3D Sensing Technology for Real-Time Quantification of Athletes' Movements

● Kazuo Sasaki ● Shoichi Masui ● Koichi Tezuka

Recently, efforts to help athletes develop skills and abilities by incorporating scientific knowledge are actively deployed in the field of sports. At the same time, as judging and scoring for increasingly sophisticated skills in scoring competitions such as gymnastics is becoming more difficult every year, the need to incorporate technologies to improve accuracy and fairness in a judging and scoring system is increasing accordingly. Therefore, Fujitsu Laboratories is engaged in research and development aimed at establishing 3D sensing technology that accurately measures and digitizes complex human movements three-dimensionally using machine learning. 3D sensing technology is comprised of 3D laser sensor technology, which generates depth images based on the contours of the body surface to represent human movements, and skeleton recognition technology, which quickly extracts the 3D coordinates of joints from the depth images. Regarding 3D laser sensors, we describe a split projection/detection optical system with a micro electro mechanical systems (MEMS) mirror for laser emission and detection, which is capable of more than 10 times the number of pixels obtained by conventional light detection and ranging (LIDAR) technology. We then demonstrate the usefulness of view-angle control technology for sports. We also describe the skeleton recognition technology, noting the fast, high-accuracy 3D joint coordinate extraction method based on a combination of skeleton recognition through machine learning and subsequent fitting. This paper illustrates the 3D sensing technology developed by Fujitsu Laboratories and presents experimental results of skeleton recognition in actual athletes' movements.

1. Introduction

The use of ICT in the field of sports has been increasing in recent years. In soccer, for example, ICT is being used to track athletes on the field from video information and to analyze the distance covered by each athlete, the formations adopted by a group of athletes, etc. In tennis, ICT is being used to capture the trajectory of the ball and record the ball's landing to make line judgments.

There is a need, however, to accurately capture an athlete's movement as 3D data to enable a detailed analysis of how the body of an athlete moves, but there are no examples of practical systems for doing so in actual competitions. One reason for this is that traditional motion capture systems require the affixing of position-indicator sensors or magnetic sensors called "markers" on the athlete's body near joints, which can hamper an athlete's intricate movements. If an athlete's movement could be accurately measured and quantified in three dimensions without attaching markers, it should be possible to apply such a markerless system to scoring competitions such as gymnastics, figure skating, etc. In addition, it can make athlete training more efficient and sports more entertaining for spectators and viewers. Fujitsu Laboratories has developed 3D sensing technology that uses 3D laser sensors and skeleton recognition technology to measure and accurately digitize human movements in three dimensions without the use of markers.

This 3D sensing technology is outlined in **Figure 1**. To perform 3D measurements of an athlete's movements without markers, this system uses a 3D laser sensor to obtain depth images representing the contours of the body's surface. It then uses skeleton recognition technology to derive the 3D positions of the body's main joints from those images. Determining the positions of

