Proceeding of the IEEE
International Conference on Robotics and Biomimetics (ROBIO)
Shenzhen, China, December 2013

Workpiece Localization with Shadow Detection and Removing

Zhengke Qin, Wenjun Zhu, Peng Wang and Hong Qiao

Abstract—**this paper presents a new approach to detect and remove the shadows for workpiece localization, which is with an extensive application in automatic assembly system. However, the shadows of workpiece will badly affect this procedure as the contour of the shadow has the same shape with the workpiece itself in the image. The localization system treats the shadow as a part of the workpiece and make incorrect decision. So removing the shadow in the image before localization is meaningful. Our approach use CAD model to estimate the pose of workpiece, and the contour of object can be drawn in the image. Gray and texture features are used to detect and remove the shadow around the workpiece, and the workpiece is localized without the disturbance of the shadow in image. Experiments have been designed and performed. The experimental results demonstrate the effectiveness of the proposed method.**

I. Introduction

There are many approaches to remove the shadow in image in previous research. These approaches are almost used in improving the performance of pedestrian or vehicle detection because both the body and the shadow are moving and they have the same feature of movement [1, 2]. During the last ten years, there are mainly four categories about shadow detection: chromacity, physical, geometry and textures [3].

Many shadow detection methods based on chromacity use color information, these methods are based on this assumption: the shadow will decrease the intensity of region but remain the chromacity of region unchanged [4] or changed follow linear attenuation model [5]. These methods often use some color models (e.g. HSV [6]) to separate the chromacity from intensity. In the condition of pure white light, when the intensity of light changes, all the points in the chromacity map change along the same direction. They found a line perpendicular to this direction and projected all the points in the chromacity map to this line. As a result, we produced a "gray image" which is independent to the changing of intensity.

The linear attenuation model assumes that the illumination is pure white light. The methods based on physical properties are similar to the methods based on chromacity. They use the linear attenuation model with consideration about multi and various illumination sources [5]. Besides the linear attenuation model, some methods use the non-linear attenuation which consider all pixel in the image [7]. This method is more accurate by learning the particular scenarios [8, 9]. Both of the methods are based on the condition: pure white light. These two kinds of methods are suitable for outdoor with natural pure white light and colorful environment such as glass. However, there is almost no pure white light in the industrial environment and almost all the elements of industrial environment is not colorful (they are gray as). Therefore it is hard to separate the color information from the intensity information. It is hard to detect the shadow from the scenario. So the methods based on chromacity and physical properties are not appropriate for industrial environment.

The methods based on geometry information are that: If the geometrical shape, the distribution of illumination and the parameters of the plane are all known, we can compute the shape of the shadow [10, 11]. These methods are used in specific conditions for specific object types such as vehicle detection [12]. The geometrical shape of the objects in the industrial environment are usually known, but the distribution of the illumination could be very complicated as there are many illumination sources.

The methods based on texture information: The shadow often has the same texture as background in the image. So the shadow can be classified base on texture correlation. There are various methods of correlation such as cross correlation [13, 14] and gradient or edge correlation [15]. The methods based on texture information have the highest performance among the four kinds of methods in shadow detection, but need the longest time for computing. Automatic assembly is a real-time system so it is better to accelerate the procedure of the shadow detection.

From the above, these four kinds of methods may not be suitable for industrial environment, so this paper propose a new approach to detect and remove the shadow of workpiece in the industrial environment. Our method uses the geometry and the texture information, which is more like combination of the methods based on geometry and texture properties.

The objects of this approach proposed in this paper are industrial workpieces. Recognizing and estimating the pose of the workpieces is an extensive application in automatic assembly system. However, the shadow of the workpiece will negatively affect this procedure because the contour of the shadow of workpiece has the same shape as the workpiece itself in the image. The shadow tends to be classified as part of the workpieces, as Fig.1 shows, which makes the automatic assembly system unreliable. The pose of the workpiece is estimated well after detecting and removing shadow, as Fig. 1 shows (the red curves in Fig. 1 represent the result of pose estimation). So removing the shadow of the workpiece in the image is meaningful and important in industrial application. Crankshaft is a component of the engine and a classical workpiece. So we choose crankshaft as the object of the experiment.

