The Impact of Geometric And Motion Features on Sign Language Translators

SARA BILAL*, RINI AKMELIAWATI**
**Department of Science in Engineering*
***Department of Mechatronics Engineering*
International Islamic University Malaysia (IIUM)
Jl Gombak 53100, Kuala Lumpur
MALAYSIA
**sarra@iium.edu.my, ** rakmelia@iium.edu.my*

Abstract: - Malaysian Sign Language (MSL) recognition system is a choice of augmenting communication between the hearing-impaired and hearing communities in Malaysia. Automatic translators can play an important role as alternative communication method for the hearing people to understand the hearing impaired ones. Automatic Translation using bare hands with natural gesture signing is a challenge in the field of machine learning. Researchers have used electronic and coloured gloves to solve mainly three issues during the pre-processing steps before the singings' recognition stage. First issue is to differentiate the two hands from other objects. This is referred to as hand detection. The second issue is to describe the detected hand and its motion trajectory in very descriptive details which is referred to as feature extraction stage. The third issue is to find the starting and ending duration of the sign (transitions between signs). This paper focuses on the second issue, the feature extraction by studying the impact of the vector dimensions of the features. At the same time, signs with similar attributes have been chosen to highlight the importance of features' extraction stage. The study also includes Hidden Markov Model (HMM) capability to differentiate between signs which have similar attributes.

Key-Words: - Malaysian Sign Language, Feature Extraction, Features Vector Dimensions, HMM

1 Introduction

Everyday communication with the hearing population poses a major challenge to those with hearing loss. Most hearing people do not know sign language (SL) and know very little about hearing-impaired people in general. Furthermore, most hearing people do not know how to communicate in a spoken language with a hearing-impaired person, even though the latest can speak and read lips. Therefore, not only the communication barriers for the hearing-impaired people appear at the bank, police station and etc, but also essential information about health, employment, and legal matters is inaccessible for them. Common current options for alternative communication modes include cochlear implants, writing, and interpreters.

Cochlear implants, handwriting and interpreters are alternatives for communicating with hearing-impaired people [1]. In fact, most of the hearing-impaired people use SL as their preferred language for communication. Therefore, Automatic SL Translators (ASLT) can be an effective solution to break the difficulties of communicating between the hearing-impaired people and hearing people.

The SL translators have various stages starting from collecting the signs from the signer, the two hands and face blobs detection, blob tracking, features extraction and finally training and testing the system. The aim of this paper is to study the system accuracy using signs that have similarities and the ASLT capability to differentiate between the signs. Besides, the key objective of this work is to study the effect of features extracted from the signs and feature vector dimensions on the overall system accuracy.

Our ASLT system has been developed using VC++ and OpenCV under Windows and VC++ under LINUX. Various features have been extracted from the hand trajectory path. Hidden Markov Model (HMM) has been used as the recognizer. The experimental results show that the extracted features and feature vector dimension have a high impact on the overall ASLT system accuracy.

2. Research Background

Diverse traditional methods exist in the field of pattern recognition to achieve hand posture and

ISBN: 978-1-61804-276-7

gesture recognition [2, 3], such as artificial intelligence techniques, statistical algorithms and other non- traditional developed algorithms. HMM is a statistical method that has been used widely in pattern recognition and HCI applications. The original difficulty of automatic SL recognition is to find transitions between signs and to differentiate between signs with similarities.

First problem can be solved by using HMM because it has the ability to segment the continuous signs. The second problem can be solved by obtaining the most descriptive features from the signs. This can be done by studying the nature of the signs.

The hand posture can carry the meaning of the sign by itself. However, some signs can be understood only if the hand posture is accompanied by the information of the motion trajectory path. Two feature vectors can be gathered based on the hand shape and motion trajectory. The hand shape features will be extracted from hand contour and statistic information will be obtained. Besides, features from the motion trajectory path will be extracted using hand gesture path.

2.1 Hand Shape Features

It is not an easy task to differentiate signs with the same trajectories but different posture shape see, Fig.1 a&b. In this case, hand shape plays an important role.

