Journal of Applied Science and Engineering, Vol. 16, No. 3, pp. 241–248 (2013) DOI: 10.6180/jase.2013.16.3.03

Real-Time Patients' Face Tracking Based on Facial Feature Matching

Hsin-Hung Chiang[1], Wei-Ming Chen[2]*, Chiou-Shan Chou[2] and Han-Chieh Chao[1,2]

[1]Department of Electrical Engineering, National Dong Hwa University, Hualien, Taiwan 974, R.O.C.
[2]Institute of Computer Science and Information Engineering, National Ilan University, Yilan, Taiwan 260, R.O.C.

Abstract

This paper will discuss how to track the faces of patients who only lie in bed with real-time manner. We propose a real-time mechanism to extract and track patients' facial features effectively. When camera facing the patient's face, we will use our DRK (Dynamic Radial Kernel) mechanism to record and match facial features. Then we will collect the information of the patient's face. It could be provided to doctors for diagnostic aiding. For patients, this mechanism can bring a track record as well as warning function. We hope that this mechanism can provide valuable features of the health-care application.

Key Words: Face Tracking, Feature Extracting, Feature Matching, Health Care

1. Introduction

In the health care, the monitoring of patient's health condition is very important. Especially the patients who can't move and only lie in the bed need it urgent. We only rely on their facial expression to determine how is going on. Because of this, we will describe how to get the patient's facial feature in real-time.

In previous research [1], we can know the colors are easily influenced by illumination [2]. We should overcome this situation to do the most detailed analysis of the image. There have been several studies about color space conversion [3–5]. In this study, we choose the YCbCr color space [6–10]. The reason is because the YCbCr uses the component Y to storage the brightness value. In addition to this, the conversion between RGB and YCbCr is not too complex. So we will not cost much time to calculate the impact of the brightness. After the processing of the color space, we should know the face's position before tracking it. In the beginning, we use the camera facing the patient's face, so what we should do is extracting the patient's facial features for matching or recording. We will modify the mechanism of previous studies to flag the facial features.

In this paper, we propose an approach to record data of patients' facial features and match another image to find the most similar region for tracking. As the previous, we set a goal to implement a system which uses our algorithm and satisfies the following requirements:

- Working in real time
- Get the features automatically
- Rotation resistance when matching
- Scaling resistance when matching
- Parallax resistance when matching
- Analysis patient's movement

2. Facial Features Recording

2.1 Features Definition

In [1], we define seven feature points which include four canthi (F1~F4 in the Figure 1), two corners of mouth (F6~F7 in the Figure 1) and the center of nose (F5

*Corresponding author. E-mail: wmchen@niu.edu.tw

in the Figure 1). In this study, our focus will be placed on the patient's face. Because they can't move and only lie in the bed need it urgent. Even some need to wear an oxygen mask. We will let the three features be additional references, include the two corners of mouth and the center of nose. When a patient wears the oxygen mask, the camera will loss them. That's mean medical person might flag the patient's features by manual. Then our mechanism will automatically do the follow-up action.

2.2 Features Extracting

At this stage, we should get the features' position before recording them. Here we use the proposed Facial Feature Filter and the Minagawa's method [11] to extract features.

Our Facial Feature Filter is divided into two steps. Firstly, we use the edited one dimension horizontal Gaussian filter to remove image noise. Because the operation of original Gaussian filter is too complex, we decrease its dimension of kernel from 2D to 1D for reducing the computing time. In the Figure 2, the left image is the original testing image and the right image is the result of testing image after 1D horizontal Gaussian Filter. It not only reduces computing time, but also emphasizes the horizontal textures because its horizontal kernel blurs the vertical texture effectively. Then we can distinguish between horizontal textures and other direction.

The second step is using a one dimension vertical local threshold. It helps to keep the region we are interested, and remove which we don't need. In general, a global threshold value may be not the most optimal value for each sub region. Local threshold could be used to binarize a facial images influenced by illumination or shadow. But the operation of local threshold is too heavy, it is 2D computing. Then we decrease the dimension of its kernel like the previous section. The result is well as non-edited, because the previous Gaussian filter already emphasizes the horizontal texture. So we can only use the 1D vertical kernel of local threshold to keep horizontal texture regions quickly. The results of images after facial feature filter are shown in the following Figure 3.

Then we use the Minagawa's projection method on the processed image to locate the features positions (Figure 4).

