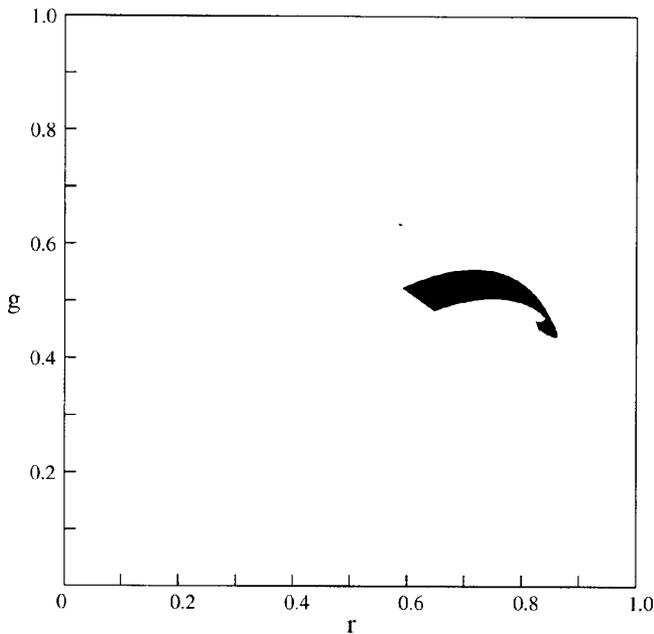


Figure 8. Skin color region.

model. Given the light source $L(\lambda)$ and the reflectance $s(\lambda)$, the R, G, and B values are computed using

$$R = \int_{\lambda} L(\lambda)s(\lambda)f_R(\lambda)d\lambda, \quad (8)$$

$$G = \int_{\lambda} L(\lambda)s(\lambda)f_G(\lambda)d\lambda, \quad (9)$$

$$B = \int_{\lambda} L(\lambda)s(\lambda)f_B(\lambda)d\lambda, \quad (10)$$

where $f_R(\lambda)$, $f_G(\lambda)$, and $f_B(\lambda)$ are the spectral sensitivities of the red, green, and blue elements in the color camera. For $L(\lambda)$, a Plankian light source²² at 6500 K is used. This source is often used to represent average daylight.^{23,24} For other lighting conditions we can use other light source models for conversion from reflectance to R, G, and B values. The $f_R(\lambda)$, $f_G(\lambda)$, and $f_B(\lambda)$ functions were obtained from the spectral sensitivity functions for the SONY video camera used in the experiments.

Evaluation by Three-Dimensional Geometry Model

Each facial features candidate is evaluated by the geometry of the extremities. This is done by comparison with a geometric template. As a result, a value E_{geom} is computed for each candidate that measures the degree of unlikelihood of the facial features.

Geometric Template. The template is determined based on medical statistics of anthropometry in Ref. 25. As shown in Fig. 9, this describes the 3-D geometry of the extremities of two eyes and a mouth. The “ex” denotes an outer extremity of an eye, “en” denotes an inner extremity of an eye, “ch” denotes an extremity of a mouth, and “sto” denotes the center of two “ch”. The data used to obtain the template are from three groups of people: North-American Caucasian, Chinese, and African-American. The dimensions used to get the template are averages for each group with most of them taken directly from Ref. 25 or

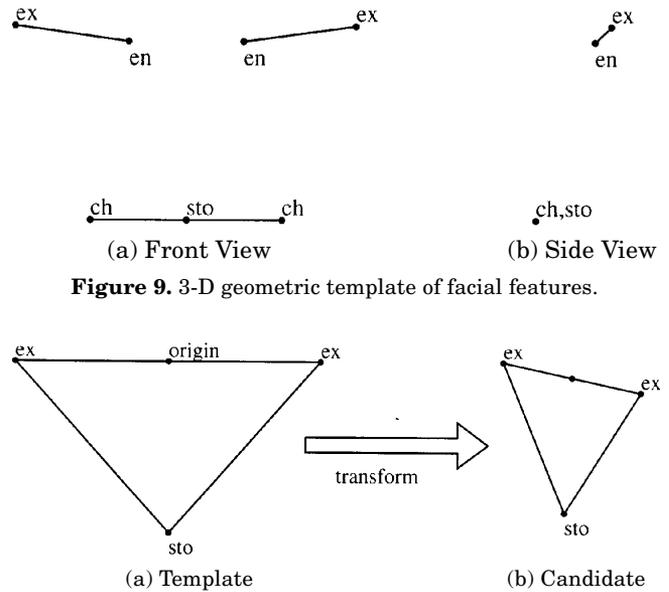


Figure 9. 3-D geometric template of facial features.

