Edge-based Object Tracking for Dynamic Projection Mapping

Yuki Morikubo
Faculty of Informatics and Engineering
The University of Electro-Communications
Tokyo, Japan
morikubo@ims.cs.uec.ac.jp

Naoki Hashimoto
Graduate School of Informatics and Engineering
The University of Electro-Communications
Tokyo, Japan
naoki@cs.uec.ac.jp

***Abstract*—We propose an edge-based object tracking for realizing dynamic projection mapping. In order to achieve projection to moving targets, we have to obtain 3D posture of its targets in real time and with high accuracy. There are a lot of approaches which tackle this issues, and among these approaches a camera-based 3D pose estimation is suitable because this approach does not impair the appearance of targets. However, the approaches using depth information or feature points has trouble processing in real-time because of their large calculation costs. Therefore, we estimate 3D posture from 2D images by using the contour of the target object that we can obtain quickly by edge detection. And at this time, by using edge continuity without complex analysis of edges, we prevent the decline of the tracking accuracy caused by the extra edges detected from target's surface and surroundings. Consequently, we realize a real-time and high accuracy tracking for dynamic projection mapping applicable to various objects.**

Keywords— dynamic projection mapping; edge-based tracking; motion prediction; edge continuity

I. Introduction

Recently, dynamic projection mapping which projects images to real target object moving freely has been attracting considerable attention as the technique proving fantastic appearance editing. In the field of entertainment such as a dance and a stage play, projection to moving target can expand the range of expressions, and they are being introduced gradually. Also, this projection technique has been expected to be applied for a wide range of fields, for example reviewing the design in the site of product development, improving education by understand internal structure of objects and supporting medical practice by high presence simulation.

However, the image projection to moving target requires estimating 3D posture of the target in real-time and with high accuracy. Although previous approaches have used a magnetic sensor or an optical motion capture system for obtaining the posture, these approaches are not suitable for projection mapping because attaching sensors or markers impair the appearance of the targets. Also, camera-based approaches using depth information or feature points of the target without attaching sensors or makers disturb real-time processing because feature point detection and alignment require large calculation time.

978-1-5386-2615-3/18/$31.00 ©2018 IEEE

Therefore, in this research, we propose a camera-based object tracking method using contours of target objects that is suitable for real-time processing because the contour is obtained at high speed by edge detection. We making this tracking more robust against the shape of the target object and the environment outside the target while maintaining the high-speed property of the edge detection. Consequently, we realize dynamic projection mapping that can be applied to various target objects and usable as realistic image expression technique.

II. Previous Worls

As a camera-based non-contact measurement approach, some researchers estimate 3D posture of target object using the 3D shape obtained by RGB-D camera [1][2]. Although RGB-D camera can obtain the 3D shape easily based on depth information, instability and time-delay of that camera cause the decline in accuracy of estimation and the misalignment of projected images on the moving targets. On the other hand, Sticky projections [3] obtains the 3D shape of the target by using feature points and estimates its posture. However, real-time processing is difficult because large processing time is required for feature point extraction and alignment in 3D. Some people use a special high-speed vision and mirror-based optical axis controller for realizing a dynamic projection mapping with small misalignment of projected images [4]. However, in this method, target objects are limited to a simple shape such as a sphere and the operator can move the object only in a plane.

As a method using edges that can be obtained at high-speed, there are classic method named RAPiD [5] and that of using line segment [6]. However, these approaches limit the target object to those having a linear shape because the target have to have clear edges that is not significantly affected by the posture of the target, in order to estimate 3D posture from 2D images. Moreover [5] declines accuracy of estimation due to influence of edges around the target object, and [6] disturbs real-time processing because of line segment extraction.

III. Proposed Method

In this research, we propose real-time object tracking using the contour of the target object and extends the target to general objects without clear edges that have trouble estimating posture in conventional methods. Contours detected from the target do not be guarantee to have clear edges that can be obtained stably without significantly affected by posture of the target object. However, in most object, there are contours corresponding to each posture of its object. Therefore, we make the contour detected from known 3D model of the target correspond to the 2D edge of target object in captured image and estimate the 3D posture from 2D image by minimizing the distance between the corresponding points.

