8th IPPR Conference on
Computer Vision, Graphics and Image Processing
Tao-Yuan, Taiwan, R.O.C.
Aug. 20~23, 1995

Improving Generalized Hough Transform for Color Image Detection and Matching

Rong-Chin Lo (駱 榮 欽)* and Wen-Hsiang Tsai (蔡 文 祥)

Department of Computer and Information Science
National Chiao Tung University
1001 Ta Hsueh Rd., Hsinchu, 300

Abstract

A new approach to improving the generalized Hough transform (GHT) for color image detection and matching is proposed. In the conventional GHT, each preprocessed pixel is regarded as a feature and the relationship of the pixel neighborhood is not used. Useful color information and lighting changes in the image usually are not considered. Furthermore, boundary detection and shape thinning, which are required before applying the GHT to an image, are not easy to perform on color images so that the conventional GHT is seldom applied to color images. The proposed approach tries to improve these drawbacks. First, it removes lighting changes by using normalized color values. Next, it extracts certain critical pixels of the input color image whose neighborhoods have larger variances of color values. A feature vector, which includes the normalized color values of the pixel as well as those of the pixel's neighbors, is then used to represent each critical pixel. A voting rule for the GHT is proposed accordingly, which is based on a similarity measure function of feature vectors. Each vote for a cell increments the cell value by the fractional value of the similarity measure instead of the fixed value of 1. High maximum peaks in the cell array are searched finally as the result. The proposed method is robust for color image detection and matching in noisy, occlusive, and lighting change environments, as demonstrated by experimental results.

Key Words: Improving generalized Hough transform, Color image, Image detection and matching, Lighting changes, Critical pixel, Feature vector, Voting rule, Color reference table.

1. Introduction

A generalized version of the Hough transform, called the generalized Hough transform (GHT), was proposed by Ballard[1] for detecting arbitrary 2-dimensional (2D) shapes.It has been used successfully in many applications of image processing and computer vision

The GHT has received a lot of improvements[3-6].

In this paper, an approach to improving the conventional GHT for color image matching is proposed. One shortage of the conventional GHT is that it regards each preprocessed binary pixel as a feature from the input image, and the useful color information existing in the input image and the relationship between each pixel and its neighbors are not utilized. Furthermore, the preprocessing step of boundary detection or shape thinning is not easy to perform on color images so that the conventional GHT is seldom applied to color images. Besides, lighting changes in images are seldom considered. It is tried to remove these shortages of the conventional GHT in this study.

First, the color values (RGB values) of each image pixel are *normalized*. Next, certain *critical pixels* in the input color image are extracted. A pixel is said to be critical in this study if the variances of both the original color values and the normalized color values in the neighborhood of the pixel are sufficiently large.

A feature vector then is constructed for each critical pixel. The feature vector includes the normalized color values of the pixel as well as those of its neighbors in the image. A voting rule for the GHT is proposed accordingly, which is based on a similarity measure function of the feature vector. Each vote for a cell increments the cell value by the fractional value of the similarity measure instead of the fixed value of 1. The location of the cell with the maximum value exceeding a preselected threshold is searched finally as the result of image detection and matching. Advantages of the proposed approach include at least the following.

(1) The normalization of image color values reduces the influence of lighting changes in environments to input images and matching results.

(2) The extraction of critical pixels transforms the input color image into a point-type image to which the GHT becomes applicable.

*Is also with the Department of Electronic Engineering, National Taipei Institute of Technology.

(3) The use of critical pixels help avoiding the preprocessing step of boundary detection or shape thinning, which usually is necessary in applying the GHT to binary images but is difficult to perform on color images.

(4) Inclusion of the color values of each critical pixel and its neighbors in the feature vector makes use of not only the color information in the images but also the relative position relationship among the pixel neighborhood.

(5) Because of the use of the feature vectors of the proposed critical pixels as well as the use of the corresponding voting rule for fractional cell value incrementation, the resulting peaks in the HCS become more prominent and easier to detect.

As a result of the above advantages, the proposed method is robust for color image detection and matching in noisy, occlusive, and lighting change environments, as shown by the experimental results.

The remainder of this paper is organized as follows. In Section 2, the GHT is first reviewed. In Sections 3, the proposed approach to improving the GHT for color image detection and matching is described in detail. Experimental results are given in Section 4. Some concluding remarks are given in Section 5.

