

Human Gait Feature Extraction based-on Silhouette and Center of Mass

Miftahul Jannah¹, Sarifuddin Madenda², Tubagus Maulana Kusuma³, Hustinawaty⁴

Faculty of Computer Science and Information Technology
Gunadarma University, Jakarta, Indonesia

Abstract—When someone walks, there is a repetitive movement or coordinated cycle that forms a gait. Gait is different, unique and difficult to imitate. This characteristic makes gait one of the biometrics to find out one's identity. Gait analysis is needed in the development of biometric technology, such as in the field of security surveillance and the health sector to monitor gait abnormalities. The center of mass is the unique point of every object that has a role in the study of humans walking. Each person has a different center of mass. In this research, through a series of processes in image processing such as video acquisition, segmentation, silhouette formation, and feature extraction, the center of mass of the human body can be identified using a webcam with the resolution of 640 x 480 pixels and the frame rate of 30 frames/second. The results obtained from this research were gait frames of 510 frames from 17 pedestrian videos. Segmentation process using background subtraction separates the pedestrian object image from the background. Silhouette gait was produced from a series of image enhancement processes to eliminate noise that interferes the image quality. Based on the silhouette, feature extraction provides the center of mass to distinguish each individual's gait. The sequence of center of mass can be further processed for characterizing human gait cycle for various purposes.

Keywords—Human gait; center of mass; silhouette; feature extraction; gait cycle; people identification

I. INTRODUCTION

Walking is a movement that allows one to move from one place to another by moving the foot forward in the correct position alternately [1]. Repeated movements or coordinated cycles form a gait. Every individual's gait is different, which makes it unique and difficult to imitate. These characteristics then make gait one of the biometrics to find out one's identity.

Biometrics is a technology of recognition and identification based on physiological or behavioral characteristics possessed by humans such as gait, face, voice, iris, fingerprint, etc. [2]. However, biometric recognition with face, sound, iris, and fingerprint cannot be done remotely and requires interaction with the subject to be observed. In contrary, gait biometric not require direct contact with the subject to be observed making the image acquisition of gait can be performed easily in public places as well as captured remotely. Gait is difficult to be hidden and engineered. This characteristic is very important in the surveillance system.

Gait identification is needed in various applications, such as in the health sector where identification is intended to identify the type of disease from abnormal gait motion. Gait

identification also has an important role in video surveillance and access control systems for supervision and security, for example in security sensitive environments such as airports, banks and certain spaces. Other information can be obtained through identification of gait such as age, race, and gender.

In order to identify human gait, image processing is needed, which consists of several stages, such as capturing gait videos using a camera. Captured videos that consist of a set of image frames are extracted so that they can be processed frame by frame to produce a silhouette image. Feature extraction of silhouette images is the key step in identifying gait. Silhouette images represent binary maps of human walking, forming strong features to represent gait because they capture the movements of most parts of the human body [3]. Feature extraction has been carried out in existing studies such as extraction of the entire human body [4] or some limbs such as the waist, hips, feet [5].

A number of research related to gait extraction features have been carried out. As proposed in [6], gait extraction based on features of distance and angle between the two legs using Hough Transform was proposed. This study produced a silhouette and skeleton gait. Different approach was shown in [5], where gait was analyzed using DGait database. Feature extraction from 2D and 3D body silhouettes for gait identification was performed. Support Vector Machine (SVM) kernels were used for classification. This research successfully compared the two features (2D and 3D). In [7], Multi-scale Principal Component Analysis (MSPCA) was proposed, which performed gait recognition based on modelling limbs using a spline curve. The feature extraction uses the CASIA-B Gait Database silhouette dataset. For classification Neuro-Fuzzy and K-Nearest Neighbors (KNN) was used. Another approach was introduced in [8], where Kinect camera sensor was used for acquisition. The feature extraction process used static features and dynamic features such as wrists, ankles, body, knees, shoulders, arms and thighs. K-Nearest Neighbors (KNN) was used for gait classification. The research output is a database containing 20 pedestrians walking from right to left.