There is a difference between the contours of the previous frame used for corresponding and the current frame estimated after corresponding, because the object contour changes very sensitively according to the posture. On the other hand, for the correct correspondence, we have to detect the contour from the 3D model whose posture is close to the posture of the object in the currently captured image. Then, we use the 3D model of which posture is estimated with the previous frame by using a motion prediction in the process of detecting the contour of the 3D model. The overview of our proposal is shown in Fig. 1.

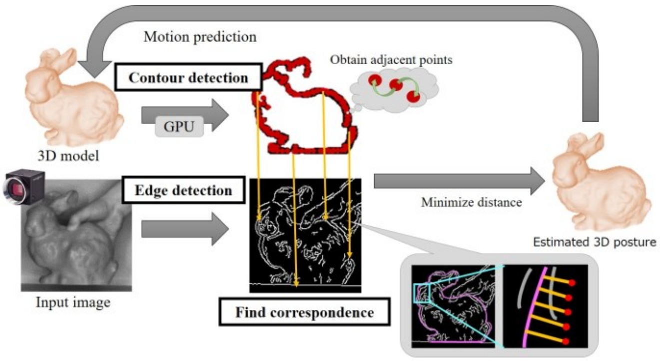

Fig.1. Process flow of our edge-base tracking

In order to realize object tracking applicable to general objects, instead of detect clear edges such as linear edges, we detect general edges. For this reason, there are not only the contour of target but also a lot of extra edges in the input image obtained by edge detection in Fig.1, and many candidates of the correspondence are generated. These candidates cause miss correspondence and decline accuracy of the estimation. Then, as shown in the find correspondence process of Fig.1, we find correct correspondence between the contour of a 3D model and the edges of the target in a input image by using the continuity of contour, while suppressing an increase in calculation processing of edge analysis. These processes will be described later in detail.

A. Detect contour point and 2D edge

In the process of detecting 3D contour points, in order to extend edge-based tracking to deal with general objects without clear edges, we generate sampling points uniformly for the whole 3D model. The contour is detected by obtaining the sampling point where the luminance gradient is larger than a threshold value because the luminance gradient is generally high at the edge of the 3D model. At that time, in order to make correspondence to 2D edges in a capture image, 3D contour points are mapped to 2D points in the 2D edge space and these points are called contour points. On the other hand, the 2D edges is extracted by using Canny edge detector. For the fast processing, the contour point detection is performed on the GPU, moreover, contour points detection and 2D edges extraction are performed in parallel.

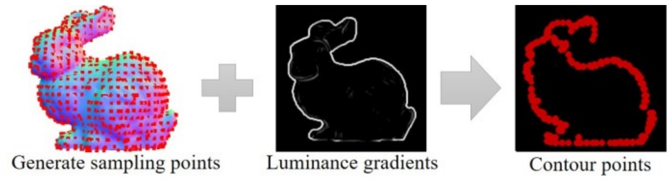

Fig.2. Detecting contour point

In order to obtain the contour of target that close to the contour in the currently captured image, we use the 3D model of which posture are estimated with the previous frame by using a motion prediction. We use sequential least squares method with forgetting factor for motion prediction. By using forgetting factor, the weight of the current frame increases. As a result, we can predict the posture robustly against rapid change of movement. This prediction is performed on the 3D position and the quaternion.

B. Find correct correspondence

We make the contour points correspond to 2D edges of the target object. Hereafter, 2D edges are defined as the edges detected from the captured image containing the target. The 2D edges also contain extra edges detected from unevenness of the target surface and external object around the target such as a user's hand. In order to find correspondence to 2D edges under the circumstances, we regard edges near the contour point as the candidates of the correspondence. Then, considering that the contour of the target is continuous, we select candidates on the same edge. For these processes, we prevent miss correspondence with extra edges without complicated analysis of the 2D edges. As a result, we find correct correspondence with the contour of the target. The process of the finding correspondence is shown in Fig.3.

