Polygonal Shape-based Features for Pose Recognition using Kernel-SVM

M.F. Abu Hassan[1,3], A. Hussain[1] and M.H.M. Saad[2]
[1]*Department of Electrical, Electronic & Systems Engineering, Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia, Selangor, Malaysia*
[2]*Department of Mechanical & Material Engineering, Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia, Selangor, Malaysia.*
[3]*Universiti Kuala Lumpur, Malaysia France Institute, Bandar Baru Bangi, Selangor, Malaysia.*
fadzil@siswa.ukm.edu.my

***Abstract*— Pose recognition is an intriguing and challenging problem particularly in surveillance, inspection, etc. that lies in computer vision. This paper presents an efficient human walking and abnormal poses recognition system based on kernel-support vector machine (KSVM) using a novel feature set based on polygonal shape generalization on the human silhouette. The Shapiro-Wilk test was conducted to assess the data distribution and it summarized that the test rejected the hypothesis of normality for all features. Therefore, an inferential Mann-Whitney *U* test was performed to evaluate the proposed feature set statistically and results showed that all features were significantly different between the groups of poses ($p < 0.001$). Three kernel models: linear, polynomial and radial based function were adopted for SVM to classify the walking and abnormal poses. Results obtained showed that all three kernels of the KSVM classifiers performed well with accuracies of more than 95%. However, further experiments proved that the polynomial KSVM yields the best accuracy of 99.96%. Thus, it can be concluded that the proposed polygonal shape-based feature set is best paired with the polynomial KSVM for abnormal pose detection task.**

***Index Terms*— Feature Extraction; Kernel-Support Vector Machine; Polygonal Shape; Pose Recognition.**

I. INTRODUCTION

In computer vision, research in pose recognition contributes to a major improvement in realizing of smart environments that able to provide assistance and analysis of subject's activities. In recent years, various recognition systems in many applications particularly in surveillance and monitoring have been proposed in which higher recognition accuracy is always essential. Those computer-aided systems are crucial to being adapted in working spaces or premises to detect abnormal activity such as slip and fall action [1][2]. This kind of abnormality action is rarely occurring; however, it may cause a serious health and safety implications on the subject.

Although three-dimensional (3D) shape analysis is considered as an effective and efficient technique in machine vision, however, the major challenge lies in designing effective high-level features and minimizing the computational time [3]. Many two-dimensional (2D) shape-based feature extraction (FE) techniques, methods, and algorithms have been developed in the past such as chain code, centroidal distance, shock graphs, etc. An algorithm for human recognition system using centroidal profile features was developed by [4] and a good classification rate was obtained using Artificial Neural Network Classifier. Whereas in [5], they performed human posture classification tasks using the Simplified Shock Graph (SSG) technique.

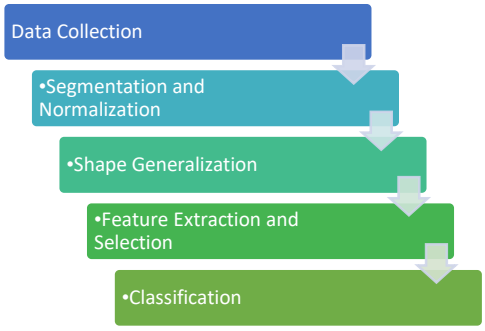


Figure 1: The framework of human pose recognition system

In this paper, a novel polygonal shape-based feature extraction technique is proposed and the framework of the human pose recognition system is as shown in Figure 1. The main steps include data collection, segmentation, normalization, shape generalization, feature extraction, feature selection and classification.

The rest of this paper is organized as follows: Section II illustrates the proposed methodology in detail. Results and discussion are presented in Section III and finally, Section IV presents the conclusion of this paper and future research.

II. METHODOLOGY

In this research, we analysed and classified two human poses: walking and other abnormal poses i.e. bending, sitting, laying and squatting poses. The human poses datasets were acquired from two different databases. The human walking dataset was obtained from the CASIA Gait database and the abnormal dataset acquired from Laboratoire d'Electronique, Informatique et Image (LE2I) database.

