A Color Channel Ratio Image Based Object Detection Method in Complex Scenario

Yungang Luo, Yi Xu, Cheng Zhi.

Institute of Image Communication and Network Engineering & Shanghai Key Laboratory of Multi Media Processing and Transmissions,Shanghai Jiao Tong University

{luoyungang, xuyi, zhicheng@sjtu.edu.cn}

***Abstract*—In complex scenario such as underwater imaging or rough weather, local textures are unavailable due to image blurring and background clutters. Global features like color and contours became the important clues for object detection. However, it is challenging to extract the object region with regard to the variances of color and massive disruptions of object contour. In this paper, we proposed a new method which used *Color Channel Ratio Image* (CCRI) for edge detection. The CCRI-based edge detection method can effectively extract the object contour in complex scenario where traditional grayscale-based edge detectors often fail. Moreover, we design a fast object detection algorithm based on the CCRI image and chamfer matching. Performing *k*-means clustering on the CCRI image can obtain the candidate regions of the target object. Thus applying directional chamfer matching only in the candidate regions can efficiently speed up the detection procedure.**

***Index Terms*—Edge detection, Chamfer matching, Object detection, *K*-means clustering.**

I. Introduction

THE edges of images are regions of interest where there is a sudden change in intensity. Image edges have always been used as primitives in the field of image analysis, segmentation and object detection in computer vision. Among many visual cues, object edges or contours are illuminant invariant and stable to present the semantic perception. Meanwhile, object detection is a crucial step for automatic object recognition. Over the decades many object detection algorithms have been proposed, among which chamfer matching [1] is widely used because of its speed and robustness. Chamfer matching algorithm matches the contour of each candidate object region with the given shape template and figures out the best match as the target object. Its matching effects extremely depend on the edge extraction results. Whether we can obtain the relatively complete contours of the objects determines whether the chamfer matching algorithm can be effectively performed.

However, in some complex scenario, we can hardly obtain the object edge using the traditional edge detectors. Fig. 1 gives an example about this. In these cases, such shape matching algorithms as chamfer matching lose their effectiveness.

Fig. 1. An underwater image and its luminance edge. The edge is obtained by Canny detectors. The luminance edge is mixed with many noisy edges. It hardly supplies any useful cues for object localization.

Human eyes are sensitive to both intensity changes and color changes. In Fig.1, the intensity change around the object region is too weak to extract the luminance edge. However, the color change is still clear and we can easily recognize the object. This inspires us to utilize the color cues to extract the object edge. In this paper, we propose a novel method to detect the edges of object with color, which is very helpful when traditional grayscale-based edge detectors fail in complex scenario. Our edge detection method is based on the Color Channel Ratio Image (CCRI), which subtly integrates the color inconsistency of original image into a single channel ratio image and can effectively preserve the color contour of the object.

There are several benefits using CCRI image for color object detection in complex scenario. Firstly, from CCRI image we can successfully extract edges of objects whose luminance edges were destroyed in the original image for some reasons, which make it possible that we can use chamfer matching to detect the object of interest. Secondly, the CCRI-based edge detection method can effectively suppress the edges from object texture and it is illumination-invariant to some extent, which would make the shape matching robust. Meanwhile, we can find the candidate regions of objects of interest through clustering the CCRI image pixels. Performing chamfer matching only in these candidate regions can effectively reduce

This work was supported in part by Research Fund for the Doctoral Program of Higher Education of China (200802481006) and NSFC-60902073.

Yungang Luo is with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China. (Phone: +86-18817558986; e-mail: luoyungang@sjtu.edu.cn Address: 800 Dongchuan RD. Minhang District, Shanghai, China.)

Yi Xu is with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China.(e-mail: xuyi@sjtu.edu.cn)

Cheng Zhi is with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China.(e-mail: zhicheng@sjtu.edu.cn)

the false matches and speed up the chamfer matching algorithm.

II. Previous Related Works

There are many edge detection methods, many of which model the problem in the monochrome channel. Canny detector [2] models edges as sharp discontinuities in brightness channel. It is an image gradient-based edge detection method.

In [3,4], the authors proposed an Oriented Energy approach, which considered the response of the image to a family of filters under multiple scales and orientations. The filter bank is composed of quadrature pairs of even and odd symmetric filters.

