

## Gait Recognition based on Inverse Fast Fourier Transform Gaussian and Enhancement Histogram Oriented of Gradient

S.Z.A. Rahman<sup>#</sup>, S.N.H.S. Abdullah<sup>#</sup>, K.A.Z. Ariffin<sup>#</sup>

<sup>#</sup> Digital Forensic Research Group, Center for Cyber Security, Faculty of Information Science and Technology,  
Universiti Kebangsaan Malaysia 43600 Bangi, Selangor Darul Ehsan, Malaysia.  
E-mail: sitizaharaharab@gmail.com, snhsabdullah@ukm.edu.my, k.akram@ukm.edu.my

**Abstract**— Gait recognition using the energy image representation of the average silhouette image in one complete cycle becomes a baseline in model-free approaches research. Nevertheless, gait is sensitive to any changes. Up to date in the area of feature extraction, image feature representation method based on the spatial gradient is still lacking in efficiency especially for the covariate case like carrying bag and wearing a coat. Although the use of Histogram of orientation Gradient (HOG) in pedestrian detection is the most effective method, its accuracy is still considered low after testing on covariate dataset. Thus, this research proposed a combination of frequency and spatial features based on Inverse Fast Fourier Transform and Histogram of Oriented Gradient (IFFTG-HoG) for gait recognition. It consists of three phases, namely image processing phase, feature extraction phase in the production of new image representation and the classification. The first phase comprises the image binarization process and energy image generation using average gait image in one cycle. In the second phase, the IFFTG-HoG method is used as a features gait extraction after generating energy image. Here, the IFFTG-HoG method has also been improved by using Chebyshev distance to calculate the magnitude of the gradient to increase the rate of recognition accuracy. Lastly, K-Nearest Neighbour (k=NN) classifier with K=1 is employed for individual classification in the third phase. A total of 124 people from CASIA B dataset were tested using the proposed IFTG-HoG method. It performed better in gait individual classification as the value of average accuracy for the standard dataset 96.7%, 93.1% and 99.6% compared to HoG method by 94.1%, 85.9% and 96.2% in order. With similar motivation, we tested on *Rempit* datasets to recognize motorcycle rider anomaly event, and our proposed method outperforms Dalal Method.

**Keywords**— Gait recognition; features; spatial; frequency; histogram; fusion; HOG, IFFT

### I. INTRODUCTION

Gait recognition becomes an inspiration as one of the biometric technology applications due to the nature of this study that allows the identification, detection, and recognition of humans through their gait. Gait is defined as the way a person walks [1]-[3]. Gait technology can also be used as human identification systems without requiring interaction with subjects and high-resolution cameras, which makes data collection more accessible.

In gait technology, silhouettes are used as an image and feature representations for identification. The process of combining silhouettes is known as fusion where the process adds and strengthens the information of image feature. Furthermore, by studying behavioral movements or human actions, it can provide a good way of identifying behavioral malfunctions for safety purposes [4]-[6]. Fig. 1 shows the studies found in video and image analysis such as human detection and recognition.

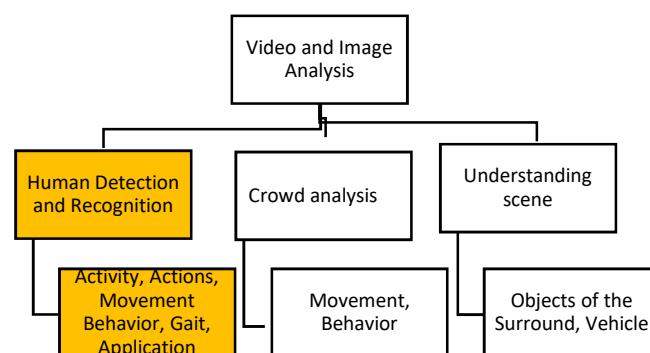


Fig. 1 Behavioral recognition movement and gait studies are under human detection and recognition.

Overall, biometric technology in human detection and recognition has begun to grow, and it is still in search of a suitable detection method. The study of gait needs to be expanded to achieve universal progress and convenience to be seen as a biometric detection and authentication tool. It is

the ability in identifying the unique characteristic between the individuals [7] and easy to deploy, as it does not require interaction with the subject has made it as one of the chosen in biometric technology. In general, the recognition can be done well based on features classification [8]. This method can differentiate the features of each object and subject. Many methods have been proposed in this study but it still a shortfall in term of effectiveness to achieve the robust precision rate for the standard dataset. Hence, the improvements or suggestions of new methods in feature representation need to be introduced to achieve a robust method and work effectively in all cases of human behavioral movement.

Feature extraction is the most essential and crucial process in people recognition based on their gait. Gait recognition is sensitive towards any changes on the individuals such as carrying bag and wearing a coat. This will result in a disorder and disruption in features and affect the accuracy rate of recognition. Most of the existing image representations only focus on the spatial gradient that can cause the rate of recognition to be less efficient [9]-[11]. Gait representation is used to represent and locate features on gait images. Even though it helps in adding useful features, most of the processes are not considering spatiotemporal information [12]. The Gait Energy Image (GEI) is proposed by [11]; it is said to be robust, but the representation only captures the side information on the silhouette image. Then continued by [13] and [14] to solve the said problem [11], but the rate of accuracy has not yet reached an acceptable rate of recognition especially for covariate cases.