A. Data Collection

The CASIA Gait Database was provided by The Institute of Automation, Chinese Academy of Sciences (CASIA) [6]. It consists of three datasets: Dataset A, Dataset B (multi-view dataset) and Dataset C (infrared dataset). However, only Dataset B was considered to be used for the purpose of this study. The selected dataset is a large multi-view gait database which was acquired from 124 subjects and the gait data was captured by several video cameras from 11 views as shown in Table 1(a)-2nd and 3rd columns.

The abnormal database consists of human sitting, bending, squatting and laying poses acquired from LE2I, The National Center for Scientific Research (CNRS) [7]. The video sets contain various illumination setting and typical difficulties like occlusions and textured background. The subjects performed various normal daily activities (walking, reading and standing) and the abnormal action; falling which consists of bending, squatting and lying poses as shown in Table 1(a)-3rd and 4th columns.

B. Segmentation and Normalization

Moving object segmentation in video frames is the most significant step in many computer vision applications including human activity/action analysis. Segmentation is required to isolate the object of interest i.e. foreground from the background image. There are various and robust segmentation algorithms [8]-[10], but in this work, the common background subtraction technique was chosen to detect and segment the moving object i.e. human from a background in the video due to its simplicity, fast processing time, and reliability of output.

Using the arithmetic calculations shown in Equation (1), the object is segmented by subtracting the corresponding gray-scaled pixel values at the same position of the current image frame, denoted by $P[I(t)]$ with the background image frame denoted as $P[B]$. In this particular work, all background pixels are considered static and object movement in the scene is at low speed level. The sample object detection result using this method as illustrated in Table 1(c).

$$P[F(t)] = P[I(t)] - P[B] \quad (1)$$

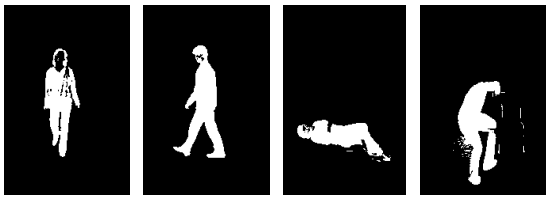
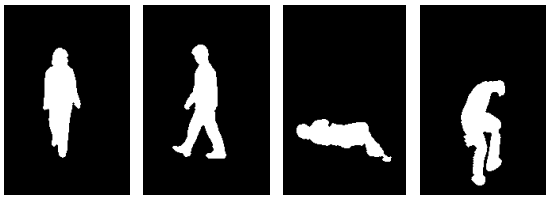
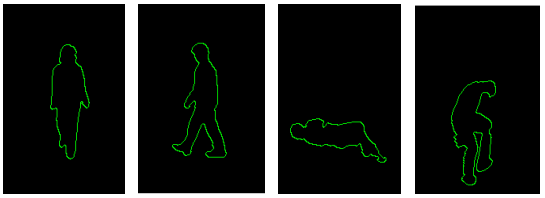
Typically, the output of background subtraction algorithm contains noise. Therefore, appropriate post-processing was conducted on the foreground before using them for further processing. To diminish the effect of noise on foreground and shape enhancement, the morphological process such as erosion and dilation were performed. The clean shapes of human silhouette samples for two poses are as shown in Table 1(e).

All raw data in both databases were captured using different stationary video cameras. For the first database, the subjects were instructed to walk straight from a starting to end position and the scenes were captured by multiple cameras in different angle views. Whilst for abnormal scenes, the subjects were ordered to act fall actions in a various manner and the scenes were recorded by a single camera. Thus, both sets of silhouettes may vary in size due to the perspective of the scene and the physical size of subjects. Therefore, all datasets have to be normalized by rescaling the silhouette size to a uniform dimension without changing the original shape. The horizontal and vertical dimension of silhouette were measured from the minimum to maximum x and y coordinates; respectively. Then, the vertical dimension was rescaled to 100 pixels, whilst the horizontal dimension was proportionally rescaled to the vertical dimension.