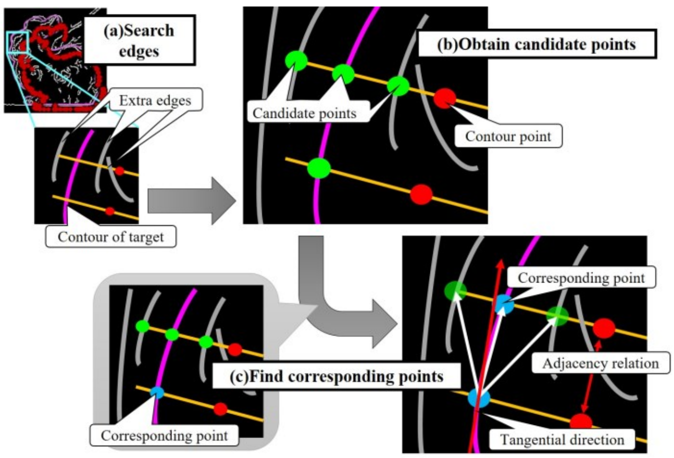

Fig. 3. Process flow of finding correct correspondence

In Fig.3, first of all, we search for 2D edges within the range close to the contour points along the normal direction of the contour point and we obtain points on the 2D edge as candidate point of corresponding. By searching this 2D edges, we can obtain candidates containing the contour of target object in the input image because contour points are projected to near the contour of the target in the input image by motion prediction. At that time, we can decline the total number of candidate points due to search range is limited to a small range. The search range is adjusted to an appropriate value according to the speed change of the target object, and perform as a mask while keeping the contour of the target of 2D edges in the range. Moreover, candidate points are restricted to the points that the angle between the search direction and the normal direction of the 2D edge at the candidate point is close. This restriction excludes extremely wrong candidates and limits the candidate points to realistic ones.

Next, the only candidate point for the contour point is decided as a correct corresponding point existing on the contour of the object. Based on this corresponding point, we select other correct corresponding point from the candidate points. As shown in the obtaining candidate points of Fig.3 (b), input image contain extra edges such as surface of object and user's hand besides the contour of the target object, and there are candidate points on each edges. The candidate point that is correct corresponding point on the contour of the target is obtained by selecting the candidate point on the same edge as edges where known corresponding points exist shown as pink line in Fig.3.

At that time, corresponding points can be found by searching along the continuous direction of 2D edges from known corresponding points. However, as mentioned before, this searching disturbs real-time processing because analysis of edge direction requires high calculation cost. Also there is a possibility that the edges may not be continuous or be continuous with extra edges due to lighting condition and environment surrounding the target. Therefore, considering adjacent contour points are acquired very closely, we approximate the direction of contiguous contour to tangential direction of the edge, and obtain the point closest to that direction as correct corresponding point, as shown in Fig.3 (c).

This process goes from the contour point having the known correspondence point to the adjacent point one after another, and we acquire corresponding points of all contour points having candidate points

IV. EXPERIMENTAL RESULTS

We implemented our proposed algorithm, and achieved real-time and robust 3D posture estimation. Also we achieved dynamic projection mapping by using that algorithm. As shown in Fig.4, we used a real mannequin object and a bunny object created by using a 3D printer. Their original 3D CAD models were used for the tracking process as a reference 3D model. A projector (1920 x 1080 pixel, 5200 lm), an IR camera (640 x 480 pixel, 337 fps, IR band-pass filter 850nm) and IR lights (800 - 900 nm) are placed at 1m away from the targets. By using the web camera (1920 x 1080 pixel), the intrinsic and extrinsic parameters of the projector and the IR camera system can be well estimated with a prior processing.