The use of fusion technique on an image or algorithm varies depending on the conditions or information required in the feature extraction. The study in [15] highlights the selection of fusion as a problem dependent. Hence, improvements in the fusion methods should be applied to reduce the loss of the required information. Histogram of Oriented Gradient (HOG) is widely used in human detection especially in the pedestrian video but the accuracy rate of detection has not yet reached the acceptable rate to discriminate the features of data sets in different scenarios such as between the standard, and with additional features (i.e., carry bag and wearing a coat). However, due to the good use of the HOG method in pedestrian detection, this method is applied as an image representation for the individual gait recognition [14]. The person gaits recognition by analyzing and identifying the unique features in their way of walking is a challenging task. In this case, an improvement of the feature extraction method is needed to reduce the complexity of computations and increase the rate of recognition accuracy rates.

The paper is organized as follows: Section II discusses the state of the art, Section III explain the methodology, and Section IV shows the result and Section V for the conclusion. The technique fusion is a process of integrating multiple data and knowledge that represents the same real-world objects into consistent, accurate and useful representations. The reason behind this technique is to combine the data from two sources so that the classification can yield better results than applying only a single data [16]. The fusion process of data can be categorized as low, middle, or high, depending on the

processing stage where the combination may occur. In the low level, several raw data sources are combined to generate new raw data. Through this process, the fusion data become more informative as image fusion techniques enable the integration of different information sources. Also, the technique of fusion can also occur by combining methods for feature extraction [17].

Gait feature representation technique is first introduced to identify people from their unique way of walking. Many researchers create and produce various techniques in deploying gait cycles as an introduction for classification processes. All the techniques discussed the recognition based on the appearance approach. The study in Table 1 below illustrates several methods based on the appearance approach and requires the enhancement of feature information.

TABLE I  
IMAGE FEATURE REPRESENTATION STUDIES

Author, Year	Features / Methods	Interest in the study
Bobick dan Davis, 2001[18]	Motion-Energy Image dan Motion-History Image	Feature representation for various types of human movement for identification purposes
Chen et al., 2009 [19]	Frame Difference Energy Image	Helps in repairing incomplete silhouette from affected gait
Wang et al., 2012 [13]	Multi-channel temporal coding technique	Preserve with temporal information but lose some dynamic information such as frequency information
Arora et al., 2015 [9]	Gradient Histogram Gaussian Image	Calculate the gradient histogram at each location where it can maintain side information for human silhouette
Burhan dan Md Jan 2015[20]	Energy-Based Image and Random Transform	Solve difficulties in geometric transformation such as translation, scale, and rotation.

Gait recognition study can be divided into three approaches, namely accumulation, introduction, and fusion approach [2]. Gait information accumulation approach extracts and combines sequence frames of silhouette image using mathematical methods of average, difference, maximum and minimum operation, and others. This approach is insensitive to silhouette side errors although it provides richer information than a single binary image. The examples of this approach are Gait Energy Image (GEI), Active Energy Image (AEI) and Gait Histogram Energy Image (GHEI).

GEI is a common gait representation technique, and it is the basis representation techniques proposed [3]. GEI is a method that is still used in the appearance-based study. This technique is simple, robust and represents all the sequences of gait into a single image. This technique is obtained by generating an average gray value for silhouette binary images on a walking subject in a complete cycle. The occurrence of body parts is described in pixel intensity. The

GEI images are said to be a robust image of gait representation because random noise has been reduced, preserve the dynamic and static information.

However, there is a reduction of some dynamic information such as movement information, and it can be quickly interrupted if the subject is in covariate condition like carrying bag [2] and [14]. A study in [21] produces an image feature extracting the active area by calculating the difference between two adjacent silhouettes in a gait cycle. The dataset used is the CASIA benchmark. This representation contains more temporal features in discriminating features for cases other than usual, but it pays no attention to static information. Work in [14] introduces a more efficient and more comfortable spatiotemporal representation of the human feature by calculating the gradient histogram at each location of the original image. This study shows that the increase in recognition rate can easily be disrupted by the presence of covariate conditions such as carrying a bag and speed.

Gait information introduction approach introduces dynamic information to the static silhouette images based on GEI by implementing the mathematical methods of average, difference and motion regions extraction and more. The examples of this approach are Frame Difference Energy Image (FDEI) and Chrono Gait Image (CGI). Work in [19] proposed the FDEI image where the gait cycle is divided into several clusters to observe the affection of incomplete silhouette in the common introduction feature. On the average image of each cluster, Dominant Energy images (DEIs) was derived from the noise removal process. The frame difference is calculated by subtracting two consecutive frames. Therefore, FDEI is a representation built as a sum of corresponding DEI clusters and a positive part of the frame difference [19].

This method requires a large cluster value because the use of a small cluster leads to the loss of relevant information for identification but the use of a large cluster causes the computer to become complicated and the performance rate to decrease [2]. In solving the problem of missing information in GEI images, [13] builds multiple temporal channels to encode the sequence of gait into multiple channel images. The CGI images are a combination of contour images with multi-channel contour images. Although there is an increase in the rate of recognition, CGI may lose some dynamic information such as frequency information.

