3D Shape Reconstruction of Human Foot using Distance Sensors

Moeko Iwasaki[1†], Suzanne Low[1], Mitsunori Tada[2], Yuta Sugiura[1], Hideo Saito[1], Maki Sugimoto[1]

[1]Keio University, Yokohama, Japan

National Institute of Advanced Industrial Science and Technology (AIST), Tokyo, Japan

Abstract: We proposed a simple, easy-to-use system to reconstruct an individual's foot shape in 3D by using an array of distance sensors and data of known foot shape collected in a database. In our system, values from the sensors will be used to generate a silhouette image, where its information will then be used as the base to search for the most similar model from the database to reconstruct a 3D foot shape. In order to verify the effectiveness of our method, we conducted experiments to evaluate the accuracy of the reconstructed foot shape.

Keywords: Shape Reconstruction, Human Foot, Distance Sensor.

1. INTRODUCTION

Walking is an important part in our daily life. On average, people walk about 5000-10000 steps per day depending on their lifestyle. Thus, having good shoes are important. Generally, in order to find a suitable product, we have to refer to various ready-made product sizes. However, we have difficulties in getting the right sized and shaped shoes as each individual's foot shape differs from one to another. For example, some people's foot length may be in between ready-made shoe sizes, or some people may have wider or narrower foot width, or some toes are straight in shape while others are curved in shape. Prolonged usage of shoes while lacking proper consideration of the foot shape may have an impact on our health in the long run. We might get blisters or bleeding toes due to prolong rubbing against the shoe, or forefoot deformity such as hallux valgus, or may even lead to long term problems such as back or joint pain.

The standard way of purchasing a shoe is to visit a shoe shop and to try the product on before purchasing. Though, due to the popularization of e-commerce, many people prefer to purchase products online, limiting their chances to test the product before purchase. Therefore, being able to measure the 3D shape of the foot gives an advantage to choose a suitable shoe, even for online shopping.

A common method to measure the foot shape is to use a 3D scanner. There are various types of 3D scanners such as an installation type, a handy type, contact or non-contact type etc. and by choosing the appropriate product, it will be easy to measure the foot shape. Nevertheless, the price of a 3D scanner itself is relatively high and the process of scanning and managing the scanned data can be quite time consuming.

Another frequently used method is to use a camera. Numerous methods to reconstruct a foot shape by comparing images from the camera with a known foot shape database have been proposed [2][8]. Although these methods are inexpensive, they work best if the installations are set up professionally such as proper camera calibration etc., which are rather challenging.

As compare to the methods above, our research focuses on creating a portable and simple-to-install system to reconstruct the 3D foot shape. To achieve this, we collect data values from an array of sensors as illustrated in Fig.1 and compare them with a known foot shape database.

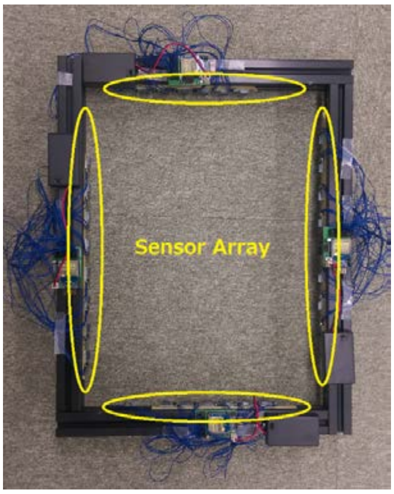

Fig.1 Proposed Device

2. RELATED WORK

2.1 Body shape reconstruction using body shape database

Prior researches have proposed methods of reconstructing 3D body shape using information from a known body shape database, without performing measurements by a 3D scanner or 3D modeling. An example is to reconstruct a body shape using information in a body shape database [5], where a small amount of measurements are taken to estimate a full set of measurements by making use of the regression formula derived from the multiple regression analysis of the body shape database. Other than direct measurements, Kinect is also frequently used as a distance sensor to collect body measurements [10].