The authors are with the Research Center of Precision Sensing and Control, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, PRC(e-mail:wenjun.b.zhu@gmail.com; qinzhengke@hotmail.com; peng_wang@ia.ac.cn; hong.qiao@mail.ia.ac.cn).

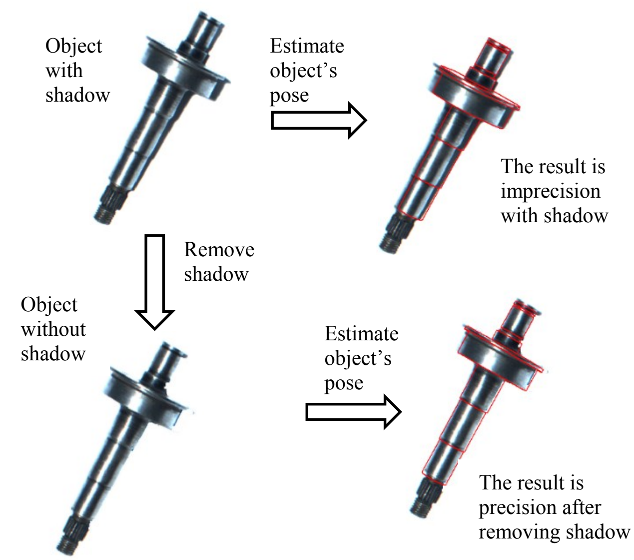

Fig.1. Estimate the pose of crankshaft before and after removing shadow.

In section 2, we will describe the overview of our approach, preprocessing, how to estimate the pose of object, the features of the shadow and how to remove detected shadow. In section 3, we will discuss the details of the experiment, including the programming environment, the robot system, the vision system, the arguments and the analyzing the result of the experiment.

II. ALGORITHM

A. Overview

According to the discussion in section 1, it is learned that: (1) detecting and removing the shadow can increase the precision of the pose estimation of the object. The shadow always appears around the object. If the pose of object is precisely known, then the contour of object can be inferred and the region where the shadow will probably appear can be limited. (2) So the increase of the precision of pose estimation of the object will increase the precision of detecting the shadow. The procedures of pose estimation of object and shadow detection are highly related and are inter-dependent on each other. Increasing one of them will increase the other. So we can improve both of them based on this logic, until neither of them can be improved any more. Because our final objective is to detect the shadow, we don't have any information about the shadow. Fortunately, we have the initial information about the pose of the object in the image (Even though the estimated pose of the object is not precise because of the existence of the shadow). Shadow of the object can be detected based on the initial location of the object. In turn the precision of localization of the object will be improved after removing the shadow. Repeating as a circle, both the precision of location of object and shadow can be improved to a high level.

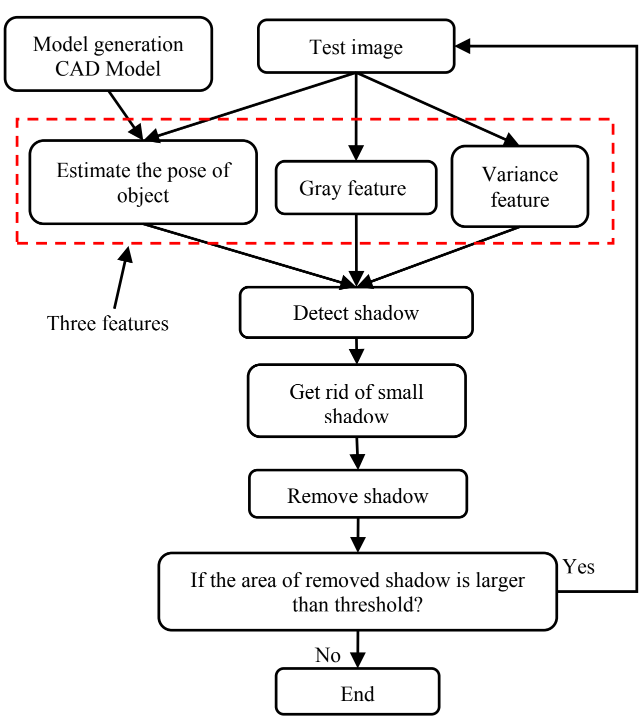

Fig. 2. Flow diagram of algorithm

B. Preprocessing

The original image contains the object which needs to be located with the shadows. Canny operator [16] is used to extract Edge of the original image. It is worth noting that we set the low and high threshold of Canny operator so low that what we get is not an image of "edge" but an image of "texture". We call result image of Canny operator "texture image". As Fig. 3 (b) shows, noting that the region of the shadow in the texture image is cleaner than other regions of the object.