C. Shape Generalization

In general, 2D geometric shapes are constructed by a set of vertices, thus a closed chain lines; connecting the vertices forming the interior space. In other words, the lines connecting the points can be referred to as the sides of the shape and the minimum number of vertices, n_v required to form a 2D shape is 3; $n_v \geq 3$. Polygon categories, for instance, are classified according to their number of edges such as triangles, quadrilaterals, pentagons, etc. Each of these major categories can be divided into several different shapes. For

Table 1
Object Detection and Post-processing of Walking and Abnormal Dataset

	Sample 1: Walking [6]	Sample 2: Walking [6]	Sample 3: Abnormal [7]	Sample 4: Abnormal [7]
(a) Raw data				
(b) Pre-processing				
(c) Object detection				
(d) Binarization				
(e) Morphological				
(f) Contour tracing				

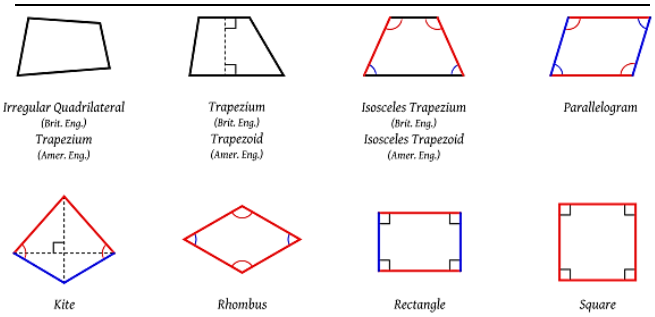


Figure 2: Quadrilaterals families with properties [15]

instance in Figure 2, a quadrilateral or quadrangle can be formed as rectangles, rhombus, trapezoid or squares.

Obviously, a 2D human silhouette shape is a non-polygon shape where the silhouette outline is composed of infinite sets of straight line segments and corners as shown in Table 1(e). In addition, these properties differ in every human shape even though they are categorized under the same pose. Therefore, a generalization of shape should be imposed on the silhouette outline.

An algorithm of contour tracing was performed to extract the closed chain outline of silhouettes from the foreground regions [11] and the samples of contour tracing are shown in Table 4(f). The contour length was then equally divided into n_v parts; thus, the location of appropriation points will represent the vertices of polygon shape, $V_i = (x_i, y_i)$. The initial vertex, V_1 was defined at the topmost left of the contour and the sequence order of the next vertices, V_{i+1} is in clock-wise rotation. Then, each pair of vertices points were virtually connected in sequence to form a 2D n_v-gon shape. Table 2 shows the polygonal shape generalization of the silhouette to various numbers of vertices, i.e. triangle (3 vertices), quadrilateral (4 vertices), pentagon (5 vertices), hendecagon (11 vertices) and pentacontagon (50 vertices) shapes.

D. Feature Extraction

Each 2D n_v-gons shape has different properties such as the distance, C_i in between the center of mass, C_m and the vertex, V_i. Another potential feature that able to discriminate the shape is the distance in between vertices, V_i to V_{i+1} a.k.a. side length, S_i and the angle in between each vertex, A_i. All of these features have unique properties to differentiate the polygonal shapes, thus able to contribute a promising classification result for object recognition. Thus, three main feature groups, G were extracted to differentiate the human silhouette shapes as shown in Figure 3.

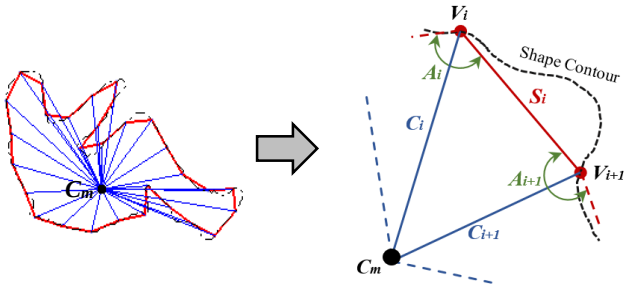


Figure 3: S_i, C_i and A_i features extracted from Polygonal shape

where: C_i = Distance in between center of mass and vertex
S_i = Length of side
A_i = Vertex angle

S_i and C_i features are defined as the distance in pixels and angle in degree for A_i. The total number of features, n_f is proportionally depending on the chosen n_v as shown in Equation (2).