It was shown that image frames and silhouettes used in the previous research were not captured and processed in real time but based on provided datasets. Feature extraction has been done on several members of the human body but has not used the Center of Mass (CoM) that has a role in the study of humans walking. Therefore, real-time process of acquisition, silhouette generation, and feature extraction based on CoM are proposed in this research. The importance and findings of this

research is the extraction of CoM sequence from human walking cycle, which can be further used for classification purpose.

II. METHODOLOGY

The feature extraction method consists of a number of stages, namely video acquisition, image segmentation, and silhouette generation/forming as shown in Fig. 1. The output of feature extraction process is the CoM and its location in the human body image.

A. Video Acquisition

The video acquisition was performed in real time when someone walks in front of the camera. Webcam was positioned 50 cm above the ground with the distance to the object of 3 meters. The viewing angle of the webcam is 90°. Gait analysis of the video files can only be done frame-by-frame. The video resolution of the gait frame extraction process is 640 x 480 pixels.

The total number of video used in this research is 17 videos, where each video contains human pedestrian. In the process of frame extraction, all frames in the video were extracted. Afterward, using the extracted frames, background images were captured. The specification of the video acquisition is shown in Table I.

Fig. 2 shows the stages in identifying the CoM is the gait video acquisition process which consists of camera calibration, video recording, and video frame extraction. The camera calibration process is used to adjust the camera position, distance, capturing angle, and adequate lighting in order to obtain biometric information of the gait cycle and the configuration of the device used for the gait video recording process. The recording process starts with recording the background and recording the person walking in front of the background. Recorded videos are saved in .avi video file format. Once the video is recorded, frame extraction is taken place to produce a series of image frames to be able to do the next process as shown in Table II.

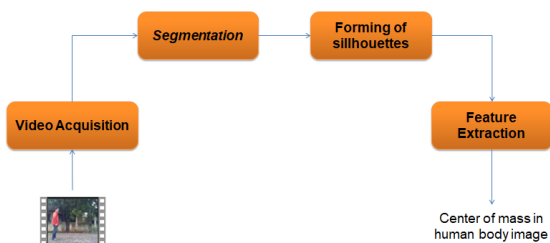


Fig. 1. Research Methodology.

TABLE I. THE SPECIFICATION OF THE VIDEO ACQUISITION

Number	Specification	Description
1	Camera	Logitech C170 Webcam
2	File Format	.avi
3	Duration	5 seconds
4	Video Resolution	640 x 480 pixels
5	Frame Rate	30 frames/second
6	Bit Depth	24-bit
7	Sensor Resolution	True 2 Megapixels

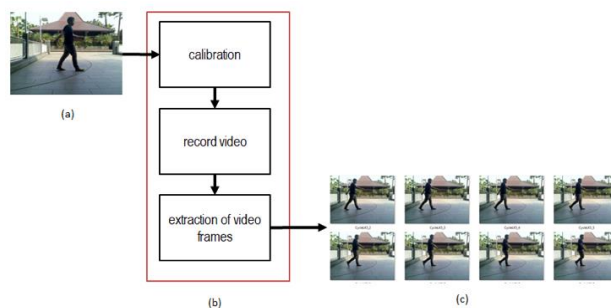


Fig. 2. Video Acquisition Process.

TABLE II. EXAMPLE RESULTS OF FRAME EXTRACTION

Frame background	Frame 1	Frame 10
Frame 20	Frame 25	Frame 30

Output frames from the video acquisition process are stored in the database and are used at the segmentation stage.