In [5], the author proposed a high-performance contour detector using a combination of local and global cues. They combine the local information derived from brightness, color, and texture signal, with global information obtained from spectral partitioning, to produce a contour detector. The contour detector is able to obtain high accuracy results. But it is high in computation complexity and is not applicable in real-time situation.

Objet detection is a challenging problem in computer vision. An extensive literature has developed in the past decades. Recently, the contour-based detection method has become popular, because shape is invariant to extreme lighting conditions and large variations in texture or color. The works of [6, 7] focus on the aspect of learning edge codebooks, where chamfer matching is used to evaluate local shape similarity. In [8, 11], the author proposed a family of scale invariant local shape descriptors which are formed by k-connected near straight contour fragments. In [9, 10, 12] the problem is cast as a matching between shape-based descriptors on local interest points. In [13], a contour segment network framework is presented where shape matching is formulated as finding paths on the network similar to model outlines.

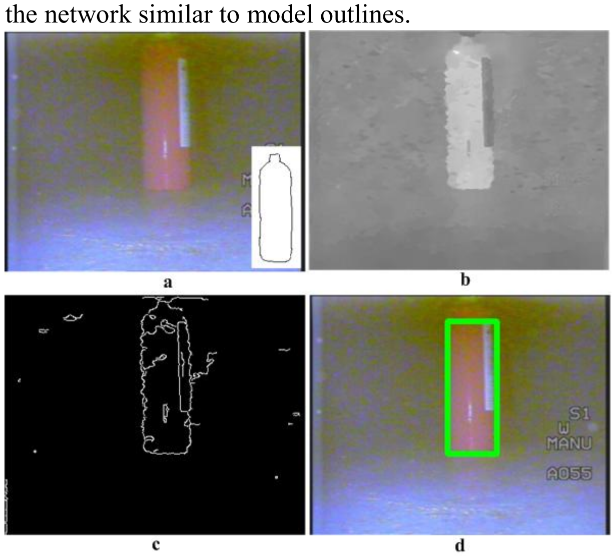

Fig. 2. Object detection result with the proposed method. a) Original image. b) Color channel ratio image. c) CCRI-based edge. d) Chamfer matching result.

These algorithms provided impressive results for shape matching problem. However, they share a common drawback of high computational complexity, which makes them unsuitable for time critical applications.

We aim to propose methods that can cope with the complex situations where luminance edge cannot be detected. Fig. 2 shows that the proposed method can effectively detect the object edge (fig. 2c), and match the edge with template to localize the object in the image (fig. 2d). In addition, the proposed algorithm is fast and robust. Firstly, the CCRI-based edge detector can suppress the texture edges and noisy edges, which is helpful to avoid false detections and omissions. Secondly, the proposed method can find the candidate regions of object, which liberates us from searching the whole image in chamfer matching step.

III. A Color Channel Ratio Image Based Object Detection Method

As discussed before, our proposed object detection method mainly cope with those cases where the luminance edges of objects were too weak to distinguish but the color information is strong enough for contour extraction. The algorithm includes two parts, one is CCRI-based edge detection, and the other is chamfer matching, as shown in Fig. 3.

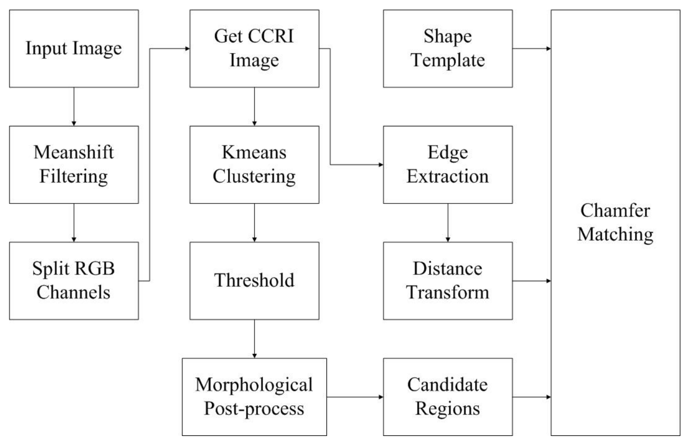

Fig. 3. Flow chart of the proposed algorithm

A. *CCRI-based Edge Detection*

Color is a powerful descriptor that often simplifies object identification and extraction from a scene. Human can discern thousands of color shades and intensities, compared to about only two dozen shades of gray. The RGB model is one of the most commonly used color space model. The model is based on a Cartesian coordinate system. The color subspace of interest is the cube shown in Fig. 4.

