International Journal of Internet, Broadcasting and Communication Vol.12 No.1 122-129 (2020)
http://dx.doi.org/10.7236/IJIBC.2020.12.1.122

IJIBC 20-1-17

A Study on the Development of a Program to Body Circulation Measurement Using the Machine Learning and Depth Camera

Dong-Gyu Choi[1], Jong-Wook Jang[2*]

[1]*Graduate student, Department of Software Convergence, Dong-Eui University, Korea*
[2]*Professor, Department of Computer Engineering, Dong-Eui University, Korea*
*E-mail : donggyu5878@gmail.com, *jwjang@deu.ac.kr*

Abstract

The circumference of the body is not only an indicator in order to buy clothes in our life but an important factor which can increase the effectiveness healing properly after figuring out the shape of body in a hospital. There are several measurement tools and methods so as to know this, however, it spends a lot of time because of the method measured by hand for accurate identification, compared to the modern advanced societies. Also, the current equipments for automatic body scanning are not easy to use due to their big volume or high price generally. In this papers, OpenPose model which is a deep learning-based Skeleton Tracking is used in order to solve the problems previous methods have and for ease of application. It was researched to find joints and an approximation by applying the data of the deep camera via reference data of the measurement parts provided by the hospitals and to develop a program which is able to measure the circumference of the body lighter and easier by utilizing the elliptical circumference formula.

Keywords: *Machine Learning, Depth Camera, Post Estimation, Realsense, Body Shape*

1. Introduction

Recent advances in video technology using cameras are providing convenience to various fields in the field of industry and living. Smartphones are equipped with two or more cameras, even with basic camera. It maximizes the effect of filming or solves the problem of replacing lenses that were not solved by digital cameras. It is seeking new ways to apply the effect of inserting characters into data taken in such diverse ways and to make inferences on images by applying AI.

Among them, a depth-recognition camera is being used to create a method that has never been seen before. Beyond using passwords by designating existing fingerprint recognition, patterns, and passwords, recognize iris and face in 2D images and use them as passwords. This means that what was previously limited to 2D images is extended to 3D data and used for practical facial recognition. The 3D recognition data is used as a password to unlock mobile phones and security systems in restricted areas. The data from the previous 3D recognition system is used as a password to unlock the security system and mobile phones in restricted areas.

Manuscript Received: January. 10, 2020 / Revised: January. 16, 2020 / Accepted: January. 21, 2020
Corresponding Author: jwjang@deu.ac.kr
Tel: +82-51-890-1709, Fax: +82-0505-182-6904
Professor, Department of Computer Engineering, Dong-Eui University, Korea

3D recognition provides depth data beyond similar colors and shapes in 2D data, and can be recognized even if change their behavior by combining deep learning, which results in higher recognition accuracy than before.

According to one media outlet, the Army Ordnance Command is conducting a "smart mover project" that provides uniforms suitable for those being measured by analyzing body types through 3D telegraph scanners. This 3D body type analysis method is also used in the hospital to provide the necessary degree of treatment to reduce pain and prevent recurrence. These expanded technologies fill the gaps and provide more accurate metrics than traditional technologies. Currently, various body type analysis tools on the market have a variety of customers and are recognized for their needs. However, it is difficult to move devices because of the large volume, which is limited to the installation environment, and because of its greater use as a medical device, it is difficult to use them in everyday life because of its high price range. In contrast, smaller 3D measurement tools are easy to measure, but do not lack the power to maintain or are highly utilized. Therefore, in this paper, we developed a program to provide as much data as necessary through the operation of changing to actual figures by utilizing a single depth camera and deep learning-based joint-tracking OpenPose library to reduce the constraints as above. The composition of the paper is as follows. In section 2, we describes open-pose libraries and real-sense depth cameras, and checks the K-means algorithm used to reduce errors in data. Section 3 describes the flow of this program and how to calculate the body type. Section 4 identifies actual program operation and result values and concludes in Section 5.

2. Related research

2.1 OpenPose

OpenPose is a deep learning-based open-source project announced by Carnegie Mellon University's artificial intelligence imaging team to CVPR 2017[1-3]. The Deep Learning framework uses a Caffe library. Based on the computer language C++ and OpenCV library, it helps to estimate the person's face and body parts and fingers and feet in real time, and jointly detects a total of 135 key points in a single image. The OpenPose library learns about 100,000 images data and gets joint weights to recognize coordinates on the highest score at Key Point. The recognition method is as follows. If image data is entered, the reliability of each part will be understood and the final decision will be made after determining whether the deduced location is appropriate. OpenPose has been produced in a variety of ways, and single-person reasoning is carried out in the following way of Fig.1. First, through camera filming, it is converted into image data that can be used for CNN and weighted in the heatmap method to detect certain parts of the body. It then gives partial preference to find association with joints, and links measured areas with deduced associative coordinates.