Figure 10. Estimation of transform.

derived indirectly from other dimensions. The template is determined by first normalizing all dimensions in each group so that the distance between “ex” of the left and right eyes becomes 1 and then taking the average of the three groups for all dimensions. For dimensions not available in certain groups the average of the data available is used.

Evaluation by Geometry. To evaluate a candidate by geometry, the rotation, scaling, and offset of the candidate from the template are estimated. The estimation of this transform is based on the triangle whose vertices are two “ex” and “sto”, as shown in Fig. 10(a). Using the coordinate system shown in Fig. 11, where the plane $z = 0$ corresponds to the input image plane, the angles of rotation α , β , γ , scale S , and offset (x_{off}, y_{off}) are estimated. The template is rotated about the x axis, y axis, and then the z axis in this order and by α , β , and γ , respectively. The S is the factor to be multiplied to template dimensions. Offset is the position of the origin of the template that is the center of the two “ex” relative to the origin of the input image which is at the upper left corner of the image. The offset in the z axis direction is not considered because the parameters are for acquiring the orthogonal projection of the transformed template to the image plane along the z axis. The parameters for this transform are calculated so that the projection of the transformed template triangle to the image plane matches the triangle of the candidate. As a result, unique values for γ , S , (x_{off}, y_{off}) , and two combinations of α and β are obtained such that when one combination is α_0 and β_0 , the other is $-\alpha_0$ and $-\beta_0$. If the results satisfy $|\alpha| > 35^\circ$ or $|\beta| > 35^\circ$ or $|\gamma| > 45^\circ$, the candidate is not considered further. This is because large rotations hide the extremities of facial features. A slightly larger value 45° is used for γ because rotation about the z axis does not affect the visibility of facial features. The limit for γ is introduced because it is not likely that a face in an image is largely tilted and by using these limits, the likelihood of misdetection can be reduced.

Once the parameters are acquired, the entire template with six extremities is transformed accordingly. The result of projection of the transformed template to the image plane is shown in Fig. 12 together with the candidate. Thick lines are used for the template and thin lines are

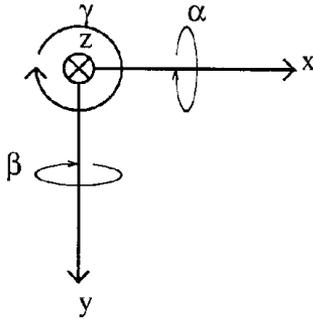


Figure 11. The coordinate system used in the estimation of transform.

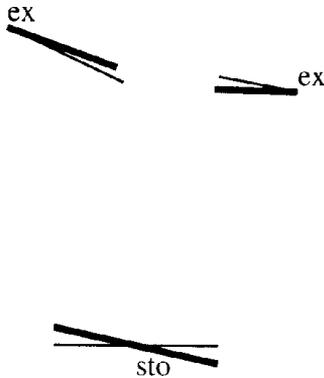


Figure 12. A candidate and the template transformed to match it.

used for the candidate. As seen from Fig. 12, two “ex” and “sto” are at the same positions for both template and candidate. Both eyes are evaluated based on the difference between the vectors “ex” → “en” of the template and the candidate, and the mouth is evaluated based on the difference between the vectors “sto” → “ch” of the template and the candidate. Denoting the difference between the length of the template vector and the length of the projection of the candidate vector on the template vector as A and the length of the projection of the candidate vector in the direction orthogonal to the template vector as B , each facial feature is given a value

$$A + P_WEIGHT \cdot B, \quad (11)$$

where P_WEIGHT controls the relative significance of A and B (see Fig. 13). This value estimates the degree of unlikelihood as a facial feature. In this work, the value of P_WEIGHT was taken to be one and system performance was not sensitive to changes in the value of this parameter. We introduce P_WEIGHT as a variable, however, to allow different components of the error to be treated differently. The sum of the values of the three features denoted E_{geom} is used to evaluate the facial features candidate. Because there are two combinations of parameters for the transformation corresponding to two combinations of α and β , E_{geom} is calculated for each set of parameters.