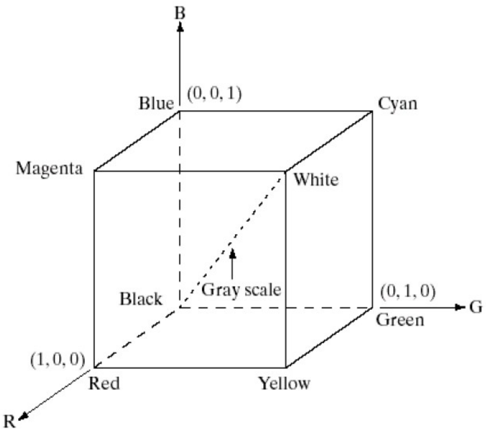

Fig. 4. The RGB color cube. The RGB primary values are at three corners. Points along the main diagonal have gray values, from black at the origin to white at point (1, 1, 1).

Images represented in the RGB color model consist of three component images, one for each primary color. According to our experiments, although one same color may have different intensities, the ratio value between channels is nearly consistent. Based on this fact, we present the Color Channel Ratio Image. Actually, the CCRI-image reflects the color distribution of the original image. Here we denote the ratio image as RI.

$$RI = \frac{I_i}{I_j + \varepsilon} \quad i, j \in \{r, g, b\}, i \neq j \tag{1}$$

In equation (1), I represents the original image matrix, i, j denotes the RGB color channels, ε is a small positive number to avoid the denominator to be zero. This equation transforms a RGB color image into a single channel image. Then we perform edge detection algorithm on the ratio image to obtain the image edge. There are three color channels, so one original image can produce six different ratio images, among which three pairs are equivalent. Which image should we use to extract edge? Generally we choose the one in which the ratio value of the target object reaches the maximum. Given a red object, for example, we should choose the r/g ratio image.

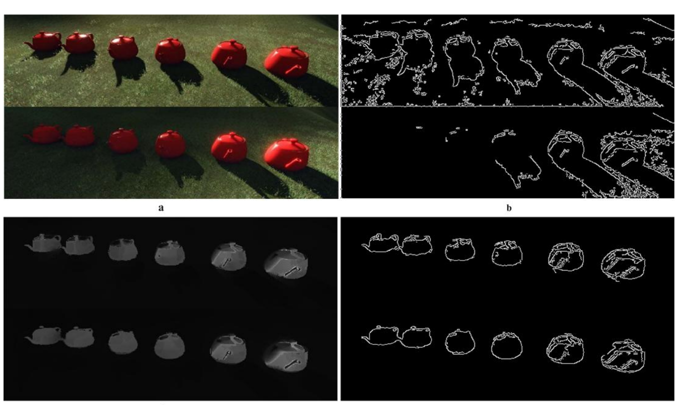

Fig. 5. CCRI-based image edge detection. a) original image; b) original luminance edge (Canny result); c) ratio image of r/g; d) edge extracted from ratio image.

Figure 5 displays an example of CCRI-based image edge. We can see that the ratio image edge (fig. 5d) is clean and neat, and the contours of pots are relatively complete, especially in the shadow areas of the pots. On the contrary, in the original luminance edge (fig. 5b), the contours of pots are surrounded with many noisy edges, and edges of the 3 pots in the bottom-left area are not detected. In addition, the edges in the bottom part of the pots are missed, but the shadow edges of pots come out. In conclusion, the proposed CCRI-based edge detection method has 3 advantages.

1) The proposed CCRI-based edge detection method is illumination-invariant. The objects with similar color may have different gray-intensities, just like the red pots in bottom left and bottom right in fig. 5a, but their ratio values between channels are very close, which we can see from the ratio image (fig. 5c).
2) The CCRI-based edge detection method can effectively suppress the noisy edge, especially those that come from the background texture. The anti-noise property is very helpful in the chamfer matching step, because heavy noise can seriously interfere the matching results, which may result in many fault detections.
3) The CCRI-based edge detection method can remove the shadow edges of objects.

