ESTIMATING A RUNNER'S STRIDE LENGTH AND FREQUENCY FROM A RACE VIDEO BY USING GROUND STITCHING

Kentaro Yagi[1], Kunihiro Hasegawa[1], Yuta Sugiura[1] and Hideo Saito[1]

Keio University, Yokohama, Japan[1]

This study estimated stride length and frequency of runners in a 100 m race video. One method for measuring stride length and frequency is using infrared sensors. However, this method is not applicable to real races since numerous markers with infrared-reflective material must be attached to the runner's entire body. Therefore, we proposed a method using a race video. We generated a panoramic image of the 100 m track to estimate the distance of each frame from the start line. We detected the positions of the runner's steps from the movement of the leg joints. We projected every step to the overview image of the 100 m track. In the experiment, we applied our method to the video of an IAAF World Championship Track and Field 100 m race and obtained data from Usain Bolt. As a result, we can automatically estimate stride length and frequency of real races.

KEYWORDS: 100 m race, stride length and frequency, image processing

INTRODUCTION: Measuring the movement of athletes is important for evaluating their performance and for improving their skills. Many researches have measured athletes' movements using image-processing techniques. Cao et al.'s (2017) method proposed a way to estimate the position of body joints in a single image using CNN (Convolutional Neural Network). Hasegawa's method proposed a way to make a strobe image from an image sequence to visualize the movement of athletes (Kunihiro & Saito, 2016). However, no researches have focused on measuring the movements of 100m runners. For track and field athletes, stride length and frequency are critical (Högberg, 1952). One way of measuring these movements is by using infrared sensors. However, small markers with infrared-reflective material must be attached to the runner's entire body. Furthermore, multiple cameras must observe the sensors. Therefore, using such a method is impractical during real races.
Our research objective is estimating the stride length and frequency of 100m runners from a race video using image-processing techniques. Our method requires only a video of the race; therefore, our proposed method is applicable to real races.

METHODS: Our proposed method has three steps. The first step is image stitching. We generate a panoramic image of the 100m track to determine which frame is taken at how many meters away from the start line. We generate it by utilizing the consistency of the track lines. The second part is counting the runner's steps. We use Cao et al.'s (2017) method for determining the position of the runner's leg joints, and this is a real-time, multi-person system to detect the human body's joints from single images. Finally, we estimate the homography matrix from the panoramic image to the overview image, obtaining the runner's step length and frequency at the 100m scale.

1. Image Stitching: To generate the panoramic image, we utilize the track lines of a 100m lane. Typically, a matching method using local feature points, such as SIFT, SURF, and AKAZE, is used for image stitching (Brown & Lowe, 2007). However, those methods do not work because several feature points are not detected on a 100m track due to the lack of texture features.
1.1 Track-line detection: For image stitching, we detect the track lines of a 100m track in each frame to determine which line corresponds to which line between the two images.
In the first frame, each track line is detected by clicking two points on the line. In other frames, RANSAC (Fischler & Robert, 1987) and the least-squared method detect the lines. First,

applying a vertical differential filter obtains the horizontal edges. In the video, the track lines are supposed to be contained in these edges since the camera moves in parallel with the track lines. Then, to differentiate each track line, we mask the peripheral region of each track line in the previous frame. Next, each line is obtained by applying RANSAC and the least-squared method to the white pixels in each image. The flow of this part is shown in Figure 1.