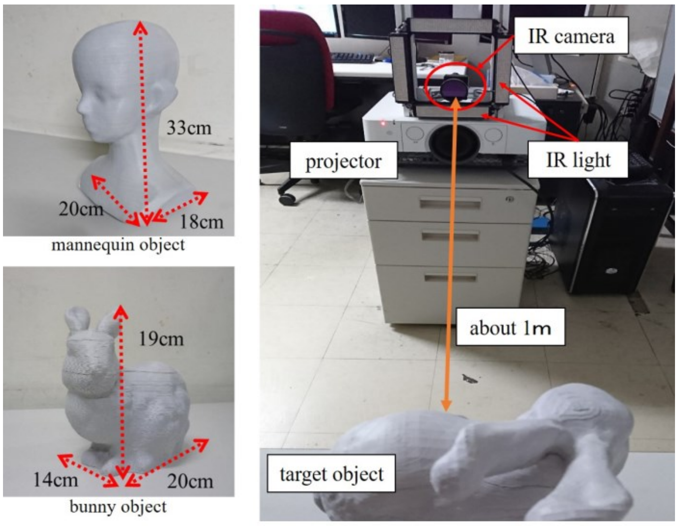

Fig.4. Experimental environment

A. Processing speed

We measured a processing time taken for the main processing in this method, and compared with [6] and our previous proposal [7]. The result is shown in Table.1. In the measurement of this method, we calculated the average of each processing time for 3000 frames when the bunny target was freely moved. As shown in Table.1, our approach is sufficiently faster than the approach of line segment extraction [6]. Also, the processing speed of our proposed method is about the half of the update rate of the projector which is 60Hz (about 16.6msec). Considering the compensation of the projection delay by using motion prediction, although the processing speed is slightly slower than the conventional method, it can be said that it is also sufficiently fast.

Table.1. Processing time of each edge-based tracking

	[6] with many line segment	[6] with few line segment	Conventional method[7]	Proposed method
Edge or contour detection (/msec)	344	37	4.1	4.5
Finding Correspond (/msec)	182	4	1.9	3.6
Posture estimation (/msec)	39	13	0.1	0.1
Total (/msec)	690	60	6.1	8.4

B. Results of tracking and projection results

We visualize the correspondence results as shown in Fig.5 and subjectively confirm whether the correspondence is correctly made. Fig.5 shows the measured results in which the target is moved freely. Red points show contour points and are connected to corresponding points by yellow lines. In spite of the situation where many extra edges are detected from the operator's hand and the surface of the object, it is find that we realize correct correspondence to the contour of target. In other words, we achieved robust tracking to the shape of the object and its surroundings

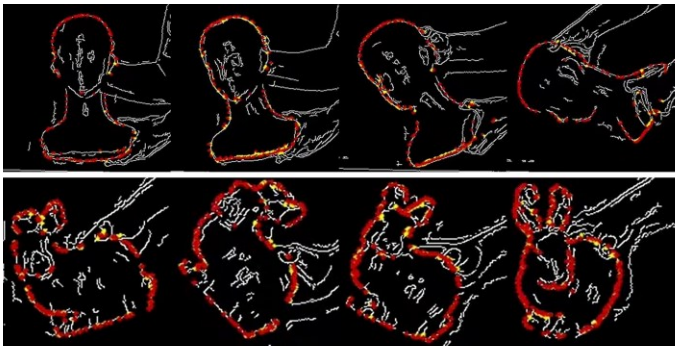

Fig.5. Results of tracking with moving object

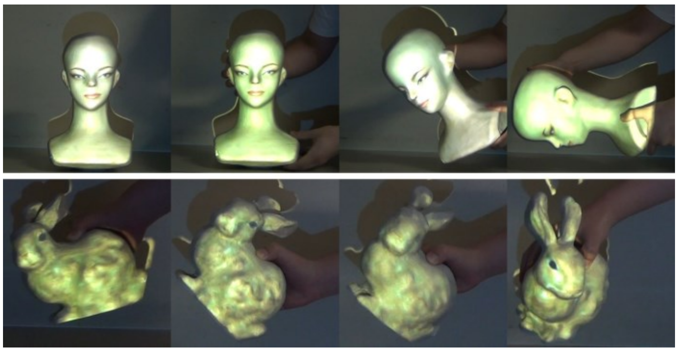

Fig.7. Results of projection onto moving objects

In Fig.6, we compare proposed method with conventional method of corresponding the nearest edges. Fig.6 shows frames that shows the results of correspondence when the target object is moved in the same way in each method. We confirmed robust tracking with fewer miss correspondence than the conventional method, under the circumstances containing many extra edges.