C. Estimating the pose of object

The method to estimate the pose of object is based on an approach which is called "Feature based Aspect-Tree, Generation and Interpretation" proposed by Olaf Munkelt[17-18].

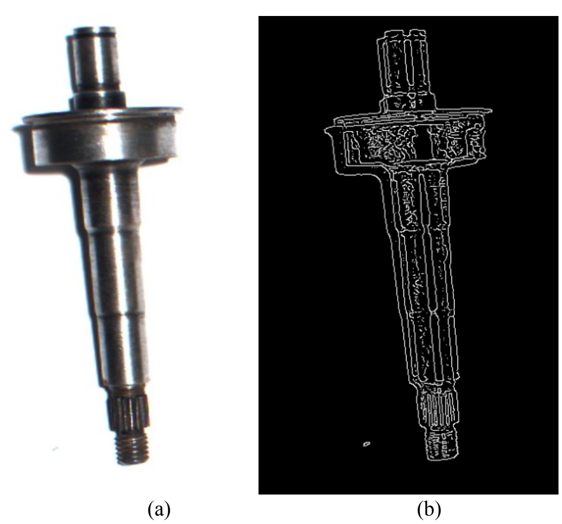

Fig. 3. (a)Original image; (b) Extract the "texture image" by the Canny operator

How to estimate the pose of object is not the main objective of this paper, so we introduce the method in a simple way.

First, the CAD model of the object must be developed. It is easy to achieve because almost all the workpiece have CAD model in industrial application. Second, image features are extracted, including regions, textures and so on. Third, the correspondence between these features and CAD model features is established. And then, from these correspondence we derive object hypotheses, which also include a rough estimation for the six degrees of freedom of the attitude of the object. At last, these hypotheses are subsequently verified and refined by a module for pose estimation. More details can be found in these references [17, 18].

D. Features of the shadow

Three features are used in our method: the average gray value of the original image, the variance of the texture image and the initial contour of the object.

1) The average gray value of original image

For each pixel of the original image, the average gray value of its $N_A \times N_A$neighborhoods is calculated. We calculate the average gray value of all the channels if the image is colorful. As shown in Fig. 3 (a), the gray value of background is very high, the gray values of the region of the object and the shadow are relatively low. Actually it is a procedure of smoothing.

$$g(x,y) = \frac{1}{{N_A}^2} \sum_{x-\frac{N_A}{2}}^{x+\frac{N_A}{2}} \sum_{y-\frac{N_A}{2}}^{y+\frac{N_A}{2}} I(x,y) \quad (1)$$

$g(x,y)$ is the average gray value in pixel (x,y), $I(x,y)$ is the gray value in pixel (x,y).

2) The variance of texture image

For each pixel of the texture image, the variance of its $N_V \times N_V$ neighborhoods is calculated. On an intuitive level, if the region of image is smooth, the variance of the region will be low. On the other hand, if the region of image is "rough", the variance of the region will be high. We can speculate that the variance of the region of the shadow is low because the region of the shadow is smooth as Fig. 3 (b) shows. This is the equation of variance.

$$v(x,y) = \frac{1}{{N_V}^2} \sum_{x-\frac{N_V}{2}}^{x+\frac{N_V}{2}} \sum_{y-\frac{N_V}{2}}^{y+\frac{N_V}{2}} [I(x,y) - g(x,y)]^2 \quad (2)$$

$v(x,y)$ is the variance value in pixel (x,y).