2. Review of Conventional Generalized Hough Transform

To use the conventional GHT to detect an arbitrary template shape, it is necessary to set up a 2D HCS (X_t, Y_t) where (X_t, Y_t) is a translation vector with respect to a reference point of the template shape and describes the location of the template shape. Each cell in the HCS has a value specifying the possibility that the reference point of the template shape to be detected is located at the cell.

Before performing the GHT, a *reference table* (R-table)[1] for the template shape is built up by the following steps: (1) select a suitable point R in the given template shape as the reference point; (2) rotate the shape 180^o with respect to R; (3) trace all the boundary or skeletal pixels of the template shape and construct an R-table consisting of the displacement vectors between all the boundary or skeletal pixels and R.

In the GHT process, all the displacement vectors of the R-table are superimposed on each pixel in the input image. The value of each cell pointed to by a displacement vector is incremented by one. If there exists any cell with its value exceeding a preselected threshold value and being the maximum in HCS, then it is determined that the template shape is detected at the location of the cell.

3. Proposed Approach to Improving Generalized Hough Transform for Color Image Detection and Matching

The proposed method extracts critical pixels from an input image and utilizes the normalized color information of each critical pixel and its neighbors to perform the shape detection and matching. By normalizing the color values of the image, the unknown but common incident lighting illumination and the sensor responsivity and gain can be removed[7-10]. Then, each critical pixel is represented by a feature vector that includes the normalized color values of the pixel itself as well as its neighboring pixels. To perform the GHT process, a *color reference table* (abbreviated as CR-table) and a corresponding voting rule are proposed. The details are described in the following.

A. Removal of lighting changes

In an RGB color image, the color values of each pixel at position (j,k), $I_i(j,k)$, where i = $R(red)$, $G(green)$, $B(blue)$, can be approximately modeled by the product of an object reflectivity function $F_i(j,k)$, an illumination function $\Phi(j,k)$, and a sensor responsivity and gain function $S(j,k)$:

$$I_R(j,k) = F_R(j,k)\Phi(j,k)S(j,k),$$
$$I_G(j,k) = F_G(j,k)\Phi(j,k)S(j,k),$$
$$I_B(j,k) = F_B(j,k)\Phi(j,k)S(j,k).$$

Note that for each pixel, both of $\Phi(j,k)$ and $S(j,k)$ are unknown but identical for all three color bands [7-10]. In order to remove the influence of $\Phi(j,k)$ and $S(j,k)$, we use the normalized color values $(r(j,k), g(j,k), b(j,k))$ defined as follows:

$$r(j,k) = \frac{255\, I_R(j,k)}{I_R(j,k)+I_G(j,k)+I_B(j,k)} = \frac{255\, F_R(j,k)}{F_R(j,k)+F_G(j,k)+F_B(j,k)}$$
$$g(j,k) = \frac{255\, I_G(j,k)}{I_R(j,k)+I_G(j,k)+I_B(j,k)} = \frac{255\, F_G(j,k)}{F_R(j,k)+F_G(j,k)+F_B(j,k)}$$
$$b(j,k) = \frac{255\, I_B(j,k)}{I_R(j,k)+I_G(j,k)+I_B(j,k)} = \frac{255\, F_B(j,k)}{F_R(j,k)+F_G(j,k)+F_B(j,k)} \quad (1)$$

From Eqs. (1), we see that the normalized image $(r(j,k), g(j,k), b(j,k))$ (abbreviated as (r, g, b) henceforth) only includes the information of the surface reflective properties of the imaged object and is not susceptible to environmental lighting changes. Therefore, it is suitable to match object images using normalized color values to avoid the influence of lighting changes, as is done in this study.

B. Extracting critical pixels from input color image

Due to the limitation of the responsivity and gain of the human visual system or sensors, an image does not contain meaningful color information at pixels with very low or very high intensity levels of the incident light[11]. Also, regions formed by such pixels usually have smaller color variance values. Thus, to extract meaningful critical pixels, we check first the variance U of the R, G, B values of the neighborhood of each pixel in the template shape or input image: if U is larger than a threshold value, then the pixel is labeled as a candidate critical pixel. Next, in order to remove the influence of lighting changes, we check further the the variance V of the r, g, b values in the neighborhood of each candidate pixel: if V is larger than another threshold value, then the pixel is decided to be critical and is extracted.