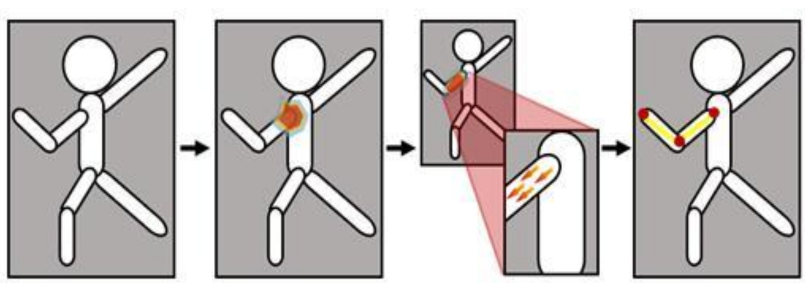

Figure 1. Overall pipeline

2.2 Realsense Depth Camera

The Realsense Depth Camera is a depth perception camera built by Intel Corp. The main specifications are as follows[4,5]. The cameras used in the study are Realsense D435 and Time-of-Flight (ToF) cameras. In the above camera image, infrared light is emitted through the IR (Infrared) sensor at the filming section. The emitted infrared rays are reflected back after touching the obstruction and the signal is received by the ToF sensor. The longer the infrared rays return, the longer the time is, so the distance is measured by the difference according to that time. Unlike other ToF cameras, D435 uses two ToF sensors, which can be used in a Stereoscopic manner, to correct the difference accepted in two places, such as the human eye, for more accurate values. A typical ToF camera has only one ToF sensor. However, D435 uses two ToF sensors in a Stereoscopic manner, which can be used similarly to the human eye to correct the difference accepted by the two values to obtain a more accurate value.

Figure 2. Intel® realsense depth camera D435 specifications

2.3 K-means Algorithm

K-means Algorithm generally has the following goals[6].

$$arg_s min \sum_{i=1}^{k} \sum_{x \in S_i} ||x - \mu_i||^2 \tag{1}$$

In (1), when n sets of d-dimensional data objects are given, K-means Algorithm divides data objects (n) into sets (S) that have smaller numbers (k) than objects that maximize cohesion between objects within each set. The goal of this algorithm is to find a set S with the least squares of the distance between objects in the set at the center point of each set, when μ is the center point of the set S.

3. System Design

3.1 Suggested Program Progress Order Overview

There was previously a Tele-MFAsT study conducted with Kinect cameras in the field of rehabilitation, but it is not being used very well, and the hospital also feels the need for its usability[7]. For other devices with similar performance, access is not easy because they are large, or because the program is classified for medical purposes and therefore expensive. Accordingly, this paper proposes a program and its method to measure a person's body shape with a single ToF method Depth camera. First, it shows the sequence of human body shots in the program proposed in Fig. 3. First, when the program runs, take a three-way picture of the person you want to measure. This is similar to the 3D scan method. Previously, the camera was rotated so that there was no shortage of data, but the proposed method simplified the process. The sequence is performed with the basic attention position, the T-shaped position with the arms open, and finally the side.

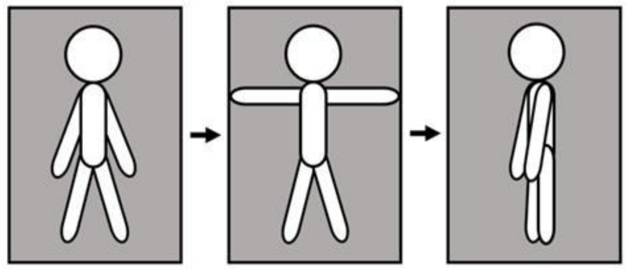

Figure 3. Flow diagram for body shape measurement

About 30 frames will be filmed for each posture, where the OpenPose Library will find human figures and use the data. In addition to the joint position coordinate data, Depth data is entered through the Realsense camera. However, due to the characteristics of the ToF sensor, the value has noise. In this part, K-means Algorithm stabilizes the value to help identify the body shape.