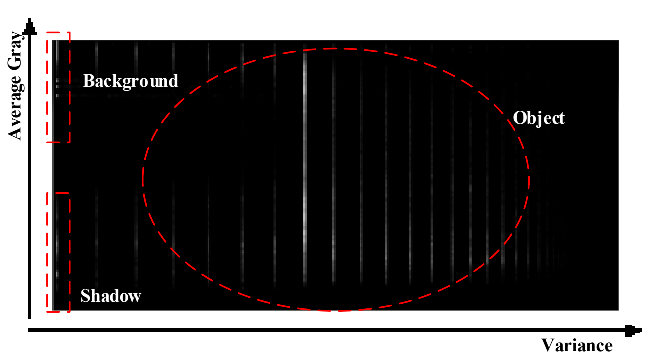

Fig. 4. Feature distribution image

Fig. 4 is the feature distribution image. Each point of this image corresponds to certain pixels in the original image. In the feature distribution image, the horizontal axis represents the variance of pixels in texture image, the vertical axis represents the average gray value of the pixels in the original image. The intensity in the feature distribution image represents the number of points which contain current feature. The more points that have certain feature, the point corresponding the feature will be lighter in the feature distribution image. Because the points of the shadow have the same feature, so we can deduce that the these points gather at a certain point in the feature distribution image, so are the points in the region of the object and the background. In the feature distribution image, we can see that the three regions are marked by rectangles and ellipses. The points in the left-top rectangle correspond to the background because the background is bright and smooth. The points in the left-bottom rectangle correspond to the shadow because the shadow is dark and smooth. The points in the right ellipse correspond to the object because the region of object is dark and rough.

3) Contour of the object in image

A rough contour of the object can be obtained by pose estimation (even though the pose is imprecise). Because the shadow will only appear around the object, we can select a region of the margin of the object as Fig. 5 (b) shows. We only search for the shadow in the selected region. The width of the region will be changed for different object.

E. Removing the Shadow

In the industrial environment, the background is relatively simple. The background of industrial image is often bright, enough as the illuminations are sufficient. We use the almost pure white color as the background. So removing the shadows after detecting them is simple: just set the gray value of the region of the shadow to white color as Fig. 1 shows. More background type would be tested in future work.

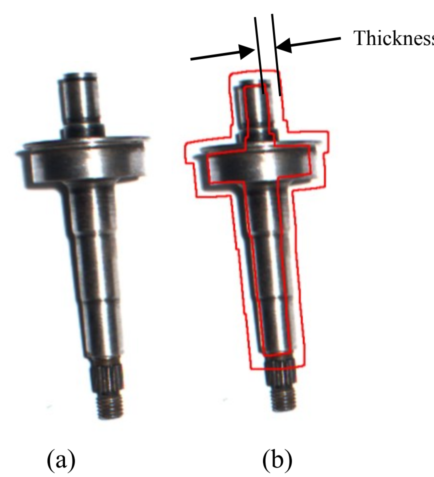

Fig. 5. (a)Original image of crankshaft; (b) Region among the contour of the crankshaft. The shadow is likely to appear in the region which is surround by red curves.

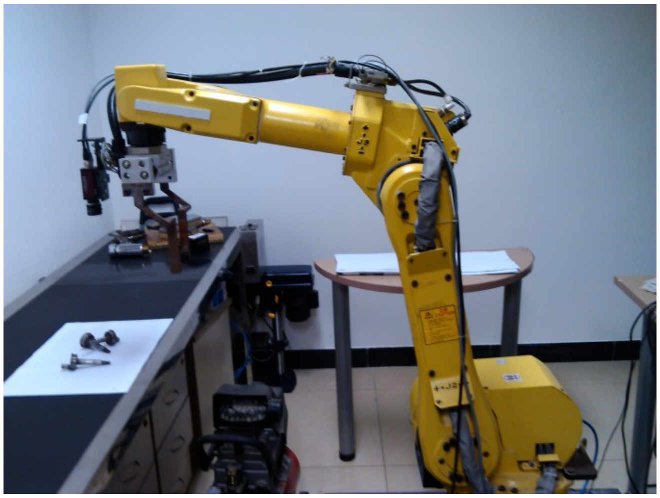

Fig. 6. The robot and vision system

III. EXPERIMENT

A. Robot and Vision System

Our objective to detect shadow is to improve the precision of estimating the pose of the object. And the pose estimation of the object is applied in robot-grab system. So it is necessary to introduce the whole robot and vision system.