More specifically, let P_i be a pixel in the input image and P_{ij} with j = 1, 2, ..., 8 be its 8-neighbors, as shown in Fig. 1(a). Let (R_i, G_i, B_i) and (R_{ij}, G_{ij}, B_{ij}) be the RGB values of P_i and P_{ij}, respectively, as shown in Fig. 1(b), and let the corresponding normalized RGB values be (r_i, g_i, b_i) and (r_{ij}, g_{ij}, b_{ij}), respectively, as shown in Fig. 1(c). The critical pixels are extracted by the following steps: (1) for each pixel P_i in the input image, compute the variance U_i of the RGB values between P_i and its 8-neighbors, i.e., compute $U_i = \frac{1}{8}\sum_{j=1}^{8}\left(\left|R_{ij} - R_i\right| + \left|G_{ij} - G_i\right| + \left|B_{ij} - B_i\right|\right)$; (2) if U_i is greater than a threshold value, then pixel P_i is labeled as a candidate critical pixel; (3) for each candidate pixel P_i in the input image, compute the variance V_i of the normalized color values between P_i and its 8-neighbors, i.e., compute $V_i = \sum_{j=1}^{8}\left(\left|r_{ij} - r_i\right| + \left|g_{ij} - g_i\right| + \left|b_{ij} - b_i\right|\right)$; (4) if V_i is greater than another threshold value, then pixel P_i is labeled as a critical pixel.

C. Color reference table

For each color template shape, we have to constitute a CR-table. The CR-table not only records the displacement vector between each extracted critical pixel P_i and the reference point but also preserves the normalized color values of both P_i and its neighboring pixels, which form a feature vector. The feature vector f_i of pixel P_i includes the normalized color values of P_i and all of its 8-neighbors P_{ij} (j = 1, 2,..., 8), i.e.,

$$
\begin{aligned}
f_i &= \left(f_{i0}, f_{inr}, f_{ing}, f_{inb}\right) \\
&= ((r_{i0}, g_{i0}, b_{i0}), (r_{i1}, r_{i2}, r_{i3}, r_{i4}, r_{i5}, r_{i6}, r_{i7}, r_{i8}), \\
&\quad (g_{i1}, g_{i2}, g_{i3}, g_{i4}, g_{i5}, g_{i6}, g_{i7}, g_{i8}), \\
&\quad (b_{i1}, b_{i2}, b_{i3}, b_{i4}, b_{i5}, b_{i6}, b_{i7}, b_{i8}))
\end{aligned}
$$

where (r_{i0}, g_{i0}, b_{i0}) is the normalized color values of P_i and $(f_{inr}, f_{ing}, f_{inb})$ includes the normalized r, g, b values of the 8-neighbors. All extracted critical pixels together with its feature vectors are traced to constitute the CR-table by the following steps: (1) select a suitable point R in the given template shape as the reference point; (2) rotate the shape 180° with respect to R; (3) trace all the critical pixels of the template shape and construct the CR-table by including the displacement vector between each critical pixel and R, as well as the feature vector of the pixel. The steps are illustrated by Figs. 2(a) through 2(d).

D. Voting rule

The similarity measure $S(P_i, P_j)$ between two critical pixels, P_i and P_j, is first defined as:

$$S(P_i, P_j) = S_1 * S_2$$

$$= \left(\frac{1}{|f_{i0} - f_{j0}| + 1} \right) *$$

$$\left(\frac{1}{w \circ \left(|f_{inr} - f_{jnr}| + |f_{ing} - f_{jng}| + |f_{inb} - f_{jnb}| \right) + 1} \right);$$

with

S_1 = the similarity measure between the two points P_i and P_j themselves

$$= \left(\frac{1}{|f_{i0} - f_{j0}| + 1} \right)$$

$$= \frac{1}{|r_{i0} - r_{j0}| + |g_{i0} - g_{j0}| + |b_{i0} - b_{j0}| + 1};$$

S_2 = the similarity measure between the two sets of 8-neighbors of two points

$$= \left(\frac{1}{w \circ \left(|f_{inr} - f_{jnr}| + |f_{ing} - f_{jng}| + |f_{inb} - f_{jnb}| \right) + 1} \right) =$$

$$\frac{1}{\left(\sum_{k=1}^{8} w_k |r_{ik} - r_{jk}| \right) + \left(+ \sum_{k=1}^{8} w_k |g_{ik} - g_{jk}| \right) + \left(\sum_{k=1}^{8} w_k |b_{ik} - b_{jk}| \right) + 1} \tag{2}$$

where

(1) $\circ$ is the operator of vector product;

(2) $w = (w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8)$ is a weighting vector computed by: $w_k = 1/n$ if neighboring pixel P_k (k=1, 2, ..., 8) exists; = 0 otherwise, where n is the total number of the neighboring pixels (note that the number of the 8-neighbors of a pixel on the image boundary is not necessarily 8);