3.2 Body type calculation method

Because the human body is not a perfect circle, a formula was used to find the circumference of the ellipse. An approximate image of the measurement is as follows. If there is data from a and b for measuring the perimeter, as in Fig. 4, then the circumference can be checked by simple computation. A is the measured length data from the front, and b uses the measured length data from the side.

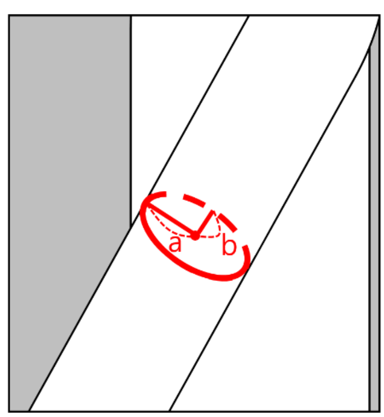

Figure 4. Image showing how to check the perimeter

A and b are the radius of long axis and shortening, and the circumference operation of ellipse using this value is as (2).

$$\mathrm{l} = \left\lceil 2\pi\sqrt{Ra^2Rb^2} \right\rceil \tag{2}$$

To find the circumference of the ellipse, a formula was used that deals with approximate values. The reason is that a correction filter was used to set the depth value of the Depth camera, but the value is not very accurate. For the above operation, you need half the depth and half the width, or Ra, the horizontal radius of the ellipse, and Rb, the radius of the vertical. To obtain these two values, it is necessary to obtain the length of the line perpendicular to the joint. Fig. 5. shows how to measure the actual length using Depth Scale of the Realsense camera at the coordinates of the required site point to be measured. P1 and p2 represent the point at both ends of the joint_dest, the perpendicular line of the skeleton being measured.

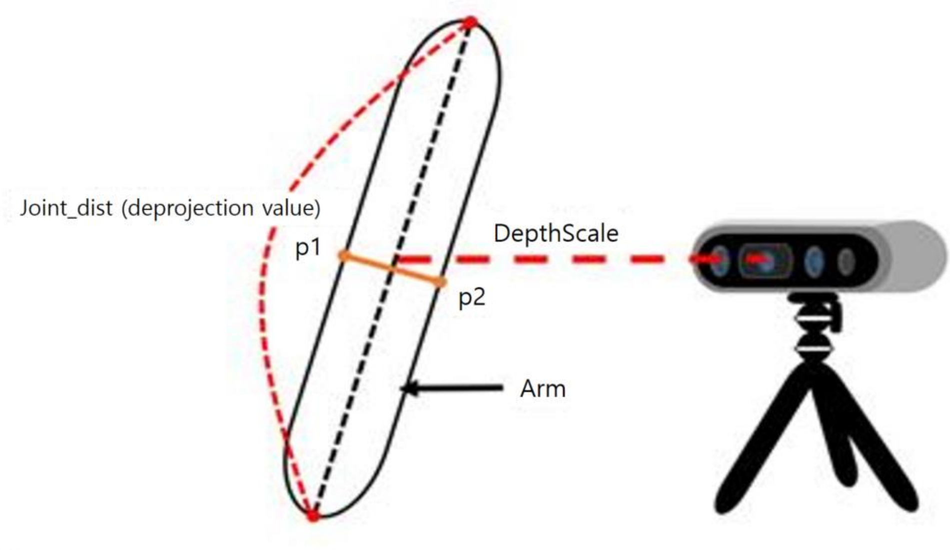

Figure 5. The process of checking the thickness at the arm

The actual length value is returned when the Deprojection function used in the SDK within Realsense is used to enter the desired line, i.e. the coordinate values of both ends of the skeleton. To find a vertical line in the joint_dist, an operation is performed to find a right angle for all coordinates in the joint_dist, and the formula is as (3).

$$\theta_n = atan\frac{y_n - y_{n+1}}{x_n - x_{n+1}} \tag{3}$$

The coordinates used for x and y are p1 and p2 and n is the joint measured first. When viewing the angle position that you want to obtain by expressing this method as an image, it can be expressed as follows. Use the above angle to select one of the most sharply different depth values at right angles. This value is the coordinates to be used to connect the lines.