Due to the fact that each face usually has a somewhat different geometry of facial feature extremities from the template, E_{geom} seldom becomes 0 even if the extremities are properly detected. Because the distance between “sto” and the origin of the template is assumed to have the greatest influence on E_{geom} , three values are used for this distance in the template to reduce the influence. The first is the average value as described earlier, the second is $(1 +$

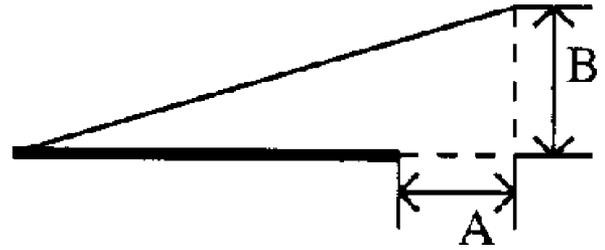


Figure 13. Dimensions used to evaluate a facial feature.



Figure 14. Original image with detected features.

DEV) times the average value, and the third is $(1 - DEV)$ times the average value, where DEV is a value to account for the deviation from the average. Faces with a longer distance between “sto” and the origin of the template will have a smaller E_{geom} using the second value and faces with a shorter distance will have a smaller E_{geom} using the third value. Here E_{geom} is calculated using two sets of transform parameters and three templates, producing six values. The smallest one is used to evaluate each candidate. The evaluation is robust against rotation about the x , y , and z axes and scale changes of the face because the algorithm takes these transforms into consideration.

Total Evaluation of Candidates. Each candidate is finally evaluated by E_{total}

$$E_{total} = C_WEIGHT \cdot E_{color} + G_WEIGHT \cdot E_{geom}, \quad (12)$$

where C_WEIGHT and G_WEIGHT are chosen so that $C_WEIGHT \cdot E_{color'} = 1$ and $G_WEIGHT \cdot E_{geom'} = 1$, where $E_{color'}$ and $E_{geom'}$ are the largest values of E_{color} and E_{geom} over the face database. Thus, C_WEIGHT and G_WEIGHT correct for the relative magnitude differences between E_{color} and E_{geom} . We note that detection performance in our experiments is not sensitive to small changes in these weights. The candidate with the smallest E_{total} is chosen to be the result of detection. The result of detection is shown in Fig. 14 with extremities of the eyes marked with “+” and extremities of the mouth marked with “x”.

Experimental Results and Discussion

The images used in the experiments were obtained using a consumer camcorder Sony CCD-TR850. The S-video signal from the camcorder was digitized by a RasterOps



Figure 15. Correctly detected facial features—1.



Figure 16. Correctly detected facial features—2.

video capture board installed on a Sun Sparc station 1+. The R, G, and B images of resolution 640×480 were acquired by the video capture board.

The images were taken outdoors in a somewhat cluttered background. Thirty-one images of three people with neutral facial expression facing different orientations were used. The three people are Indian, Caucasian, and Japanese, and the number of images of each person is 10, 10, and 11, respectively. The face orientations were arranged so that all six extremities of facial features were clearly visible.

In these experiments, facial features were detected correctly in 26 images out of 31 (83.9%) and some of them are shown in Figs. 15 and 16. The average execution time on a Sparc station 20 was about 17 s per image. For the images where facial features were not correctly detected, the failure was caused either by the fact that the eye region was not successfully separated from the eyebrow region or by the fact that the mouth region was too small because it was divided or eroded during segmentation. There was no failure of detection in images for which segmentation was successful, which demonstrates the effectiveness of the skin color model and the geometry model. In fact, in this set of experiments E_{color} was zero for the face region for 29 of the 31 images. Given the large number of regions generated during the segmentation of cluttered scenes, the face color model significantly reduces the combinatorics and improves the accuracy associated with finding the facial features. In these experiments, the facial features are only unambiguously detected in about 60% of the images without using the skin color model. This underlines the importance of including the skin color model with the geometric features. Although the overall detection rate is not quite as high as that achieved by other systems (e.g., Refs. 11 and 12), the experiments in this work consider a wider range of skin colors.