(3) $f_{i0} = (r_{i0}, g_{i0}, b_{i0})$ is the normalized color values of P_i;

(4) $f_{j0} = (r_{j0}, g_{j0}, b_{j0})$ is the normalized color values of P_j;

(5) $f_{inr} = (r_{i1}, r_{i2}, r_{i3}, r_{i4}, r_{i5}, r_{i6}, r_{i7}, r_{i8})$ is the neighboring red information of pixel P_i;

(6) $f_{jnr} = (r_{j1}, r_{j2}, r_{j3}, r_{j4}, r_{j5}, r_{j6}, r_{j7}, r_{j8})$ is the neighboring red information of pixel P_j;

(7) $f_{ing} = (g_{i1}, g_{i2}, g_{i3}, g_{i4}, g_{i5}, g_{i6}, g_{i7}, g_{i8})$ is the neighboring green information of pixel P_i;

(8) $f_{jng} = (g_{j1}, g_{j2}, g_{j3}, g_{j4}, g_{j5}, g_{j6}, g_{j7}, g_{j8})$ is the neighboring green information of pixel P_j;

(9) $f_{inb} = (b_{i1}, b_{i2}, b_{i3}, b_{i4}, b_{i5}, b_{i6}, b_{i7}, b_{i8})$ is the neighboring blue information of pixel P_i; and

(10) $f_{jnb} = (b_{j1}, b_{j2}, b_{j3}, b_{j4}, b_{j5}, b_{j6}, b_{j7}, b_{j8})$ is the neighboring blue information of pixel P_j.

From the definition of $S(P_i, P_j)$, we see that when the feature vectors, f_i and f_j, of two pixels are identical, $S(P_i, P_j)=1$; on the contrary, the larger the difference between the feature vectors of the two pixels is, the smaller the value of $S(P_i, P_j)$ becomes. The proposed voting rule for color images is just to use the value of $S(P_i, P_j)$, which is fractional, as the increment value instead of the value of 1 in the cell value incrementation stage of the GHT. Note that due to consideration of the influence of the pixel itself as well as the neighboring pixels, the measure $S(P_i, P_j)$ includes two parts S_1 and S_2.

E. Improved GHT for detecting and matching color object shapes

In the following, we describe the algorithm of the proposed improved GHT. The algorithm performs shape matching by detecting a given color template shape in an input color image.

Algorithm 1. Improved GHT for color image matching

Input.

1. A given color template shape and an input color image containing a translated object shape partially or fully identical to the given template shape, taken under lighting change environments.
2. A threshold value t_v.

Output.

A location in the input image where the template shape appears (or more specifically, where the reference point of the template shape is located).

Steps.

1. *Removing lighting changes:* compute the normalized color images of the template shape and the input image by Eq. (1).
2. *Initialization*: form a 2D HCS $H(X_t, Y_t)$ and set all the values of the cells in H to zero.
3. *Extracting critical pixels and feature vectors*: extract critical pixels from the template shape and the input image, and build the feature vectors of these pixels.
4. *Building CR-table*: select the centroid point (x_c, y_c) of the template as the reference point, rotate the template shape through 180° with respect to the reference point, and trace all extracted critical pixels to build the CR-table which includes the displacements between all the traced pixels and the reference point, as well as the feature vectors of these pixels.
5. *Cell value incrementation*: for each displacement vector V_i of the CR-table which is formed by pixel P_i in the template shape, perform the following steps:
   5.1. superimpose V_i on each extracted critical pixel P_j of the input image;
   5.2. calculate the similarity measure value $S(P_i, P_j)$ by Eq. (2);
   5.3. increment by $S(P_i, P_j)$ the value of the cell in H which is pointed to by the displacement vector V_i.
6. *Maximum cell value detection*: find out as output the location of the cell with its value exceeding t_v and being the maximum in H.
7. End.

Fig. 2(e) illustrates the superimposition of all the displacement vectors of a CR-table on an extracted critical pixel, in which the value of each cell pointed to by the displacement vectors of the CR-table is incremented by the similarity measure $S(P_i, P_j)$. Fig. 2(f) illustrates the superimposition of the displacement vectors of the CR-table on all the extracted critical points.

As a result of including the color information of the pixel neighborhood for computing the similarity measure S, the peaks formed in the H become more obvious than those of not including the color information of the neighborhood, and are easier to detect. This can be seen from the comparative experimental results shown in Figs. 3 and 5.

