Journal of Physics: Conference Series

PAPER • OPEN ACCESS

A simple vision based anthropometric estimation system using webcam

To cite this article: Xu Yang Gan *et al* 2020 *J. Phys.: Conf. Ser.* **1529** 022067

A simple vision based anthropometric estimation system using webcam

Xu Yang Gan, Haidi Ibrahim, and Dzati Athiar Ramli

School of Electrical and Electronic Engineering, Engineering Campus, Universiti Sains Malaysia, 14300 Nibong Tebal, Seberang Perai Selatan, Pulau Pinang, Malaysia

E-mail: `haidi_ibrahim@ieee.org`

Abstract. Anthropometry can be defined as the science of dimensional measurement of the size and proportions of the human body. Anthropometric measurement is used in a lot of fields such as medical, forensics and clothing. The modern anthropometry techniques are expensive and takes up a lot of space while traditional anthropometry techniques are slow and susceptible to human error. Hence, there is a need for the development of a vision-based estimation system that is cheap, fast, accurate and portable. The study in this project proposes a noble way of using Raspberry 3B+ with Logitech C310 camera to perform estimation of chest, waist, hip circumference and body height through 2D images. The system works by implementing Pixel Density Method, HSV thresholding method, OpenPose deep neural network, bounding box method, pixel counting method and Ramanujans ellipse circumference approximation method to achieve objectives of this study. The percentage error between the estimated result and the measured result of chest, waist, hip circumference and body height are only 2.11%, 4.66%, 4.31% and 1.74%. In conclusion, a vision based anthropometric estimation system is successfully developed with high accuracy using all the proposed methods.

1. Introduction

Anthropometry can be defined as the science of dimensional measurement of the size and proportions of the human body, whether living or dead [1, 2, 3]. Anthropometry measurement techniques nowadays can be classified into two types, namely traditional and modern [3, 4]. Traditional technique uses tools such as callipers and measuring tape to perform measurement directly on the body of the person [1, 3]. On the other hand, modern techniques use advanced machines such as Magnetic Resonance Imaging (MRI), X-Ray and Coordinate Measuring Machine (CMM) which can obtain anthropometric measurements instantly and accurately [4].

Lately, there has been significant interest in the development of a system which estimates human anthropometry using images. Compared to traditional measurement techniques, image-based anthropometric measurement system is contactless in which data are extracted from the image captured and the process is said to be simple and accurate [3]. The act of extracting and analysing information from digital images can be defined as a concept of computer vision [5].

Development of a computer vision based anthropometric system has a lot of potential and has been used in a lot of sector, such as forensic, health and clothing [6]. In forensics, apart from physical traits such as colour of hair, facial features and body build, surveillance images can be used to estimate the height of a person which aids in identifying suspects [7]. According to Parihar et al. [5], through volume estimation using a prediction model based on regression analysis for human body, the weight of a person can be estimated, which serve as another important clue in identifying fugitive or criminal in forensics. In the health sector, the weight and height of a person can be quickly and easily obtained to calculate Body Mass Index (BMI) which serves as an important health indicator that determines whether a person is underweight, normal or obese [5, 6]. In the clothing sector, it has been used in cloth fitting. There exists commercial application of vision-based anthropometry in the form of virtual trial room in which users can try on different clothes without putting them on physically [2, 8, 9]. However, those implementation uses 3- Dimensional (3D) cameras in the form of Microsoft Kinect [10] which is expensive and not affordable to small and medium businesses.

Compared to modern arthrometric measurement technique using advanced machines, development of image-based measurement system is cost effective and relevant as the advanced machines are expensive and hard to use [4]. Besides, the advanced machines usually take up a lot of space and is not portable. The development of vision-based system solves the issues of the clothing industry where small tailor shop owners cannot afford and accommodate the space required for advanced machines used for body measurements.

Hence, there is need for the development of a low-cost, fast and accurate vision based anthropometric estimation system using 2D images and this paper proposes the way of developing it. However, the scope of project is limited to just measurement of chest, waist, hip circumference and body height.

The rest of the paper is organized as follows. In Section 2 describe related works similar to the method proposed in this paper. Section 3 covers the various steps involved in the hardware and software development of vision based anthropometric estimation system. Section 4 provide the discussions and results. Lastly, conclusion and future works are reviewed in Section 5.

2. Related works

Hung et al. [11] proposes a way in anthropometric estimation using 2D image and their research is able to yield a 3cm difference between the result through image processing and result from actual measurement for the chest circumference measurement. It is noted manually selects the body landmarks to measure which is prone to human error and is inconsistent.