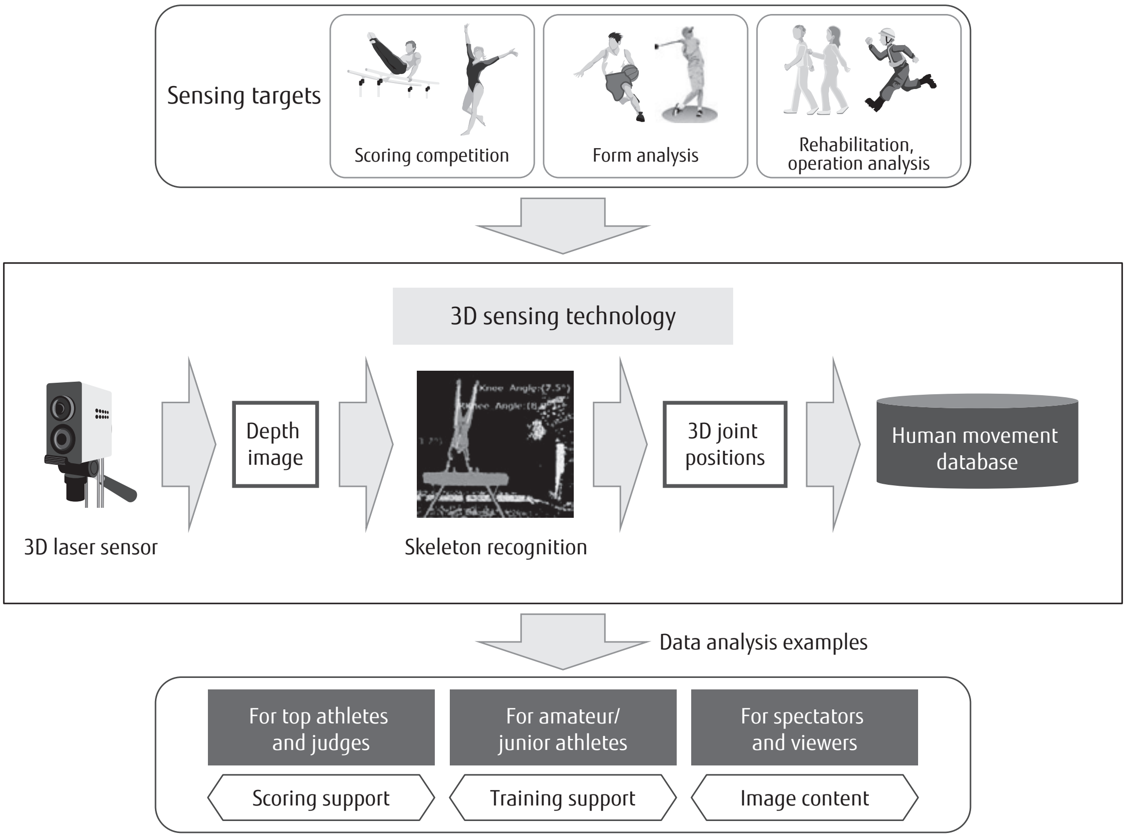


Figure 1
Application of 3D sensing technology to sports.

the joints in this way makes it possible to accurately calculate angles associated with the elbows, knees, spine, etc. and to analyze in detail body movements based on the temporal change in those angles. For example, the angle of an athlete's spine or knee joints can be used to support judges such as when having to decide whether the athlete's back is straight or whether the knees are bent. Furthermore, since this technology provides body movements in the form of numerical data, it can also be used for objective scientific training. The technology can also be used to provide the configuration, difficulty, etc., of an athlete's technique to spectators or viewers in real time so as to visually convey an athlete's outstanding physical feats and make sports all the more interesting.

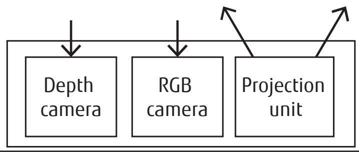
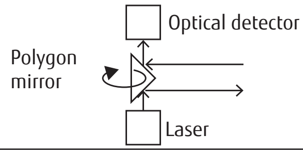
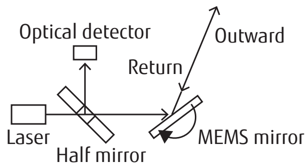
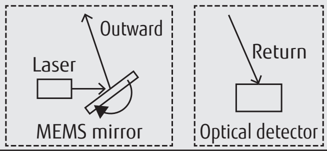
Finally, the application of 3D sensing technology is not limited to sports—it is expected to find use in a variety of fields such as visualizing the restoration of motor functions in rehabilitation programs, operation analysis at manufacturing sites, etc. The following sections describe the 3D laser sensor and skeleton recognition technologies that make up the 3D sensing technology.[1)]

2. 3D laser sensor technology

The task of accurately capturing the rapid movements of an athlete requires a high frame rate and a depth-image acquisition method that can capture an athlete at high resolution and at long range. **Table 1** compares various types of depth-image acquisition methods. A depth camera analyzes the light reflected from projected light and obtains depth information through an integrated image sensor.[2)] Cameras of this type have been commercialized for gaming applications.[3)] Although a depth camera can obtain depth information at high speed and high resolution, it cannot support distances greater than 5 m and is, therefore, limited to some degree in sports applications.

Laser sensors adopting light detection and ranging (LIDAR) technology have received much attention as a means of obtaining depth images from as far away as 15 m making it applicable to sports on the whole.

Table 1
Comparison of depth-image acquisition systems.