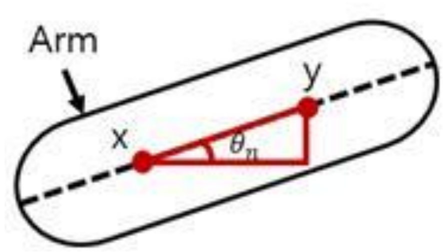

Figure 6. Angle between points

Once the coordinate value is found, K-means Algorithm is used to stabilize the depth value. This helps to differentiate the values and find more accurate lengths. When the bottom and top end coordinates are found, they should be converted to the size of reality. Using DepthScale and PointDist, which are the length values of the joints and the distance from the Depth camera, find the alignment coordinates perpendicular to the joint coordinates that existed in Fig. 5. and obtain the thickness values using formula (4).

$$\text{Thickness} = ((|(\mathrm{p}_1[0] - \mathrm{p}_2[0])| + |(\mathrm{p}_1[1] - \mathrm{p}_2[1])|) * (JointDist * DepthScale)) \quad (4)$$

4. Implementation

Fig. 7. was filmed performing a program designed for actual testing. The T-positions and aspects from the front and the front are measured according to the sequence of progress in Section 3 design. The location and personnel were changed for each shoot, and the last shot confirmed that the opposite direction of the side was also applicable.

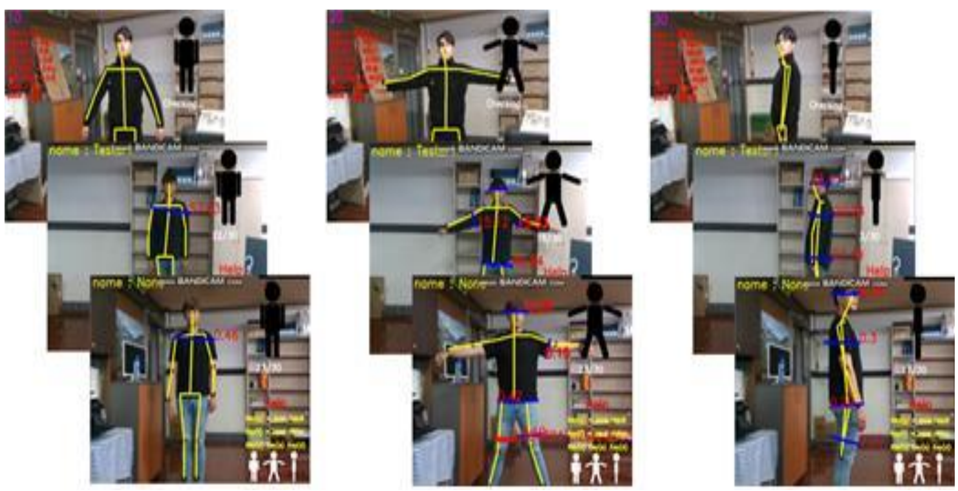

Figure 7. Image of the program operation

The images were 30 frames per shot, or 30 scenes. The computer used in the experiment was used with the Nvidia GeForce GTX 1080 TI graphics card configured as SLi, and the source of the Dataset was referred to CoCo[8]. Python was used as a development language and the Tensorflow Deepening Framework was used to use deep learning. The resulting total measurement time was about three minutes, as it was operated when all the joints were deduced by each frame. The 30 frames of data recorded were immediately stored and averaged and used as the radius value corresponding to the horizontal and vertical lines to be performed in the elliptical circumference operation based on that value. As shown in Fig. 8, all around the body involved were actually measured to compare program measurement data with how different the actual results were.

Figure 8. Body measurements and reference images for program testing

The measurement index was determined using data from the Busan University Hospital Department of Rehabilitation Medicine, which was written in reference to the Korean Human Body Index survey in 2016[9]. The measurement of the program also used the above metrics, and the corresponding actual data and program data were shown in the following table 1. Overall, similar values were found to be generated and, on average, within ±10% of the body shape was not significantly different from the actual measurement site used in hospitals. For areas with motion, such as the right arm, a better outcome could be obtained if the arm and body were to be dressed in distinct colors or if there was no interference in the background.

Table 1. The value for proposed measurement and actual measurement

First test – Round(cm)					
Proposed	Chest	120.9	Actual Measurement	Chest	116
	Head	63		Head	63
	Arm(R)	57.8		Arm(R)	42
	Arm(L)	39.5		Arm(L)	41
	Thigh(R)	51.6		Thigh(R)	56
	Thigh(L)	51.7		Thigh(L)	55
	Waist	106.4		Waist	111
Second test – Round(cm)					
Proposed	Chest	97.9	Actual Measurement	Chest	98.5
	Head	56.4		Head	59
	Arm(R)	37.3		Arm(R)	32
	Arm(L)	33.9		Arm(L)	31
	Thigh(R)	54.3		Thigh(R)	53
	Thigh(L)	53.7		Thigh(L)	53
	Waist	98.1		Waist	100
Third test – Round(cm)					
Proposed	Chest	103	Actual Measurement	Chest	101
	Head	62.5		Head	60
	Arm(R)	36.1		Arm(R)	33
	Arm(L)	36.4		Arm(L)	34
	Thigh(R)	60.6		Thigh(R)	58
	Thigh(L)	61.3		Thigh(L)	58
	Waist	97.2		Waist	96

