Swimmer Position Estimation by Lane Rectification

Takayuki Tsumita[a], Hidehiko Shishido[b], Itaru Kitahara[b], Yoshinari Kameda*[b]
[a]Graduate School of Systems and information engineering, University of Tsukuba,
1-1-1 Tennoudai, Tsukuba, Ibaraki, Japan 3058573;
[b]Center for Computational Sciences, University of Tsukuba,
1-1-1- Tennoudai, Tsukuba, Ibaraki, Japan 3058577

ABSTRACT

A new method of estimating swimmer position in swimming pool video is proposed. The video of swimming games is taken from a higher seat row in audience seat area. It can cover the whole field of a swimming pool. The swimming pool video is transformed so that each lane can be analyzed along with the lane direction. The foreground region that includes both the swimmer and their water splash is extracted by adaptive background modeling and by setting the mask region to cope with the influence of the non-planer water surface. Then, based on the color analysis on water splash, swimmer region can be successfully extracted. The position is estimated as the center of the Gaussian distribution of the swimmer region. The proposed method was applied to a nationwide swimming game.

Keywords: computer vision; water splash; pool; video analysis; sport; tracking

1. INTRODUCTION

Precise swimmer position on swimming games across time is one of the fundamental measurements in swimming sport. Since the body part of swimmers are mainly underwater, conventional human tracking methods cannot be directly applicable to swimming situation. In addition, top swimmers may make splashes, and it makes the estimation of swimmer position more difficult.

We propose a new position estimation method that is practically available on swimming games. What we need is just a video(Fig.1) that covers the whole water surface of the pool. It can be obtained by setting a camera at a higher seat row in the audience seat area. We define the body position of swimmers along the lane direction in this article since the swimming instructors and athletes traditionally think the position in lane direction is helpful for their further analysis and discussion.

The technological key points of our approach are the rectification and swimmer region extraction based on the color distribution that is unique to the swimming pool area.

In order to support swimming sport, computer vision has been applied to detect drowning people[1,2]. Precise swimmer tracking can be helpful, but it is not easy to realize it. Swimmer position can be estimated with wearable sensors[3]. This would be useful for the training period, but it will not be applicable to swimming games because of their regulations.

In addition to the position estimation, advanced analyses for specific issues on swimming sport have already been reported[4,5,6,7]. Our approach could be coupled with them.

2. ESTIMATION OF SWIMMER POSITION

2.1 Rectification

Input video should be taken so as to cover the whole area of pool surface. If a lane is set to be processed, four corners should be found in the image. Figure 1 shows a snapshot of a pool where international swimming games are sometimes held.

* kameda@iit.tsukuba.ac.jp , kameda.yoshinari@image.iit.tsukuba.ac.jp Phone +81-29-853-5256

Figure 1. A video image that covers the pool area. (This image is blurred for publication.)

Before starting the proposed video process, camera distortion should be removed. From now on, the video image should be treated as the image taken by an ideal perspective camera. This indicates that all separating lines between lanes are in parallel and the starting ends are on one line and the other ends are on the other line as shown in Figure 2. Note that lanes separating lines may not be parallel to horizontal axis of the video image.

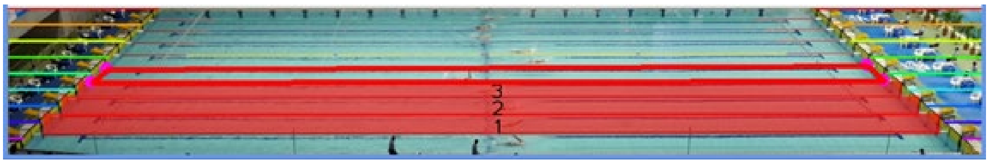

Figure 2. Annotated lanes in a video image taken by a perspective camera

Then by applying a homography transformation that can be defined by giving the four corners of a specific lane in the video image and the four corners of a rectangle, a rectified lane image is obtained. The resultant rectangle might not be similar to the real shape as the horizontal analysis and vertical analysis on the lane image can be independently scaled. Figure 3 shows an example of the lane image. Note that lines marked on the bottom of the bottom seems to be bent because of the refraction of non-planner water surface formed by little waves.