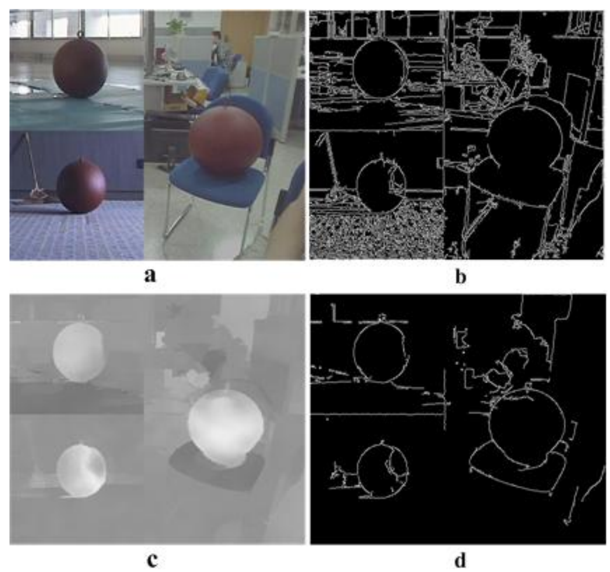

Fig. 6. Edge detection with the proposed method. a) original image; b) original luminance edge (Canny result); c) ratio image of r/g; d) edge extracted from ratio image. The red ball is the target which we want to detect.

Another example is shown in fig. 6. The red balls in original image (fig. 6a) are the objects of interest. The original image is a montage of three images of the same object which are taken in different illumination and different scenes. Fig. 6d shows that our proposed method can effectively suppress the noisy edge and extract the relatively completely edges of objects at the same time.

B. K-means Clustering to Localize Candidate Regions

K-means [14] clustering is a method of cluster analysis which aims to partition *n* observations into *k* clusters in which each observation belongs to the cluster with nearest mean. Given a set of observation $\{x_1, x_2, \cdots, x_n\}$, where each observation is a m-dimension vector, k-means clustering aims to partition the n observations into k sets ($k \le n$), $S = \{S_1, S_2, \cdots, S_k\}$ so as to minimize the within-cluster sum of squares (WCSS),

$$\arg\min_{S} \sum_{i=1}^{k} \sum_{x_j \in S_i} \left\| x_j - \mu_i \right\|^2 \tag{2}$$

where μ_i is the mean of points in S_i.

The ratio values of objects with similar color are concentrated. So if we perform k-means clustering algorithm on the CCRI image, those pixels with same color would be partitioned into the same group with high probability. After k-means clustering, we segment the candidate regions by image thresholding.

Denote the k-means image as $kms(x, y)$, the segmented binary image as $seg(x, y)$, then

$$seg(x, y) = \begin{cases} 1 & if\ kms(x, y) > T \\ 0 & if\ kms(x, y) \le T \end{cases}, \tag{3}$$

where $T \in \{\mu_1, \mu_2, \cdots, \mu_k\}$. T is not a constant, and it depends on the cluster number k and the CCRI-image. Generally, the cluster number is less than 10 and we choose the best ratio image in which the object region has the maximal value. After segmentation, the foreground regions are indicated as the candidate object regions. Performing chamfer matching algorithm only in these regions will effectively speed up the detection procedure. In practice, the chamfer matching search region is the minimum enclosing up-right rectangle region of the corresponding foreground. To avoid omission, the actual search region is set as 1.5 times larger than the minimum enclosing rectangle. Fig. 7 illustrates the k-means results and the candidate regions.