$$n_f = n_v \text{ X } G \tag{2}$$

C_m is an object physical and geometric property in the image, where it is measured based on the assumption that the pixels have equal masses [12]. Thus, it is favorable to choose C_m as a reference point where it will consistently be computed based on the region of silhouette, O_s. The center of mass, $C_m = (x_c, y_c)$ is computed as follows:

$$x_c = \frac{\sum_j^k x_j}{k} \quad , \quad y_c = \frac{\sum_j^k y_j}{k} \tag{3}$$

Where, k is the number of pixels in silhouette region, O_s. Each Euclidean distance from the C_m to V_i are computed as in Equation (4), thus:

$$C_i = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2} \tag{4}$$

The second feature group, S_i is simply calculated as below:

$$S_i = V_i - V_{i+1} \qquad \forall\ i \in \{1, 2, \dots, n_v - 2\} \tag{5}$$

It is defined as the 2D Euclidean distance between points V_i and V_{i+1} or the length of the virtual line connecting them, $(\overline{V_i V_{i+1}})$. If $V_i = (x_i, y_i)$ and $V_{i+1} = (x_{i+1}, y_{i+1})$ in cartesian coordinates are the two vertices in Euclidean two-space, then the distance, S_i from V_i to V_{i+1}, or vice versa, from V_{i+1} to V_i is given by the Pythagorean formula as in Equation (6).

$$S_i = \sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2} \tag{6}$$

The third feature group, A_i is the interval angle of V_i, measured in between each two sides of n_v-gons. This feature is calculated from the law of cosines in Equation (7) where one knows the two sides' distances: S_i and S_{i+1} as previously computed in Equation (4).

$$A_i = \arccos\left(\frac{{S_i}^2 + {S_{i+1}}^2 - (V_i - V_{i+2})^2}{2 S_i S_{i+1}}\right) \tag{7}$$

E. Feature Selection

This process is to identify and remove the irrelevant and redundant variables from the dataset that may decrease the accuracy of the predictive model [13]. Thus, it will help by the same time, requiring less data. A small number of variables is essential because it may reduce the complexity of the model and overfitting. Therefore, it will improve the classifier performance; besides, provide faster and cost-effective predictors.

In our study, we are going to analyse whether there is statistical evidence that the associated polygon shape-based features are significantly different by comparing the means of the two independent pose groups, walking and abnormal.

The normality test was conducted as the prerequisites to ensure the data is suitable to be analysed using a particular statistical examination. The Shapiro-Wilk (SW) test was conducted to validate the normality assumption [14]. All interpretations of the statistical results are set at significance level, $\alpha = 0.05$.

The Mann-Whitney U (MWU) test was used to compare the data distribution between the two unrelated groups on the same continuous-level. MWU is a non-parametric alternative test to the parametric test; independent sample t-test. Unlike the t-test, it does not assume any assumptions related to the

distribution of data. The MWU is significantly more efficient than the *t*-test by considering a larger sample size and distributions sufficiently far from normal [14]. The test is used to compare the medians of two samples as shown in Equation (8) that come from the same population whether the equality exists between them; the null hypothesis (H_0).

$$U = N_1 N_2 + \frac{N_2(N_2+1)}{2} - \sum_{i=N_1+1}^{N_2} R_i \quad (8)$$

where: U = Mann-Whitney U test
N_1 = 1st sample size
N_2 = 2nd sample size
R_i = Sample size rank

F. Classification

Generally, the Support Vector Machine (SVM) family was used as the classifier as it promises a high accuracy result and good processing time. In addition, the sequential minimal optimization (SMO) was selected as the solver to train the SVM. In principle, this type of classifier is a linear machine that creates a hyperplane as a level of decision-making, thereby enabling it to separate between the positive and the negative samples. Recently, multiple improvements on the traditional SVM had been achieved; among which the kernel SVM (KSVM) is the most popular and effective. This extended SVM allows us to fit the maximum-margin hyperplane in a transformed feature space. The kernel mappings, such as linear (Lin-KSVM), polynomial (Pol-KSVM) and radial basis function (RBF-KSVM) are able to classify both linear and non-linear data. These kernels can be attained by the following models:

$$\text{Lin-KSVM: } K(x_m, x_n) = x_m^T x_n \quad (9)$$

$$\text{Pol-KSVM: } K(x_m, x_n) = (x_m^T x_n + c)^d \quad (10)$$

$$\text{RBF-KSVM: } K(x_m, x_n) = exp\left(-\frac{\|x_m - x_n\|}{2\sigma^2}\right) \quad (11)$$

where: K = Kernel function
σ = Scaling factor
x_m, x_n = Vectors in the input space
d = Degree of polynomial
c = Soft margin constant