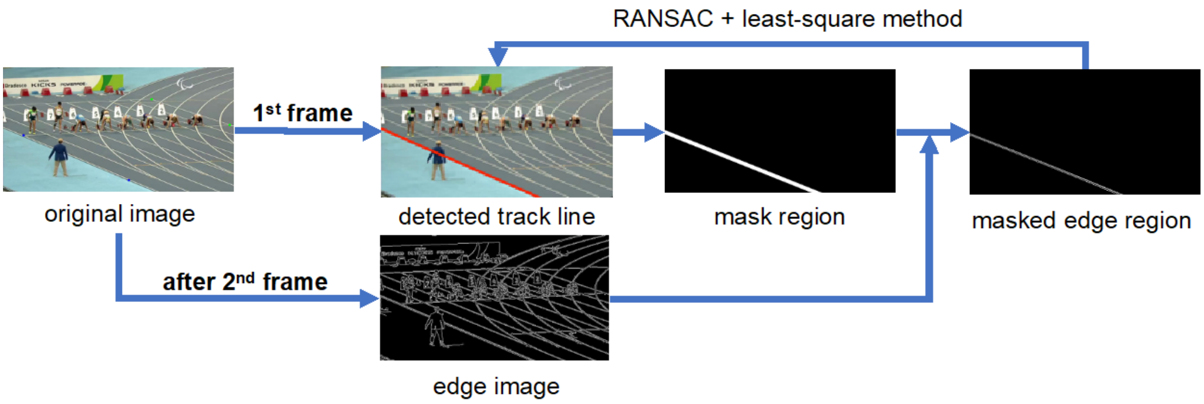

Figure 1: The flow of track-line detection.

1.2 Estimation of homography matrix: The next step for image stitching is estimating the homography matrix for the ground plane between each frame and the first frame. First, we estimate the homography matrix between adjacent frames by matching two images. Then, the homography matrix from each frame to the first frame is obtained by multiplying homography matrixes until the first frame.

For estimating the homography matrix, we suppose the camera movement between two frames can roughly be approximated by a 2D translation since the camera moves with the track-line direction only a short distance within the adjacent frames. Then, we estimate the 2D translation, which maximizes the matching score computed by SSD (Sum of Squared Difference) between two adjacent frames. In computing SSD, a hue-scale image is used to reduce the effect of the runners' shadows. Also, detected human regions and captions are masked to correctly estimate the translation of the 100m track plane.

The next step is finding corresponding points for computing the homography matrix that provides an accurate 2D position relationship for each pixel, which is needed for image stitching. To calculate the homography matrix, at least four corresponding points are needed. In this method, overlapped points on the track lines after translation are selected as corresponding points, as shown in Figure 2. As a result, we can obtain panorama image of 100m track, as shown in Figure 3.

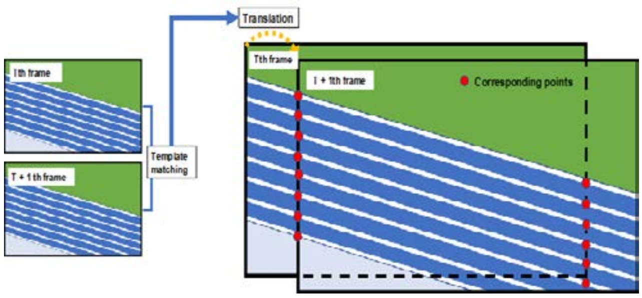

Figure 2: Flow of image matching.

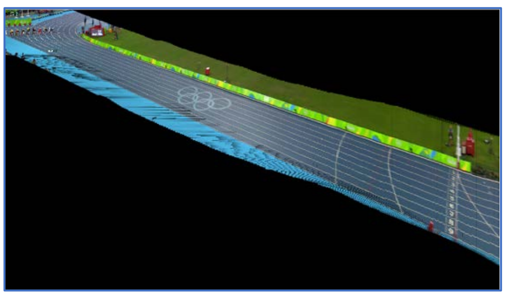

Figure 3: Result of image stitching.

2. Counting the Runner's Steps: For counting the runner's steps, we employ Cao et al.'s (2017) method for real-time joint detection of multiple individuals' body, hand, and facial key points from a single viewpoint image. As shown in

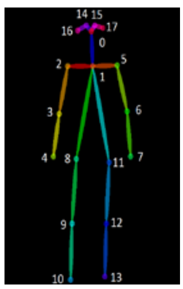

Figure 4: Body joint parts list estimated by Cao et al.' (2017) method.

Figure 4, we use the position of the leg joints (10, 13) from 18 body joints detected by Cao et al. (2017) for judging the runner's footsteps in the captured images.

