Unnormalized Correlation Method Based Object Finding System

Thuzar Nwe, Aung Kyaw Soe
University of Computer Studies, Mandalay, Myanmar
mathuzar.nwe@gmail.com

Abstract

This proposed system aims to develop an object finding system from complex background at random position that finds the Halal mark, ISO and Recycle marks from the products by means of pattern matching using correlation technique. Only the products with the Halal mark are consumable by the Muslims. ISO means International Standard Organization and Recycle means the product can be reusable. Pattern matching is used to locate an object of interest within a larger image. The image is scanned with the given pattern to locate sites on the image that match or bear a strong visual resemblance to the pattern. If the intended mark is found on a certain product, the brand of that product is stored in the database. By this way, the databases of the brands those are consumable by the Muslims and ISO remarked and Reusable products can be achieved.

1. Introduction

Recognizing categories or classes of objects in an image such as cars, humans, horses etc. as opposed to specific instances is a complex problem due to their variation of appearance in the image. This variation has multiple sources. By design, an object category can vary in shape, color, texture etc. Any system that tries to find an object category therefore has to deal with this variation. This can be done by identifying what is common to all instances and neglect the variation. It is generally considered that object shape is the property that is most preserved among instances. With this, it is meant that variation of shape is relatively smaller than the variation of color. This rule of course has exceptions, such as fruit categories where color can be a strong invariant.

The projected shape variation of a certain category can therefore be substantial making the recognition problem very complex. This means that it is needed an excessive number of shape templates to find a match. Template matching is a simple filtering method of detecting a particular feature in an image. Provided that the appearance of this feature in the image is known accurately, one can try to detect it with an operator called a template. This template is, in effect, a subimage that look just like the image of the object. In order to reduce the complexity of shape variation, a shape matching algorithm that is tolerant to the projected variation to a large degree can be used [1]. In this system, a pattern matching using correlation is applied to find a Halal mark, ISO and Recycle marks from the complex background at random position. This system is implemented by MATLAB.

2. System design and implementation

In the system of this paper, first of all, the color image to be tested is converted to intensity image. Then, edge detection is performed by Sobel Edge detection method. To determine the area of interest is or is not the dedicated mark, pattern matching is carried out by unnormalized correlation method. A large number of templates with various shapes and sizes are collected and stored in template database for pattern matching to be carried out.

2.1. Conversion into intensity image

The input image to be tested is a 8-bit RGB color image. The 8-bit RGB images contain 2563 different colors, that is, 2563 different RGB combinations. This immense number makes the system implementation harder to perform the pattern matching. Therefore, the acquired color image is transformed into gray or intensity image. The 8-bit gray image contains only 256 different gray values. The color to gray conversion is carried out by traditional color to gray conversion formula. This formula is based up the fact that the red, green and blue values stands for representative different gray levels [2].

$$\text{Gray color}=0.299*R+0.5876*G+0.114*B \quad (1)$$

The next image processing stage of the system is edge detection stage. The gray image is more suitable for the next edge detection image processing stage than the color image.

2.2. Edge detection stage of the system

First of all, only the gray value matching is applied in the proposed system to find out the dedicated mark. The gray value templates of the marks are stored in the template database. Each template is a window that slides pixel by pixel upon the tested gray image horizontally and then vertically. While sliding the gray template window upon the tested image, gray value matching is performed to find out the dedicated mark. Although the gray variations of the mark in the template as well as on the tested image are similar, the gray value of each pixel in the template is certainly different with the gray value of each correspondence pixel in the tested image. This fact is the major bottleneck of using the gray level detection for pattern matching.