Figure 3. A lane image in rectangle shape.

2.2 Foreground region

The swimmer region cannot be obtained clearly by conventional background removal. Initial foreground region that includes swimmer region should be obtained by estimating the adaptive background mixture model[8] (Figure 4). There are some obstacles including referees, floating line separators that slightly shake by waves, and apparent bent of lines drawn at the bottom of the pool. The mask image is set to locate the regions where these obstacles are frequently observed. The intermediate foreground region is obtained by removing the mask region from the initial foreground region. Figure 5 shows an example of masked region and Figure 6 shows an example of the intermediate foreground region. Note that foreground region may include not only body part of the swimmer but also water splash made by the swimmers, and some part of the body are missed because they are hard to see.

The intermediate foreground region may include some unexpected noise. It may be caused by the bigger shape of the lane separators, or specular reflection of the lighting equipment rarely visible when waves make the water surface change accordingly. On the other hand, we can expect that swimmers should move along the lane direction at a certain speed. A swimmer window is set by estimating the former position of the swimmer and the speed over a short time (Figure 7). The final foreground region is extracted by cutting off the region outside the swimmer window.

Figure 4. Initial foreground region. Lines drawn at the bottom of the pool are also extracted because of their apparent shake caused by the little waves in the bool.

Figure 5. A masked region. The white lines inside corresponds to the lines drawn at the bottom of the pool. Areas close to the top and bottom of the lane image corresponds to the shake of the lane separators. Areas close to the left and right end of the lane image corresponds to the motion of referees.

Figure 6. An example of the intermediate foreground region.

Figure 7. An example of swimmer window in correspondence with Figure 6.

2.3 Swimmer region

The foreground region still includes non-swimmer region that is water splash. The apparent color of the water splash is defined by the color distribution of the light source because the splash is a collection of very small water drops and its result is diffuse reflection. We can assume the lighting condition is usually constant during one race in large swimming games. This means strong water splash can be removed by investigating the color of the light units. Weak water splash that overlaps the body of the swimmer may weaken the body color of the swimmer. To specify the threshold on separating water splash region, we first enhance the color histogram so as to enlarge the gap between the color of the light source and the other colors that are found on swimmer body.

In the second stage, the swimmer region is extracted by the analysis of color distribution. Non-swimmer region in the swimmer window mainly corresponds to water splash and the bottom of the pool, with some reflection of the water surface. It means their colors can be segmented in color space. The swimmer region is extracted by removing the water splash region by the threshold value. The swimmer region is given by a binary description. Figure 8 shows a snapshot of extracted swimmer region. From left to right, a segment that includes main part of the foreground region, its binary representation, and its color representation. The swimmer position is estimated by finding the center of the Gaussian distribution of swimmer region. The horizontal size and the vertical size of the Gaussian distribution is set to fit the size of the swimmer.

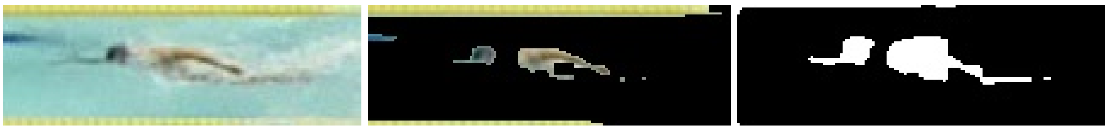

Figure 8. Swimmer region. From left to right, a segment that includes main part of the foreground region, its binary representation, and its color representation.