2.1 Tracking runner: First, we must track the target runner. The body joints obtained by Cao et al.'s (2017) method are also used for this step. At the first frame, the runner closest to the clicked point is chosen as the tracking target. From the second frame, the target is tracked by finding a runner whose joints are the closest to the runner tracked in the previous frame. We compute the sum of each body part's distance from the body's trunk (0, 1, 14, 15, 16, and 17 in Figure 4) for evaluating the distance of the candidate runners to the detected runner in the previous frame.

2.2 Estimation of step timing: To determine if the runner steps in the frame, we calculate the distance from the leg joints to the outside line of the track. Since the distance to the line becomes smaller when the runner steps, we judge that the runner steps if the distance is the smallest in five continuous frames.

After that process, we improve the resulting step estimations. If the step length is less than half the average stride length, that step is regarded as an error. Furthermore, if the step length is longer than 1.5 times the average stride length, we add another step between the two steps.

Finally, to estimate the step length in a 100m scale, we estimate the homography matrix from the panoramic image to the overview image. Corresponding points are obtained by clicking the four corners of the 100m track in the panoramic image.

RESULTS: We applied our method to the video of an IAAF World Championship Track and Field 100m race in which the current world record for the Men's 100m race was set. We visualized the steps of Usain Bolt, as shown in Figure 5. In addition, we compared his 20m interval times, step lengths, and step frequency with the data from the Biomechanics Report World Championships 2009 Berlin, as shown in Figure 6, 7 and 8.

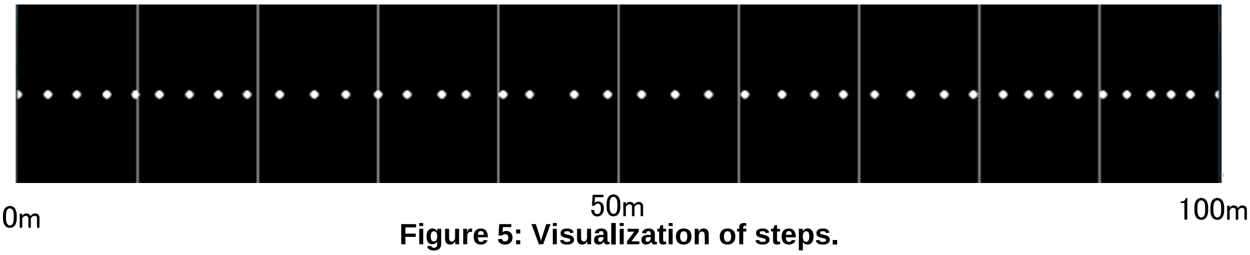

Figure 5: Visualization of steps.

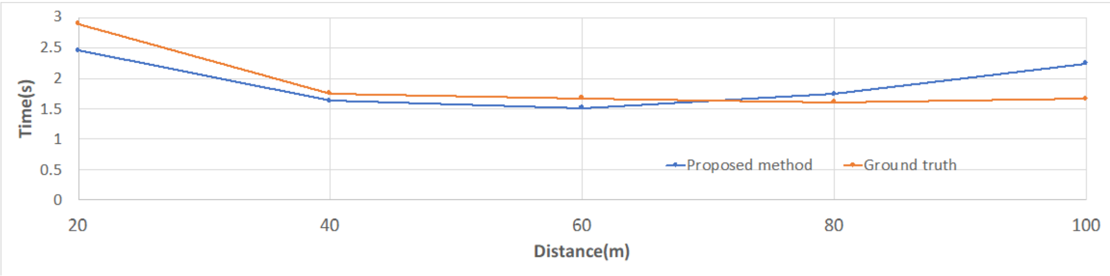

Figure 6: 20m interval times.