Gait information fusion approach incorporates feature layers and methods or combinations of decision layers to achieve a combination of temporary, static, and dynamic information. The example of feature representations for this approach is the Color Gait History Image (CGHI). A study in [22] creates an image consisting of three color channels, i.e., R, G, and B. The R and G channels are images that have a standing view on one and two times as the starting point

and the channel B is the energy image [11]. This new image has more dynamic, static, and temporal information that performs better than other image representations but improvements need to be made to make this representation robust on various types of covariate conditions [12].

The examples of all representation images are as below:

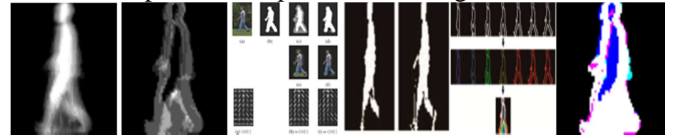


Fig. 2 (a) GEI, (b) AEI, (c) GHEI, (d) DEI and FDEI, (e) CGI, (f) CGHI

## II. MATERIAL AND METHOD

HOG method is improved by adding frequency features because this method only considers the spatial features. The improvement is done by applying the calculation of gradient's magnitude using based on the Chebyshev distance to substitute the Euclidean distance. Preliminary studies also show that by adding the image filter can enhance the feature information as well.

Figure 3 below shows the process to improve the rate of recognition accuracy. This method is used to extract the features for an average silhouette object image. To get the average image, the image should be in one sequence and a cycle for movement behavior.

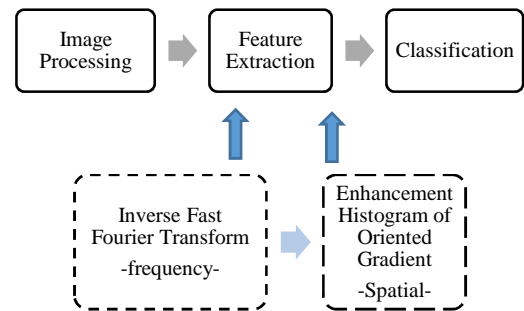


Fig. 3 The process involved in this study

The dataset image used in the gait recognition is the sequence of silhouette images in a cycle of a complete cycle. At the beginning of the framework, the image goes through the image processing to obtain an average image by using the average calculation method. After getting an average image, the next process is to figure out the common image feature by using the proposed method; IFFTG-EHoG method. The features obtained from the feature extraction process are used to classify the gait using the K-NN method. The overall framework for these three processes can be seen in Fig. 4 below:

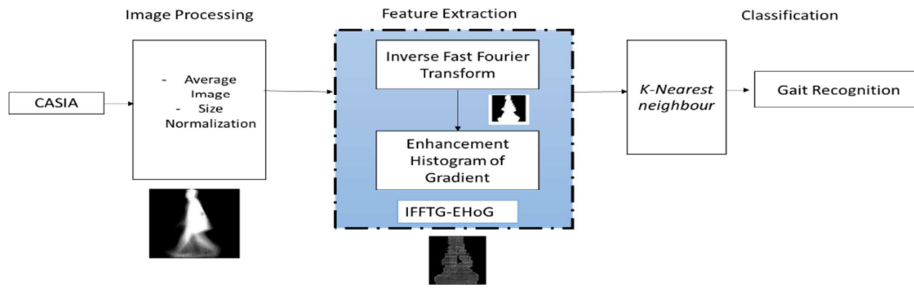


Fig. 4 Gait recognition framework

The proposed method is based on the Inverse Fast Fourier Transform and enhancement Histogram of oriented gradient, which is called IFFTG-EHoG. The EHoG is an improved Histogram of Oriented Gradient (HOG) method based on magnitude calculations using Chebyshev distance. Dalal and Triggs introduced the HOG method in 2005, and it is useful in the early phase of object detection. This method is also used in the study of model-based model deformation which is an effective object tracking method for several years around 2008 to 2016 [10] and [23]-[25]. This increasingly recognizable method looks at features from all locations in the image and calculates based on the gradient information. This method is quite simple by merely converting representations from pixel values to gradient values by searching and calculating the magnitude and direction of the gradient. Therefore, HOG method has been improved in this study, which is called EHoG.

In image processing, the data used is the standard dataset of CASIA set B that used the average image for 124 individuals. The average image is calculated using the following equation:

$$F(x, y) = \frac{1}{N} + \sum_{t=1}^N B(x, y) \quad (1)$$

$N$  is the sum of the frames in a silhouette image while  $x$  and  $y$  are the values of the 2D coordinate images. The average method of this image is often used by [9], [11], and [13]. Fig. 5 shows the CASIA set B that provides an average image dataset with three types of gaits that are normal, carrying bag and wearing coat:

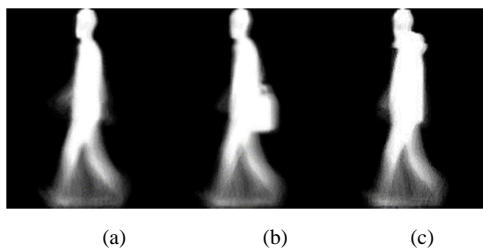


Fig. 5 Gait image (a) normal, (b) carrying bag and (c) wearing a coat.