(a) (b)

Figure 1: Malaysian SL (MSL) signs for (a) 'We' (b) 'She'

There are mainly three approaches to describe the hand shape. They are hand model based approach, appearance-based approaches and hand shape analysis where more details can be found in Bilal et al. [4]

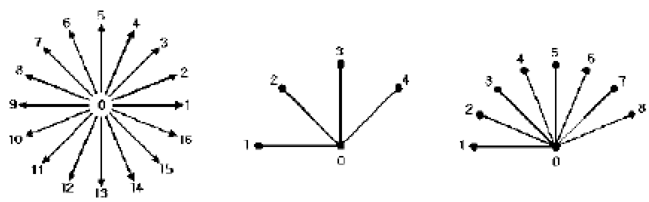
2.2 Hand Trajectory Features

Many of the applications that use hand gesture trajectory are matching and comparing the input gesture trajectory to each of the prototype gesture trajectories contained in the database for recognition. To accomplish that, there are two ways, either by direct trajectory shape matching and/or by trajectory feature matching [5]. The later one has proven its efficiency over the former one.

To capture the dynamic characteristics of the hand posture, sequence of frames with the hand motion must be available. The main issue is to track the hand in the video frame and extract the center point *(x, y)* of the hand posture over a sequence of successive frames. Then either trajectory-shape matching and/or trajectory-feature matching techniques can be developed as following:

2.2.1 Direct trajectory shape matching

Suk et al., [6] has represented the whole motion trajectory of the two hands by a sequence of motion vectors, each of which is in turn encoded by a direction code using the scheme as shown in Fig. 2(a). The central code '0' denotes 'no motion'. In this work, two chain codes were extracted from each hand but with separate chain coding for each hand. As ambiguities can arise between gestures, they represented the motion using only the chain code, and they introduced two more features: the relative position of the two hands (Fig. 2(b)) and the position of the each hand relative to the face (Fig. 2 (c)). The code '0' implies that two hands or a hand and a face are overlapping.



(a) (b) (c)

Figure 2: Features: (a) 17 direction codes for hand motions, (b) hand-hand positional relation, (c) face-hand positional relation. [6]

2.2.2 Trajectory feature matching

Several hand gesture recognition systems have been developed using various features such as raw position, velocity, chain code, mesh code, Zernike moments [7], image geometric parameters[8] , 2D edge and time edge features proposed by [9].

Mainly, all the aforementioned features extracted from the motion trajectory are based only on three basic information; position from origin, angle and velocity as shown in Fig.3 [10]

ISBN: 978-1-61804-276-7

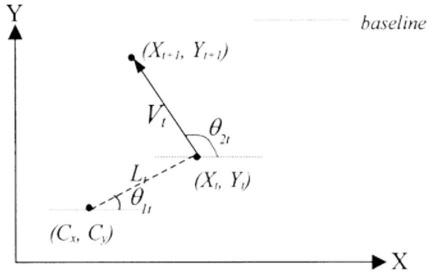


Figure 3: Basic features between two points of a gesture sequence [11]

Furthermore, for feature based trajectory matching, Bhuyan, et al. [5] have categorized features into static and dynamic features extracted from a given gesture trajectory. Static features relate to the shape of the trajectory while the dynamic features relate to the nature of hand motion during gesturing.

A. Static Features

Fig.4 shows the static features that they have proposed to use which are extracted based on center point *(x, y)*:

a. Trajectory length:

$$l = \sum_{i=0}^{N} \left\{ \left(x_i - x_{i+1}\right)^2 + \left(y_i - y_{i+1}\right)^2 \right\}^{\frac{1}{2}} \qquad (1)$$

b. Number of significant curves on the trajectory is calculated as follows

$$\theta_i = \tan^{-1}\left(\frac{y_i - y_{i-1}}{x_i - x_{i-1}}\right), \quad i = 1,2,\ldots\ldots,N \qquad (2)$$

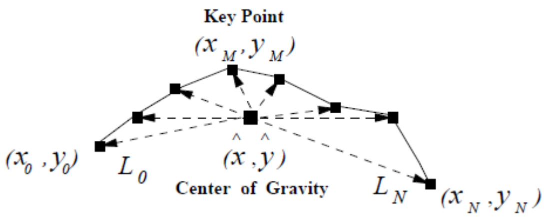


Figure 4: Distance from the key trajectory points from the center of gravity [5]