4. Experimental Results

The proposed improved GHT algorithm has been implemented on a SUN SPARC 10 workstation and several images have been tested. The size of each input image is 256x256 pixels. Some experimental results are shown in Fig. 3 through Fig. 5.

Fig. 3(a) is a color image containing a telephone card to be used as a template. Fig. 3(b) is the telephone card template segmented from Fig. 3(a). Fig. 3(c) is an input image containing two telephone cards fully or partially identical to the telephone card template and several other kinds of cards. Fig. 3(d) and 3(e) show the extracted critical pixels of Fig. 3(b) and 3(c). After the proposed improved GHT is performed, two obvious peaks exist in the resulting HCS, as shown in Fig. 3(f) or 3(g), which locate the positions of two detected telephone cards. Fig. 3(f) shows the resulting HCS with the neighboring color information being considered, and Fig. 3(g) shows the resulting HCS without considering the neighboring color information, i.e., with $S(P_i, P_j) = S_1 S_2$. Due to using the different measure $S(P_i, P_j)$, the scales of counting values in Figs. 3(f) and 3(g) are different but the detected peaks in the HCS's are equally prominent.

Figs. 4(a) through 4(k) show another set of experimental results using juice box images. Fig. 4(a) is a color image containing a juice box template. Fig. 4(b) is the template segmented from Fig. 4(a), and its extracted critical pixels are shown in Fig. 4(e). Fig. 4(c) and 4(d) are two input images, each containing two boxes fully or partially identical to the template and several other kinds of juice boxes under different lighting intensities. Fig. 4(f) and 4(g) show the extracted critical pixels of Figs. 4(c) and 4(d), respectively. After the improved GHT is performed on the images, Fig. 4(h) and 4(i) show the two resulting HCS's, each of which includes two obvious peaks indicating the positions of two detected templates.

Fig. 5(a) is another input image same as Fig. 4(c) except that some gaussian noise and spots are added. Fig. 5(b) is the extracted critical pixels. After the improved GHT is performed on the image with and without considering neighboring color information, the results of the two approaches are shown in Figs. 5(c) and 5(d), respectively. From the results, we see that the peaks in the HCS of the proposed approach which takes into consideration the neighboring color information are much sharper than those of the other approach.

5. Conclusions

A new approach to improving the GHT for color image detection and matching in lighting change environments has been proposed. Using normalized color values, lighting changes in input images can be removed. Critical pixels instead of conventional boundary or skeletal pixels are extracted from color images, and the color information of each pixel as well as that of its neighbors is utilized. A voting rule in which cell value incrementation is based on fractional similarity measure values has also been proposed accordingly. The experimental results showed that the proposed approach is robust for color image detection and matching in noisy, occlusive and lighting change environments, and has high potential for practical applications.

References

[1] D. H. Ballard, Generalizing the Hough transform to detect arbitrary shapes, *Pattern Recognition*, **13**, 111-122 (1981).

[2] L. S. Davis, Hierarchical generalized Hough transforms and line-segment based generalized Hough transform, *Pattern Recognition*, **15**, 277-285 (1982).

[3] S. C. Jeng and W. H. Tsai, Fast generalized Hough transform, *Pattern Recognition Letters*, **11**, 725-733 (1990).

[4] H. L. Wang and A. P. Reeves, Three-dimensional generalized Hough transform for object identification, *J. Soc. Photo-Opt. Instrum. Eng.*, **1192**, 363-374 (1990).

[5] S. C. Jeng and W. H. Tsai, Scale- and orientation-invariant generalized Hough transform -- a new approach," *Pattern Recognition*, **24**, 1037-1051 (1991).

[6] Z. N. Li, B. Yao, and F. Tong, Linear generalized Hough transform and its parallelization, *Image and Vision Computing*, **11**, 11-24 (1993).

[7] W. K. Pratt, *Digital Image Processing*, Second Edition, Wiley Interscience, New York, 316-320 (1991).

[8] C. E. Borges, Trichromatic approximation method for surface illumination, *J. Opt. Soc. Am. A*, **8**, 1319-1323 (1991).

[9] H. C. Lee, E. J. Breneman, and C. P. Schulte, Modeling Light Reflection for computer color vision, *IEEE Trans. on Pattern Anal. and Machine Intell.*, **12**, 402-409 (1990).

[10] R. M. Haralick, and L. G. Shapiro, *Computer and Robot Vision*, Vol II, Addison Wesley, New York, 1-15 (1993).

[11] R. C. Gonzalez, and R. E. Woods, *Digital Image Processing*, Addison Wesley, New York, 24-73 (1992).