Though the geometry template proposed here is very simple, it works well with neutral facial expression. We believe that it may also work for faces with some facial expression, provided the expression does not change the position of the facial feature extremities significantly. As is clear from the algorithm, α , β and γ can be used for rough estimation of facial pose that is sometimes important in image coding of the face. By using more facial features, such as eyebrows, or using a complex method for evaluation using the geometry template, it might be possible to further improve the effectiveness of the algorithm. Due to the fact that all failures of detection were caused by inaccurate segmentation results, the rate of correct detection

will be increased significantly by improving the segmentation method.

Conclusion

In this report, two models applicable to facial image processing were proposed. The first model is for skin color which is based on human skin optics and evaluates colors using the properties of human skin. This model is valid for almost all skin colors. The second model describes the 3-D geometry of human facial features using medical statistics and evaluates the geometry of points in terms of their likelihood as human facial features. This model uses three facial features, eyes and mouth, and works for faces with any orientation and scale as long as the extremities of the three features are visible. Combining these two models with a segmentation method, a facial feature detection method was established. In experiments with images of faces in cluttered backgrounds, the features were correctly detected in 83.9% of the images.

Appendix: Color Segmentation Algorithm

Searching through the image in raster scan order, the algorithm attempts to make a new segment each time it encounters a pixel that does not yet belong to a segment. First, there is only the single pixel in the segment. Then, at each iteration, the algorithm adds pixels to the segment that surrounds a pixel which is already in the segment provided these pixels meet some requirements. Eight connected neighbors are assumed. A segment is grown until no more pixels can be included. The requirements that must be met are as follows:

- The pixel does not belong to a segment.
- Denoting the colors of two adjacent pixels by (R_1, G_1, B_1) for the pixel that belongs to the segment and (R_2, G_2, B_2) for the other, the following is met:

$$(R_1 - R_2)^2 + (G_1 - G_2)^2 + (B_1 - B_2)^2 \leq \text{sqr_th}. \quad (13)$$

- The angle between the two vectors \mathbf{V}_1 and \mathbf{V}_2 is less than or equal to ang_th , where $\mathbf{V}_1 = (R_1, G_1, B_1)$ and $\mathbf{V}_2 = (R_2, G_2, B_2)$.

Here sqr_th and ang_th are certain thresholds. The use of the third requirement makes the segmentation sensitive to a change of chromaticity.

If the number of pixels in a segment is less than a threshold area_th , the segment will be considered as a small region and will not be processed in the following stages.

This prevents small regions caused by noise from being interpreted as facial features. ▲

Acknowledgements. The authors are grateful to Drs. H. Harashima, K. Aizawa, Y. Nakaya, and A. Nagata and the people in Harashima Laboratory who helped us begin the project, to Dr. C. Novak who kindly offered an image database, and to D. Slater and A. Jain whose pictures were used in the experiments. This work has been supported in part by the Office of Naval Research under grant N00014-93-1-0540.