5. Conclusion

Even at this time of technological advancements, Microsoft-produced Depth camera Kinect continues to be used in the ToF-type Camera market, and the capabilities of OpenNI and NiTE also continue to show consistent performance. The performance and brevity of Camera Calibration with frame, Depth, and RGB values in Camera photography are better handled by Kinect. However, Kinect did not have any problems finding the joints, but in the course of this study, it was confirmed that there were variations depending on the environment and that the accuracy and location of the joints became ambiguous. In fact, hospitals need real joint angle data and accurate indicators that are not provided by conventional devices.

Therefore, in this study, the body circumference measurements were made by using machine learning to track joints and perform simple calculations to improve existing methods. The processing speed of joint tracking using machine learning is somewhat slow. However, the angle of the joint is measured more accurately because there is less interference with the background and more accurate positioning than the way in which video processing is used, such as NiTE. Exceptions were applied to identify possible joints deviating significantly from their existing positions. This has confirmed a significant reduction in joint tracking problems. Difficulties and time spent on measurements that were pointed out as other problems also improved. Because smaller cameras are used and easier to connect than other devices in measuring the actual size and circumference of the human body, there is no difficulty in using them for the first time, even without assistance. It is simply very easy to measure a body shape, and there is no contact compared to using a tape measure as a measurement. In terms of current hospital operating costs, the overall purchase price of hardware used in the study is very low. Thus, if the program is improved and provided at an appropriate price, the efforts used to measure the hospital's body will be utilized in patient care, the patient's hospital reliability will be improved and the quality of medical care will be improved. There are 'Video stabilization algorithm of shaking image using deep learning' and 'Image recognition based on adaptive deep learning' that have been studied with deep learning using image processing[10,11]. It is expected that this implemented program can be applied to many other areas if it is further improved by utilizing these various machine learning techniques, new deep learning models, and physical measurements.

Acknowledgment

The research was supported by the BB21+ project in 2019.

References

[1] Z.E. Wei, V. Ramakrishna, T. Kanade, and Y. Sheikh, "Convolutional pose machines," *CVPR*,. 2016.

[2] Z. Cao, T. Simon, S.E. Wei, and Y. Sheikh, "Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields," *CVPR*,. 2017.

[3] Z. Cao, G. Hidalgo, T. Simon, S.E. Wei, and Y. Sheikh, "OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields," *arXiv*,. 2018.

[4] S. C. Kim, M. Song, S. Y. Shin, J. J. Kim, and T. M. Shin, "Development of System for Prescreening of Scoliosis Patient Using RealSense and PID Control," *Korea Institute Of Communication Sciences*, pp.658-659, 2017.

[5] Intel Realsense Depth Camera D400-Series, *https://software.intel.com/en-us/realsense/d400*

[6] K-means. *http://stanford.edu/~cpiech/cs221/handouts/kmeans.html*

[7] Gregorij KURILLO, Jay J HAN, Alina NICORICI and Ruzena BAJCSY, "Tele-MFAsT : Kinect-Based Tele-Medicine Tool for Remote Motion and Function Assessment," Medicine Meets Virtual Reality 21, pp.215-221, 2014.

[8] T.Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C.L. Zitnick, and P. Dollár, "Microsoft COCO: Common Objects in Context," *arXiv*, 2015

[9] Korean Human Body Size Survey 2016. *http://kostat.go.kr/portal/korea/kor_ pi/8/6/2/index.board?bmode=read&aSeq=358293*

[10] Kyung-Min Lee and Chi-Ho Lin, "Video Stabilization Algorithm of Shaking image using Deep Learning", *The Journal of The Institute of Internet, Broadcasting and Communication(JIIBC)*, Vol.19, No.1, pp.145-152, Jan 2019. DOI: https://doi.org/10.7236/JIIBC.2019.19.1.145

[11] Jin-Woo Kim and Phill-Kyu Rhee, "Image Recognition based on Adaptive Deep Learning", *The Journal of The Institute of Internet, Broadcasting and Communication(JIIBC)*, Vol. 18, No.1, pp.113-117, Feb 2018. DOI: https://doi.org/10.7236/JIIBC.2018.18.1.113