Classification System for Golf Ball Initial Conditions using CNN Based on High-Time Resolution Images

Tomomasa Yamasaki\(^a\)*, Takashi Kaburagi\(^b\), Toshiyuki Matsumoto\(^a\), Satoshi Kumagai\(^a\), Yosuke Kurihara\(^a\)

\(^a\)Aoyama Gakuin University, 5-10-1 Fuchinobe, Chuo-ku, Sagamihara-shi, Kanagawa, 252-5258, Japan
\(^b\)International Christian University, 3-10-2 Osawa, Mitaka-shi, Tokyo 181-8585, Japan

* Corresponding Author: tomomasa.yamasaki@gmail.com

Abstract

To accurately estimate the trajectory of a golf ball after a swing impact, the initial conditions—i.e., initial velocity, angle, and backspin—need to be considered. In this paper, we propose an estimation system that can classify these initial conditions of the golf ball by applying a convolutional neural network (CNN) to high-time resolution images captured by a line scan camera. In the proposed system, to characterize the combination of the three parameters on the high-time resolution images, a golf ball is marked with a black line around its great circle. Three CNNs are constructed for each initial condition, and each CNN is trained using simulation images. By applying high-time-resolution images, the three initial conditions unknown to the trained CNNs could be estimated. To evaluate the proposed system, we conducted a validity experiment using simulation images. The initial velocity was ranged from 10–80 m/s at 5 m/s intervals, the angle was ranged from 0–0.79 rad at 0.087 rad intervals, and the backspin was ranged from 100–450 rad/s at 50 rad/s intervals. The results were evaluated in terms of accuracy and the RMSE per range of each parameter (RPR): the initial velocity, angle, and backspin showed accuracies of 100.00%, 98.44%, and 73.13%, and RPR values of 0, 0.013, and 0.079, respectively.

Keywords: Golf, CNN, high-time resolution image, line scan camera

1. Introduction

Golf is a famous sport in the world\(^1\). As of 2017, the number of golf players was estimated to be approximately 55 million. This is because golf has some merits for players: it reduces their mental stress and provides enjoyment irrespective of age\(^2\). Recently, playing golf indoors has become popular because it allows players to enjoy an on-screen golf course via simulation regardless of the weather; moreover, players can analyze their playing data statistically and try to improve their performance\(^3\). In golf simulation, estimating the ball trajectory after impact is an important factor because the trajectory directory affects the golf score. Additionally, simulations are useful for practicing golf.

To estimate the golf ball trajectory with high accuracy, the initial conditions of the ball (initial velocity, angle, and backspin) need to be considered\(^6\). Therefore, many studies have proposed methods to estimate them; for instance, estimation methods using infrared (IR) scanning sensors\(^5,6\) and high-speed cameras\(^7\) have been proposed. An estimation method using the following integrated sensors: IR sensor, acoustic sensor, and condenser microphone has also been developed\(^8\). Additionally, a visual tracking system\(^9-14\), an estimation method based on line sensors\(^15\), and an estimation method using stroboscopes have been proposed\(^16\).

In this paper, we proposed a system that could estimate the initial conditions after impact using high-time resolution images captured by a line scan camera. The three initial conditions are then estimated using a convolutional neural network (CNN).

2. Proposed System

Fig. 1 shows the entire scheme of the proposed system, which consists of three phases: the measurement phase, where the high-time resolution image is captured using the line scan camera; the learning phase, where three CNNs corresponding to the three initial conditions are trained by the simulation images; and the classification phase, where the initial conditions for the captured images are classified by the three trained CNNs.
pixels of the line scan image. By capturing line scan images for $N$ frames and combining $N$ images, we obtain an $M \times N$ image for a golf swing. Fig. 3 shows sample images of a swinging ball after impact, obtained by the line scan camera with the above settings. Figs. 3(a) and 3(b) show a captured image in the actual environment and the simulation image, respectively. In the actual environment, a line scan camera requires a high-intensity light source to obtain sharp images; the lack of such a source explains the apparition of light reflection in Fig. 3 (a). Similar to pre-signal processing, the images are binarized, and the binarized image is compressed along the vertical direction to render an image with a size of $N \times N$ that can be input to the CNN during the learning phase. Fig. 4 shows two sample simulation images after conducting the above pre-signal processing with different initial conditions.

![Fig. 1. Schematic of the proposed system](image)

### 2.1 Line Scan Camera Images of Swinging Ball with Marked Black Line

- **(a) Real ball**
- **(b) Simulation ball**

Fig. 2. Golf ball with a line mark

- **(a) Line scan camera image**
- **(b) Simulation image**

Fig. 3. Images from line scan camera and simulation

A ball used in the proposed system is marked with a black line around the great circle of the golf ball, as shown in Fig. 2. This figure also shows both an actual ball and a simulation ball with the marked line. In the proposed system, the marked ball is set to the initial position before swing impact. A line scan camera is set $L$ m away from the center of the ball so that the scan plane aligns vertically to the marked line, as shown in Fig. 1. The line scan camera captures $M \times 1$ images per frame, where $M$ is the number of