Images from the camera combined with information from the body shape database are commonly used to reconstruct a 3D body shape. Ito et al. compared images from a handy camera to a face shape database to reconstruct a 3D face shape [11]. They utilized the Principal Component Analysis (PCA) of the database to express the face shape in a small number of parameters. By mounting a hairband containing a marker, they reconstructed the face shape by estimating the parameters obtained from the continuous image sequence captured. Similarly, Takeuchi et. al reconstructed a face shape by utilizing asynchronous cameras to capture the image of the same person and compare it with a face shape database [12]. Their system optimizes the evaluation based on the silhouette shape, appearance and the position of the feature point of the input multi-view images, to obtain 3D posture of the face, which will be used to estimate the parameters representing the face shape, and by repeating this estimation, they were able to reconstruct the face shape. Suzuki et al. proposed a method to reconstruct a 3D face shape and camera poses from freehand multi-viewpoint snapshots [13].

Moghaddam et al. also reconstructed the 3D face shape by utilizing silhouette images from multiple cameras and the face shape database [4]. Although the advantage of using silhouette images is that the images are unaffected by light or texture, the cameras used require to be calibrated and the setting up of the environment is quite challenging. Saito et al. reconstructed the torso by using captured silhouette images of the front and side of the torso with a torso shape database [7]. Using their method, it is possible to reconstruct the torso easily, but as a prerequisite, it is necessary for the front and side silhouette images to be in orthogonal plane, and when capturing the image, it is essential to add different adjustments such as keeping the camera fixed. Therefore, it takes time and effort to build the necessary environment for the system.

Similar to the methods mentioned above, the reconstruction of a 3D foot shape is possible by fitting with the model using image information obtained from multiple cameras and a small number of parameters reduced by the PCA of the database [2][8]. Although this method enables highly accurate restoration of the foot shape, the requirement to calibrate the camera used in the experimental environment beforehand makes it quite a challenging method to be used.

2.2 Measurement from multiple sensors

Combination of multiple sensors can produce better quality data as compare to just using an individual sensor. Dementyev et al. developed a tape-type sensor network composed of distance sensors and an inertial measurement unit (IMU) that connects nodes [3]. As each node is connected to each other, it is possible to detect the self-deformation of the tape from each piece of information. Their prototype can be used as a wearable application or home sensing application.

Richardson et al. developed a sensor network to detect the pressure on the bottom surface by creating a tile-type sensor module containing multiple pressure sensors [6]. Their prototype can detect the pressure regardless of the size and shape of the floor surface. Miyata et al. developed a system to accumulate grip data using a band-type sensor composed of multiple distance sensors [9]. It works by placing the band on the body to be measured, supporting data collection without mounting anything onto the hand. These researches utilize multiple distance sensors to obtain 3D point clouds that are unattainable with a single sensor alone.