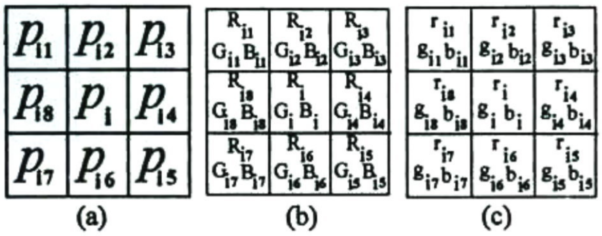


Figure 1. The color information of the neighborhood of a pixel P_i. (a) Pixel P_i and its 8-neighbors, P_{ij} for j = 1, 2, ..., 8. (b) The RGB color values (R_i, G_i, B_i) of P_i and (R_{ij}, G_{ij}, B_{ij}), for j = 1, 2, ..., 8, of 8-neighbors. (c) The corresponding normalized color values of (b).

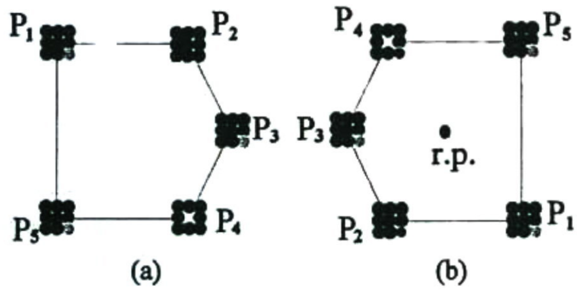


Figure 2. Illustration of building CR-table. (a) Extracted template shape consisting of five color pixels associated with their 8-neighbors. (b) Template shape rotated through 180^o with respect to the reference point. (c) Feature vector of pixel P_i constructed by using normalized color *r, g, b* values. (d) CR-table. (e) Superimposition of the CR-table on an extracted critical pixel. (f) Superimposition of the CR-table on all extracted critical pixels.

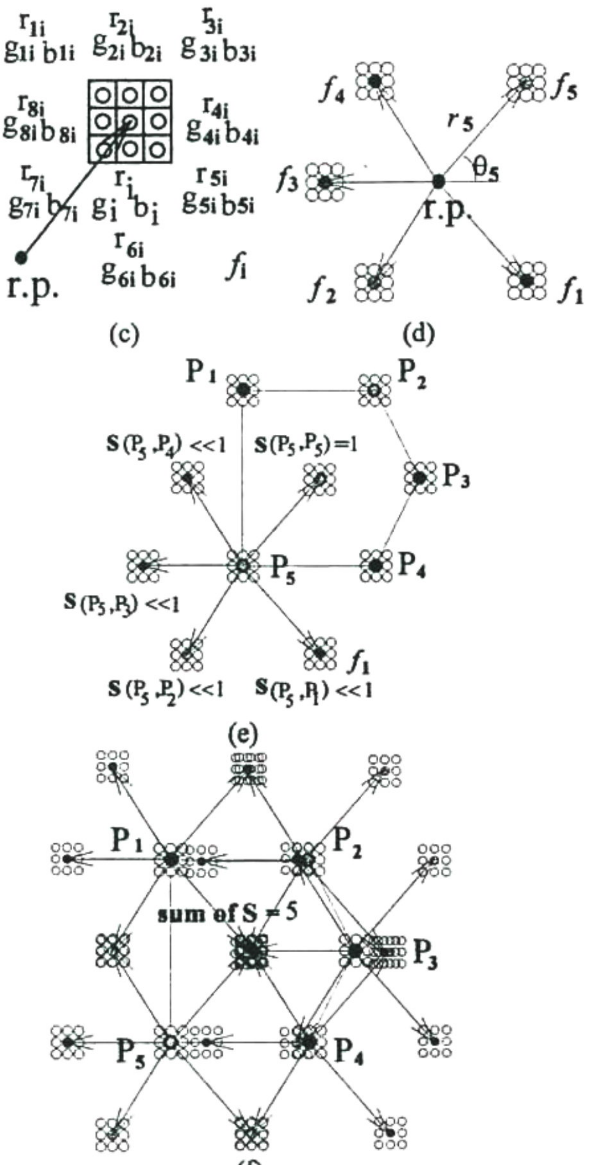
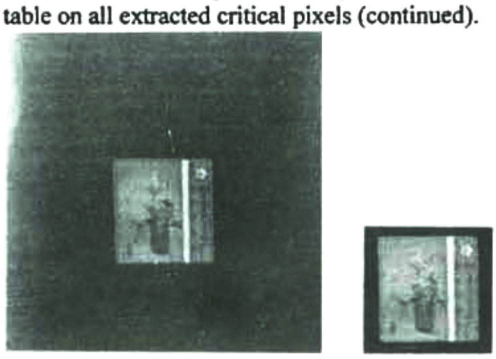
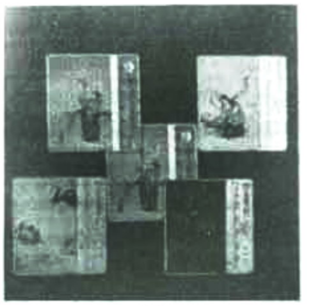