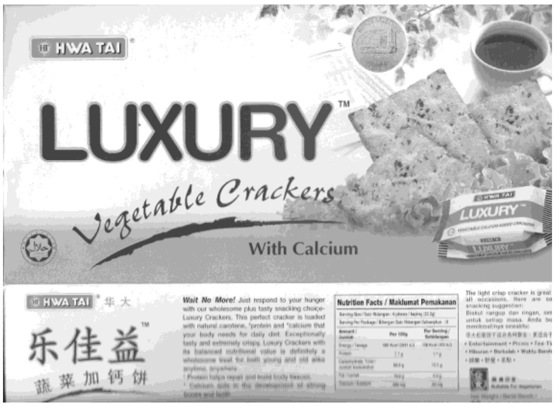

Figure 1. The resultant image of color to gray conversion stage

Therefore, to get the gray variation, that is, the abstractive graph (shape), edge detection is carried out upon both templates and tested images using Sobel edge detector.

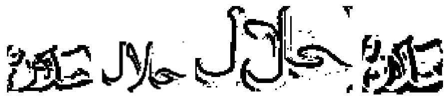

Figure 2. Edge images of Halal templates

Figure 3. Edge images of ISO templates

Figure 4. Edge images of recycle templates

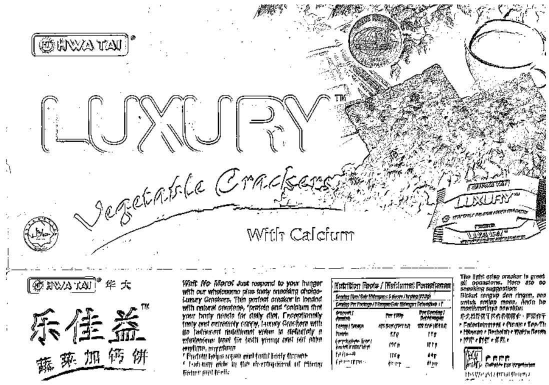

Figure 5. Edge images of tested image

2.2.1. Sobel edge detector. The Sobel edge detector is a nonlinear edge enhancement technique. It is simple variation of the discrete differencing scheme for enhancing edges.

Let a ∈ RX be the source image, and a0, a1, …, a7 denote the pixel values of the eight neighbors of (i,j) enumerated in the counterclockwise direction as follows:

a_3	a_2	a_1
a_4	(i, j)	a_0
a_5	a_6	a_7

Figure 6. Neighboring pixels of (i,j)

The Sobel edge magnitude image $m \in R^X$ is given by

$$m(i,j)=(u^2+v^2)^{1/2} \quad (2)$$

where

$$u=(a_5+2a_6+a_7)-(a_1+2a_2+a_3) \quad (3)$$

and

$$v=(2a_0+2a_1+a_7)-(a_3+2a_4+a_5) \quad (4)$$

The gradient direction image d is given by

$$d(i,j)=\arctan\left(\frac{u}{v}\right) \quad (5)$$

Let $a \in R^X$ be the source image. By image algebra formulation, the gradient magnitude image $m \in R^X$ is given by

$$m:=\left[(a \oplus s)^2+(a \oplus t)^2\right]^{1/2} \quad (6)$$

where the templates s and t are defined as follows:

$$s_{(i,j)}(x)=\begin{cases} -1 & \text{if } x=(i-1,j-1) \text{ or } (i-1,j+1) \\ 1 & \text{if } x=(i+1,j-1) \text{ or } (i+1,j+1) \\ -2 & \text{if } x=(i-1,j) \\ 2 & \text{if } x=(i+1,j) \\ 0 & \text{Otherwise} \end{cases} \quad (7)$$

$$t_{(i,j)}(x) = \begin{cases} -1 & \text{if } x = (i-1, j-1) \text{ or } (i+1, j-1) \\ 1 & \text{if } x = (i-1, j+1) \text{ or } (i+1, j+1) \\ -2 & \text{if } x = (i, j-1) \\ 2 & \text{if } x = (i, j+1) \\ 0 & \text{Otherwise} \end{cases} \quad (8)$$

The Sobel edge detection emphasizes horizontal and vertical edges over skewed edges [3].