B. Segmentation

Segmentation was used to separate object of interest from its background. In this process, the background subtraction is carried out. Prior to the subtraction process, the color space of the image frame was converted into a grayscale image, as shown in (1) [9]. The process of background subtraction was performed by subtracting each pixel of the background image with each pixel of the gait image, foreground image $F(x, y)$ is obtained by subtracting the complete image $G(x, y)$ with background image $B(x, y)$ as shown in (2).

$$Gray = 0.299 * R + 0.587 * G + 0.114 * B \quad (1)$$

$$F(x, y) = |G(x, y) - B(x, y)| \quad (2)$$

The flowchart of segmentation process is shown in Fig. 3.

The results of image segmentation process are shown in Table III.

C. Forming of Silhouettes

Silhouette formation is a very important stage in gait identification. In this research, the formation of silhouettes produced binary image after going through several enhancement processes from the previous results.

Image enhancements are needed to improve image quality to make it easier for the next process. The first image enhancement was to eliminate noise in the image using median filter with an 8x8-dimensional matrix. The second process was image morphology using dilation operations as shown in (3) and erosion as shown in (4). However, prior to

the morphological process, grayscale images were converted into binary images using the thresholding as shown in (5).

As stated in Eq. 3, the dilation operation closed the gap between two objects by adding pixels around object A to the size of the structure of element B.

$$A \oplus B = \{z | (B)_z \cap A \neq \emptyset\} \quad (3)$$

Erosion operation eroded or reduced the area of the object according to the size of the structure of element B.

$$A - B = \{z | (B)_z \subseteq A\} \quad (4)$$

Thresholding process is used to convert the image into binary image as presented in Eq. 5. The threshold (T) requirements and desired values were adjusted based on the needs.

$$g(x, y) = \begin{cases} 1 & \text{if } f(x, y) \geq T \\ 0 & \text{if } f(x, y) \leq T \end{cases} \quad (5)$$

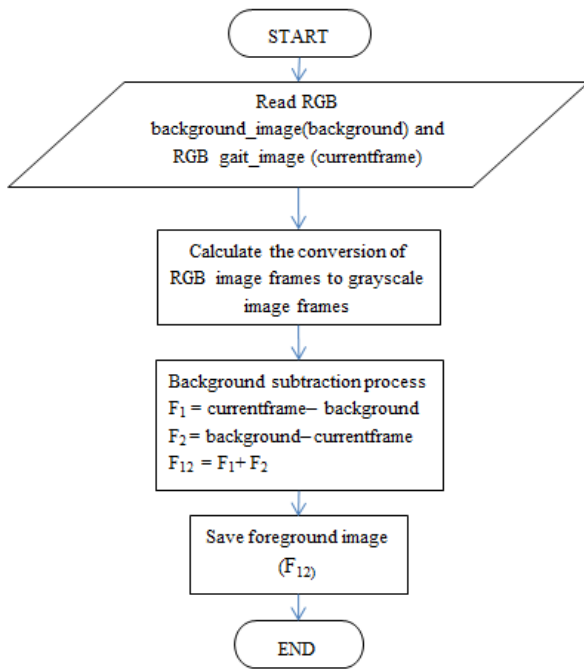


Fig. 3. Flowchart of Segmentation Process.

TABLE III. THE RESULT OF SEGMENTATION PROCESS

Frame	RGB	Grayscale	Background subtraction
Background			--

The final stage in silhouette formation is cropping or cutting to produce image frames that focus on gait objects. This cropping process requires the position of x_{min} , y_{min} , width, height. The position of x_{min} was obtained by finding the minimum value of the column, the position of y_{min} obtained by finding the minimum value of the row. The width value was obtained through reducing the maximum column value to the minimum column value, and the height value obtained by reducing the maximum row value to the minimum row value.

Table IV shows the silhouette formation process which consists of four processes. The first column is the sequence of image processing, and the second column is the result of each enhancement process to improve image quality, namely filtering, thresholding, morphology, and cropping.