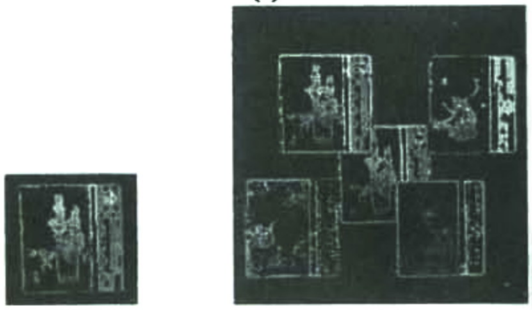
Figure 2. Illustration of building CR-table. (a) Extracted template shape consisting of five color pixels associated with their 8-neighbors. (b) Template shape rotated through 180^o with respect to the reference point. (c) Feature vector of pixel P_i constructed by using normalized color r, g, b values. (d) CR-table. (e) Superimposition of the CR-table on an extracted critical pixel. (f) Superimposition of the CR-table on all extracted critical pixels (continued).



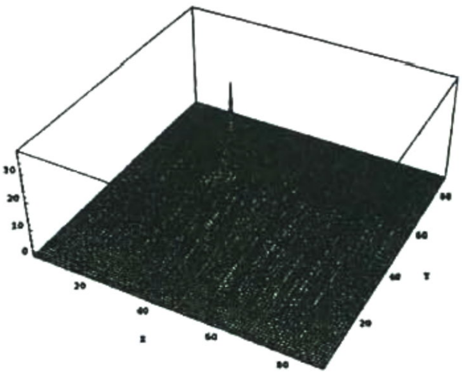
(a) (b)



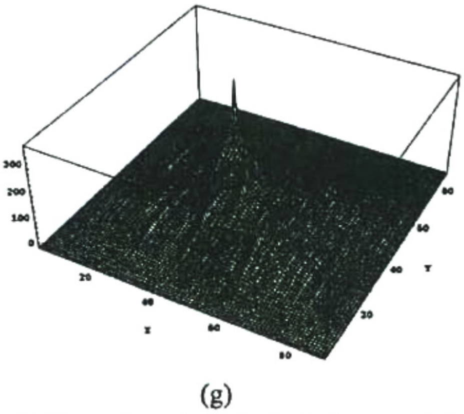
(c)



(d) (e)



(f)



(g)

Figure 3. Illustration of practical telephone card detection. (a) An image containing a telephone card template. (b) The telephone card template segmented from (a). (c) An input image containing two cards fully or partialiy identical to the card template and several other kinds of cards under different lighting intensities. (d) and (e) Extracted critical pixels from (b) and (c). (f) and (g) The resulting HCS's after performing the improved GHT with and without considering the neighboring color information, respectively. The HCS in (f) has two more obvious peaks, which locate the positions of the template cards, than those in (g).