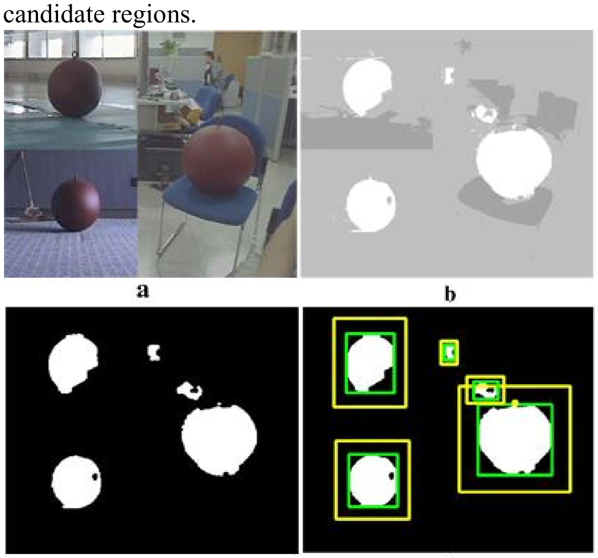

Fig. 7. *K*-means and the candidate regions. a) original image; b) *k*-means result (k=3); c) binary image after thresholding; d) the candidate object region. The green rectangle is the minimum enclosing up-right rect. The yellow one is the search region for chamfer matching. The width and height of yellow rectangle are 1.5 times larger than the green one and they share the same center.

C. Chamfer Matching

Chamfer matching [1] is a popular technique to find the best alignment between two contours. Let $U = \{u_i\}$ be the sets of template, and $V = \{V_j\}$ be the image edge maps. The chamfer distance between U and V is defined as the average of distances between each point $u_i \in U$ and its nearest edge in V.

$$d_{CM}(U,\ V) = \frac{1}{n} \sum_{u_i \in U} \min_{v_j \in V} \left| u_i - v_j \right| \tag{4}$$

where $n = |U|$.

Chamfer matching can tolerate small rotations, occlusions, misalignments and deformations. The matching cost can be computed efficiently via distance transform [15].

$$DT_V(x) = \min_{v_j \in V} \left| x - v_j \right| \tag{5}$$

The pixel value in distance transform image is the distance from each pixel to the nearest edge pixel in V. Using distance transform the equation (4) can be represented as:

$$d_{CM}(U,\ V) = \frac{1}{n} \sum_{u_i \in U} DT(u_i) \tag{6}$$

In [15], a fast distance transform algorithm is proposed and using this fast algorithm the cost function (6) can be evaluated in linear time *O(n)*.

In cluttering or noisy scene, chamfer matching becomes unreliable. Several improvements have been introduced to enhance the performance of traditional chamfer matching by incorporating edge orientation information into the matching cost.

Based on the algorithm proposed in [16], we proposed an optimal search strategy to further improve the performance of this shape matching algorithm.

In [16], the image edges are represented with a collection of m line segments. $m << n$, n is the cardinality of edge points. Then we can scale, translate and rotate the edge maps in a very convenient way. For a line-segments set $L = \{l_1, l_2, \cdots, l_m\}$, which is the line-representation of some edge image, its linear transform is,

$$T(L) = \{T(l_i)\} \quad i = 1,2,\cdots,m \tag{7}$$

$$T(l_i) = \alpha R l_i + t \tag{8}$$

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{9}$$

where T denotes the linear transform, l_i is one of the line segments, R is a rotation matrix, α is a scale factor, t is a translation factor.

In practice, translation is realized by scanning with a sliding window. The size of sliding window is determined by the following equation,

$$\alpha_{\max} = \max\left\{ \frac{w_{\max}}{w_t}, \frac{h_{\max}}{h_t} \right\} \tag{10}$$

$$\alpha_{\min} = \min\left\{\frac{w_{\min}}{w_t}, \frac{h_{\min}}{h_t}\right\} \tag{11}$$

where w_t, h_t is the template width and height, $w_{\max}, h_{\max}$ is the maximum width and height of the candidate region, while $w_{\min}, h_{\min}$ is the minimum width and height of the candidate region. Generally, the maximum size of the candidate region is 1.5-2 times larger than the size of minimum bounding up-right rectangle of the seed region (see fig. 5d), and the minimum size is 0.5 times of the minimum bounding rectangle.

Thus we get the maximum and minimum scale factor. Then we completely search in the range of $[\alpha_{\min} \ \alpha_{\max}]$ for the optimal match in the candidate region.

IV. Results

The proposed object detection method has been implemented with C++ programming on MS Visual Studio 2010 platform with embedded OpenCV2.3. This method can detect the object of interest with encouraging speed and efficiency. In some complex scenario where the brightness based edge detectors fail, the chamfer matching would lose its effectiveness. But if the object of interest is colored, our CCRI-based edge detection method can still extract the edge.