After the average image is created, the next process is the feature extraction. The proposed feature extraction method used is a combination of spatial and frequency features called as IFFTG-EHoG. Next, in the process of classification, the chosen method is the Nearest K-Neighbor (K-NN). K-NN is one of the best and practical classification methods in

identifying and detecting objects. There are two sets of data involved namely training data as well as test data. In this study, the data set is a vector value of frequency and spatial features based on the Inverse Fast Fourier Transform and enhancement Histogram of oriented Gradient. The K-NN method makes prediction and classification of tested feature equations and trained features using the value of  $K$ , which is the closest or nearest to a point called as neighborhood value. In this study, Euclidean distance is used to calculate the distance between two points.

The Euclidean equation is as follows:

$$D(x, y) = \sqrt{(x - y)^2} \quad (2)$$

$X$  is the training features while  $y$  is tested features for the targeted predictive value in which the target is the subject. The best neighborhood value,  $K$  for K-NN algorithm depends on the data. Neighborhood values,  $K$  are used to predict target classes, and this study uses the default value of the neighborhood,  $K = 1$ . If the score of equality is least meaningful, the target of the class is getting closer.

The K-NN method is a non-parametric statistical classification technique that supposes (Ba1) no statistical distribution fits to the data of each class. Hence, this method tries to predict the class of new data points based on the nearest neighbors. The value of  $K$  is hugely training-data dependent, changing the position of a few training data may lead to a significant loss of performance. Hence, this method is unstable, particularly in class borders. However, the K-fold cross-validation should be useful to find the  $K$  value, which led to the highest classification generalizability. Applying statistical methods and applying K-NN classifier achieved high classification accuracy rates against another classifier. Alternatively, estimating the test error, the study using the cross-validation method with the folding value is  $k = 10$ .

Our proposed gait features provide better results to identify individuals based on their gait with CASIA set B sets for all three styles; standard, carrying bags and wearing a coat.

#### A. Frequency Feature

The frequency feature is based on the Inverse Fast Fourier Transform, and Gaussian blur called as IFFTG. For the IFFTG, the input feature is the average image obtained from the smoothing process and the reconstruction of the image transform where the image through the diffusion process uses a Gaussian blurs method. The output image from the diffusion process is combined with the average image that

undergoes the frequency feature transformation process using the Inverse Fast Fourier Transform method. Method of fusion is obtained by multiplying the value of the features to add frequency feature information. The equation for this frequency feature method is:

$$jSFG = \left(\frac{1}{N}\right) \left[\sum_{n=0}^{N-1} x(k)e^{\frac{2\pi i k n}{N}}\right] * \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (3)$$

Where  $x(k)$  is the frequency domain sample,  $x(n)$  the time domain sample,  $N$  is the Fourier transform size and  $k = 0, 1, 2, \dots, N-1$ . The value  $x$  is the distance from the origin in the horizontal axis, and  $y$  is the distance from the origin in the vertical axis of the image and  $\sigma$  is the deviation of the Gaussian distribution. The sigma and kernel values affect the blurry rate in the image where sigma usage in this study is 2, and the size of the 3 kernels is the default value. Sigma controls the kernel function where large sigma sizes require large kernel matrix sizes in the process of the energy function. After the image goes through the frequency feature process, the next process is to use the EHoG method. The IFFTG image results are as below. The example is of an image carrying a bag:

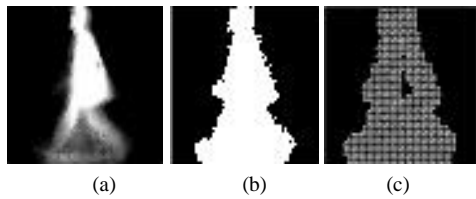


Fig. 6 Gait image (a) average image (b) frequency features image (IFFTG) and (c) frequency and spatial feature combination (IFFTG-EHoG)

### B. Spatial Feature

Enhancement Histogram of Oriented Gradient (EHoG) method is the global encoding method of an image. This global approach gives more discriminative features in recognition rate [27]. Gait images are divided into a cell,  $6 \times 6$  pixels. The gradient is calculated per pixel vertically and horizontally. Each pixel has the orientation and magnitude of the edges of the image in the cell. Calculation of the magnitude of the gradient, which is in the form of a vote, uses Chebyshev distance while the orientation of the discrete bin is correlated from  $0^0$  to  $180^0$ . This method will ultimately produce a one-dimensional representation of a histogram called orientation histogram as shown in the picture below:

