Object Detection in Color Images Using Gradient Edge Estimates

Saleh Basalamah

Umm Al-Qura University, Saudi Arabia

Summary
This paper addresses the problem of object detection in color images, the method proposes incorporating color information in addition to size and shape for a template matching technique. This is accomplished by finding the color difference estimates of the three-color channels RGB and performing the contrast marginalized gradient template matching algorithm on the image intensity.
Key words:
Object Detection, RGB, Hough Transform, Template Matching.

1. Introduction

Gradient field estimates are found to play a major role in biological vision and in particular to distinguish shapes. Gradient field estimates are also used as the foundation of the compact Hough transform that is used to detect closed shapes [1].

In [2], a generic probabilistic framework was introduced to locate planar shapes from gradient field observations, which lead to (i) a new form of the Hough transform filters that were introduced by Kerbyson and Atherton [4], and new methods for medial axis detection [5] introduced by Bharath [3] (ii) a new development in the concept of contrast invariant shape measures that is founded on non-linear combinations of the outputs of linear spatial filters.

In this paper, we introduce color to contrast marginalized gradient template matching [2] by finding the color difference estimates of the three-color channels RGB and multiplying it by the contrast marginalized gradient template matching of the image intensity. Hence, we incorporate color information [6] that are otherwise disregarded in the grayscale gradient fields. We compare the original contrast marginalized gradient template matching [2] results with the proposed output method.

2. Template Matching

In [2], contrast marginalized gradient template matching has been proposed to address the problem of locating compact planar shapes from gradient fields of images. The technique reduces the problem of finding the position of a known shape to finding its point of origin in spirit of Hough transform. The statistical formulation of the problem is defined in form of conditional probability where the most probable shape location, $\mathbf{t}_{opt}$, given the gradient field estimate β at position X.

$$\mathbf{t}_{opt} = \underset{\mathbf{t}\in R_I}{\operatorname{argmax}}\{f_t(\mathbf{t}|\beta,X;\mathbf{\Theta})\} \tag{1}$$

In a scalar intensity image, the values of the gradient vector field that is located at $\mathbf{x}^{(i)}$, $\mathbf{b}(\mathbf{x}^{(i)})$, is essentially found by running a convolution of gradient masks with the image. Given that a shape S_0 is present at location $\sim t$ in an image and given that $\mu(\mathbf{x}^{(i)}|\mathbf{t},S_0)$ is the mean gradient vector field (over experimental space), the auxiliary random vector field conditional variable is introduced

$$\beta(\mathbf{x}^{(i)}|\mathbf{t},S_0) = \mathbf{b}(\mathbf{x}^{(i)}) - \mu(\mathbf{x}^{(i)}|\mathbf{t},S_0) \tag{2}$$

A bivariate Gaussian version of the statistics of $\beta(\mathbf{x}^{(i)})$ has the form

$$f_\beta(\boldsymbol{\beta}(\mathbf{x}^{(i)})|\mathbf{t},S_0,\boldsymbol{\mu}(\mathbf{x}^{(i)})) = \frac{\alpha_\beta}{\pi}e^{-\alpha_\beta|\boldsymbol{\beta}(\mathbf{x}^{(i)}|\mathbf{t},S_0,\boldsymbol{\mu}(\mathbf{x}^{(i)}))|^2} \tag{3}$$

Between any two gradient images, there will be some degree of contrast uncertainty, so re-writing the model for auxiliary random vector field

$$f_\beta(\boldsymbol{\beta}(\mathbf{x}^{(i)})|\mathbf{t},S_0,\boldsymbol{\mu}(\mathbf{x}^{(i)}),A) = \frac{\alpha_\beta}{\pi}e^{-\alpha_\beta|\boldsymbol{\beta}(\mathbf{x}^{(i)}|\mathbf{t},S_0,\boldsymbol{\mu}(\mathbf{x}^{(i)}),A)|^2} \tag{4}$$

Analytic marginalization of Equation (4) over contrast leads to a set of terms C_1, C_2, C_3, C_4, C_5 that involves cross correlation or spatial convolution between the gradient field components and the required mask.

$$C_1 = \sum_{i=1}^{N} \frac{b_x(\mathbf{x}^{(i)})\mu_x(\mathbf{x}^{(i)}|\mathbf{t}) + b_y(\mathbf{x}^{(i)})\mu_y(\mathbf{x}^{(i)}|\mathbf{t})}{Z_0^{(i)}} \tag{5}$$

$$C_2 = \sum_{i=1}^{N} \frac{b_x(\mathbf{x}^{(i)})b_y(\mathbf{x}^{(i)})\mu_x(\mathbf{x}^{(i)}|\mathbf{t})\mu_y(\mathbf{x}^{(i)}|\mathbf{t})}{Z_0^{(i)}} \tag{6}$$