References

1. K. Aizawa, Y. Yamada, H. Harashima and T. Saito. Model-based synthesis image coding system—Modeling a person's face and synthesis of facial expressions, in Proceedings of the *IEEE Global Telecommunications Conference*, GLOBECOM '87, Vol. 1, IEEE, New York, 1987, pp. 45–49.
2. D. Terzopoulos and K. Waters, Analysis and Synthesis of Facial Image Sequences Using Physical and Anatomical Models, *IEEE Trans. Pattern Anal. Machine Intel.* **15**, 569–579 (1993).
3. R. Forchheimer and T. Kronander, Image coding—From waveforms to Animation, *IEEE Trans. Acoust., Speech, Signal Proc.* **37**, 2008–2023 (1989).
4. V. Bruce and M. Burton, Eds., *Processing Images of Faces*, Ablex Pub. Corp., Norwood, NJ, 1992, pp. 3–19.
5. G. Yang and T. S. Huang, Human face detection in a scene, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, New York, 1993, pp. 453–458.
6. X. Song, C. W. Lee, G. Xu, and S. Tsuji, Extraction of facial organ features using partial feature template and global constraints, *Trans. D-II of IEICE*, **J77-D-II**, 1601–1609 (1994).
7. C. L. Huang and C. W. Chen, Human facial feature extraction for face interpretation and recognition, in *Proceedings of the 11th IAPR International Conference on Pattern Recognition*, Vol. 2, IAPR, IEEE Computer Society Press, Los Alamitos, CA, 1992, pp. 204–207.
8. D. J. Beymer, Face recognition under varying pose, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, New York, 1994, pp. 756–761.
9. T. C. Chang, T. S. Huang and C. Novak, Facial feature extraction from color images, in *Proceedings of the 12th IAPR International Conference on Pattern Recognition*, 2, IAPR, IEEE Computer Society Press, Los Alamitos, CA, 1994, pp. 39–43.
10. Q. Chen, H. Wu, and M. Yachida, Face detection by fuzzy pattern matching, in *Proceedings of the IEEE International Conference on Computer Vision*, IEEE, New York, 1995, pp. 591–596.
11. D. Sanger, H. Haneishi and Y. Miyake, Method for light source discrimination and facial pattern detection from negative color film, *J. Imaging Sci. Technol.* **39**, 166–175 (1995).
12. D. Sanger, Y. Miyake, H. Haneishi, and N. Tsumura, Algorithm for face extraction based on lip detection, *J. Imaging Sci. Technol.* **41**, 71–80 (1997).
13. T. Kohonen, P. Lehtio, J. Rovamo, J. Hyvarine, K. Bry, and L. Vainio, A principle of neural associative memory, *Neuroscience*, **2**, 1065–1076 (1977).
14. T. J. Stonham, Practical face recognition and verification with WISCAD, in *Aspects of Face Processing*, H. D. Ellis, M. A. Jeeves, F. Newcombe, and A. Young, Eds., Nijhoff, Dordrecht, Boston, 1986, pp. 426–441.
15. T. Sakaguchi, O. Nakamura and T. Minami, Personal identification through image using isodensity lines, in *Proc. SPIE* **1199**, 643–654 (1989).
16. S. Shafer, Using color to separate reflection components, *Color Res. Appl.* **10**, 210–218 (1985).
17. T. B. Fitzpatrick, M. M. Wick and K. Toda, Eds., *Brown Melanoderma*, University of Tokyo Press, Tokyo, 1986, pp. 3–7.
18. R. Marks and P. A. Payne, Eds., *Bioengineering and the Skin*, MTP Press, Boston, MA, 1981, pp. 253–265.
19. R. R. Anderson and J. A. Parrish, The optics of human skin, *J. Invest. Dermatol.* **77**, 13–19 (1981).
20. J. A. Parrish, M. L. Kripke and W. L. Morison, Eds., *Photoimmunology*, Plenum Medical Book Co., New York, 1983, pp. 61–76.
21. S. Wan, R. R. Anderson and J. A. Parrish, Analytical modeling for the optical properties of the skin with in vitro and in vivo applications, *Photochem. Photobiol.* **34**, 493–499 (1981).
22. G. Wyszecki and W. S. Stiles, *Color Science—Concepts and Methods, Quantitative Data and Formulas*, John Wiley & Sons, New York, 1967.
23. D. Judd, D. MacAdam and G. Wyszecki, Spectral distribution of typical daylight as a function of correlated color temperature, *J. Opt. Soc. Am.* **54**, 1031–1040 (1964).
24. H. Kang, *Color Technology for Electronic Imaging Devices*, SPIE Press, Bellingham, WA (1997).
25. L. G. Farkas, Ed., *Anthropometry of the Head and Face*, Raven Press, New York, 1994.