3. EXPERIMENT

We conducted an experiment on a nationwide swimming games. The video was taken under the permission of the organizer. The camera is Panasonic DMC-GH4 and video was taken at the resolution of 3840 by 2180 at 29.97 fps (Figure 1). After the camera distortion is removed, we apply the proposed method on four styles. Figure 8 shows the result of a free style and a backstroke. Figure 9 shows the result of breaststroke and butterfly. The ground truth is given by manually pointing the body center on the video. The original error is counted by pixel, and it is converted into real unit (centimeter) in the figure. The error tends to be larger on breaststroke and butterfly. We think the reason is the larger action during their strokes may harm the visibility of the swimmer body. On free style and backstroke, not a few body part cannot be visible over strokes, but that does not make a bad influence on the result.

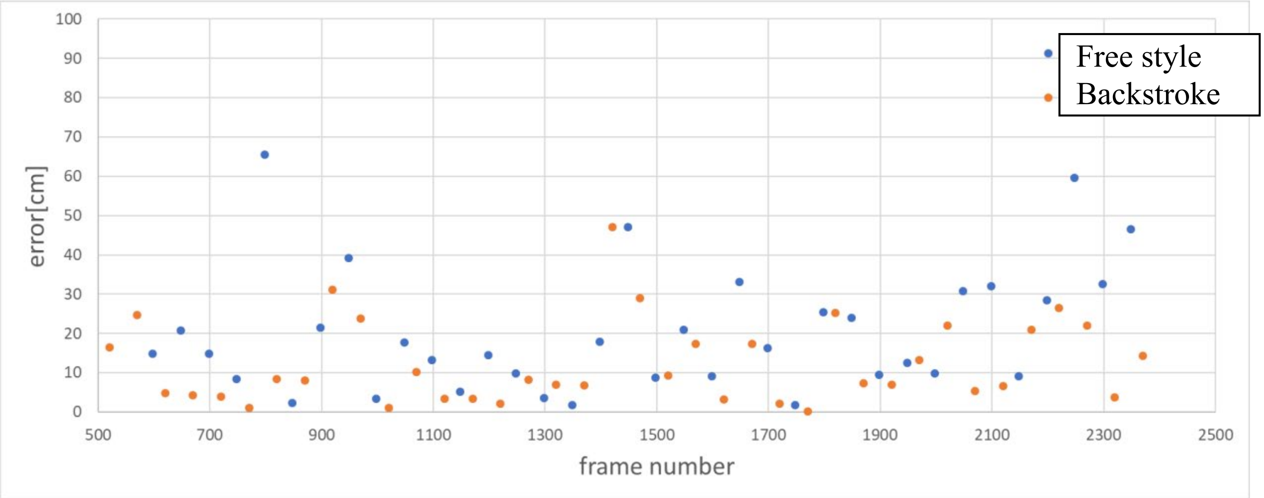

Figure 8. Result on free style and backstroke.