![Fig. 4. Simulation images with different initial conditions](image)

**Fig. 4. Simulation images with different initial conditions**

<table>
<thead>
<tr>
<th>Initial velocity</th>
<th>Angle</th>
<th>Backspin</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 m/s</td>
<td>0.79 rad</td>
<td>450 rad/s</td>
</tr>
<tr>
<td>75 m/s</td>
<td>0.09 rad</td>
<td>200 rad/s</td>
</tr>
</tbody>
</table>

(a) (b)

### 2.2 Learning Phase

![Fig. 5. CNN architecture for initial velocity and angle](image)

In the learning phase, the CNN is trained by $N \times N$-size binarized simulation images. In the proposed system, three CNNs are applied; and each CNN corresponds to an initial condition. Fig. 5 shows the CNN architecture for the initial velocity and angle with 12 layers, where conv($3 \times 3$, 8) denotes a $3 \times 3$ convolution layer with 8 neurons and BN indicates a batch normalization layer. Fig. 6 shows the CNN architecture for backspin with 30 layers, where conv($3 \times 3$, 8) and BN have the same meaning as depicted in Fig. 5, and
Drop(0.3) indicates a drop layer that is used to randomly set the input element to zero with a probability of 0.3. One can refer to the following papers for more information on CNN networks: “Batch normalization”<sup>17</sup>, “Wide Residual Networks”<sup>18</sup>, and “A Simple Way to Prevent Neural Networks from Overfitting”<sup>19</sup>.

The line scan images are generated in a certain range from \( \alpha \) to \( \beta \) for all the initial conditions. Each CNN is trained by the images using a combination of the abovementioned ranges for the three initial conditions as a label.

![CNN architecture for backspin](image)

Fig. 6. CNN architecture for backspin

2.3 Classification Phase

In the classification phase, the three initial conditions for an image captured by a line scan camera are estimated by classifying the captured image into a label for the CNN.

The captured image is applied to the same pre-signal processing for the simulation images. The \( N \times N \) binarized image is then input to each trained CNN. When the image is input into the CNN for initial velocity, the classified initial velocity for the image is obtained. Similarly, the classified angle (and backspin) is obtained by inputting the image into the CNN for angle (and backspin). Therefore, the proposed system estimates the initial conditions of the golf ball after impact.

3. Validity Experiment

To evaluate the proposed system, we conducted an experiment based on simulation images. Fig. 7 shows the environment in the simulation space for generating the simulation images using Unity (Unity Technologies). In the experiment, we used a golf ball marked with a black line and a light source on a PC. A golf ball was captured one frame at a time. In the simulation space, line scan camera images were generated.

![Environment in the simulation](image)

Fig. 7 Environment in the simulation

3.1 Generation of Simulation Images for Learning CNN and Test CNN

Table 1. Ranges for each initial condition of images on the simulation

<table>
<thead>
<tr>
<th>Initial conditions</th>
<th>Range ( (\alpha-\beta) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial velocity</td>
<td>10–80 m/s at 5 m/s intervals</td>
</tr>
<tr>
<td>Angle</td>
<td>0–0.79 rad at 0.087 rad intervals</td>
</tr>
<tr>
<td>Backspin</td>
<td>100–450 rad/s at 50 rad/s intervals</td>
</tr>
</tbody>
</table>

We established a system on Unity to generate \( 2048 \times 1 \) (\( M = 2048 \)) line images, as simulation images. \( N \) was set to 200 in this experiment. Table 1 shows the ranges from \( \alpha \) to \( \beta \) for each initial condition; the initial velocity was ranged from 10–80 m/s at 5 m/s intervals, the angle was ranged from 0–0.79 rad at 0.087 rad intervals, and the backspin was ranged from 100–450 rad/s at 50 rad/s intervals. Hence, there are 1200 combinations of initial conditions. The diameter of the simulated golf ball was set to 0.0428 m; this is typically the size of a common ball. Table 2 shows the parameter settings for each CNN.
Table 2. CNN parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Solver</th>
<th>Momentum</th>
<th>Initial learn rate</th>
<th>Learn rate drop period</th>
<th>Learn rate after drop period</th>
<th>Epochs</th>
<th>Batch size</th>
<th>Regularization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial velocity</td>
<td>SGDM</td>
<td>0.9</td>
<td>$10^{-3}$</td>
<td>8</td>
<td>$10^{-4}$</td>
<td>10</td>
<td>120</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>Angle</td>
<td>SGDM</td>
<td>0.9</td>
<td>$10^{-3}$</td>
<td>8</td>
<td>$10^{-4}$</td>
<td>10</td>
<td>120</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>Backspin</td>
<td>SGDM</td>
<td>0.9</td>
<td>$10^{-3}$</td>
<td>160</td>
<td>$10^{-4}$</td>
<td>350</td>
<td>60</td>
<td>$10^{-4}$</td>
</tr>
</tbody>
</table>

3.2 Evaluation Criteria

In the experiment based on the simulation images, we evaluated the proposed system in terms of accuracy and root mean square error (RMSE) per range of each initial condition (RPR), which were defined as follows.