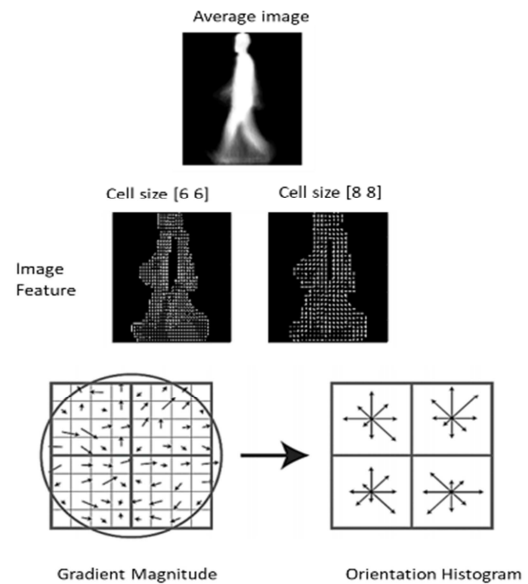


Fig. 7 Image gradient and orientation for feature extraction method

For this feature, the direction of the gradient is used as a feature. The gradient is from  $x$  and  $y$  derivatives using the kernel  $-1 \ 0 \ 1$  vertically and horizontally in an image. The magnitude of the gradient becomes large around the edges and corners. The area is a sudden change of intensity and gives more information about the shape of the image from the flat area. The experiment also looks at the parameters found in this method, i.e., cell size, block size, and bin number. The second experiment is to see the magnitude calculation of the gradient in the method that previously used the distance of the Euclidean. Improvements have been made by using Chebyshev distance for calculation of magnitude and changing the cells size to get better features. A number of the bin is also studied, and it shows the best is nine bin. Bin is the savings for pixel representation from voting calculations of magnitude and direction.

Below shows the calculation of magnitude and direction of the gradient using the following equation:

$$\text{magnitude of gradient, } g = \max(|x_1 - x_2|, |y_1 - y_2|) \quad (4)$$

$$\text{Direction, } \theta = \arctan \frac{g_x}{g_y} \quad (5)$$

The values of  $x$  and  $y$  are the gradient of the image. The image gradient at any location has a change in intensity, and because the accuracy is in vector quantity, the gradient is a function of magnitude and direction. The calculation of the gradient magnitude uses Chebyshev distance as shown in figure 8 below. Within this distance, all eight adjacent cells can be reached using one unit.

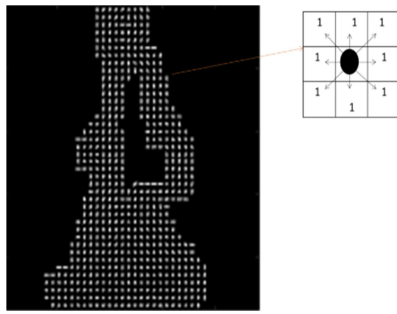


Fig. 8 Calculation of the magnitude of the image gradient using Chebyshev distance

For the vector, size calculation is to multiply the bin number with the cell number in the block. Vector normalization occurs to restriction the features from affected by the lighting variation.

### C. Dataset

The dataset used is CASIA set B. CASIA set B has a large multiview walkable database. Dataset B was taken in January 2005. There were 124 subjects captured from 11 views. Data contains normal gait (NM), carrying a bag (BG) and wearing a coat (CL). For this study, the study only uses data sets at 90-degree views.

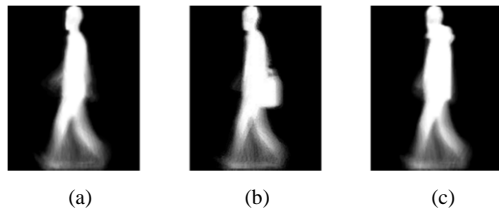


Fig. 9 Dataset CASIA B for (a) NM, (b) BG and (c) CL

## III. RESULT AND DISCUSSION

This section discusses the results of the proposed method by combining the frequency and spatial features to complement essential features. The study is divided into two experiments; the first experiment is tested proposed method with Histogram of oriented gradient (HOG) method. The second experiment is tested together with the image of the other gait representation techniques. Proposed method using cross-validation method with fold value  $k = 10$  in classification while for other comparative method, having 10 sequences of gait (6 normal labeled as NM, 2 carrying a bag (BG) and 2 wearing a coat (CL), the study uses one sequence as a training and 9 other sequences as test [2]. The table below shows the comparative results of proposed methods, with other gait feature representation methods. The table below is read in diagonal for a set of training and test set for the same set of datasets.

The study proposed an enhanced HOG method based on Chebyshev distance in the calculation of the magnitude of the image gradient. Average results for this comparison use the Nearest K-Neighbor (K-NN) classifier. The proposed method is tested on each data set for normal gait data labeled as NM, carrying the bag (BG) and wearing the coat (CL) to see the effectiveness of the method in each data set. The table below shows the preparation of a review:

TABLE II  
EXPERIMENTAL PREPARATION

Dataset	CASIA set B		
Test label	NM (Normal)	BG (Carrying bag)	CL (Wearing Cloth)
Probe size	Has 124 subjects and each subject has 6 average images of normal gait	Has 124 subjects and each subject has 2 average images of carrying bags	Has 124 subjects and each subject has 2 average images of wearing a coat
Total	744	248	248

Dataset CASIA set B has 124 individuals. This video was obtained from a USB camera (Model: Fametech 318SC) for indoor environment [26]. The silhouette quality of this dataset is high. Each has three types of gait movements which are six average images walking normal, two average images walking with a bag and two average images walking on a coat labeled in the database as nm-01 to nm-06, bg-01, bg-02, cl-01, cl-02 and each is captured on eleven different of views. In this study, it only uses 90-degree views for analyzing existing research and comparison methods. This dataset provides average images after background subtraction and through image processing and normalization. In image processing, the horizontal alignment focuses on the top of the silhouette as the centroid of the image as shown in Fig. 10 (a) below [2] and all silhouette images in a cycle such as (b) are normalized with both horizontal and vertical directions set to the same size [26].