s =

-1	-2	-1
1	2	1

t =

-1		1
-2		2
-1		1

Figure 7. Horizontal and vertical kernel

2.3. Pattern matching using unnormalized correlation

Pattern matching is used to locate an object of interest within a larger image. The pattern, which represents the object of interest, is itself an image. The image is scanned with the given pattern to locate sites on the image that match or bear a strong visual resemblance to the pattern. The determination of a good match between an image 'a' and a pattern template 'p' is usually given in terms of the metric

$$d = \Sigma\,(a-p)^2 = \Sigma(a)^2 - 2\Sigma(ap) + \Sigma(p)^2 \quad (9)$$

in detail

$$d = \sum_{i=1}^{m} \sum_{j=1}^{n} (a(i,j) - p(i,j))^2 \quad (10)$$

$$= \Sigma(a(i,j))^2 - 2\Sigma(a(i,j)p(i,j)) + \Sigma(p(i,j))^2$$

where m and n is the size of pattern.

The value of *d* will be small when 'a' and 'p' are almost identical and large when they differ significantly. It follows that the term £ap will have to be large whenever *d* is small. Therefore, a large value of £ap provides a good measure of a match. Shifting the pattern template p over all possible locations of a and computing the match £ap at each location can therefore provide for a set candidate pixels of a good match. Usually, thresholding determines the final locations of a possible good match.

$$C = \frac{\sum_m \sum_n (A_{m,n} - \bar{A})(B_{m,n} - \bar{B})}{\sqrt{\left(\sum_m \sum_n (A_{m,n} - \bar{A})^2\right)\left(\sum_m \sum_n (B_{m,n} - \bar{B})^2\right)}} \quad (11)$$

where $\bar{A} = \frac{1}{N}\sum_m \sum_n A_{m,n}$ and $\bar{B} = \frac{1}{N}\sum_m \sum_n B_{m,n}$

The method just described is known as unnormalized correlation, matched filtering, or template matching. There are several major problems associated with unnormalized correlation. If the values of a are large over the template support at a particular location, then it is very likely that £ap is also large at that location, even if no good match exists at that location. Another problem is in regions where 'a' and 'p' have a large number of zeros in common (i.e., a good match of zeros). Since zeros do not contribute to an increase of the value £ap, a mismatch may be declared, even though a good match exists. To remedy this situation, several methods of normalized correlation have been. One such method uses the formulation [3].

2.4. Variable thresholding

Each constructed edge template is scanned or slides upon the tested edge image horizontally and then vertically pixel by pixel. At each position, neighbor by neighbor subtraction is carried to identify there is similarity or not between the template and tested image. Each pixel and its eight neighbors form a 3 x 3 pixel matrix to perform this subtraction. For subtraction, the pixel values (the black pixels) are added to get the total number of black pixels in each 3 x 3 pixel matrix.

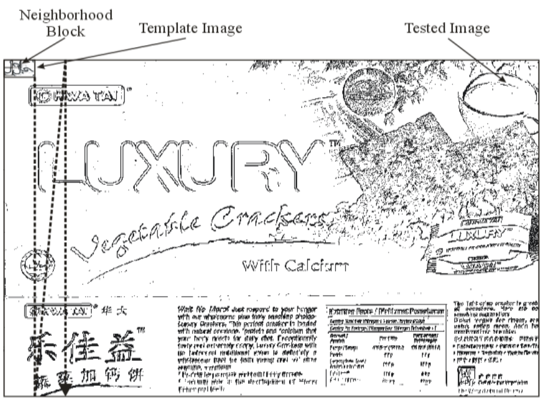

Figure 8. Pattern matching using sliding window

Therefore, while sliding the template upon the tested image, at each position, the 3 x 3 pixel matrix again slides upon the current template area [4]. Pattern matching using subtraction is performed at each 3 x 3 pixel matrix positions. The lesser in subtraction result, the more similar the two images are. The higher in subtraction result, the lesser similar the two images are. The difference from subtraction is representing the similarities in the two images. The differences from 3 x 3 pixel matrix matching within each current template area is again added together and this total difference is representative similarity value for the tested image to be the dedicated mark.