\[
\text{Accuracy} = 100 \times \frac{\text{Correctly classified images (CCI)}}{\text{Total testing images (TTI)}} \tag{1}
\]

\[
\text{RPR} = \frac{1}{\beta - \alpha} \sqrt{\frac{1}{\text{TTI}} \sum_{k=1}^{\text{TTI}} (c_k - \hat{c}_k)^2} \tag{2}
\]

In Eq. (2), $c_k$ denotes the true value of the initial condition, and $\hat{c}_k$ is the estimated value of the initial condition. Hence, the root part in Eq. (2) represents the RMSE. The accuracy term expresses how accurately the test image can be classified. The RPR expresses how inaccurate the label of a classified image is. We divided all simulation images for each initial condition into training images (75%) and test images (25%). The initial velocity had 15 combinations in the range of 10–80 m/s at 5 m/s intervals; thus, each combination had 80 simulation images because there were a total of 1200 simulation images. Therefore, 60 images from each combination were used to train the CNN to estimate the initial velocity. The angle had 10 combinations in the range of 0–0.79 rad at 0.087 rad intervals; thus, each combination had 120 simulation images and, hence, 90 images from each combination were used to train the CNN to estimate the angle. The backspin had 8 combinations in the range of 100–450 rad/s at 50 rad/s intervals; thus, each combination had 150 simulation images and, therefore, 112 images from each combination were applied to train the CNN for backspin estimation.

4. Experimental Results

Table 3. Accuracy of classification

<table>
<thead>
<tr>
<th>Parameter</th>
<th>CCI</th>
<th>TTI</th>
<th>Accuracy [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial velocity</td>
<td>75</td>
<td>75</td>
<td>100.00</td>
</tr>
<tr>
<td>Angle</td>
<td>443</td>
<td>450</td>
<td>98.44</td>
</tr>
<tr>
<td>Backspin</td>
<td>439</td>
<td>600</td>
<td>73.13</td>
</tr>
</tbody>
</table>

Table 4. RPR of classification

<table>
<thead>
<tr>
<th>Parameter</th>
<th>RMSE</th>
<th>$\beta - \alpha$</th>
<th>RPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Velocity</td>
<td>0.000</td>
<td>70</td>
<td>0.000</td>
</tr>
<tr>
<td>Angle</td>
<td>0.011</td>
<td>0.79</td>
<td>0.013</td>
</tr>
<tr>
<td>Backspin</td>
<td>27.540</td>
<td>350</td>
<td>0.079</td>
</tr>
</tbody>
</table>

Tables 3 and 4 show the results for Accuracy (defined by Eq. (1)) and RPR (defined by Eq. (2)), respectively, for each initial condition. For the initial velocity, the Accuracy was 100.00%; for the angle, the Accuracy was 98.44%. The reason for these high Accuracy values for the initial velocity and angle is because the length, in the horizontal direction, and the tilt of the image of the ball correspond to the initial velocity and angle, respectively. For the backspin, the Accuracy was comparably lower. Hence, the CNN was not trained sufficiently to classify the backspin to a high degree of accuracy. According to Table 4, however, the RPR of the backspin was almost the same as that of the initial velocity and angle. Therefore, increasing the number of simulation images used for training can improve the classification accuracy and RPR.

5. Discussion

In the proposed system, the backspin is classified using the black mark on the ball. In some cases, however, the mark makes it difficult to classify the backspin. Fig. 8 shows cases where most classified backspins may be considerably different from the actual backspin. Both images in Fig. 8 capture only a part of the golf ball, and the afterimage of the mark on the golf ball becomes a nearly straight line. Hence,
even if the backspin is large, the pattern difference between a larger backspin value and lower backspin value cannot be indicated using the images. This explains the larger backspin value, juxtaposed with the error of the initial velocity and angle. In most cases where the initial velocity was very high, the same characteristics were observed.

![Images with very high initial velocity](image)

6. Conclusions

In this paper, we presented a novel classification system using a line scan camera to estimate the initial conditions of a golf ball after swing; these initial conditions are important for calculating the golf ball trajectory for simulation purposes. In the proposed system, three CNNs were applied to the line scan images to estimate each initial condition. To evaluate the proposed system, we carried out an experiment based on simulation images. From the obtained results, the accuracy of the initial velocity was 100.00%, that of the angle was 98.44%, and that of the backspin was 73.13%. Additionally, we evaluated the system in terms of an index defined as the RMSE per range of each parameter (RPR), which was used to measure how inaccurate the label of a classified image was. According to the RPR of each initial condition, the initial velocity was 0, the angle was 0.013, and the backspin was 0.079.

For future work, the proposed system needs to be evaluated using actual line scan images. In addition, we aim to refine the estimates for these initial conditions by using regression methods. We should reconsider the mark a ball has.

References

(8) SW Seo, M Kim, Y Kim, "Optical and Acoustic Sensor-Based 3D Ball Motion Estimation for Ball Sport Simulators," The Eighth International Conference, 2018.
(13) Zhengtao Zhang, De Xu, Min Tan, "Visual Measurement