Tomi et al. [10] on the other hand uses Microsoft Kinect 3D camera to perform anthropometry. The result obtained by Tomi et al. [10] has an error average mean of 1.33cm for the hip circumference measurement. It is noted that it uses Kinect camera specific SDK to identify the body parts to be measured. It is also stated that Ramanujans ellipse circumference approximation equation is the best formula to estimate the circumference of chest, waist and hip.

3. Proposed methodology

In this work, the chosen hardware to implement the vision based anthropometric estimation systems are Raspberry Pi 3B+ and Logitech C310 camera due to their cheap cost and good performance. The operating system chosen for the Raspberry Pi is Raspbian because it is the official operating system and due to its stability and performance. OpenCV is selected as the image processing library because of it supports Raspberry Pi and it is widely documented. Python is chosen as the programming language due its simplicity, speed and being the official programming language of the Raspberry Pi. The data acquisition using camera is conducted in a brightly illuminated environment with bright coloured background. The camera is oriented in a vertical position when taking images. The general overall flow of proposed system is illustrated in Fig. 1.

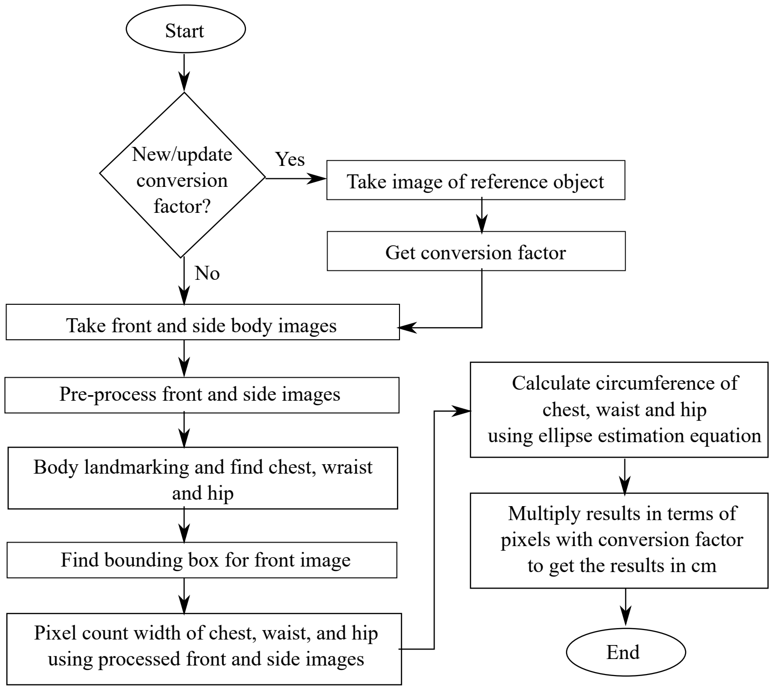

Figure 1. The overall flow of the vision based anthropometric estimation system.

3.1. Pixel Density Method

This system proposes Pixel Density Method [10] to estimate real world object dimensions using 2D images. The Pixel Density works by using a reference object to obtain a conversion factor that relates pixel size measurement with real-world size measurement. The following equations best describes the above theory:

$$X = Hx \tag{1}$$

where X is the actual dimension of the object, H is the conversion factor while x is the image dimension of the object. If the actual width of the reference object is W (in cm or m), and the width of that reference object is w (in pixels), then, the conversion factor H can be described using the following equation:

$$H = \frac{w}{W} \tag{2}$$

An A4 paper with width 29.7cm and height 21.0cm is used as reference object. The A4 paper is captured in a horizontal position hence 29.7cm is the known width of the reference object. Contour detection is used to get the width of the reference object.

3.2. Human subject data acquisition

Two images are captured namely the front view of the human body and the side view of the human body because of its importance in estimating the circumference of the chest, waist and hip.

3.3. Image pre-processing

HSV thresholding is chosen as the method of background removal. The images of human subject are pass through a HSV thresholder which allows all values of Hue and Value to pass through but limits the Saturation value to only 50-255. This results in a white human silhouette with black background called a mask. The mask is saved for later step to use.