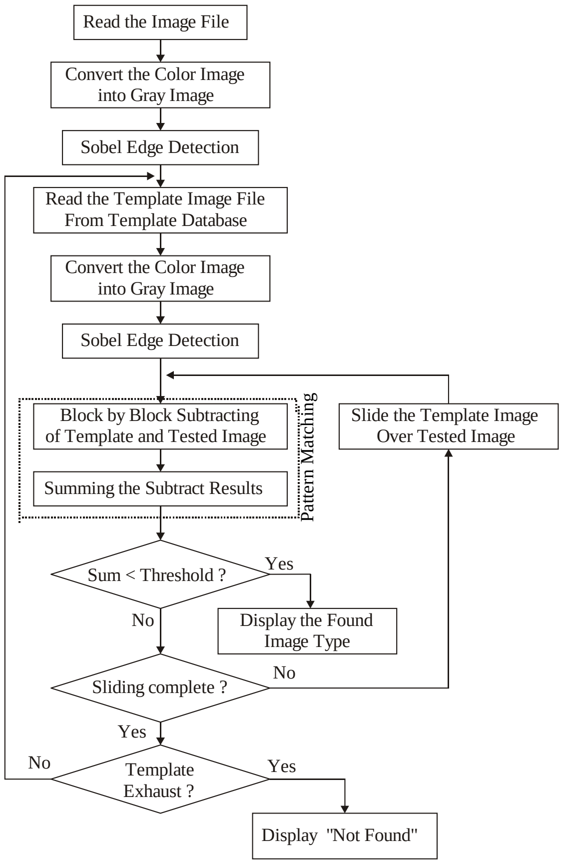

Figure 9. **Flowchart of the system**

The total difference is being quantized with the predetermined threshold value to find similar or not between the template and tested image. Since the templates in the template database are of various sizes, the predetermined threshold value can not be a fixed value. The threshold value used for similarity decision making should also be adjusted with the template size. The larger the template size, the larger the threshold value should be. The smaller the template size, the smaller the threshold value should be. The respective ratio of template size and the threshold value is analyzed and applied with trial and error approach to reach the accuracy of the system to become higher. Such thresholding is known commonly as variable thresholding.

If the total difference acquired from the pattern matching for a template position is smaller than the predetermined threshold value, the tested image portion is said to be the dedicated mark, otherwise, it is not. When the tested image portion is determined as dedicated mark, this portion in the original color image is highlighted by a square box.

3. Conclusion of the system

It is observed that the system is able to perform the dedicated pattern matching efficiently. Another fact to be pointed out is that the tested image must not be somehow rotate or tilt. If so, the pattern is assumed to be distorted. Experimental results illustrate the accuracy of detection was approximately 0.5 degrees in the rotation angle. Firstly, each edge template is scanned upon the tested edge image horizontally and then vertically pixel by pixel. At each position, pixel by pixel (0, 1) subtraction is carried out to identify there is similarity or not between the template and tested image. Hue value was firstly used when subtracting pixel by pixel rather than gray value, it is not efficiently usable. Its values are not match between template and image because of intensity changes. Therefore, the system was adopted to use subtraction of grouped pixels (3 x 3 pixel matrix).

Execution of the system takes about 5 second for each template with 28 x 26 size over 300 x 300 sized image. Both are 200dpi. It takes 2 minute for template of 103 x 105 size. The time taken to find an intended mark also depends on the amount of templates in the template databases.

4. References

[1] A. Rosenfeld and A. C. Kak, *Digital Picture Processing*, New York, Academic Press, 1976.

[2] C. OZ, F. Ercal and Z. Demir, “Signature Recognition and Verification with ANN”, *Sakarya University Computer Eng Department*, 2003.

[3] G.X. Ritter, J.N. Wilson, *Handbook of Computer Vision Algorithms in Image Algebra*, 1996.

[4] G.A. Schiavone, K.C. Hardis and S. Appadwedula, “Detecting Datum Shifts and Coordinate Transform Errors in DIS Terrain”, *University of Central Florida*.