Based on these three kernels, we trained and validated the KSVM classifiers with the proposed feature set and seek the best KSVM classifier that can provide the highest accuracy, *Acc* of classification. The *Acc* can be derived using Equation (12).

$$Accuracy: Acc = \frac{TP+TN}{TP+TN+FP+FN} \text{ x } 100\% \quad (12)$$

$$Precision: Pr = \frac{TP}{TP+FP} \text{ x } 100\% \quad (13)$$

$$Specificity: Sp = \frac{TN}{TN+FP} \text{ x } 100\% \quad (14)$$

where:
True positives (*TP*) = number of walking pose correctly detected;
False positives (*FP*) = number of walking pose detected as abnormal;
True negatives (*TN*) = number of abnormal pose correctly detected;
False negatives (*FN*) = number of abnormal pose detected as walking.

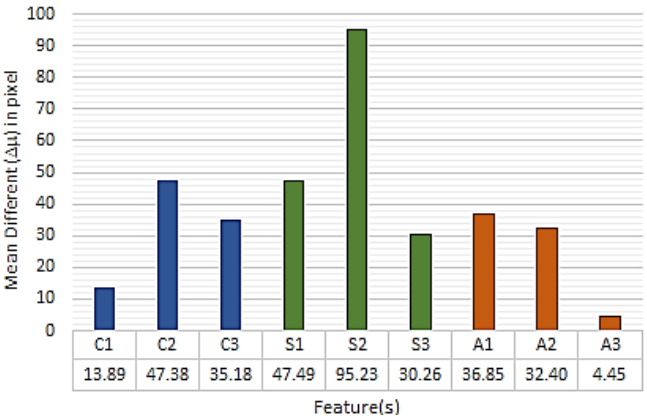


Figure 4: Means different between walking and abnormal features groups for n_v=3 i.e. $C_{1\text{-}3}$, $S_{1\text{-}3}$ and $A_{1\text{-}3}$

The *k*-fold cross-validation was employed on each KSVM models where the dataset was randomly divided into *k* approximately equal size subsets. The training and validation sets were comprised of *k*-1 subsets and the remaining subset; respectively. This procedure was repeated *k* times, so that each subset was used once for validation. Then, a single estimation of the whole dataset was calculated from the combination of *k* folds result. In our case, *k*=10 which is considered as the best compromise value in computational cost and reliable estimates.

III. Results and Discussion

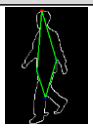
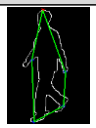
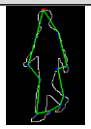
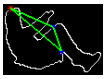
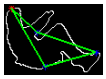
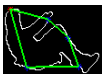
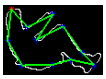
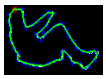
All tasks were done in MATLAB® R2015a and Statistical Package for the Social Science (SPSS) V22 software, which are embedded in a notebook computer: Intel i7 processor, running Windows 10 OS, with 16GB of RAM.

A. Feature Analysis

Generally, the total number of features; C_i, S_i and A_i is based on the number of vertices ($n_v \geq 3$) selection. The consequence of this factor may affect the generalized shape of the silhouette (refer to Table 2). Each generalized shape of silhouette will contribute $3n_v$ features i.e. C_1-C_{nv}, S_1-S_{nv} and A_1-A_{nv}. For features analysis, we examined all groups of features. However, only features extracted from the minimum n_v, i.e. 3 vertices will be highlighted in this paper. All features were measured on 10,000 walking pose samples, N_1 and 10,000 abnormal pose samples, N_2.