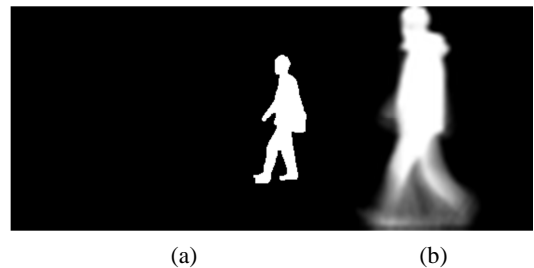


Fig. 10 (a) Image before average image process (b) and Average image

For the proposed IFFTG-EHoG method, these are the parameters used in image processing, features extraction and classification process.

Image processing	<ul style="list-style-type: none"> <li>• Comparison with HOG: size image is 240x240</li> <li>• Comparison with other methods: size image is 64x64 (follow the paper)</li> </ul>
Features extraction	<ul style="list-style-type: none"> <li>• Gaussian blur filter, kernel =3;sigma=2</li> <li>• Calculation gradient horinzotal and vertical use kernel -1 0 1</li> </ul>
Classification	<ul style="list-style-type: none"> <li>• Value of neighbor in K-NN classifier, K=1</li> <li>• Value of folded in cross-validation , k=10</li> </ul>

Fig. 11 Experiment preparation

The table below shows the overall accuracy rate for the performance of IFFTG-EHoG proposed method compared to the HOG method. The results are as follows:

TABLE III  
HOG AND IFFTG-EHoG METHOD PERFORMANCE RESULTS (%)

Methods/ Type	NM	BG	CL	Average (%)	Standard Deviation
HOG (Dalal Method)	94.10	85.90	96.2	92.1	5.4
EHoG	96.51	91.90	98.40	95.6	3.3
IFFTG-EHoG	96.5	92.6	99.1	96.1	3.3
G-EHoG	96.5	92.5	99.1	96.0	3.3
Proposed IFFTG-EHoG	<b>96.70</b>	<b>93.10</b>	<b>99.60</b>	<b>96.5</b>	<b>3.2</b>

Remark IFFT- Inverse Fast Fourier Transform, G- Gaussian blur technique.

The table above shows an increase in all the improvements in the HOG method called IFFTG-EHoG with an average value of 96.5, which is an increase of 4.4%, and the deviation value decreased by 2.2%. Analysis using the ANOVA test shows the value of F as much as 17.21. Since the statistical test is larger than the critical value, the null hypothesis is rejected. This shows there is a difference between the above methods. The study continued with the test of the classification error for each set of data that was conveyed through the proposed method. For BG data sets, the error rate is 0.30, and CL data is 0.33 while the error value for the NM data set is 0.04. Classification error tests indicate there a classification error of about 3% of training observations for BG data sets while for CL data sets, the error from the observation of classification training is about 3.3%. For NM data, there was 0.4% of the classification error from training observation.

Overall, the gait recognition accuracy of wearing a coat is higher than normal and carrying a bag. This is because when someone wears a coat adds burden to the movement and causes the movement to be limited and not too different to the target. For the carrying, a beg recognition has not reached 95% on this because the bag features itself undermines the features of the silhouettes as well as the movement. The study then continued tested the proposed method, IFFTG-EHoG with several methods that have advantages regarding dynamic, static, and temporal feature representation.

TABLE IV  
RESULT OF COMPARISON OF PROPOSED METHOD, IFFTG-EHoG WITH RECENT METHOD

Testing	Training			
	Methods/Dataset	NM	BG	CL
Normal (NM)	GEI [11]	<b>90</b>	44	34
	FDEI [19]	<b>94</b>	68	42
	CGI [13]	<b>87</b>	63	53
	CGHI [22]	<b>95</b>	58	60
	Proposed, IFFTG-EHoG	<b>96</b>	58	57
Carrying bag (BG)	GEI [11]	31	<b>90</b>	17
	FDEI [19]	39	<b>98</b>	44
	CGI [13]	45	<b>95</b>	40
	CGHI [22]	69	<b>98</b>	64
	Proposed, IFFTG-EHoG	28	<b>93</b>	26
Wearing	GEI [11]	27	20	<b>96</b>

coat (CL)	FDEI [19]	22	56	<b>98</b>
	CGI [13]	43	41	<b>95</b>
	CGHI [22]	48	55	<b>98</b>
	Proposed, IFFTG-EHoG	27	20	<b>99</b>

The rows and columns for the above table show the difference in data sets for training and testing. There are six experiment set environments. For example, experiments with normal gait training datasets are tested with normal gait datasets, carrying a bag and wearing a coat. This study looks at the strengths in the same set of data or different sets of data for training and testing. For a comparative study, the GEI image representation is a method of information accumulation approach, for the FDEI and CGI image representation are introduction information approach and CGHI is categorized as an image representation of fusion information approach. For the method proposed, IFFTG-EHoG can be categorized infusion information approach, which combines two features of information from different methods. Overall in the environment of training and test set is the same, the study shows that CGHI method dominates the high accuracy rate of 95% in NM and 98% in BG and CL followed by FDEI method that is 94% for NM and 98% and CL. For the proposed method, IFFTG-EHoG is the third best method of 96% in NM, 93% in BG and CL is 98%. The CGI method and GEI methods in the fourth and fifth places, for the CGI method the accuracy rate for the NM is 87%, the BG and CL is 95%, and the GEI method with a 90% accuracy rate in NM and BG whereas for the CL is 96%. This difference in accuracy depends on the type of feature available. In the case of different environments, CGHI demonstrates robust which highlights the environment especially when the training set is NM and CL, and the test set is BG with an average of 67% accuracy. This is because of the rich information that is dynamic, static, and temporal information that creates an image consisting of three-color channels. For GEI methods, this method has little dynamic information and much static information, for the FDEI method, this method has no temporal information yet rich with dynamic and static information.