TABLE IV. THE RESULT OF SILHOUETTES FORMATION PROCESS

Image Processing	Result
Grayscale imagery, there is noise that can be a nuisance and must be repaired.	
Median image filtering with an 8x8 dimension matrix to disguise the remnants of background images that are considered as noise in the subtraction background image..	
The image thresholding with a threshold value of 32 produces a binary image with the intensity of the color of the background image worth 0 (black) and the color intensity of the gait image value 1 (white).	
Morphological images with dilation and erosion operations close the gap between two objects or holes contained in the image thresholding.	
Cropping on the image frame shows that the frame is more focused on the image of a white object, namely the gait image.	

D. Feature Extraction

Feature extraction is the process of extracting features from each silhouette image to get the CoM of each silhouette image. The CoM (centroid) was generally obtained by using the average coordinate (x, y) value of each pixel composes the object [10]. The center of mass value was stored in the matrix in the form of .mat. As shown in (6), CoM of an object was obtained by calculating the number of pixel in the silhouettes.

$$N = \sum_{x=1}^n \sum_{y=1}^m B(x, y) \tag{6}$$







Eq. (7) and Eq. (8) shows the calculation the average CoM on the axis of coordinates (\bar{x}, \bar{y}) based on the number of pixels [10].














$$\bar{x} = \frac{1}{N} \sum_{x=1}^n \sum_{y=1}^m B(x, y) \tag{7}$$

























$$\bar{y} = \frac{1}{N} \sum_{x=1}^n \sum_{y=1}^m B(x, y) \tag{8}$$













The results of the silhouette feature extraction and the CoM are shown in Table V.






TABLE V. EXAMPLE OF THE CENTER OF MASS (X,Y) BASED ON FEATURE EXTRACTION PROCESS

Number	Frames	Center of Mass	
		x	y
1		89	178
2		93	175
3		97	179
4		99	179
5		98	179
6		94	181

7		85	182
8		76	177
9		63	189
10		51	185
11		37	177
12		39	186
13		41	194
14		52	196
15		57	189
16		64	186
17		71	185
18		81	185
19		86	185

20		88	186	32		54	185
21		89	184	33		64	177
22		82	186	34		70	175
23		81	180	35		79	171
24		79	182	36		86	175
25		73	191	37		91	181
26		62	208	38		95	180
27		50	185	39		93	185
28		38	172	40		89	180
29		39	171	41		88	171
30		40	174	42		82	164
31		44	184	43		77	164

44		67	179
45		60	178
46		52	170
47		41	176
48		41	197
49		46	206
50		53	213
51		58	198
52		67	201
53		73	201
54		79	197
55		83	197

56		89	193
57		90	191
58		94	195
59		89	198
60		84	197

As the person start walking, the position of CoM dynamically shifts according to the walking motion. The graph showing the CoM movement on the x-axis and y-axis at 60 frames of silhouette images is shown in Fig. 4 and Fig. 5, which the CoM data were plotted from Table V.

Fig. 4 shows a graph of changes in CoM movement on the x-axis horizontally occurring as a result of changes in position or place when walking. While in Fig. 5 shows a graph of changes in the CoM movement on the y-axis vertically with respect to the coronal plane because when walking, one foot will be on the ground and one foot in the air and ends when the same foot returns to the ground again. This process is known as the gait cycle.

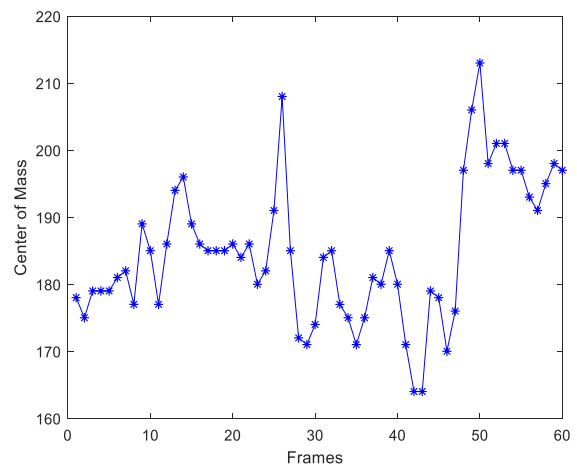


Fig. 4. Movement of the Center of Mass on the X-Axis.