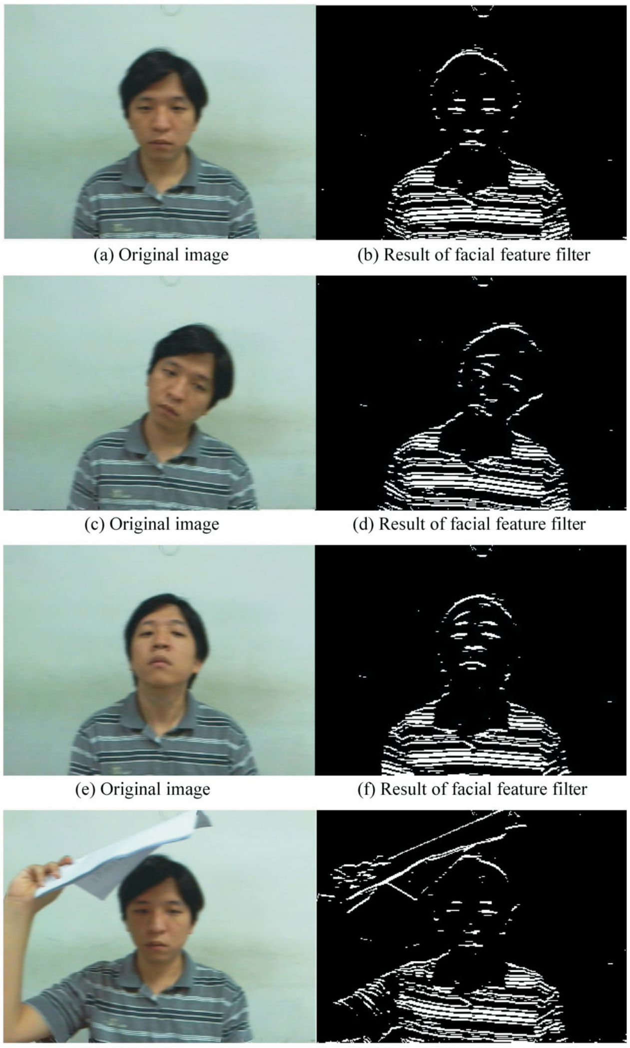
Figure 1. Seven feature points (F1~F7) we used.

2.3 Feature Recording

Most of approaches for facial feature detecting assume that patient is facing the camera directly. It might be the warning case when he turns or nods his face slightly. The features might disappear. Therefore, we use the proposed novel feature recording method, Dynamic Radial Kernel (DRK), to solve it. After getting a feature's coordinate, we will record pixels near the feature point by DRK methods.

During the experiment, we find some rules of DRK. There are totally N radiations with length R in a DRK (Figure 5 (a)). Each radiation will record R/D pixels (Figure 5 (b)). The distribution of radiations will change with different conditions, and the recorded pixels are used to store feature information. But the amount of recorded pixels is limited. It will depend on the location of the recorded pixels. If it is in significant area, the feature information will be rich and vice versa. We will use dif-

Figure 2. Original image (left) and after 1D Gaussian filter (right).