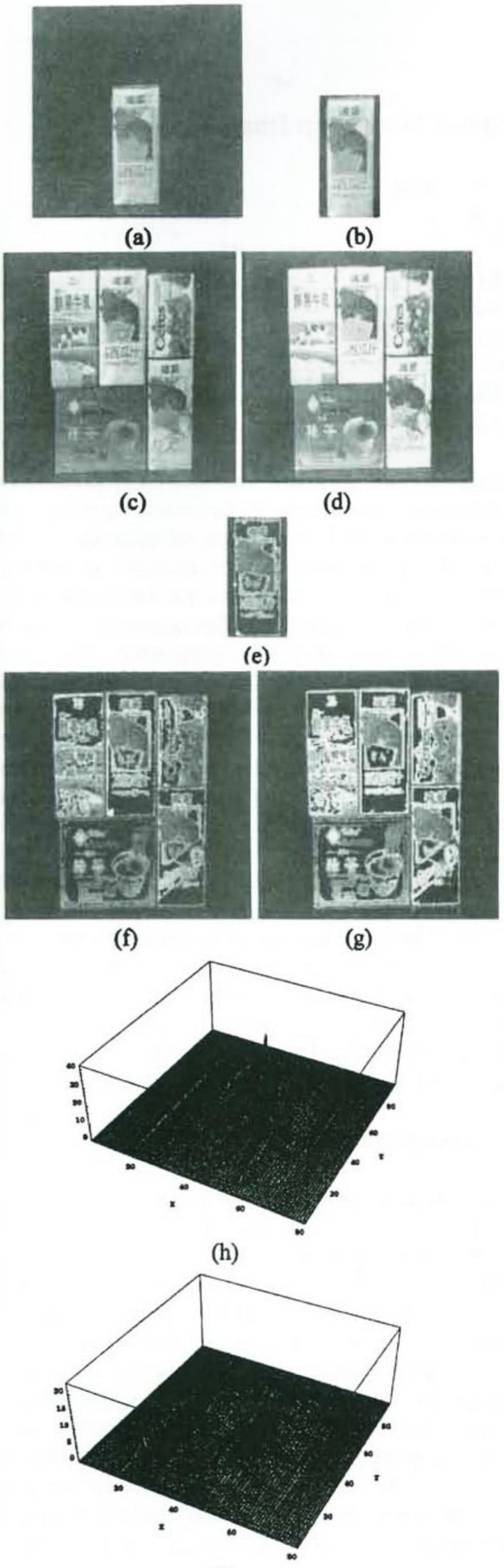


Figure 4. Illustration of practical juice box detection. (a) The template image. (b) The segmented template. (c) An input image containing two juice boxes fully or partially identical to the template. (d) Same as (c) but taken under a different lighting intensity. (e), (f) and (g) Extracted critical pixels from (b), (c) and (d), respectively. (h) and (i) Resulting HCS's after performing the improved GHT on (f) and (g), respectively, with consideration of neighboring information. Both HCS's in (h) and (i) respectively have two obvious peaks which locate the positions of two detected template.

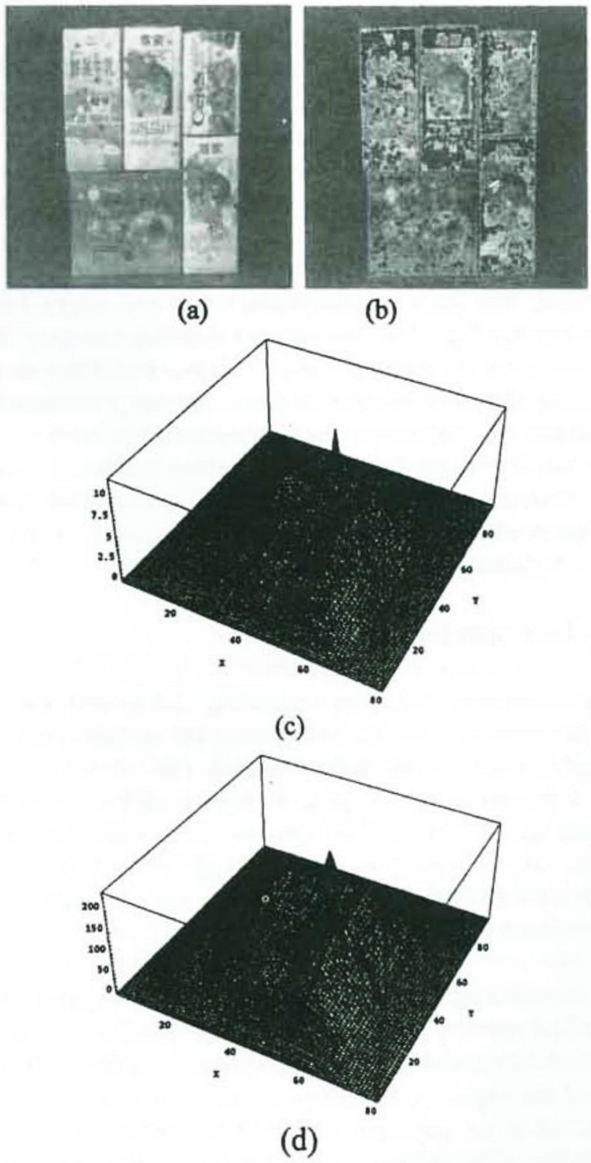


Figure 5. Comparing results of the proposed improved GHT for images with added gaussian noise and spots. (a) An input noisy image with added spots. (b) Extracted critical shape. (c) The resulting HCS with consideration of neighboring color information. (d) The resulting HCS without consideration of neighboring color information.