Category			Features	Issues	Configuration
Depth camera			High speed High resolution	Short distance (max. 5 m) Ambient light stability	
LIDAR	Rotating motor type		Long distance Light resistant	Low speed (low resolution)	
	MEMS mirror type	Coaxial projection/ detection	Long distance Light resistant	Low speed (low resolution)	
		Split projection/ detection	Long distance High speed High resolution	Ambient light stability	

LIDAR uses the reflection of pulse-irradiated laser light to analyze properties of a target such as the distance to it and its contour. The speed and resolution of a LIDAR system varies depending on the way its scanning system on the projection side and optical system on the detection side are configured. For example, in a rotating motor type of scanning system that uses a polygon mirror, after one line of scanning, the system needs to wait for the mirror to rotate to a specific location until it can start the next scanning process, which makes it difficult to achieve high-speed scanning. A micro electro mechanical systems (MEMS) mirror that performs mechanical scanning only within a limited range is, therefore, an effective means of overcoming this problem.

In sports applications, there is a need for more than ten times the number of scanning points compared with existing LIDAR to achieve high-resolution results, so the scanning speed of the MEMS mirror must be increased. This requires the downsizing of the MEMS mirror by using a scanning-angle magnifying lens. If projection and detection of the light is performed in a coaxial manner, downsizing of the MEMS mirror that is also used for detection will prevent all of the light reflected from the target, thereby, reducing the amount of light at the photodetector. To secure a sufficient amount of detected light, we adopted an optical system with split projection and detection units.

Figure 2 shows the configuration of the 3D laser sensor developed by Fujitsu Laboratories featuring a split projection/detection optical system using a MEMS mirror.[4] To measure the distance to a target, this system uses the time of flight (ToF) method that measures the time from the projection of a laser pulse to the detection of its reflection. Denoting the time taken for a laser pulse to be projected, reflected off of the target, and detected at the detection unit as ΔT and the speed of light as c (approximately 300,000 km/s), the distance d to the target can be given by the following equation:

$$d = \frac{c \times \Delta T}{2}$$

Furthermore, to support diverse types of sports competitions and achieve high-resolution 3D sensing across a wide range of distances and areas, there must be a sufficient amount of positional freedom as to where the 3D laser sensor can be installed. As shown in **Figure 3 (a)**, depth information can be obtained at high resolution for a nearby target, but in the case of a faraway target, resolution drops for the same view angle. Consequently, the case of a faraway target makes view-angle control technology essential. As shown in **Figure 3 (b)**, this technology narrows the view angle by controlling the MEMS mirror, thereby,

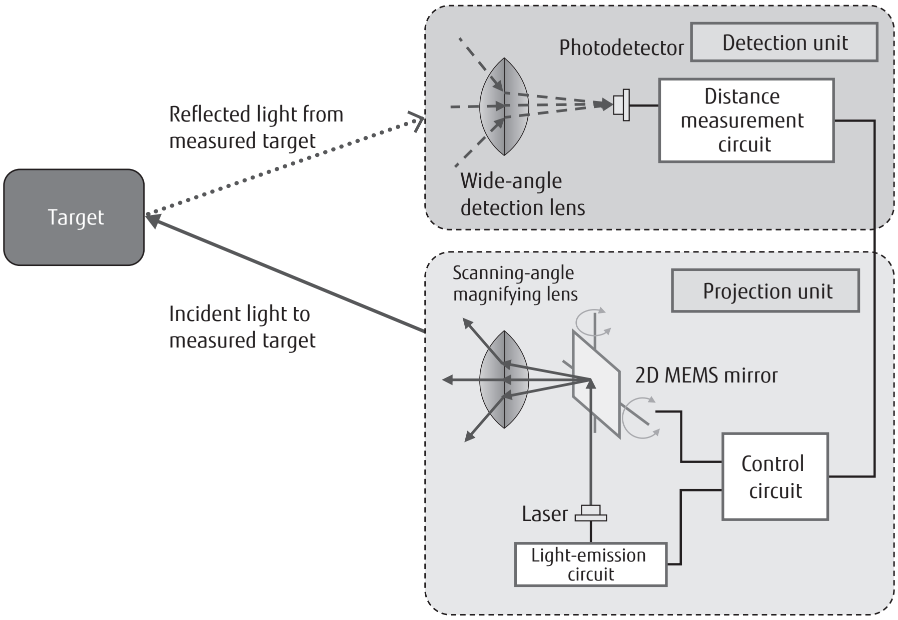


Figure 2
Configuration of 3D laser sensor.