The difference in means, $\Delta\mu_{\{C,\ S,\ A\}}$ were measured, and generally defined as the absolute difference between the mean value, μ in two different groups. Figure 4 summarized the $\Delta\mu_{\{C,\ S,\ A\}}$ between walking and abnormal groups' features for the triangle shape; which composing nine features in total. In general, it will shows the distinction of $\mu_{\{C,\ S,\ A\}}$ between corresponding features in both groups.

Table 2
Human Silhouette Polygonal Shape Generalization

Pose	$n_v = 3$	$n_v = 4$	$n_v = 5$	$n_v = 11$	$n_v = 50$
Walking:					
Abnormal:					

By feature groups, the average of Δμ for *C*, *S* and *A* groups are 32.15, 57.66 and 24.57 pixels; respectively. Thus, the *S* group contributes the highest Δμ and gives a good impact in distinguishing the poses, rather than feature *C* and *A* groups. Whereas, the average of mean different, $\Delta\mu_{average}$ for all features, regardless to feature groups is 38.13 pixels.

Therefore, the mean difference results show a significant value to differentiate the pose groups in a descriptive sense for all features. The side by side box plots; grouped by feature typed-*C* in Figure 5 shows the alternative way to visualize the distribution of data in discriminating between the two pose groups and to observe the outliers. This descriptive statistic results show that most of the features might be prominent in distinguishing the two different groups. But it does not, however, allow us to make a decision beyond this preliminary data analytic to reach a conclusion regarding the hypotheses mentioned above. Therefore, an inferential statistic is performed for judgments of probability that the observed samples between groups are different in this study and be prominent features in classification.

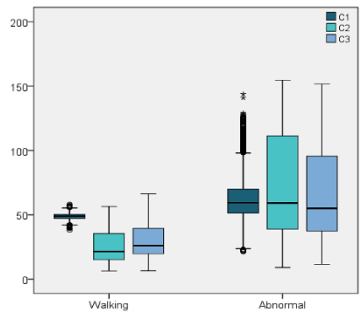


Figure 5: Graphical rendition of *C* feature group for walking and abnormal dataset

The SW test was conducted as numerical means of assessing normality. The normality test results for all features are summarized in Table 3 and the it shows that all probability result, *p*-value (Sig. row) to correspond features are less than 0.001 (N = 10,000). Therefore, the test rejects the hypothesis of normality for all features due to the conducted test resulting *p*-value is less than 0.05 and failing the normality test states with 95% confidence, the data significantly deviates from a normal distribution. Therefore, the non-parametric MWU test was conducted to compare differences between the two independent groups.

The MWU test is based on ranks or medians, where the ranks represent the relative position of an individual in comparison to others and robust to outliers. The selected test is to analyse all corresponding features extracted from the generalized triangle shape of walking pose should be distinct from those abnormal pose features.

The detailed statistical results of MWU are shown in Table 4. It shows the actual significance value of the test; specifically, *U* statistic as well as the asymptotic significance (2-tailed), *p*-value. From this table, it shows all probabilities values, *p* were below 0.001, rejecting the null hypotheses for all features ($\alpha < 0.05$). Thus, the mean rank between the groups for all nine features were associated statistically significantly different median latencies in groups of walking and abnormal ($N_1 = N_2$ = 10,000).

B. Pose Classification

The dataset for classification consists of 10,000 samples for each pose groups. We performed the classification of walking and abnormal poses using three selected kernel SVM models; namely Lin-KSVM, Pol-KSVM and RBF-KSVM. The evaluation was made on all attributes *C*, *S* and *A*; extracted from the range of three to 25 vertices. Table 5 presents the classification accuracy performance of our proposed features based on different numbers of vertices.

From the range of vertices number undertaken, all highest average classification accuracy (highlighted in red font) were obtained from the Pol-KSVM, with the minimum accuracy rate, Acc_{min}=98.16% (vertex-3) and the maximum accuracy rate, Acc_{max}=99.96% (vertex-20). However, all KSVM models performed well (Acc > 95%) with a minimum of 95.98% accuracy rate (Lin-KSVM at vertex-3). The average performance rate for Pol-KSVM was 99.65% and the classifiers performance were saturated at vertex-20 onwards (refer to Figure 6). Therefore, in the selection of 20 vertices, particularly for Pol-KSVM was sufficiently good to gain the optimum classification rate (99.96%).