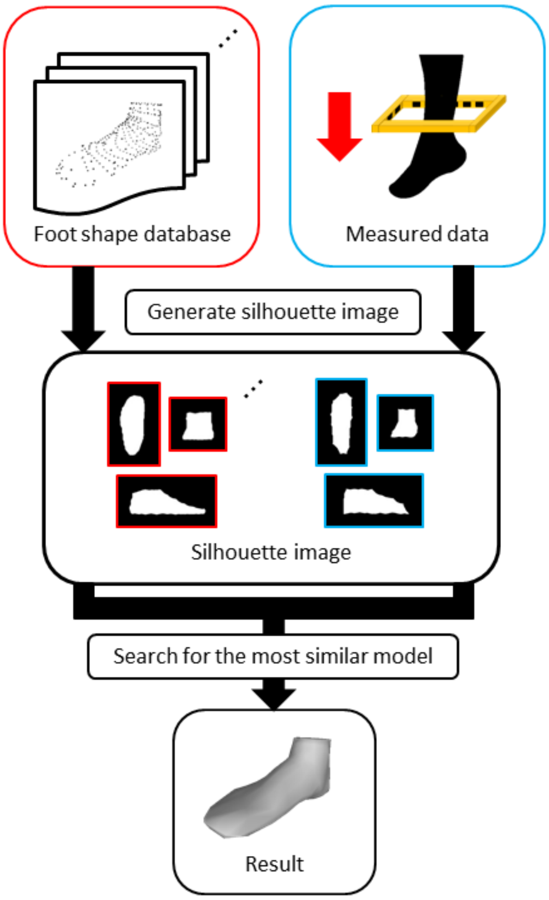

Fig.2 Framework of Proposed Method

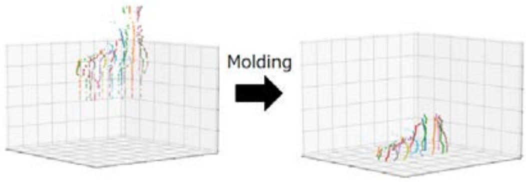

Fig.3 3D Point Cloud of the Measured Data

3. PROPOSED METHOD

Fig.2 demonstrates the framework of our proposed method. The system first measures the foot with the sensors and then uses these values to generate a silhouette image. Similarly, the system generates silhouette image from the foot shape database. The system then compares the silhouette images of measured data and the foot shape database to find the most similar model from the database. When a similar model is found, the system uses this model to reconstruct the 3D shape of the foot.

3.1 Retrieve 3D point cloud from distance sensors

In order to measure the foot, the foot must be placed on a stand with known height, parallel to the long side of the device. The device must be moved up and down horizontally to prevent the deviation of the center of

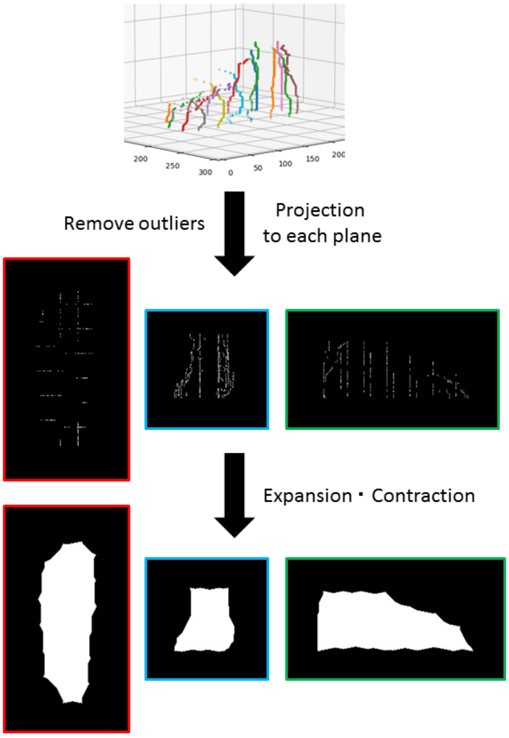

Fig.4 The Generation of Silhouette Images from the Measured Data

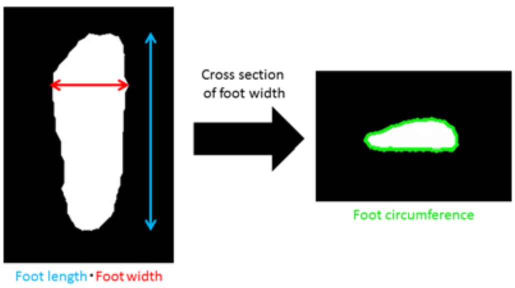

Fig.5 Foot Shape Database to Calculate Foot Size

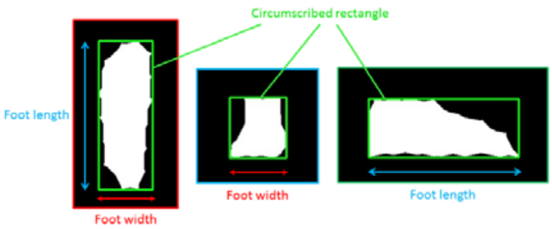

Fig.6 Estimation of Foot Length and Width using Measured Data

gravity. After measuring the foot, the system changes the sensor values from the foot direction sensor to distance information. Here, a median filter is added to reduce unwanted noises when obtaining the sensor values. The system then removes obvious outlier values and combines these results with the base sensor values to create a 3D point cloud. It then removes values of the stand and values above the ankle from the point cloud, leaving only values of the foot, followed by removing duplicated values in the height direction, to form a post-processing of the 3D point cloud as shown in Fig.3 right.