3.4. Body landmarking and finding the chest, waist and hip
OpenPose deep neural network using MPII dataset is used for body landmarking purposes. The front and side view of the human subject image is past through the neural network producing 15 keypoints for each image. The keypoints are namely Head – 0, Neck – 1, Right Shoulder – 2, Right Elbow – 3, Right Wrist – 4, Left Shoulder – 5, Left Elbow – 6, Left Wrist – 7, Right Hip – 8, Right Knee – 9, Right Ankle – 10, Left Hip – 11, Left Knee – 12, Left Ankle – 13, Chest – 14, Background – 15. Chest region is found by calculating the midpoint of Neck – 1 and Chest – 14. The hip region is found by calculating the midpoint of Right Hip – 8 and Left Hip – 11. The waist region is calculated by finding the midpoint of the calculated waist region and Chest – 14.

3.5. Estimating height using bounding box
The height of the human body is found by first finding the contours of the front view human body image which result in a bounding box. Then the height of the bounding box is identified and used as the estimated height.

3.6. Pixel count width of chest, waist and hip
The HSV thresholding resulting mask for the front view and side view of the human body is firstly processed using dilation and erosion to eliminate black holes inside the white silhouette of the masks. Then, Canny edge detector is used on them resulting in edge maps as final image.

Using the edge maps and the chest, waist and hip region identified previously, the pixel counting is performed sideways from the region of interests. This is done by comparing each of the pixels sideways until it detects the white edge then it will stop counting.

3.7. Calculating the circumference of chest, waist and hip
Ramanujans ellipse circumference approximation equation [10] is used to approximate the circumference of chest, waist and hip. The equation is shown in the following:

$$p \approx \pi \left[3(a+b) + \sqrt{(3a+b)(a+3b)}\right] \tag{3}$$

where a and b are the radiuses to estimate an ellipse as shown in Fig. 2. Equation (3) is used to calculate the circumference of chest, waist and hip by designating the front view width as $2a$ and side view width as $2b$. $2a$ is the width of the front view chest, waist or hip obtained previously through pixel counting and $2b$ is the width of side view chest, waist or hip obtained previously through pixel counting.

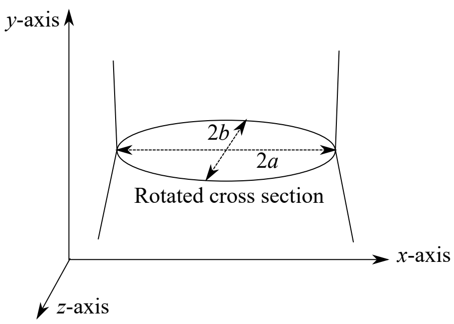

Figure 2. Approximation of an ellipse on tranverse plane.

3.8. Getting estimation result in terms of centimeter

Finally, all the pixel result of chest, waist, hip circumference and body height in terms of pixel is multiplied by the conversion factor obtained previously to get chest, waist, hip circumference and body height estimation in centimetre as stated in Equation (1).

4. Results and discussions

Figures 3 to 6 show the data acquisition results and the final width measurement results of the front view and the side view of the human body. To determine the proposed systems accuracy, the chest, waist, hip circumference and body height are measured physically using measuring tape and height measurer. The result of both the measured and estimated value are shown in the following table.

Figure 3. The front view of the human body.

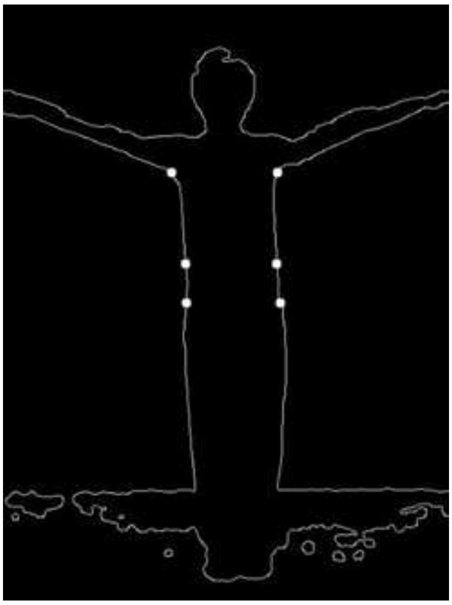

Figure 4. The front view width measured.

Based on the result in Table I, it can be concluded that the vision based estimation system developed in this project is very accurate as it is similar to the accuracy done by Hung et al. [11] using 2D image but still not on par with Tomi et al. [10] which uses 3D camera.

Table 1. Comparing the measured and the estimated results of the anthopometric measurement.

Description	Measured (cm)	Estimated (cm)
Chest circumference	85	86.7976
Waist circumference	75	78.4959
Hip circumference	80	83.4483
Body height	175	172.018

Figure 5. The side view of the human body.