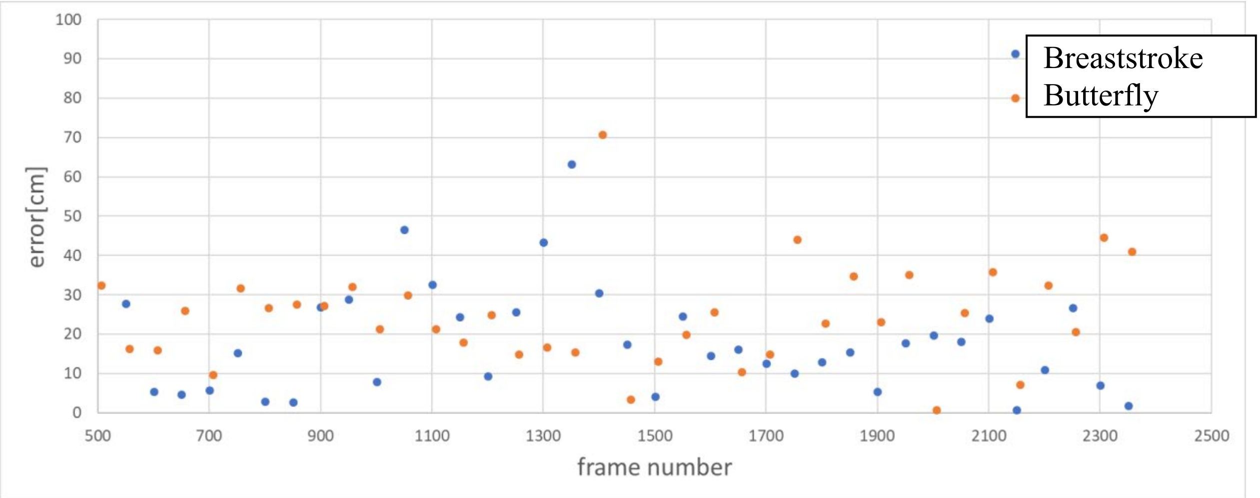

Figure 9. Result on breaststroke and butterfly.

Further evaluation should be conducted on open video database. A small set of swimming videos with annotated poses has been released[9] but it does not include the swim of multiple strokes on a real pool in the size of official swimming games.

4. CONCLUSION

We propose a new method of estimating swimmer position in swimming pool video. The video of swimming games is taken from a higher seat row in audience seat area. As It covers the whole field of a swimming pool, we can analysis the entire races with only the single video. The swimming pool video is transformed so that each lane can be analyzed along with the lane direction. The foreground region that includes both the swimmer and their water splash is extracted by adaptive background modeling and by setting the mask region to cope with the influence of the non-planer water surface and resultant apparent shakes of the lines drawn at the bottom of the pool. Swimmer region can be successfully extracted by the color analysis of water splash. The position is estimated as the center of the Gaussian distribution of the swimmer region. The proposed method was verified on a nationwide swimming game. For further evaluation we need to prepare more videos with grand-truth annotation.

A part of this research is supported by JSPS research grant Kakenhi 15H01825.

REFERENCES

[1] Lu, W. and Tan, Y. P., "A camera-based system for early detection of drowning incidents," Int. Conf. Image Processing, III-445-III-448 (2002).

[2] Eng, H. L., Toh, K. A., Kam, A. H., Wang, J. and Yau, W. Y., "An automatic drowning detection surveillance system for challenging outdoor pool environments," 9th IEEE Int. Conf. Computer Vision, 532-539 (2003).

[3] Marshall, J., "Magnetic field swimmer positioning," IEEE Sensors Journal, 15(1), 172-179 (2015).

[4] Christian X. Ries, C. X. and Lienhart, R., "Automatic pose initialization of swimmers in videos," Technical Report, Institute of Computer Science, University of Augsburg, Report 2009-19, 12 pages (2009).

[5] Zecha, D., Greif, T. and Lienhart, R., "Swimmer detection and pose estimation for continuous stroke-rate determination", The Int. Soc. for Optics and Photonics, Electronic Imaging, 830410–830410, (2012).

[6] Zecha, D. and Lienhart, R., "Key-Pose Prediction in Cyclic Human Motion," IEEE Winter Conference on Applications of Computer Vision, 86-93 (2015).

[7] Victor, B. He, Z., Morgan S. and Miniutti, D., "Continuous Video to Simple Signals for Swimming Stroke Detection with Convolutional Neural Networks," IEEE Conf. Computer Vision and Pattern Recognition Workshops (CVPRW), 122-131, 2017.

[8] KaewTraKulPong, P. and Bowden, R., "An improved adaptive background mixture model for real-time tracking with shadow detection", European Workshop Advanced Video Based Surveillance Systems, 5 pages (2001).

[9] Greif, T. and Leinhart, R., "An annotated data set for pose estimation of swimmers," Technical Report, Institute of Computer Science, University of Augsburg, Report 2009-18, 6 pages (2009).