3.2 Formation of silhouette images

The system generates silhouette images by undergoing a repetition process of expansion and contraction on the projected images of each planes; the front, beside and overhead of the 3D point cloud of the measured data and data from the foot database. Fig.4 illustrates the flow of the process. To generate these images, all data undergo a similar process with the exception that the foot shape database undergoes an extra step of removing the outliers. Taking into consideration the distance between two points and the general form of the silhouette image, and making use of a 25x25 oval kernel, the system will repeat the expansion process followed by the contraction process 5 times each. One pixel corresponds to 1mm in the generated silhouette image.

3.3 Foot shape database

The data in the foot shape database used in this research is the 3D measured data of the human body by the Digital Human Research Group of the Human Informatics Research Institute. In the database model, each individual data consists of the same number of the same topology data points and is defined to have the same anatomical meaning. In this research, a homologous model of a total of 68 people consisting of 295 data points is used for each individual data.

Our system calculates the measurement related to the foot size for each model in the database by using the silhouette images. Fig.5 illustrates how the foot size is being measured. The typical measurements are the length from the heel to the tiptoe, the width of the widest part of the foot and the circumference around the foot breadth.

As a result of computing each size in the foot shape database and checking the correlation between the measurements, we observed a strong correlation with a correlation coefficient of 0.88 between the foot width and circumference. The system performs a single regression analysis with the foot width as explanatory variable and foot circumference as purpose variable.

3.4 Similar model search

The system estimates the foot size measurements by using the silhouette images of the measured data. First, the system uses bounding rectangles on the region of the silhouette images to obtain the estimated value of the foot width and length as shown in Fig.6. Subsequently, the system obtains an estimate value of the foot circumference from a regression equation calculated by using the estimated value of the foot width and the foot shape database.

The following are the description of the method to search for the most similar model of the measured foot from the foot shape database. The system compares the foot length, width and circumference and silhouette region from the silhouette images of the homologous model and the measured data. The system first narrows down the models that have the smallest differences in foot length, width and circumference between the data of

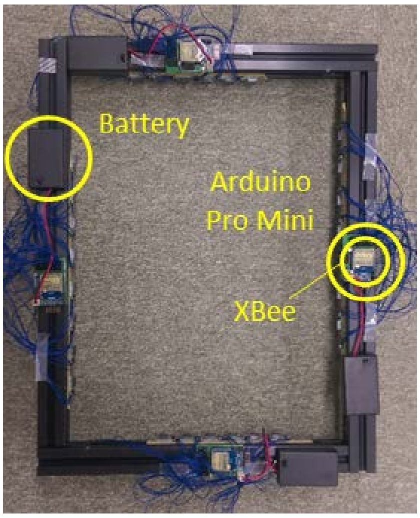

Fig.7 General View of the Device

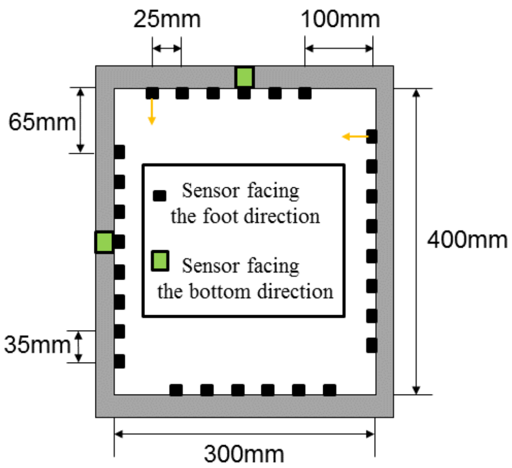

Fig.8 Size and Sensor Placement of the Device

Fig.9 Experiment View

the homologous model and the measured data. Then, the system outputs a model with the smallest difference in the area of the silhouette image region corresponding to the side of the foot from the narrowed down models as the search result, which will be used as the restoration result of the 3D foot shape.