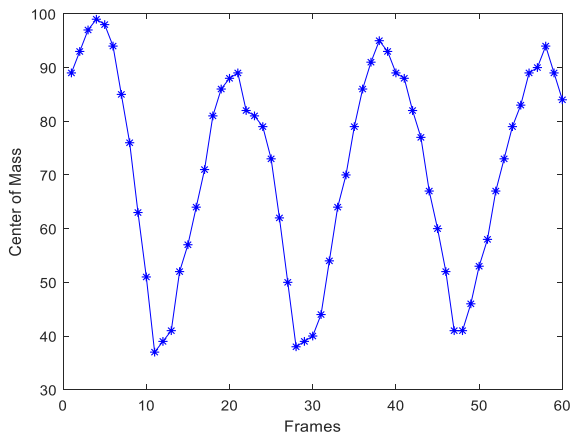


Fig. 5. Movement of the Center of Mass on the Y-Axis.

III. CONCLUSION

In this research, a frame extraction method based on silhouette and center of mass is presented. The results obtained from this research are 510 frames that were extracted from 17 pedestrian videos. Background subtraction process was successfully separate the gait images from the background. The gait silhouette images were acquired after performing a number of stages by processing grayscale image starting from the noise reduction process to cropping. Based on the silhouette image, feature extraction was performed to obtain the coordinates of the CoM (x, y) for each gait silhouette. The results have shown that the CoM in all image frames were successfully identified. For future work, the CoM can be used as a feature in conducting gait classifications.

REFERENCES

- [1] Carpentier, J., Benallegue, M., and Laumond, J. P., "On The Centre Of Mass Motion In Human Walking," International Journal Of Automation And Computing Vol. 14, Issue 5, October 2017. Pp 542-551. 2017.
- [2] Alsaadi, I., M., "Physiological Biometric Authentication Systems, Advantages, Disadvantages And Future Development: A Review," International Journal Of Scientific & Technology Research. Vol. 4, Issue 12, October 2015. Pp 285-289. 2015.
- [3] Boulgouris, N., Plataniotis, K., and Hatzinakos, D., "Gait Recogniton Using Linear Time Normalization," Pattern Recognition. Vol. 39, Issue 5, May 2006. Pp 969-979. 2006.
- [4] Sudha, L.R. and Bhavani, R., "Gait Based Gender Identification Using Statistical Pattern Classifiers," International Journal of Computer Applications. Vol. 40, Issue 8, Pp.30-35, 2012.
- [5] Borràs, R., Lapedriza, À., and Igual, L., "Depth Information In Human Gait Analysis: An Experimental Study On Gender Recognition". International Conference Image Analysis and Recognition. Vol. Part II, June 2012. Pp 98-105. 2012.
- [6] Hustinawaty., "The Prototype Of Non-Intrusive Skeleton Detection & Feature Extraction Software & Gait Man In Real Time," International Journal Of Sports Science & Engineering. Vol. 07, No. 01, September 2012, Pp. 003-022. 2012.
- [7] Sai, S. R., Ravi, R., "Multi-Scale Principal Component Analysis Based Gait Recognition," 1st IEEE/IIAE: International Conference On Intelligent Systems And Image Processing 2013 (ICISIP2013).
- [8] Ahmed, M., "Kinect-Based Human Gait Recognition Using Static and Dynamic Features," (IJCSIS) International Journal of Computer Science and Information Security". Vol. 14, No. 12, December 2016.
- [9] Gonzalez, R. and Woods, R., Digital image processing. 4th ed. Pearson, 2018.
- [10] M.Sayed, "Biometric Gait Recognition Based On Machine Learning Algorithms," Journal of Computer Science. Vol.14, Issue 7, July 2018, Pp 1064-1073, 2018.