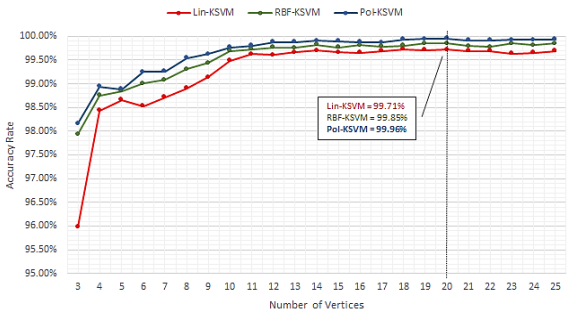


Figure 6: Classification accuracy performance correspond to various vertex number

C. Feature Extraction Techniques Comparison

For further performance evaluation of our proposed FE technique, we designed an extra experiment to compare them with other shape-based FE techniques: [4] and [5]. The experiment was following the same framework as shown in Figure 1 except we neglected the feature selection process. The performance was evaluated base on three criteria: accuracy (*Acc*), precision (*Pr*) and specificity (*Sp*) as derived in Equation (12)-(14). From the classification performance results in Table 6, our proposed FE technique outperformed.

Table 3
Shapiro-Wilk test result for all features extracted from triangle shape (N=10,000)

Feature:	C_1		C_2		C_3		S_1		S_2		S_3		A_1		A_2		A_3	
Pose:	W	A	W	A	W	A	W	A	W	A	W	A	W	A	W	A	W	A
Statistic:	.062	.106	.121	.137	.104	.119	.086	.199	.105	.127	.080	.176	.087	.076	.039	.158	.073	.101
Sig.:	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

W: *Walking* A: *Abnormal*

Table 4
Statistical result of Mann Whitney U Test (N=10,000)

Feature:	C_1	C_2	C_3	S_1	S_2	S_3	A_1	A_2	A_3
U:	17946019.500	10481969.000	15026983.00	17431236.000	2487242.000	22623075.000	2280210.000	20887694.00	46097398.000
Asymp. Sig. (2-tailed):	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 5
The KSVMs' Accuracy Rate Performance (%)

No of Vertices:	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	**20**	21	22	23	24	25
Lin-KSVM:	**95.98**	98.43	98.66	98.53	98.71	98.90	99.14	99.48	99.62	99.60	99.66	99.70	99.66	99.65	99.69	99.72	99.71	99.71	99.68	99.69	99.63	99.65	99.68
Pol-KSVM:	98.16	98.93	98.88	99.25	99.26	99.54	99.62	99.76	99.81	99.88	99.88	99.91	99.90	99.88	99.87	99.93	99.94	**99.96**	99.92	99.92	99.93	99.93	99.94
RBF-KSVM:	97.93	98.75	98.84	99.00	99.07	99.30	99.44	99.69	99.73	99.76	99.75	99.82	99.76	99.81	99.78	99.80	99.85	99.85	99.79	99.77	99.85	99.82	99.85

Table 6
Average Percentage (%) Performance of Different Feature Extraction Techniques on POL-KSVM classifier

Feature Extraction Techniques	*Acc*	*Pr*	*Sp*
Centroidal profile [4]	99.94	99.89	99.89
SSG [5]	97.18	95.04	94.81
Polygonal shape-based features (proposed method)	**99.96**	**99.94**	**99.94**

both earlier techniques with above 99.94%.

IV. Conclusion and Future Research

We have proposed new features to be extracted from the polygonal shape generalization on the human silhouette. The Shapiro-Wilk test was conducted to assess the normality of data distribution and it summarized that the test rejected the hypothesis of normality for all features ($p < 0.001$). Thus, to analyse whether all corresponding features of walking pose are statistically significantly different with abnormal pose features, the Mann-Whitney U test was conducted. The results showed that all proposed features could discriminate the human walking and abnormal poses at the significant level of $p < 0.05$. Based on classification accuracy, the Pol-KSVM performed well (Acc_{max} = 99.96%) at vertex-20 than Lin-KSVM and RBF-KSVM for all range of vertices number i.e. three to 25 vertices. Overall, all KSVMs were giving high achievement performance (Acc > 95%). In the future, we are keen to evaluate our proposed features on other types of supervised classifiers such as Discriminant analysis and Naïve Bayes. Such that the best classification model will be tested on other online databases and real-time application; particularly in surveillance system.