(a) Original image (b) Result of facial feature filter

(c) Original image (d) Result of facial feature filter

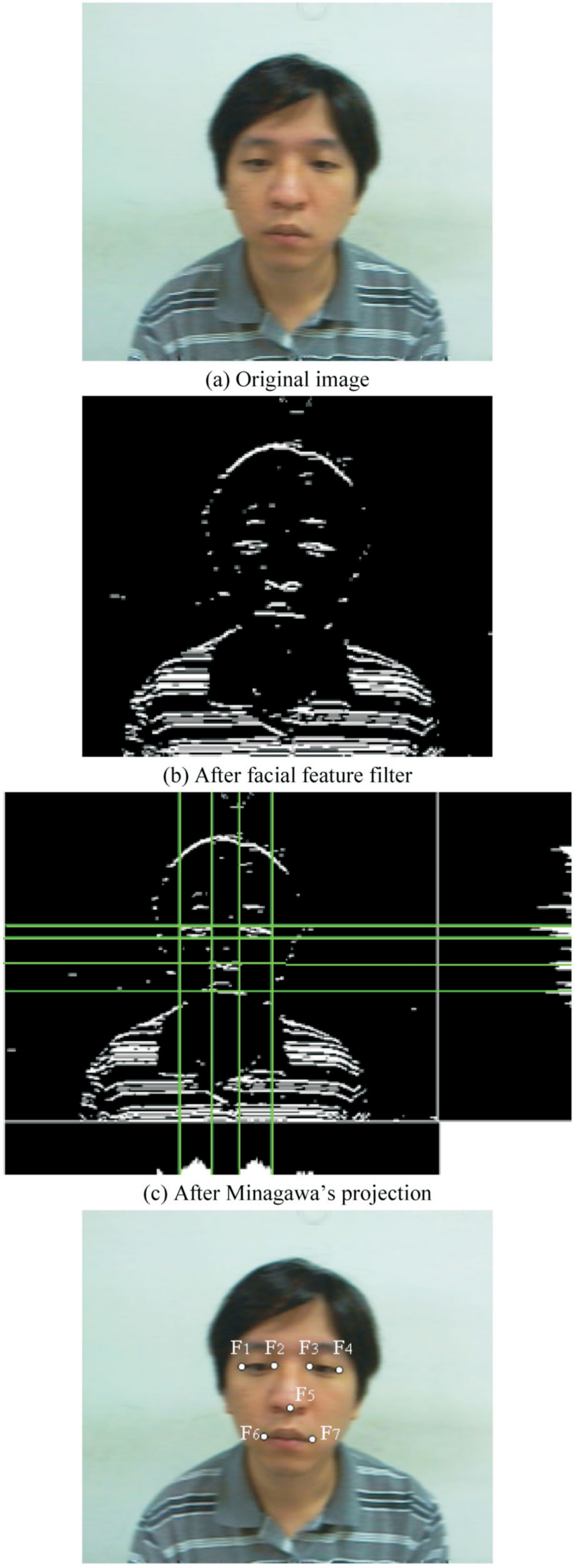
(e) Original image (f) Result of facial feature filter

(g) Original image (h) Result of facial feature filter

Figure 3. Original image (left) after facial feature filter (right).

ferent radiation distribution on different feature points to ensure the recorded pixels will be meaningful. The three kinds of condition are shown in the Table 1.

Each pixel in DRK will be recorded its color information. We mainly use the RGB's 3 color components to describe a color. However, RGB could be influenced by environmental illumination easily. We use an additional color space, YCbCr, to improve the robustness. Sometimes, because of the illumination around the patient, there will be some shadow on his face, and it may bring errors when feature matching. YCbCr has good ability of correcting error derived from shadows. But the different brightness value will only affect Y value, so we just record the Cb and Cr value to prevent error.



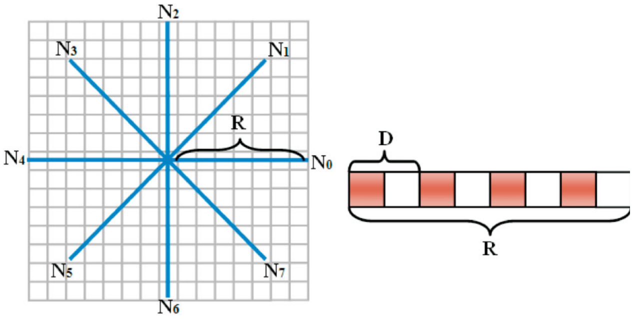
(a) Original image

(b) After facial feature filter

(c) After Minagawa's projection

(d) Extraction result

Figure 4. Extract seven facial feature points we need.



(a) A whole DRK. (b) One radiation of DRK.

Figure 5. Structure of DRK.

Table 1. 3 kinds of DRK (The location number is reference from Figure 4)

Distribution mode	R	D	Degree	Location
	14	2	45°	F5
	Left: 4 Right: 14	2	Left: 90° Right: 30°	F1, F3, F6
	Left: 14 Right: 4	2	Left: 30° Right: 90°	F2, F4, F7

Besides the color information, we also need to record distance between each of the feature points. It would help system to know the face's scaling ratio value, and the positional relationship between the each two feature points of them.

3. Facial Features Matching

Once DRK records nearing pixels' color information as we mentioned above, it has ability to match up other camera frames. The pseudo code shown in Figure 6 summarizes the matching algorithm. In the Matching function, we use two parameters to adjust the impact weights. W1 is the error value weight of the YCbCr, and W2 is another one of the RGB. In static cases, we set the values of them are 50%. But in dynamic cases, they will be set different value. This is because when the patients moving, the influence of light will large then stationary. In this case, W1 will more important than W2.