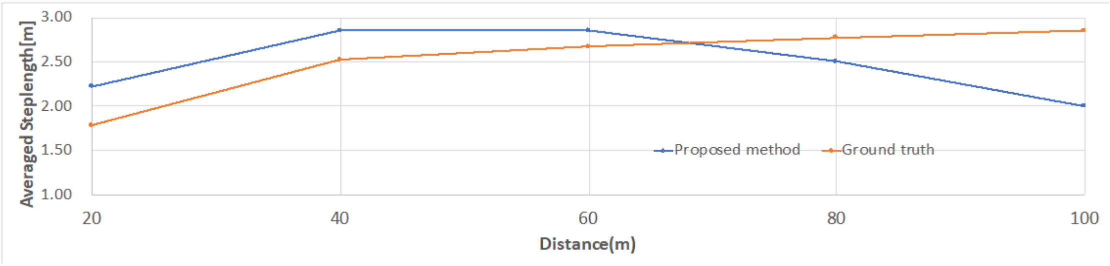

Figure 7: 20m interval average step length.

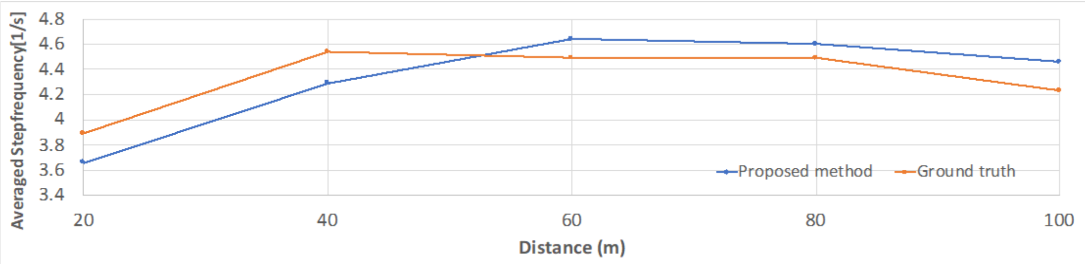

Figure 8: 20m interval average step frequency.

DISCUSSION: According to the figures shown, the first and last 20 m had more difference from ground truth than the other terms. For the first 20 m, one of the possible causes of the noise is the runner's posture. The accuracy of Cao et al.'s (2017) method depends on the individual's posture in an image. If all the body's joints are captured in an image, the accuracy becomes higher. However, in the first part of the acceleration section, which is the first 20 m, the runner bends the upper body. Therefore, the accuracy is reduced since every joint is not shown in an image. For the last 20 m, we assume that the accumulated noise of the homography matrix affects the result. During the image stitching, the homography matrix from each frame to the first frame is obtained by multiplying the homography matrices until the first frame. Therefore, the later the frame is, the more noise is accumulated. We assume this is why the estimated step length is too low in the last 20 m. Currently, we match two images by regarding the camera movement between the two frames as a translation. This might be the reason for the noise; therefore, we must consider alternative methods for image stitching.

CONCLUSION: We proposed a way to estimate the stride length and frequency of 100 m runners from a race video using image-processing techniques. In the experiment, we applied our method to the video of an IAAF World Championship Track and Field 100m race and obtained the data from Usain Bolt. In conclusion, we have demonstrated that our proposed method is applicable to the real race video without manual detection of footstep positions.

REFERENCES

Brown, M. & Lowe, D. G. (2007). Automatic panoramic image stitching using invariant features. *International Journal of Computer Vision, 74*(1), 59-73.

Cao, Z., Simon, T., Wie, S., & Sheikh, Y. (2017). Real time multi-person 2D pose estimation using part affinity fields, *Computer Vision and Pattern Recognition*, 1(2), 7291-7299.

Högberg, P. (1952). How do stride length and stride frequency influence the energy-output during running? *European Journal of Applied Physiology and Occupational Physiology, 14*(6), 437-441.

Kunihiro, H. & Saito, H. (2016). Synthesis of a stroboscopic image from a hand-held camera sequence for a sports analysis, *Computational Visual Media, 2*(3), 277-289.

Fischler, M. A., and Robert, C. B. (1987). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, *Readings in Computer Vision*, 726-740.

ACKNOWLEDGEMENTS: The authors are grateful to Dr. Hiroaki Hobara and Dr. Yoshinobu Kobayashi of the National Institute of Advanced Industrial Science and Technology (AIST) for suggesting this paper's topic and data.