For the proposed method, IFFTG-EHoG still preserves all three types of information namely dynamic, static, and temporal information but there is a lack of minimal dynamic information and for the CGI method although preserving temporal information and generating local features but still have not achieved high precision rates for each case of gait. Besides, a preliminary study of *Rempit* recognition has also been conducted. We tested on a set of *Rempit* video collection consisting of normal and abnormal riders from YouTube in the total number of 11 riders as stated in Table V.

TABLE V  
EXPERIMENTAL PREPARATION FOR *REMPIT* RECOGNITION

Dataset	Youtube	
Test label	MNM (Normal)	MR (Abnormal)
Probe size	5 video: 3 video has 2 average images rider 2 video has 3 average images rider	5 video: 3 video has 2 average images rider 2 video has 3 average images rider
Total	12	14

The figure below shows the images for the *Rempit* recognition.

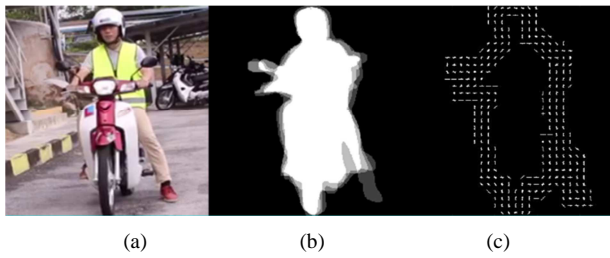


Fig. 12 Normal rider (a) extracted image from video (b) average image (c) IFFTG-EHoG image

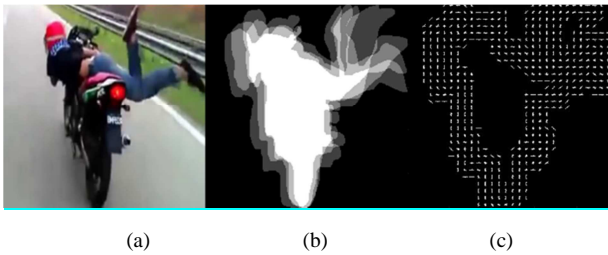


Fig. 13 Abnormal rider (a) extracted image from video (b) average image (c) IFFTG-EHoG image

We execute in similar processes average and with IFFTG-EHoG image as a benchmark as mentioned in earlier datasets. We also conduct cross-validation with 10 K-fold onto self-collection datasets. With similar motivation, we found that our proposed method was able to identify, recognize, and differentiate the human-style of riding a motorcycle in normal and abnormal known. Our finding on *Rempit* activity (video) able to raise the accuracy rate up to 75.11% in comparison to Dalal Method approximately 73.06% as dictated below:

TABLE VI  
IFFTG-EHOG METHOD PERFORMANCE RESULT FOR *REMPIT* RECOGNITION

Method	Accuracy (%)
Dalal Method	73.06
Proposed IFFTG-EHoG	<b>75.11</b>

#### IV. CONCLUSIONS

Gait recognition has begun to develop as biometric technology due to the nature that gait being said to be unique and does not require interaction with the subject. This technology is intended to make an identification, or from the perspective of criminal cases are identifying suspects according to distinctive gait. Overall, the study was conducted to improve the existing methods using frequency and spatial methods through the gradient of their image. Because of the sensitivity of changes such as carrying things and speed in image appearance, make this study continue from time to time. The experiments are the use of this proposed method in features extraction and classifications' using the KNN method shows that the accuracy rates of recognition for the dataset CASIA set B higher than the comparing method. The advantage of this method is a high computational process, and huge dataset is not required. In

addition, this method can be applied to the other human behavior movement with some modification of features extraction parameter and preprocessing process.

#### ACKNOWLEDGMENT

The author would like to present his gratitude towards the Faculty of Information Science and Technology, the National University of Malaysia for the help and support that had been accorded during the research period. This work was funded by research grant AP-2017-005/2 entitled "Using Stem Data through Smart Self-Crime Prevention at Schools for Open Data Readiness."