However, if external objects like hand shield the contour of the target, and that of edges exists in a direction close to the contiguous direction of the contour, miss correspondence may occur. In the future we consider shielding of external objects by effectively using the luminance gradient.

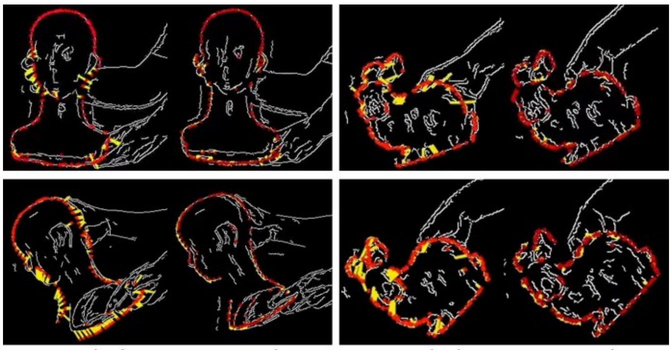

Fig.6. Comparison with conventional method

Moreover, we applied this tracking method to target objects and projected onto the moving target as shown in Fig.7. In Fig.7, the target is moved freely by user's hand and projected the texture according to the posture of themselves. It is find that this projection is sufficient to be used for actual image expression technique because the projection performed without being disturbed by the surroundings of the object and applicable to target object of various shapes.

However, our method cannot adapt to object of completely objective shape because the contour does not change according to the posture. In order to apply to more various target, we consider detecting edges from large undulating parts in 3D model and use the correspondence of them.

V. CONCLUSION AND FUTURE WORK

We proposed real-time and robust object tracking using contours of target objects for the dynamic projection mapping. This method can provide more robust projection to moving target than conventional method. In near future, we will achieve more robust tracking against the shielding of the contour by external objects. Also we will evaluate the tracking accuracy with objective indices.

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number JP16K00267.

REFERENCES

[1] Kazuna Tsuboi, Yuji Oyamada, Maki Sugimoto and Hideo Saito, "3D object surface tracking using partial shape templates trained from a depth camera for spatial augmented reality environments," the Fourteenth Australasian User Interface Conference (AUIC), Vol. 139, pp. 125-127 , (2013).

[2] Daisuke Kobayashi and Naoki. Hashimoto, "Spatial augmented reality by using depth-based object tracking," ACM SIGGRAPH 2014 Posters, p.33:1 (2014).

[3] Christoph Resch, Peter Keitler and Gudrun Klinker, "sticky projections: A new approach to interactive shader lamp tracking," Mixed and Augmented Reality (ISMAR), pp. 151-156 (2014).

[4] Kohei Okumura, Hiromasa Oku and Masatoshi Ishikawa, "Acitve projection AR using high-speed optical axis control and appearance estimation algorithm," IEEE International Conference on Multimedia and Expo (ICME), pp. 1-6, (2013).

[5] Chris Harris, "Tracking with rigid models," Active vision, chapter 4, pp. 59-73 (1992).

[6] Yusuke Nakayama, Hideo Saito, Masayoshi Shimizu and Nobuyasu Yamaguchi, "Marker-Less Augmented Reality Framework Using On-Site 3D Line-Segment-basedModel Generation," Journal of Imaging Science and Technology,pp.020401-1-020401-24 (2016).

[7] Naoki Hashimoto, Ryo Koizumi and Daisuke Kobayashi "Dynamic Projection Mapping with a Single IR Camera," International Journal of Computer Games Technology, vol. 2017, Article ID 4936285, 10 pages (2017).