In the Find_Feature procedure (Figure 7), Img is the processing frame, and its pixels are described by YCbCr color space. After extracting a facial feature point, its coordinates will be recorded. The x and y will be stored in

```
function Matching(x, y)
  error←0
    for n←0 to DRK.N-1
      for p←0 to DRK.R / DRK.D
        errR←Abs( DRK.color_info[n, p].R –
                    Img[x+ posx(n, p), y+ posy(n, p)].R )
        errG←Abs( DRK.color_info[n, p].G -
                    Img[x+ posx(n, p), y+ posy(n, p)].G )
        errB←Abs( DRK.color_info [n, p].B -
                    Img[x+ posx(n, p), y+ posy(n, p)].B )
        errorYCbCr←errorYCbCr+ errR+ errG+ errB

        errCb←Abs( DRK.color_info[n, p].Cb -
                    Img[x+ posx(n, p), y+ posy(n, p)].Cb )
        errCr←Abs( DRK.color_info [n, p].Cr -
                    Img[x+ posx(n, p), y+ posy(n, p)].Cr )
        errorRGB←errorRGB+ errCb+ errCr
    end for
  end for
  return errorYCbCr * W1 + errorRGB * W2
end function
```

Figure 6. Pseudo code of DRK feature matching.

variance parameter pre_x and pre_y. Then, we define a detecting window according to them. The variance z is the scaling ratio of face. It will impact size of the detecting window. When the z value increases, the window will get bigger, and vice versa. All pixels in this window will be processed using the Matching function. The Matching function computes the difference between each pixel and its facial feature's color information recorded in DRK.

In the Matching function, we compare the original color information and the processing pixel's nearing pixel. The difference will be computed in variance error, and returned to Find_Feature procedure.

```
procedure Find_Feature(pre_x, pre_y)
  min_error←∞
  for i← (Window_h-1)*z / -2 to (Window_h-1)*z / 2
    for j←(Window_w-1)*z / -2 to (Window_w-1)*z / 2
      this_error←Matching(pre_x+i, pre_y+j)
      if this_error < min_error then
        min_error←this_error
        new_x←i
        new_y←j
      end if
    end for
  end for
  pre_x←new_x
  pre_y←new_y
end procedure
```

Figure 7. Pseudo code of DRK find feature.

Therefore, the key point when matching is to know which pixel in the Img needs to be compared. In this case, we propose two functions and use the value of posx and posy to get the compared pixel's coordinate. In functions, the parameters n represents the n-th radiation of DRK, and p represents the p-th recorded pixel in the radiation. Then we can use the equation (1) and (2) to get the compared pixel's coordinate.

$$posx(n,p) = \cos\left(\frac{360^\circ}{N} \times n + t^\circ\right) \times D \times p \times z \tag{1}$$

$$posy(n,p) = \sin\left(\frac{360^\circ}{N} \times n + t^\circ\right) \times D \times p \times z \tag{2}$$

Variance *N* and *D* have been explained in the Figure 8. Because the DRK is dynamic, variance z determines its size. When *z* increases, the computed result is much bigger, and vice versa. Variance *t* is the degree of face tilting. If system detects the user's face is rotating, the compared region can change as well.

Every pixel of the detecting window computes its error value by DRK. After all pixels are processed, the point with the smallest error value is the new feature which we determined. The Find_Feature procedure will repeat six times to get six feature points and replace the original feature points.

4. Movement Analysing

Video is a series of continuous images. So their contents are also continuous and there is the relationship of them. Then we can simply analyzing user's movement. We define four kinds of movements, and they will be explained in the following paragraph.

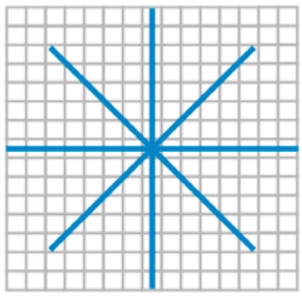
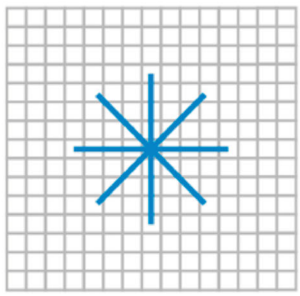
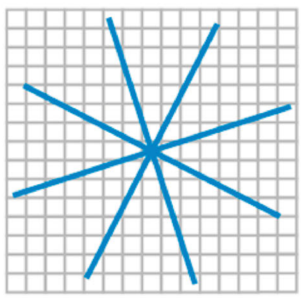
DRK region	t	z
	0	1
	0	0.5
	20	1

Figure 8. Dynamic region.