Manuscript received September 5, 2020
Manuscript revised September 20, 2020

DOI: 10.22937/IJCSNS.2020.20.09.17

$$C_3 = \sum_{i=1}^{N} \frac{\alpha_A + \alpha_\beta b_y^2(\mathbf{x}^{(i)})\mu_x^2(\mathbf{x}^{(i)}|\mathbf{t})}{Z_0^{(i)}} \tag{7}$$

$$C_4 = \sum_{i=1}^{N} \frac{\alpha_A + \alpha_\beta b_x^2(\mathbf{x}^{(i)})\mu_y^2(\mathbf{x}^{(i)}|\mathbf{t})}{Z_0^{(i)}} \tag{8}$$

$$C_5 = \sum_{i=1}^{N} \frac{b_x^2(\mathbf{x}^{(i)}) + b_y^2(\mathbf{x}^{(i)})}{Z_0^{(i)}} \tag{9}$$

3. Method

The method we have introduced here is shown in Figure 1, it is composed of two main algorithms; the first is a color difference estimation and second is a gradient edge estimation.

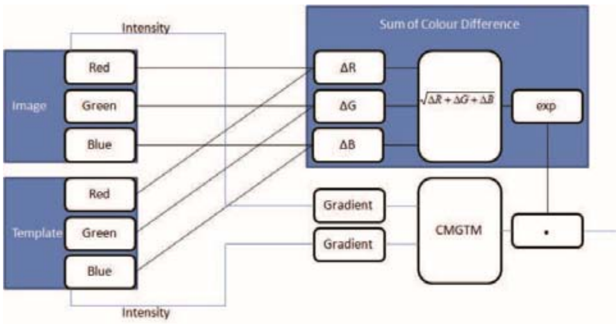


Fig. 1: The proposed system which uses color difference estimation to incorporate color into detecting objects using CMGTM.

3.1 Color Difference Estimation

The system needs two inputs: an RGB color image and a color template that resembles the object to be detected. We first extracted the intensities of each RGB color component from the image and found the sum of color difference in each channel with the same color channel from the template. In order to find the sum of color value difference between the image and the much smaller in size template, we convolve each image channel with an all-one matrix that is of the template size and we convolve an all one matrix that is of the image size with each image channel matrix. We then subtract the results of the two convolutions for each color channel to find the sum of difference. Our method then uses the measurement of color differences between the color component values of the images with the color component values of the template using the distance measure:

$$D = \sqrt{(R_I - R_T)^2 + (G_I - G_T)^2 + (B_I - B_T)^2} \tag{10}$$

An estimation of the areas of the image where the color values with higher similarity to the color values of the template is then calculated using the following equation

$$O = e^{-2D} \tag{11}$$

and as the distance D (the color difference) becomes 0, the similarity value becomes 1 and as the difference in color D increases (maximum of $\sqrt{3}$), the similarity value decreases exponentially (reaching 0.0313 when $D = \sqrt{3}$).

3.2 Gradient Edge Estimation

In order to apply the contrast marginalized gradient template matching algorithm, we need to find the intensity maps of the RGB color image and template. We do this by converting the RGB colourmap to HSV and taking the V channel as our intensity map. Next, we convolved them with a 3×3 Prewitt mask to find the gradients of the image.

$$\begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

The same Prewitt masks are also convolved with the intensity map of the object template to find its gradients.
The gradients of the image and the template are then used by the contrast marginalized gradient template matching algorithm, the algorithm is applied once in the system. The output of the algorithm is normalized so that is it always positive. This results in an accumulator space with a peak at center of the detected object's location.
The accumulator space is finally multiplied with the color similarity estimation which results in the peaks as the most probable location of the color object in the color image.

4. Experimental Results

The performance of the method has been tested on a variety of images. Figure 2(a) is a 241×241 synthetic image used to test the technique's ability to detect objects based on the color information even in the presence of other objects that have the same size and shape. The circular objects in this image have a diameter of 30 pixels and have different colors (red, blue, white, cyan, magenta, green and yellow).
Figure 2(b) shows the color similarity are based on the white circle. The accumulator space shows the locations where a white object is most probable to be. Figure 2(c) is

the accumulator space of the contrast marginalized gradient template matching applied on the intensity map of the image and the template. The shape and size of the object is taken into account here since the algorithm relies of the gradient edge estimates of the intensity map. Although color is not accounted for in Figure 2(c), the multiplication of the two accumulator spaces generates an accumulator space that is sensitive to size, shape and color as shown in Figure 2(d) where the location of the white circles is detected.