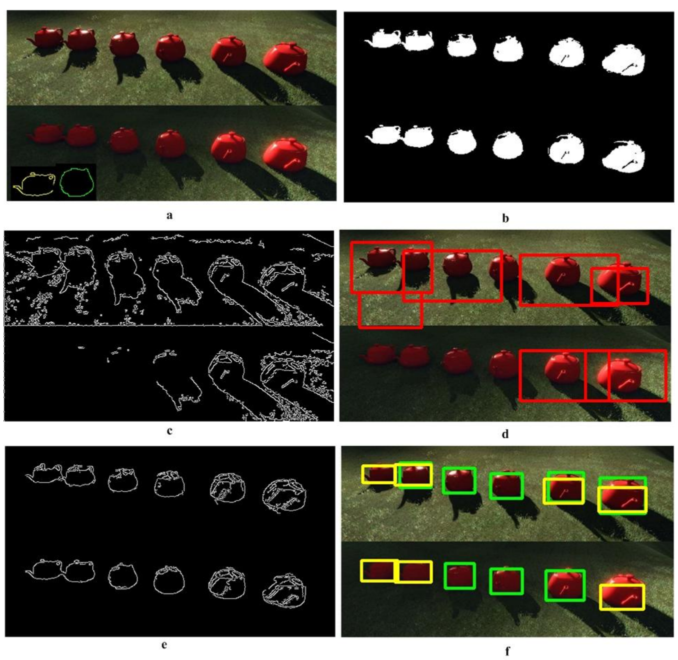

Fig. 8. The experiment result. a) Original image, the two shapes at the bottom-left are two pot template. b) Seed regions. c) Canny edges of original image. d) Chamfer matching result with fig. 8c. e) Proposed edge detection result. f) Chamfer matching result with fig. 8e.

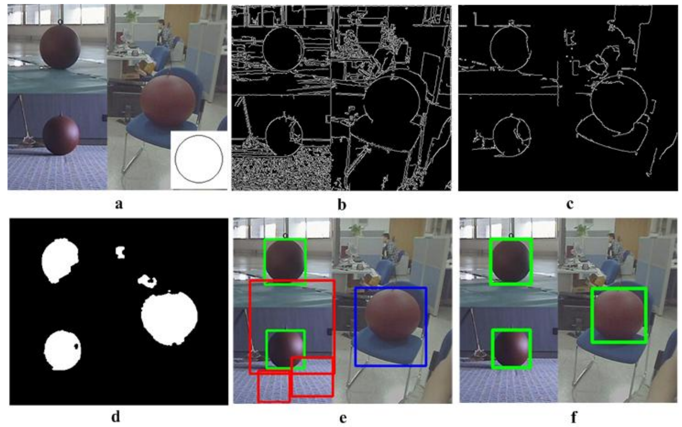

Fig. 9. The experimental results. a) Original image with two templates at the bottom-right. b) Canny edges of original image. c) Proposed edge detection result. d) Seed regions. e) Chamfer matching result with fig. 9b. f) Chamfer matching result with fig. 9c.

Figure 8, figure 9 and figure 10 shows some experiment results. In fig.8 we can see that matching the canny edge map with template results in a lot of false detections and omissions (fig. 8d). But for the CCRI-based edge, the matching result is perfect. Note that here we use two templates to match the 10 pots with different posture. The template with green color in fig. 8a produces the detection results with green rectangle in fig. 8e, and the yellow template corresponds to the yellow ones in fig. 8e.

In fig. 9, the luminance edge-based detection results have 3 false detections (the red rectangles in fig. 9e). The blue rectangle in fig.9e is an inaccurate detection. The CCRI-edge based detection results are perfect in comparison.

In fig. 10, the original image was obtained in foggy environment. Fig. 10b is the gray edge extracted with Canny detectors from the corresponding gray-scale image of original image, and fig. 10c is the proposed CCRI-based edge detect result. We can see that in the gray edge image, the object (red balls in original image) contours are missed. But in the CCRI edge image, the object contours are completely extracted. Thus we can successfully detect the object with chamfer matching method (see fig. 10d).