Acknowledgment

The authors acknowledge the financial support from the following: UKM (DIP-2015-012) & MOSTI (01-01-02-SF1386).

References

[1] D. Triantafyllou, S. Krinidis, D. Ioannidis, I. N. Metaxa, C. Ziazios, and D. Tzovaras, "A real-time fall detection system for maintenance activities in indoor environments," *IFAC-PapersOnLine*, vol. 49, no. 28, pp. 286–290, 2016.

[2] S. S. Khan and J. Hoey, "Review of fall detection techniques: A data availability perspective," *Med. Eng. Phys.*, vol. 39, pp. 0–1, 2016.

[3] M. F. Abu Hassan, A. Hussain, M. H. Md Saad, and K. Win, "3D distance measurement accuracy on low-cost stereo camera," *Sci. Int.*, vol. 29, no. 3, pp. 599–605, 2017.

[4] N. M. Tahir, A. Hussain, S. L. A. Samad, and H. Husain, "Pengecaman insan berasaskan kaedah profil sentroid dan pengelas rangkaian neural buatan," *J. Teknol.*, vol. 53, no. September, pp. 69–79, 2010.

[5] A. Hussain, S. Shahbudin, H. Husain, S. A. Samad, and N. M. Tahir, "A simplified shock graph for human posture classification using the adaptive neuro fuzzy inference system," *J. Inf. Comput. Sci.*, vol. 9, no. 8, pp. 2035–2048, 2012.

[6] T. T. Shuai Zheng, Junge Zhang, Kaiqi Huang, Ran He, "Robust view transformation model for gait recognition," in *18th IEEE International Conference On Image Processing*, 2011, pp. 2117–2120.

[7] I. Charfi, J. Miteran, J. Dubois, M. Atri, and R. Tourki, "Optimized spatio-temporal descriptors for real-time fall detection: comparison of support vector machine and Adaboost-based classification," *Electron. Imaging*, vol. 22, no. 4, pp. 752907-752907–9, 2013.

[8] S. Aslani and H. Mahdavi-nasab, "Optical flow based moving object detection and tracking for traffic surveillance," *Int. J. Electr. Comput. Energ. Electron. Commun. Eng.*, vol. 7, no. 9, pp. 963–967, 2013.

[9] J. E. Santoyo-Morales and R. Hasimoto-Beltran, "Video Background subtraction in complex environments," *J. Appl. Res. Technol.*, vol. 12, no. 3, pp. 527–537, 2014.

[10] M. Oral and U. Deniz, "Centre of mass model - A novel approach to background modelling for segmentation of moving objects," *Image Vis. Comput.*, vol. 25, no. 8, pp. 1365–1376, 2007.

[11] F. van der Heijden, *Image based measurement systems: Object recognition and parameter estimation*. 1994.

[12] E. Şaykol, U. Güdükbay, and Ö. Ulusoy, "A histogram-based approach for object-based query-by-shape-and-color in image and video databases," *Image Vis. Comput.*, vol. 23, no. 13, pp. 1170–1180, 2005.

[13] P. Zhu, Q. Xu, Q. Hu, C. Zhang, and H. Zhao, "Multi-label Feature Selection with Missing Labels," *Pattern Recognit.*, 2017.

[14] E. Han and A. V. M. Ines, "Downscaling probabilistic seasonal climate forecasts for decision support in agriculture: A comparison of parametric and non-parametric approach," *Clim. Risk Manag.*, no. November 2016, 2017.

[15] Jamadagni, "Illustration of various quadrilaterals," *Wikimedia Commons*, 2016. [Online]. Available: https://commons.wikimedia.org/wiki/File:Quadrilaterals.svg. [Accessed: 19-Jul-2017].