4.1 Face Closing or Leaving

When the user closes to the camera, his face would get bigger, and vice versa. Therefore, if the distances of each feature points become bigger or smaller at the continuous detecting and the same ratio, we could define the status which is closing or leaving. In our mechanism, we use the two outer canthi to compute the face's scaling ratio value (Figure 9). First, it records the distance between outer canthi as a reference distance D1 after extracting the facial features. In the tracking process, it records the distance between tracked outer canthi as current distance D2. Then we get the scaling ratio z, D2/D1.

4.2 Head Tilting

In general case, the four canthi on human face are arranged in a straight line. The slope of this line defines the degree of tilting, and it also determines the variance t (Figure 10).

4.3 Head Shaking

Assume that human head is a sphere, the facial features are points on the surface of this sphere. According to sine and cosine function, we understand a phenomenon that the ratio of length of left eye to length of right eye changes with the rotation angle of user's head. Figure 11 is a top view, and it characterizes this movement.

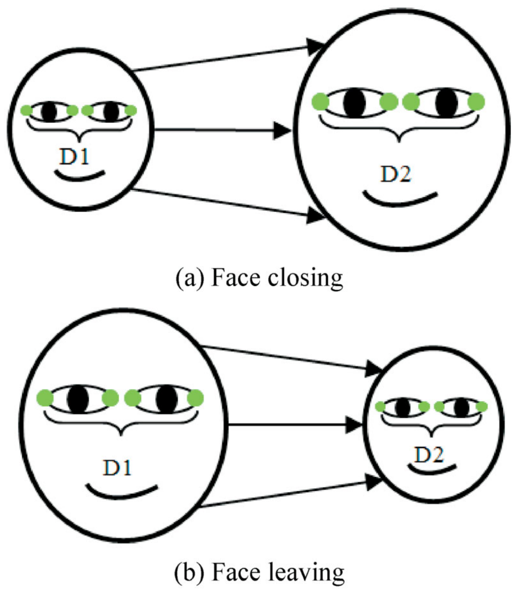


Figure 9. The scaling rate is simply computed from D2/D1. The green points represent the tracked canthis.

4.4 Up, Down-Bending (Optional)

Human facial features arranged in horizontal lines. In the previous movement analyzing, we can calculate it with single feature, but up, down-bending is a vertical movement, so we need to use more than one feature to analyze it. Here we use a vertical line from eyes to nose and mouth. As the head shaking, the analyzing is calculated by two length value of eye to nose and nose to mouth shown is the Figure 12. When user face bends up or down, the ratio to these two length value will change, and then we can predict the bending angle of it. But if the patient wears the oxygen mask, this method will not have any result.

5. Experimental Result

Our prototype system uses a stand-alone PC with 1.96 GB DDR2 RAM and Intel 2.66 GHz CPU. The captured device is a Microsoft Webcam, its captured frame size format is 320×240, and frame rate is 30 frames per second. Although this resolution is not high, it is high enough to tracking facial features in mostly situations. However, there are still some restricts of DRK matching.

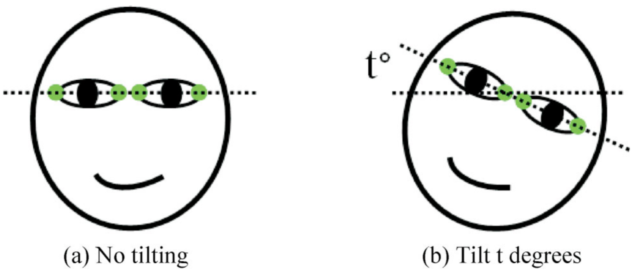


Figure 10. The slope of canthal line determines the tilting angle. The four canthi are marked as green points.

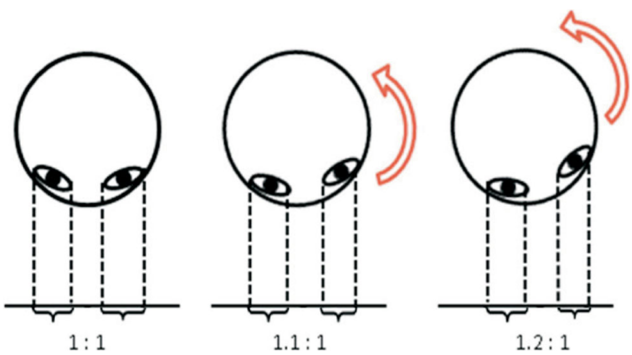


Figure 11. The ratio changes when user's head shakes.

If the degree of rotation is too high, DRK can't match features accurately.