#### REFERENCES

- [1] Afsar, P., Cortez, P. & Santos, H., "Automatic visual detection of human behavior: A review from 2000 to 2014", *Expert Systems with Applications* 42(20): 6935–6956, 2015.
- [2] Lv, Z., Xing, X., Wang, K. & Guan, D., "Class energy image analysis for video sensor-based gait recognition: A review," *Sensors (Switzerland)* 15(1): 932–964, 2015.
- [3] Lee, C. P., Wee, A., Tan, C. & Lim, K. M., "Review on Vision-Based Gait Recognition: Representations, Classification Schemes, and Datasets," *American Journal of Applied Science* 14(2): 252–266. doi:10.3844/ajassp.2017.252.266, 2017.
- [4] A. Saadoon and M. J. Nordin, "An automatic human gait recognition system based on joint angle estimation on silhouette images," *Journal of Theoretical and Applied Information Technology*, vol. 81, p. 277, 2015
- [5] R.A.K. Noaman, M. A. M. Ali, N. Zainal, F. Saeed, "Human Detection Framework for Automated Surveillance Systems." *International Journal of Electrical and Computer Engineering (IJECE)*, 2016; 6(2): 877–886.
- [6] Rabiah Adawiyah Shahad, Mohamad Hanif Md Saad and Aini Hussain, "Activity Recognition for Smart Building Application Using Complex Event Processing Approach," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 8, no. 2, pp. 315-322, 2018
- [7] Basel Alshaiikhdeeb and Kamsuriah Ahmad, "Biomedical Named Entity Recognition: A Review," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 6, no. 6, pp. 889-895, 2016.
- [8] S. N. H. S. Abdullah, M. Khalid, and K. Omar, "Performance Comparison of License Plate Recognition System Using Multi-Features and SVM," *Asia-Pacific Journal of Information Technology and Multimedia*, pp. 53-62, 2011.
- [9] Arora, P., Srivastava, S., Arora, K. and Bareja, S., "Improved Gait Recognition Using Gradient Histogram Gaussian Image," *Procedia Computer Science* 58(408-413), 2015.
- [10] Dalal, N. & Triggs, B., "Histograms of oriented gradients for human detection," *Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005 I*: 886–893, 2005.
- [11] Han, J., Bhanu, B., "Individual recognition using gait energy image," *IEEE Transition Pattern Analysis Machine Intelligent*, 28, 316-322), 2006.
- [12] Chen, J. & Liu, J., "Average gait differential image based human recognition," *Scientific World Journal*, pp. 1-8, 2014.
- [13] Wang, C., Zhang, J., Pu, J., Yuan, X. and Wang, L., "Chrono-gait image: a novel temporal template for gait recognition," *Computer Vision–ECCV*, page: 257-270. Springer, 2010.
- [14] Hofmann, M. & Rigoll, G., "Improved Gait Recognition using Gradient Histogram Energy Image," *Proceedings - International Conference on Image Processing, ICIP 1389–1392*, 2012.
- [15] D. K. Sahu and M. P. Parsai, "Different Image Fusion Technique- A Critical Review," *International Journal of Modern Engineering Research (IJMER)*, Vol. 2, No.5, pp. 4298-4301, 2012.
- [16] Luo, J., Zhang, J., Zi, C., Niu, Y., Tian, H. & Xiu, C., "Gait Recognition Using GEI and AFDEI," *International Journal of Optics*, 2015.
- [17] Eleyan, A. & Demirel, H., "Co-occurrence matrix, and its statistical features as a new approach for face recognition," *Turk J Elec Eng & Comp Sci* 19(1): 97–107, 2011.



- [18] A. F. Bobick J. W. Davis, "The recognition of human movement using temporal templates," *Pattern Analysis and Machine Intelligence IEEE Transactions*, vol. 23 no. 3 pp. 257-267, 2001.
- [19] Chen C., Liang J., Zhao H., Hu H., Tian J., "Frame difference energy image for gait recognition with incomplete silhouettes," *Pattern Recognition Letter*,30:977-984, 2009.
- [20] I. M. Burhan M. J. Nordin, "Multi-View Gait Recognition Using Enhanced Gait Energy Image and Radon Transform Techniques," *Asian Journal of Applied Sciences*,8 no. 2 pp. 138-148, 2015.
- [21] Zhang, E., Zhao, Y., & Xiong, W., "Active energy image plus 2DLPP for gait recognition", *Signal Processing*, 90, 2295-2302., 2010.
- [22] Chen S., Ma T., Huang W., Gao Y., "A multi-layer windows method of moments for gait recognition," *J. Electron. Inf. Technol.*, 31:116-119, 2009.
- [23] Antony, J. J. & Suchetha, M., "Vision-Based Vehicle Detection: A Literature Review," *International Journal of Applied Engineering Research ISSN* 11(5): 973-4562, Retrieved from <http://www.ripublication.com>, 2016.
- [24] Cai, Z., Yu, P., Liang, Y., Lin, B. & Huang, H., "SVM-KNN Algorithm for Image Classification Based on Enhanced HOG Feature," *Proceedings of the 4th IIAE International Conference on Intelligent Systems and Image Processing 2016* (3): 111-117. doi:10.12792/icisip2016.023, 2016.
- [25] Rahman, S. Z. A., Abdullah, S. N. H. S. and Nazri, M. Z. B. A. 2016, "The analysis for Gait Energy Image based on statistical methods," 2016 *International Conference on Advances in Electrical, Electronic and Systems Engineering (ICAEEES)*, pp. 125-128, 2016.
- [26] Shuai Zheng, Junge Zhang, Kaiqi Huang, Ran He and Tieniu Tan., "Robust View Transformation Model for Gait Recognition," *Proceedings of the IEEE International Conference on Image Processing*, pp. 2073-2076,2011.
- [27] M. Naeimizaghiani, S. N. H. S. Abdullah, F. Pirahansiah, and B. Bataineh, "Character and object recognition based on global feature extraction," *Journal of Theoretical and Applied Information Technology*, vol. 54, pp. 109-120, 2013.