B. Dynamic Feature

Motion features are computed from the spatial positions of the hand in the gesture trajectory and the time interval between two prominent hand positions. The two motion features are the velocity (v_i) and the acceleration (a_i) of the hand during signing and are computed as below:

$$v_i = \left\{ \frac{x_{i+1} + x_i}{t_{i+1} - t_i}, \frac{y_{i+1} - y_i}{t_{i+1} - t_i} \right\}, \quad i = 0,1,2,\ldots\ldots N-1 \qquad (3)$$

$$a_i = \frac{v_{i+1} - v_i}{t_{i+1} - t_i}, \quad i = 0,1,2,\ldots\ldots N-2 \qquad (4)$$

Fig.5 shows the typical velocity profiles for gestures representing “Circle” and “Square” respectively.

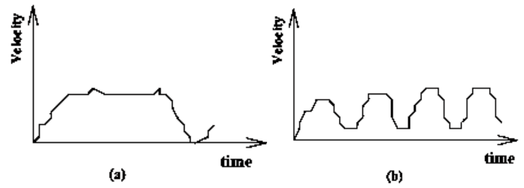


Figure 5: Typical velocity profile of gestures representing (a) ‘Circle’ and (b) ‘Square’. [5]

3. The MSL Translator

The MSL translator is a system developed to capture and recognize the hearing-impaired hands’ sign. There are four preparation stages before the final recognition stage as shown in Fig.6.

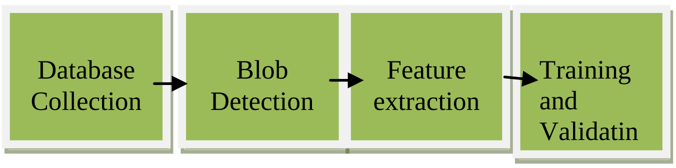


Figure 6: Overall Gesture Recognition System

3.1 MSL Database

The first step toward SL recognition is to collect the database from signers. A basic structure of our system consists of a web camera and a computer. The signer must face the camera and his/her upper body must be in the field of the camera view. The signer starts signing and the system will captured images and store the video. Depending on the system developer requirement the database can be collected in the form of isolated signs or continuous signs. In this work the criteria of collecting the database is as follows:

a. Many signs have been collected from MSL datasets [12].
b. The signer has to repeat each sign 10 to 15 times.
c. Signs with similarities have been chosen for study.

3.2 Blob Detection

A preprocessing stage is performed on the database to segment and extract the two hands and face. Bilal

ISBN: 978-1-61804-276-7

et al. in [13] have applied a hybrid method to detect skin blobs followed blob labeling. Then, the segmented hand and face were stored separately because features are to be extracted individually from each detected blob.

3.3 Feature Extraction

The performance of recognition system first depends on obtaining efficient features to represent pattern characteristics [11]. There is no algorithm which shows how to select the representation or choose the features [14]. Therefore, for our proposed system, the decision was made to select features which are trajectory and shape based features.

3.3.1 Trajectory Feature

Eight features have been extracted from the extracted blob trajectory. These eight features are stored in a feature vector. A feature vector is an *n*-dimension vector which represents numerical features of an object.

a. Centre point of the hand (x_c, y_c). *(2 features)*
b. Difference between consecutive (x_c, y_c): *(2 features)*
c. ra_x and ra_y are the average sum of x_c and y_c respectively. *(2 features)*
d. Velocity
e. Angle

3.3.2 Geometric Features

Contours or edges are somewhat universal features which can be used in any model-based technique [15] as well as non-model ones. The aim is to have similar values of features for similar hand shapes and distant values for different shapes. It is also required to have scale invariant features so that images with the same hand shape but different size would have the same feature values [16]

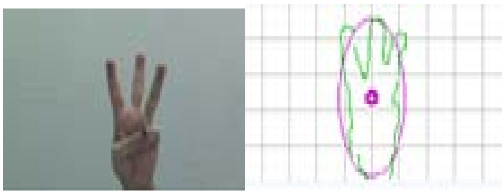
(a)  (b)

Figure 7: (a) Posture "W" from MSL Database (b) Contour of (a)

This is done by choosing eight features which are scale invariant. These features are moment based features from the gray scale image. The definition of moments of the gray value-function *f(x, y)* of an object is the following: [17]

$$m_{p,q} = \iint x^p y^q f(x,y)dxdy \qquad (5)$$

Moments are generally classified by the order of the moments. The order of a moment depends on the indices *p* and *q* of the moment $m_{p,q}$ and vice versa. The sum $p + q$ of the indices is the order of the moment $m_{p,q}$. Based on that, seven shape features have been derived from the orders of the moment to describe the face and the two hands blobs.