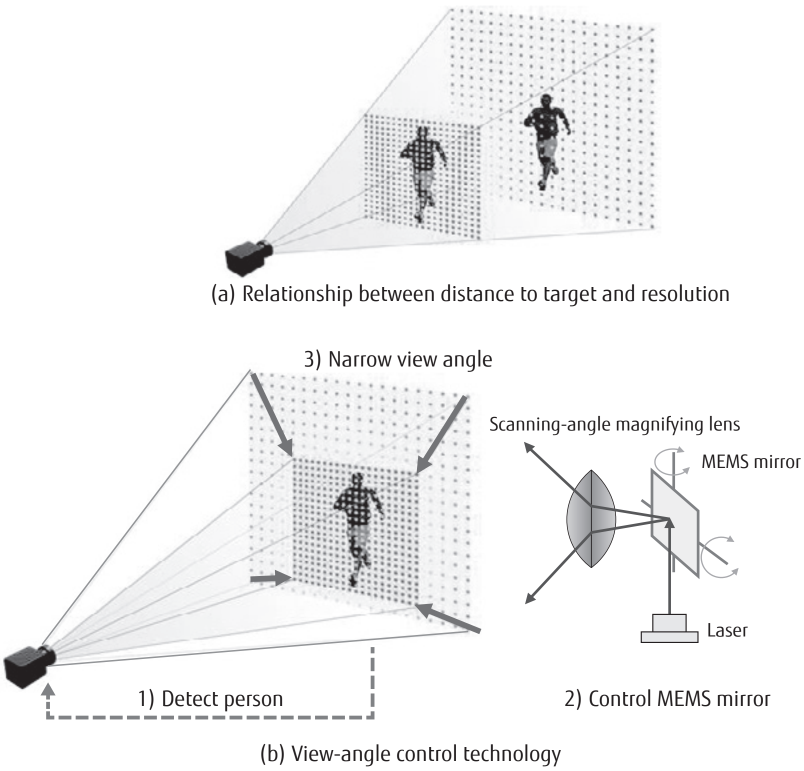


Figure 3
Overview of view-angle control technology.

increasing resolution to the same level as that of a nearby target. **Figure 4** compares the cases of sensing a faraway target with and without view-angle control (ON and OFF, respectively) using depth information for a nearby target as reference. As shown, using view-angle control enables the movement of a faraway target to be sensed at the same (high) level of resolution as that of a nearby target.

Due to the use of a wide-angle detection lens, the MEMS-mirror split projection/detection optical system is easily affected by ambient light such as sunlight or indoor lighting, as shown in **Figure 5 (a)**. Because of this

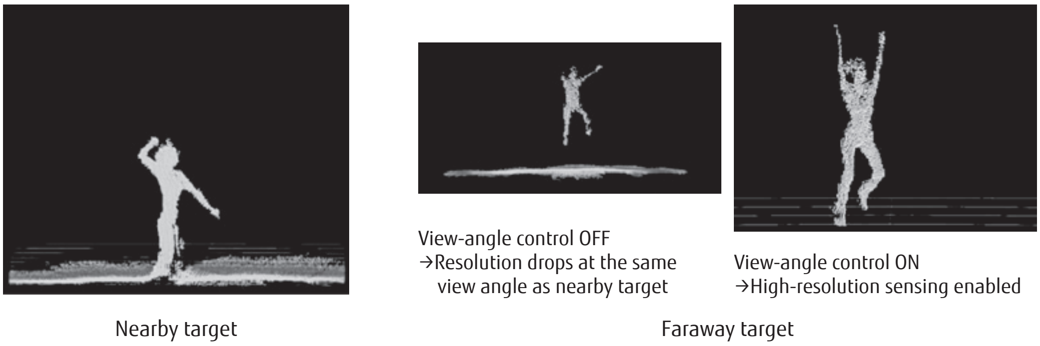


Figure 4
Effect of view-angle adjustment.