1) The robot system

The FANUC 6-DOF (DOF: degree of freedom) industrial robot whose product model is R-J3iB is used in experiment. The repeat accuracy of the robot is ± 0.08mm. The maximum load is 6kg. All the joints are driven by servo motor. There are 6 degree of freedoms with 3 of them are translational and other 3 are rotational. The control system of the robot consists of a microcomputer, servo system and the I/O system. The robot is used in assembling, welding, carrying, spraying and so on.

2) The vision system

The vision system is consist of an industrial camera and a microcomputer. They are connected with gigabit Ethernet. The product model of the camera is Manta201c. The resolution of the camera is 1624×1234. The focal length of the camera lens is 5mm.

B. Programming Environment

The program languages are C++, Halcon, together with OpenCV 2.4 Computer Vision Library. The Operate System is Linux (Ubuntu13.04).

C. Arguments in Experiment

There are some arguments in the experiment that we need to consider about.

1) The high threshold T_H and the low threshold T_L of Canny operator.

As described in section 2, if the high and low threshold are very low, the result image of Canny operator can reflect the degree of "texture" which is used as a feature of the shadow. So we set the value of the two thresholds in a low level, see the value in Tab 1.

TABLE 1
VALUE OF ARGUMENTS

Symbol	Quantity	Value	Unit
$\mathrm{T_H}$	High threshold of Canny operator	15	Gray
$\mathrm{T_L}$	Low threshold of Canny operator	10	Gray
N_A	Neighbor in calculating the average gray	5	Pixel
N_V	Neighbor in calculating the variance	5	Pixel
T_K	Thickness of region where shadow may appear	20	Pixel
T_A	Threshold of area of shadow	120	Pixel

Tab. 1. The unit Gray use 256 levels to present the intensity from the darkest to the brightest.

2) The number of neighbors which is used in calculating the average gray N_{A} and variance NV.

The average gray value of pixels in the image is used to reflect how dark the region is. The background of industrial image is often bright enough as the illuminations are sufficient. So in most conditions, the shadow is darker than the background.

Treating the current pixel as the center, the average gray value of $N_A \times N_A$ neighbors around it is calculated. This average gray is one of the features of the shadow. If N_A is too small, the result will be influenced by noise, and if N_A is too large, the feature will be lack of discrimination. The value of N_A is specified by experience. See Tab 1.

Treating the current pixel as the center, the variance of gray value of $N_V \times N_V$ neighbors around it is calculated. This variance is another one of the features of the shadow. If N_V is too small, the result will lost the texture information (we can't get texture information in only a few pixels), and if N_V is too large, the feature will be lack of discrimination. The value of N_V is specified in experience. See Tab 1.

3) The Thickness T_{K} of the region where the shadow probably appear.

As Fig. 5 (b) shows, the shadow always appears around the object. So after the initial pose of the object was estimated, we can extend the contour of the estimation to a region. The argument thickness T_K describes how thick the region will be. If T_K is too small, we will lose some shadow and if T_K is too large, the effect of region limit will be weakened. The value of T_K is chosen by experience, see Tab 1.

4) The threshold of area of the shadow T_A:

The area of the shadow in image is usually not too small, so we can use a threshold T_A to filter some small "shadow" which is actually noise but not shadow. The value of T_A relies on the area of the object.

D. Result of Experiment

50 images have been tested and 47 of them are successful. The detection rate is 94%. Two groups of samples are selected as Fig. 7 shows. The red contour represent the result of pose estimation. The precision of pose estimation is high if the red contour is consistent with the crankshaft. It is easy to notice that the pose estimation of crankshaft was imprecise before with the shadow in image. Obviously, the estimation is improved after detecting and removing the shadow.

Fig. 7. The result of detecting the experiment. There are two groups. For each group, from left to right, the first one is the original image, the second one is pose estimation of crankshaft before removing shadow, the third is detection of the shadow, and then the image after removing shadow, and the last one is pose estimation of crankshaft after removing shadow.

Fig. 8. The result of other samples. The left one of each picture is original image, the middle one of each picture shows detection of the shadow and the right one of each picture shows image after removing shadow

Some results of the detection are erroneous as Fig. 8 shows, certain regions in the image are regarded as shadow, and actually, these regions are part of crankshaft. The reason may be some regions of crankshaft has the same features and these regions are big enough. Fortunately, this shadow over-detection does not affect to the estimation of object because the estimation uses the edge information.