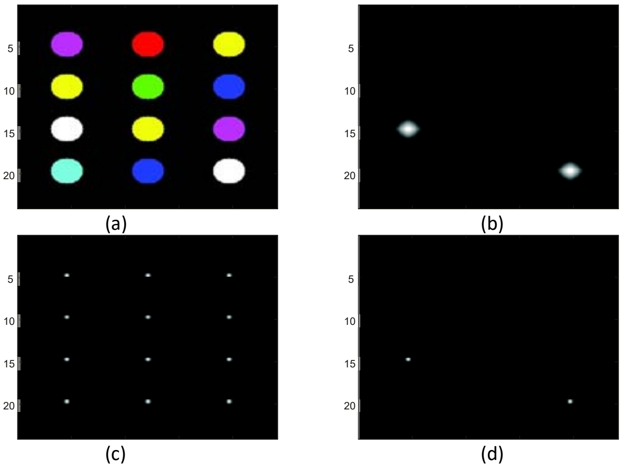
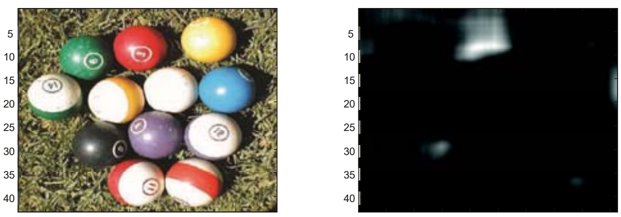
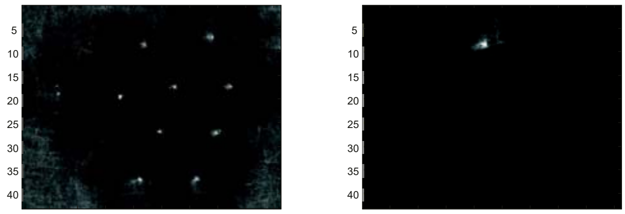


Fig. 2: (a) Synthetic test image. (b) The colour similarity area based on the white colour circle template. (c) The detected objects based on the CMGTM algorithm. (d) The detected white circles.

In order to test the performance of the technique in real situations; two real-image tests were performed on the technique. The first is Figure 3(a) which is a set of different colored billiards balls put on grass. Some balls have lines and writing, and others have stripes. The red ball is to be located on the image and therefore a template of a red ball with a diameter of 100 pixels is synthesized. Figure 3(b) shows the red color similarity locations and Figure 3(c) shows the accumulator space of CMGTM on the intensity map of the image. The result is shown in Figure 3(d) where the location of the red ball is.



(a) (b)



(c) (d)

Fig. 3: (a) Different coloured billiards balls on grass. (b) The probable location of red objects in the image. (c) The location of balls with a diameter of 100 pixels using CMGTM. (d) The detection of the red ball.

The second real-image test is to detect a more complex object. We try to detect faces in an image that has been used in [2]. Although the particular issue of face detection using this method has to be further investigated in future work, we use this example here to prove the algorithm and to compare it to the original CMGTM [2]. Figure 4(c) is the accumulator space from the original CMGTM and Figure 4(d) is the accumulator space with color information incorporated. Note that since color is incorporated, the face of the author (last to the right) has not been detected because the facial color is different than the other persons. The template used has the following color values: R=1, G=1 and B=0.5.

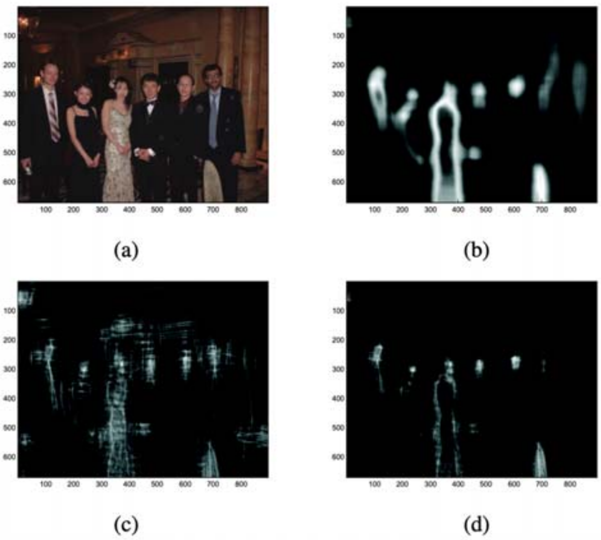


Fig. 4: (a) Image with faces. (b) The colour similarity estimation. (c) The detection of faces with no colour information. (d) The detection of faces with colour information.

5. Conclusion

The work in this paper employs contrast marginalized gradient template matching for object detection in color images. Color difference information and intensity values has been incorporated into the process of object detection for this system. In the experiments using synthesized images, we showed that our method selectively detects objects according to size, shape and color. We also showed that our results work in real images.

References

[1] D.H. Ballard. Generalising the hough transform to detect arbitrary shapes. *Pattern Recognition*, 1981.

[2] S. Basalamah, A. Bharath, and D. McRobbie. Contrast marginalised gradient template matching. *8th European Conference on Computer Vision*, 2004.

[3] A.A. Bharath and C.J. Huberson. Obtaining medial responses from steerable filters. *IEE Proceedings on Vision, Image and Signal Processing*, 1999.

[4] J.H. Kerbyson and T.J. Atherton. Circle detection using hough transform filters. *Fifth International Conference on Image Processing and its Applications*, 1995.

[5] Stephen M. Pizer, Christina A. Burbeck, James M. Coggins, Daniel S. Fritsch, and Bryan S. Morse. Object shape before boundary shape: Scale-space medial axes. *Journal of Mathematical Imaging and Vision*, 1994.

[6] M. Ruzon and C. Tomasi. Edge, junction, and corner detection using color distributions.