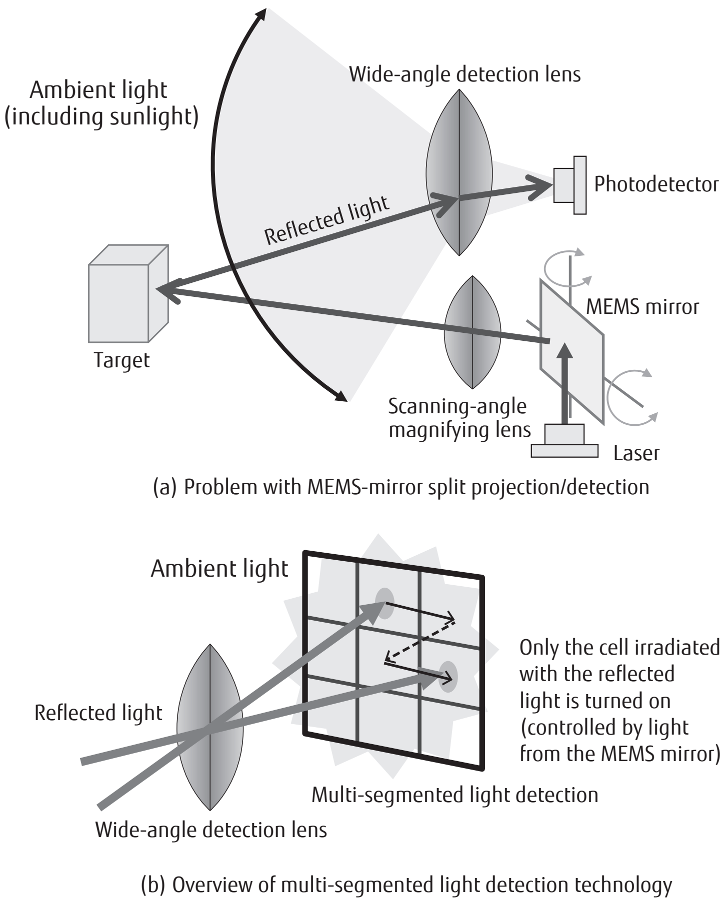


Figure 5
Overview of multi-segmented light detection technology.

property, we developed multi-segmented light detection technology as shown in **Figure 5 (b)**. This system synchronizes with control signals from the MEMS mirror to selectively turn on only the photodetector receiving the most light reflected off of the target while turning off all other photodetectors affected by ambient light.

In addition to the technologies described so far, multiple units of this 3D laser sensor can be synchronized to enable accurate detection of movement in areas that turn out to be blind spots when using only one laser sensor. In addition to improving the accuracy and fairness of judging and scoring, these technologies combined can also make it easier to learn new skills. In this way, Fujitsu Laboratories has developed view-angle control technology, multi-segmented light detection technology, and multi-unit linking technology for 3D laser sensors with the aim of increasing the accuracy of 3D sensing in sports.

3. Skeleton recognition technology

Skeleton recognition technology extracts the positions of various joints making up the human body from depth images obtained by the 3D laser sensor. In a judging support system used in gymnastics or other competitive sports, providing judges with real-time assistance in making decisions requires the output of joint positions and joint angles as 3D data. There is, therefore, a need for an extremely high level of accuracy compared with gaming applications. **Figure 6** provides an overview of high-speed, high-accuracy skeleton recognition technology that satisfies these requirements for 3D sensing. To make use of machine learning in skeleton recognition from depth images, a learning phase is needed prior to the recognition phase.

The learning phase creates prediction models that output inferred values of joint coordinates using depth images as input. This requires the creation of depth images by computer graphics from previously obtained movements with joint coordinates to prepare a training set for machine learning.

The recognition phase, in turn, inputs multi-viewpoint depth images obtained from multiple 3D laser sensors and uses a prediction model generated in the learning phase to derive 3D joint coordinates (skeleton recognition). This phase then uses these joint coordinates as initial values to apply a human model to the point cloud corresponding to the depth images obtained from each sensor. This process is called "fitting." To enable point-cloud coordinates to match as close as possible the surface coordinates of the human model used for fitting, we define "degree of matching" (likelihood) and then search for coordinates with maximum