In the results, we have successfully demonstrated that the system can track user's facial features under various movements and environments. We try to let the head be in the tilting, closing, leaving or shaking. Figure 13 characterizes the experimental results, and we use the green pint to locate the features in the face.

6. Conclusion

In this paper, we implement a real-time system which can record human facial features, and search the most similar region in another frame using our DRK. Most of the face tracking approaches can only track face when user is facing to camera directly. To overcome this problem, DRK provides a flexible structure to optimize matching region in each frame. Through analyzing user's movement, we can compute what is the most suitable structure of DRK in a specific frame. It increases the accuracy of our system.

In the health care, we can use this mechanism to monitor the faces of patients who could just lie in the bed. We also hope that mechanism will do a favor for them. Through this mechanism, we can carefully observe the expressions and move the state of these patients. The results would help the doctor for a follow-up medical action.

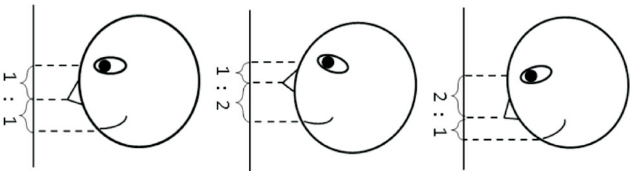


Figure 12. The ratio changes when user's head bending.

(a) Basic type

(b) Tilting

(c) Closing

(d) Leaving

(e) Head Shaking

Figure 13. Experimental results. The results of matching marked in green point.

References

[1] Chen, W. M., Chou, C. S., Lin, Y. L. and Yeh, C. H., "Real-Time Face Tracking Based on Facial Feature Matching," *in Proc. of APSIPA Annual Summit and Conf.*, Special Session on Image/Video Processing and Analysis, Xi'an, China, Oct. (2011)

[2] Maxwell, B. A., Friedhoff, R. M. and Smith, C. A., "A Bi-Illuminant Dichromatic Reflection Model for Understanding Images," *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8 (2008). doi: 10.1109/CVPR.2008.4587491

[3] Soriano, M., Martinkauppi, B., Huovinen, S. and Laaksonen, M., "Skin Color Modeling under Varying Illumination Conditions Using the Skin Locus for Selecting Training Pixels," *Proc. IEEE Nordic Signal Processing Symposium*, Kolmården, Sweden (2000).

[4] Soriano, M., Martinkauppi, B., Huovinen, S. and Laaksonen, M., "Skin Detection in Video under Changing Illumination Conditions," *Proc. 15th International Conference on Pattern Recognition*, Barcelona, Spain (2000). doi: 10.1109/ICPR.2000.905542

[5] Gonzalez, R. C. and Woods, R. E., *Digital Image Processing*, 2nd ed. Prentice Hall Press (2002).

[6] Garcia, C. and Tziritas, G., "Face Detection Using Quantized Skin Color Regions Merging and Wavelet Packet Analysis," *Proc. IEEE Trans. On Multimedia*, Vol. 1, No. 3, pp. 264–277 (1999). doi: 10.1109/6046.784465

[7] Chen, A. P., Pan, L. A., Tong, Y. B. and Ning, N.,

"Face Detection Technology Based on Skin Color," *Proc. Second International Workshop on Education Technology and Computer Science*, Vol. 2, pp. 708–711 (2010). doi: 10.1109/ETCS.2010.465

[8] Kryszczuk, K. M. and Drygajło, A., "Color Correction for Face Detection Based on Human Visual Perception Metaphor," *Proc. of the Workshop on Multimodal User Authentication*, pp. 138–143 (2003).

[9] Lin, Y. P., Chao, Y. P., Lin, C. C. and Chen, J. H., "Webcam Mouse Using Face and Eye Tracking in Various Illumination Environments," *Proc. 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, Vol. 1–4 (2005). doi: 10.1109/IEMBS.2005.1617296

[10] Soriano, M., Martinkauppi, B., Huovinen, S. and Laaksonen, M., "Using the Skin Locus to Cope with Changing Illumination Conditions in Color-Based Face Tracking," *Proc. IEEE Nordic Signal Processing Symposium*, Vol. 13–15, pp. 383–386 (2000).

[11] Minagawa, T., Saito, H. and Ozawa, S., "Face-Direction Estimating System Using Stereo Vision," *Proc. International Conference on Industrial Electronics, Control and Instrumentation*, pp. 1454–1459 (1997). doi: 10.1109/IECON.1997.668533

Manuscript Received: Feb. 16, 2013
Accepted: Jun. 28, 2013