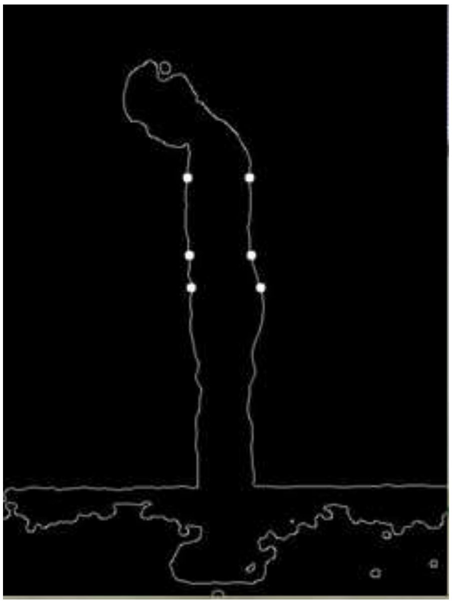

Figure 6. The side view width measured.

The percentage error is calculated between the measured and estimated result to judge its accuracy. The following equation is used for calculation:

$$\text{Percentage error } (\%) = \frac{|Measured - Estimated|}{Measured} \times 100\% \tag{4}$$

Table II shows the percentage error between the measured and estimated result. Based on result from Table II, it can be concluded that the system is very accurate due to its small percentage error.

Table 2. Percentage error between the measured with the estimated result.

Description	Percentage error (%)
Chest circumference	2.1149
Waist circumference	4.6612
Hip circumference	4.3104
Body height	1.7041

5. Conclusion and future works

The implementation a low-cost, compact and accurate vision based anthropometric estimation using Raspberry Pi 3B+ and Logitech C310 camera is successfully achieved. The percentage errors between the measurement given by the proposed system with the normally measured result for chest, waist, hip circumference and body height are only 2.11%, 4.66%, 4.31% and 1.74% respectively. This means that the proposed system can provide measurements with high accuracy.

The automatic body landmarking algorithm is successfully implemented using OpenPose deep neural network with the MPII dataset. The region for chest, waist and hip are successfully found through calculations using the output keypoints of OpenPose as reference.

A recommendation for future work inspired from modern smartphones having time-of-flight (ToF) camera to supplement its primary camera for depth sensing is proposed that the Pixel Density Method used in this project is replaced with a distance sensor such as ultrasonic sensor, infrared sensor or ToF laser sensor to realise the Focal Length Method.

Through this approach, the reference object can be removed, and the human subject can stand at a flexible distance from the camera and not fixed specifically as the same distance when taking the reference object image.

Acknowledgments
The author gratefully acknowledges the School of Electrical and Electronic Engineering, Engineering Campus, Universiti Sains Malaysia, Nibong Tebal, Penang, Malaysia as this work was supported in part by the Universiti Sains Malaysia Research University Grant 1001/PELECT/8014052.

References
[1] Kainz O, Cymbalk D, and Jakab F 2015 *Int. J. Model. Optim.* **5** 300
[2] Estrada JE, Benedito AA, Belicano JLD, Galang MRR, and Cabrera JKC 2017 *2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP)* (Singapore: IEEE) p 262
[3] Abdelkader BC, and Yacoob Y 2008 *Computational Forensics* (Berlin: Springer-Verlag) 200
[4] Stancic I, Supuk T, and Ceci M 2009 *WSEAS Transactions on Systems* **8** 430
[5] Parihar AS, Gupta M, Sikka V, and Kaur G 2017 *2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (Delhi, India: IEEE) p 1
[6] Nguyen LT, and Nguyen TH 2016 *BRAIN. Broad Research in Artificial Intelligence and Neuroscience* **7** 5
[7] Criminisi A 2002 *Joint Pattern Recognition Symposium* Berlin, Heidelberg: Springer p 224
[8] Saha S, Ganguly P, and Chaudhury S 2014 *Proceedings of the 2014 Indian Conference on Computer Vision Graphics and Image Processing* (Bangalore, India: ACM New York, USA) p 1
[9] Markiewicz L, Witkowski M, Sitnik R, and Mielicka E 2017 *Expert Syst. Appl.* **85** 366
[10] Tomi B, Sunar MS, Mohamed F, Saitoh T, Mokhtar MKB, and Luis SM 2018*2018 IEEE Conference on e-Learning, e-Management and e-Services (IC3e)* p 56
[11] Hung PCY, Witana CP, and Goonetilleke RS 2004 *Computing Systems* **29** p 764