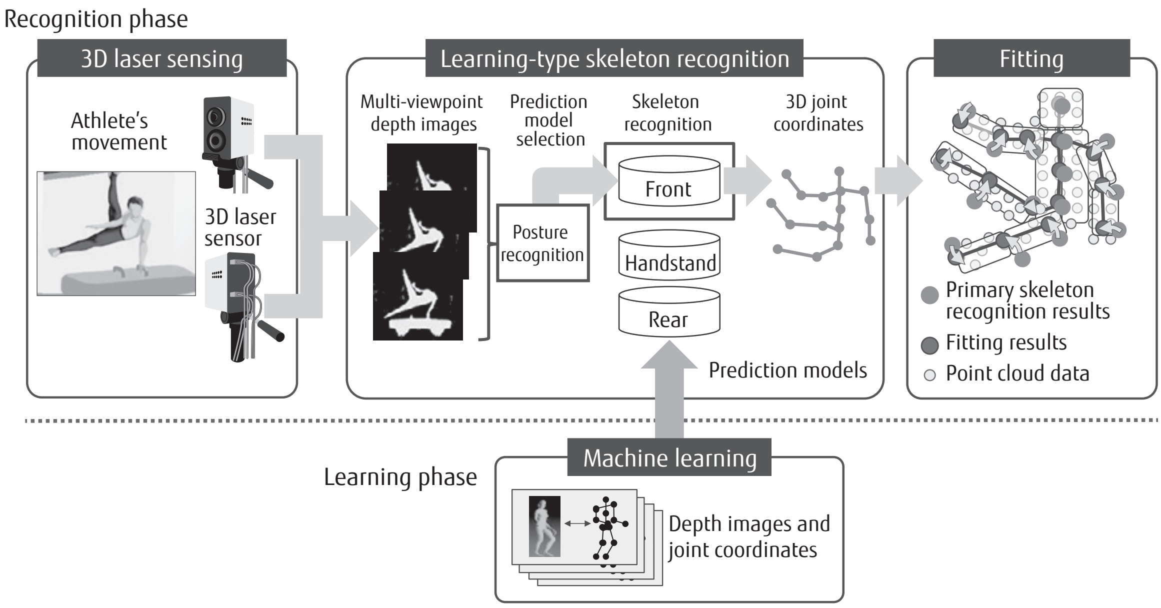


Figure 6
Overview of high-speed, high-accuracy skeleton recognition technology.

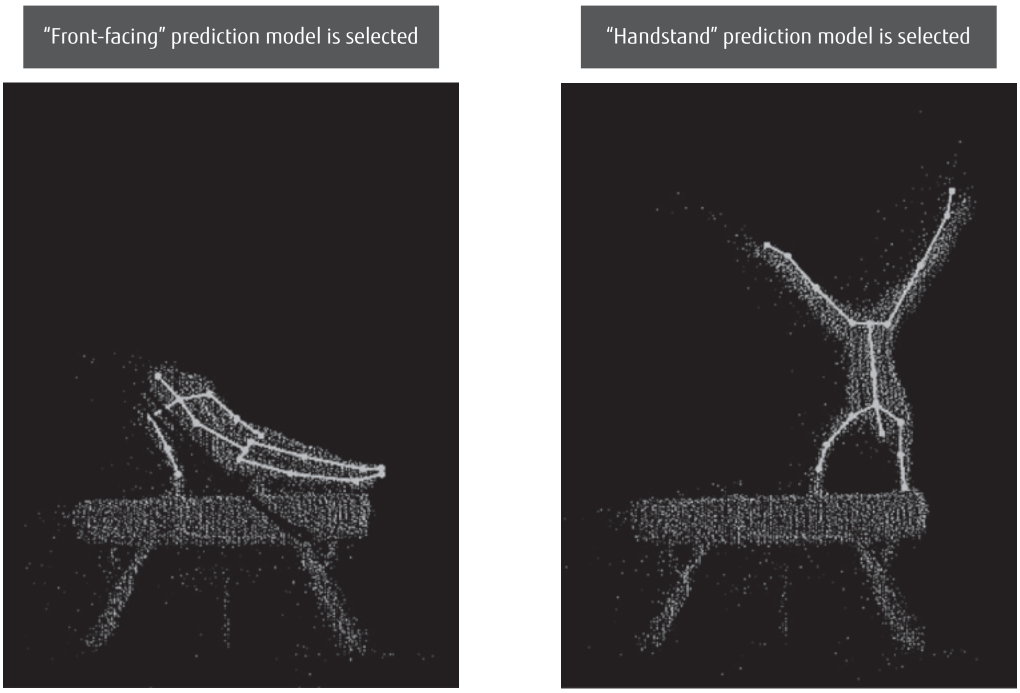
likelihood to determine final 3D joint coordinates.

In skeleton recognition using machine learning, accuracy is generally low since joint positions are inferred on the basis of a prediction model. However, this subsequent fitting process improves accuracy by matching joint positions with actual measured values in accordance with the point clouds from multiple 3D laser sensors. At this time, the accuracy of measured values in machine learning-based skeleton recognition determines the range of fitting and, therefore, affects the accuracy of final skeleton recognition results and processing time. To improve accuracy in machine learning-based skeleton recognition, we prepare beforehand multiple prediction models that consolidate postures such as front, handstand, and rear as shown in Figure 6 and adopt a technique that selects the optimal prediction model by detecting posture before skeleton recognition. Compared with the method of consolidating all movements in a single prediction model, this technique greatly improves recognition accuracy by limiting the movements to be learned in a prediction model.