Table 1 Error in each evaluation item of reconstructed model and true values

	Foot Length (mm)	Foot Width (mm)	Foot Circumference (mm)	Average (mm)
Participant 1	5.0	6.0	5.1	5.4
Participant 2	5.0	9.0	1.1	5.0
Participant 3	2.0	4.0	3.2	3.1
Participant 4	2.0	8.0	17.8	9.3
Participant 5	1.0	6.0	1.8	2.9
Participant 6	5.0	15.0	19.1	13.0
Average	3.3	8.0	8.0	6.4

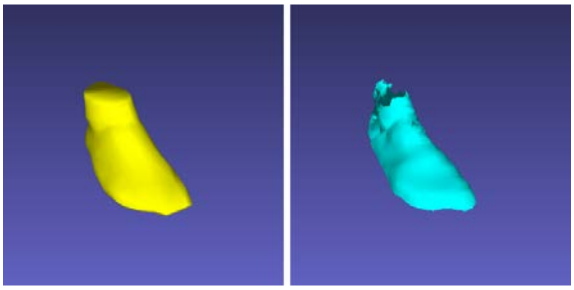

Fig.10 Reconstructed Model (Left), Participant's Foot-Shape (Right)

3.5 Hardware

Fig.7 illustrates the general view of our proposed device. The base of the device is an aluminum frame with a cross section of 30mm x 30mm. The system is composed of batteries as the power source and 4 Arduino Pro Mini as the microcontrollers, with a wireless communication with the PC, connected by XBee.

The device utilizes two different sensors; an inner-facing foot direction sensor (GP2Y0E02A by Sharp) as the distance sensors and sensors facing to the bottom direction (VL53L0X by ST Micro). Fig.8 illustrates the dimensions of each device and the arrangement of the sensors. In order to prevent grounding of the bottom direction sensors while maintaining it in a horizontal position, a foot of 50mm is attached to all four corners of the device.

For the foot direction sensor, as it is necessary to calculate the real distance from the sensor data, we added a mapping of the sensor data and distance to each sensor. The system then can obtain the relational expression between the sensor values and distance from the correspondence of the sensor values when an object is placed at a known distance with respect to each side of the device.

4. EVALUATION

We conducted an experiment to accurately evaluate the reconstructed model from our proposed method, where we evaluated the foot length, width and circumference of 6 male participants. True values of each evaluation item of the participants' foot are taken as the actual measured values. In the experiment, participants will place their right leg on a fixed stand of 20cm and the measurer will move the device from above to a stand of 7cm in height. The measurer will move the device horizontally along the supporting pole to prevent the deviation of the center of gravity of the device. Fig.9 shows the view of the experiment.

Table 1 shows the error in each evaluated item of the reconstructed model and true values of each participant and the average error of each evaluated item. We observed that the average error for each participant was roughly less than 10mm and the average error for foot length, width and circumference was 3.3mm, 8.0mm and 8.0mm respectively, showing that the average error was also less than 10mm.

We then visually compared the foot shape model reconstructed by the proposed method and the 3D shape of the participant's foot acquired by a 3D scanning software included in Google Tango[1] installed in ASUS ZenFone AR as the true value of the participant's foot shape. Fig.10 illustrates an example of the reconstructed model and participant's foot shape, where the reconstructed model was confirmed to match the participant's foot shape and general form.

5. LIMITATIONS AND FUTURE WORK

When measuring the data, the device has to be moved horizontally in order to prevent any deviation in the centre of gravity of the device. In order to simplify the measuring process in the future, we have to take into account how to overcome any deviation or inclination of the device. In addition, at present, there is a limitation to reconstruct an accurate model as the number of homologous model for comparison is as little as 68 and the probability of having the exact homologous model as the user is quite low. In the future, improvement in the accuracy can be expected by increasing the reference model variation through modifying the homologous models using valid parameters on the averaged models as the base.