IV. CONCLUSION

This paper presents a new approach to detect and remove the shadow in the image while estimating the pose of the object. This approach is suitable for industrial environment, especially in automatic assembly system which is different from previous work. This paper proposes three features of the shadow: the average gray value of original image, the variance of texture image and the contour of object in image. We perform an experiment using crankshaft as sample. 50 images have been tested and the detection rate is 94%. This approach can detect the shadow of crankshaft and improve the precision of estimating the pose of the object after removing the shadow.

ACKNOWLEDGMENT

This work was partly supported by the NNSF (National Natural Science Foundation) of China under the grants 61100098, 61379097, 61033011 and 61210009.

REFERENCES

[1] S. Nadimi and B. Bhanu. “Physical models for moving shadow and object detection in video”. IEEE Transactions on Pattern Analysis and Machine Intelligence, 26(8):1079–1087, 2004.

[2] B. Mitra, R. Young, and C. Chatwin. “On shadow elimination after moving region segmentation based on different threshold selection strategies”. Optics and Lasers in Engineering, 45(11):1088–1093, 2007.

[3] A.Sanin, “Shadow detection: A survey and comparative evaluation of recent methods”. Pattern recognition, 2012, 45(4): 1684-1695.

[4] T. Horprasert, D. Harwood, and L. Davis. “A statistical approach for real-time robust background subtraction and shadow detection”. In IEEE ICCV’99 Frame-Rate Workshop, 1999.

[5] J.B. Huang and C.-S. Chen. “Moving cast shadow detection using physics-based features”. In IEEE Conference on Computer Vision and Pattern Recognition, pages 2310–2317, 2009.

[6] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati. Detecting moving objects, ghosts, and shadows in video streams. IEEE Transactions on Pattern Analysis and Machine Intelligence, 25(10):1337–1342, 2003.

[7] N. Martel-Brisson and A. Zaccarin. “Kernel-based learning of cast shadows from a physical model of light sources and surfaces for low-level segmentation”. In IEEE Conference on Computer Vision and Pattern Recognition, pages 1–8, 2008.

[8] A. Joshi and N. Papanikolopoulos. “Learning to detect moving shadows in dynamic environments”. IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(11):2055–2063, 2008.

[9] Z. Liu, K. Huang, T. Tan, and L. Wang. “Cast shadow removal combining local and global features”. In IEEE Conference on Computer Vision and Pattern Recognition, pages 1–8, 2007.

[10] C.-C. Chen and J. Aggarwal. “Human shadow removal with unknown light source”. In International Conference on Pattern Recognition, pages 2407–2410, 2010.

[11] L. Z. Fang, W. Y. Qiong, and Y. Z. Sheng. “A method to segment moving vehicle cast shadow based on wavelet transform”. Pattern Recognition Letters, 29(16):2182–2188, 2008.

[12] J.-W. Hsieh, W.-F. Hu, C.-J. Chang, and Y.-S. Chen. “Shadow elimination for effective moving object detection by Gaussian shadow modeling”. Image and Vision Computing, 21(6):505–516, 2003.

[13] Y. Tian, M. Lu, and A. Hampapur. “Robust and efficient foreground analysis for real-time video surveillance”. In IEEE Conference on Computer Vision and Pattern Recognition, volume 1, pages 1182–1187, 2005.

[14] O. Javed and M. Shah. “Tracking and object classification for automated surveillance”. In European Conference Computer Vision, volume 4, pages 343–357, 2002.

[15] A. Sanin, C. Sanderson, and B. Lovell. “Improved shadow removal for robust person tracking in surveillance scenarios”. In International Conference on Pattern Recognition, pages 141–144, 2010. DOI: 10.1109/ICPR.2010.43.

[16] C. John. "A computational approach to edge detection.” IEEE Transactions on Pattern Analysis and Machine Intelligence, 6 (1986): 679-698.

[17] M. Olaf. "Aspect-trees: Generation and interpretation." Computer Vision and Image Understanding 61.3 (1995): 365-386.

[18] M. Olaf, C. Zierl. "Fast 3-D object recognition using feature based aspect-trees." Pattern Recognition, 1994. Vol. 1-Conference A: Computer Vision & Image Processing, Proceedings of the 12th IAPR International Conference on. Vol. 1. IEEE, 1994.