Figure 7 shows the results of machine learning-based skeleton recognition using multiple sensors in the pommel horse event in gymnastics. For the swirling motion of Figure 7 (a), the process uses the prediction model corresponding to a front-facing posture, and for the dismount of Figure 7 (b), it uses the prediction model corresponding to a handstand. These results show that switching between prediction models for different types of postures enables high-accuracy skeleton recognition even for complex movements typical of gymnastics.

4. Steps toward practical use and future outlook

Fujitsu is developing a judging support system based on 3D sensing technology in collaboration with International Gymnastics Federation (FIG) and Japan Gymnastics Association (JGA). For details, please see the paper titled ICT-Based Judging Support System for Artistic Gymnastics and Intended New World Created Through 3D Sensing Technology in this issue. Gymnastics events feature much variation in human movement including advanced somersaults and twists. Skeleton recognition technology established to support



(a) Swirling motion (b) Dismount

Figure 7
Results of learning-type skeleton recognition in the pommel horse event.

such events should, therefore, be applicable to all human movements. On the basis of the collaborations mentioned above, Fujitsu Laboratories successfully conducted a verification experiment of this technology in the still rings event at the 47th FIG Artistic Gymnastics World Championships in Montreal, Canada held in 2017. This achievement demonstrated the possibility of applying this technology to all gymnastics events. **Figure 8** shows the installation and preparation of 3D laser sensors for the first time at an international gymnastics championship.

As described above, 3D sensing technology can be applied to a variety of competitive sports other than gymnastics and to fields other than sports. In basketball, for example, it could be used to analyze an athlete's shooting form to improve shooting percentage or obtain pointers for breaking out of a slump. Additionally, in a fitness gym, it could be used to analyze training form to assess correct movement of a person's frame, and in a rehabilitation center, it is expected to be used to visualize the degree to which joint movability has recovered.

Figure 8
Installing and preparing sensors at the 47th FIG Artistic Gymnastics World Championships.

In this way, it should be easy to support other competitive sports and fields by replacing the prediction models in the skeleton recognition process shown in Figure 6 in accordance with the types of human movement in each application field.

5. Conclusion

In this paper, we described 3D sensing technology for quantifying athletes' movements in real time.

This technology holds the possibility of significantly changing training methods that up to now have been based largely on experience and intuition. It can also help grow the sports industry by making sports more entertaining for spectators and viewers and enlarging the fan base.

To this end, Fujitsu Laboratories aims to put this technology into practical use as quickly as possible and to expand its application to a variety of competitive sports as a way of contributing to the further development of the world of sports.

References

1) S. Masui et al.: Changing Sports Activities by 3D Sensing Technology. Journal of the Institute of Electronics, Information and Communication Engineers, Vol. 100, No. 11, pp. 1182–1188 (2017). (in Japanese)
2) A. Payne et al.: A 512x424 CMOS 3D Time-of-Flight Image Sensor with Multi-Frequency Photo-Demodulation up to 130MHz and 2GS/s ADC. IEEE International Solid-State Circuits Conference, No. 7.6, pp. 134–135, San Francisco, U.S.A., Feb., 2015.
3) J. Han et al.: Enhanced Computer Vision with Microsoft Kinect Sensor: A Review. IEEE Trans. on Cybernetics, Vol. 43, No. 5, pp. 1318–1334, Oct., 2013.
4) K. Iida et al.: Development of 3D Range Sensor with Super-wide Angle Detection to Observe Vehicle Surrounding. 19th ITS World Congress, No. AP-00079, Vienna, Austria, Oct., 2012.

Kazuo Sasaki
Fujitsu Laboratories Ltd.
Mr. Sasaki is currently engaged in overseeing research and development in the sports field and in researching the analysis of athletes' skills.

Shoichi Masui
Fujitsu Laboratories Ltd.
Mr. Masui is currently engaged in research related to skeletal and movement recognition.

Koichi Tezuka
Fujitsu Laboratories Ltd.
Mr. Tezuka is currently engaged in research and development of optical and mechanical systems for LIDAR.