6. CONCLUSION

In this research, we proposed a method to reconstruct the 3D shape of the foot by utilizing array of distance sensors and comparing it with a homologous model in a known foot shape database. The system first creates a silhouette image for the measured data and for each data in the homologous model. The system then compares these silhouette images to find the most similar model from the database with the measured data and will use this model to reconstruct the 3D foot shape. We also conducted an experiment to evaluate the accuracy of our proposed method where we observed that the average error for the foot length, width and circumference were 3.3mm, 8.0mm and 8.0mm respectively, averaging to 6.4mm. Due to the limited data in the database, our current proposed method may still be challenging for order-made shoes, but it can be used to choose ready available shoes.

ACKNOWLEDGEMENTS

This work was supported by JST PRESTO Grant Numbers JPMJPR17J4.

REFERENCES

[1]Tango — google developers. https://developers.google.com/tango/.

[2] Edmee Amstutz, Tomoaki Teshima, Makoto Kimura, Masaaki Mochimaru, and Hideo Saito. PCA-based 3d shape reconstruction of human foot using multiple viewpoint cameras. International Journal of Automation and Computing, 5(3):217-225, 2008.

[3]Artem Dementyev, Hsin-Liu Cindy Kao, and Joseph A Paradiso. Sensortape: Modular and programmable 3d-aware dense sensor network on a tape. In Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology, pages 649–658. ACM, 2015.

[4]Baback Moghaddam, Jinho Lee, Hanspeter Pfister, and Raghu Machiraju. Model-based 3d face capture with shape-from-silhouettes. In Analysis and Modeling of Faces and Gestures, 2003. AMFG 2003. IEEE International Workshop on, pages 20–27. IEEE, 2003.

[5]Ryuki Nohara, Yui Endo, Akihiko Murai, Hiroshi Takemura, Makiko Kouchi, and Mitsunori Tada. Multiple regression based imputation for individualizing template human model from a small number of measured dimensions. In Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the, pages 2188–2193. IEEE, 2016.

[6] Bruce Richardson, Krispin Leydon, Mikael Fernstrom, and Joseph A Paradiso. Z-tiles: building blocks for modular, pressure-sensing floorspaces. In CHI'04 extended abstracts on Human factors in computing systems, pages 1529–1532. ACM, 2004.

[7]Shunta Saito, Makiko Kouchi, Masaaki Mochimaru, and Yoshimitsu Aoki. Simple system for 3D body shape estimation, IEEE 1st Global Conference on Consumer Electronics (GCCE 2012), pp.203–205. 2012.

[8]Jiahui Wang, Hideo Saito, Makoto Kimura, Masaaki Mochimaru, and Takeo Kanade. Human foot reconstruction from multiple camera images with foot shape database. IEICE TRANSACTIONS on Information and Systems, 89(5):1732–1742, 2006.

[9]Natsuki Miyata, Yuta Sugiura, Takehiro Honoki, Yusuke Maeda, Yui Endo, Mitsunori Tada. Wrap and Sense: Grasping Data Collection System by a Band Sensor. Information Processing Society of Japan Interaction2016 proceedings, (2C49):681–682, 2016.

[10]Ippei Samejima, Satoshi Kagami, Hiroshi Mizoguchi and Makiko Kouchi. A Body Dimensions Estimation Method of Subject from a few Measurement Items using Kinect. Journal of the Robotics Society of Japan, 31(8):761–768, 2013.

[11]Yosuke Ito, Hideo Saito, and Masaaki Mochimaru. Face Shape Reconstruction from Image Sequence Taken with Monocular Camera Using Anatomical Face Shape Database. MIRU2006 Meeting on Image Recognition and Understanding, pages 764–769, 2006.

[12]Toshio Takeuchi, Hideo Saito, and Masaaki

Mochimaru. 3D-Face model Reconstruction from Asynchronous Multi-view cameras. The 10th Meeting on Image Recognition and Understanding (MIRU2007) proceedings, pages 1289–1294, 2007.
[13]Seiji Suzuki, Hideo Saito and Masaaki Mochimaru. Reconstruction of Facial Shape from Freehand Multi-viewpoint Snapshots. Journal of the Institute of Image Information and Television Engineers, 63(4):506–515, 2009.