V. Conclusions

In this paper, we proposed a CCRI-image based object detection method for real time requirement. This method can cope with some complex environment where traditional brightness based edge detection methods often fail. From the CCRI image we can extract relatively complete edges for chamfer matching. In addition, performing k-means clustering and thresholding can find the object seed regions, which limit the search range to some candidate regions and thus speed up the matching procedure.

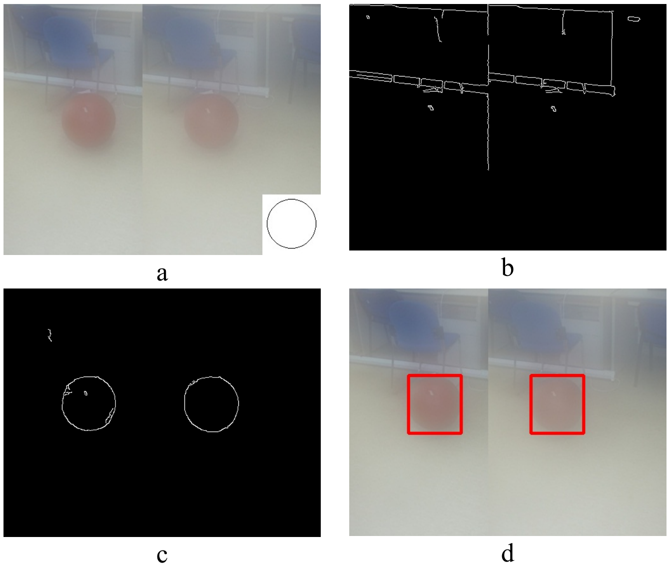

Fig. 10. The experiment result. a) Original image. b) Canny edges of original image. c) Proposed edge detection result. d) Chamfer matching result .

REFERENCES

[1] H. Barrow, J. Tenenbaum, R. Bolles, and H. Wolf. *Parametric correspondence and chamfer matching: Two new techniques for image matching*. In Int'l Joint Conf. of Artif. Intel., pages 659–663, 1977.

[2] J. Canny, *A Computational Approach to Edge Detection*. PAMI, 1986.

[3] M. Morrone and R. Owens, *Feature detection from local energy*. Pattern Recognition Letters, 1987

[4] P. Perona and J. Malik, *Detecting and localizing edges composed of steps, peaks and roofs*. ICCV, 1990

[5] M. Ruzon and C. Tomasi Edge, *Junction, and corner detection using color distributions.* PAMI, 2001

[6] Opelt, A., Pinz, A., Zisserman, A., *A boundary-fragment-model for object detection*. In: ECCV. (2006)

[7] Shotton, J. Blake, A. Cipolla, R, *Contour-based learning for object detection*. In: ICCV. (2005)

[8] Ravishankar, S., Jain, A., Mittal, A. *Multi-stage contour based detection of de-formable objects*. In: ECCV. (2008)

[9] Maji, S., Malik, J.: *Object detection using a max-margin hough transform*. In: CVPR. (2009)

[10] Ommer, B., Malik, J.: *Multi-scale object detection by clustering lines*. In: ICCV.(2009)

[11] Ferrari, V., Jurie, F., Schmid, C.: *From images to shape models for object detec-tion*. In: IJCV. (2009)

[12] Berg, A., Berg, T., Malik, J.: *Shape matching and object recognition using lowdistortion correspondences*. In: CVPR. (2005)

[13] V. Ferrari, T. Tuytelaars, and L. V. Gool. *Object detection by contour segment networks.* In Proc. European Conf. on Comp. Vis., volume 3953 of LNCS, pages 14–28. Elsevier, June 2006.

[14] MacKay, David . "Chapter 20. An Example Inference Task: Clustering". *Information Theory, Inference and Learning Algorithms.* Cambridge University Press. 2003. pp. 284–292.

[15] P. Felzenszwalb and D. Huttenlocher. *Distance transforms of sampled functions.* Technical Report TR2004-1963, Cornell Computing and Information Science, 2004.

[16] Ming-Yu Liu, Oncel Tuzel, Ashok Veeraraghavan, Rama Chellappa. *Fast Directional Chamfer Matching*. CVPR 2010.