The integration is calculated over the area of the object. Generally each other pixel based feature instead of the gray value could be used to calculate the moments of the object. Using binary images the gray value function *f(x,y)* becomes

$$f(x,y)\begin{cases} 1 & Object \\ 0 & Background \end{cases} \qquad (6)$$

and can be neglected in the subsequent formulas.

1. Blob area: The zero order moment describes the area *A* of the object
$$m_{0,0} = \iint f(x,y)dxdy \qquad (7)$$
2. Semi major and semi minor axes denoted by *a* & *b* respectively
3. Orientation angle in which the blob has its biggest extension θ as in Fig. 7:
$$\theta = \frac{1}{2}\arctan\left(\frac{2m_{1,1}}{m_{2,0} - m_{0,2}}\right) \qquad (8)$$
4. Compactness C= (Area / Perimeter^2) (9)
5. Roundness. $k = \dfrac{P^2}{A}$ (10)

Because a circle has the maximal Area *A* within a given perimeter *P*, a scaling of roundness *k* is performed:

$$k = \frac{\text{Perimeter}^2}{2\pi A} \qquad (11)$$

Therefore $k = 1$ for circle and $k > 1$ for other objects

6. The eccentricity ε can directly derived from the semi-major and semi-minor axes *a* and *b* of the object. Also ε can be directly calculated from the central moments of second order by
$$\varepsilon = \frac{\sqrt{(m_{2,0} - m_{0,2})^2 - 4m_{1,1}{}^2}}{m_{2,0} + m_{0,2}} \qquad (20)$$

ISBN: 978-1-61804-276-7

The eccentricity ε can have values from 0 to 1. It is 0 for a perfectly round object and 1 for a line shaped object.

3.4 Training and Testing

Training of the HMM-based recognizer was done with GT²k [18]. The tool is developed using C++ under LINUX. Many parameters and variables have to be modified inside the GT²k tool to fit our requirement for classifying MSL gestures.

3.4.1 GT²k Grammar

GT²k needs grammar to perform training and recognition. It is possible to generate it automatically or can be specified manually in a text file. If it is auto set then the grammar will only allow gestures to be recognized one at a time and for continuous recognition a special grammar setting which is set manually is required. An example of grammar used for testing the two isolated gestures ‘big’ and ‘dirty’ is shown as follows:

```
$gesture =  big | dirty;

 ( $gesture )
```

3.4.2 Type of training

GT²k provides two types training/validation techniques and they are *cross validation* and *leave-one-out validation*. Basically, cross-validation randomly divides the data into two sets, 66.6% as training set and 33.3% validation set. To select the options for the above techniques, the main settings in GT²k have changed during HMM training and set to “CROSS”.

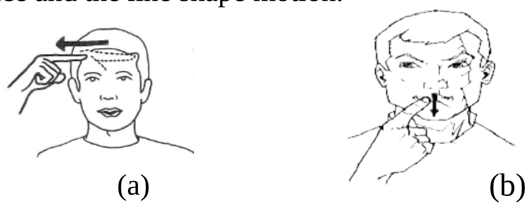
4. Experimental Results

A set of training and testing sets were created using a randomly selected 90% of the data for training and the remaining 10% for model validation. The model was then trained with the automatic trainer. Sign boundaries were re-estimated over several iterations to ensure better training. After training, the models were tested against the remaining 10% of the data. Recognition accuracy was determined, with the standard penalties for substitutions (S) , insertions (I) , and deletions (D) errors and Total number of gesture *(N)* classified and total number of gesture *(H)* recognized. Substitution errors arises when the system classifies the gesture incorrectly and the insertion error occurs when the system imagine the occurrence of a gesture. The deletion error happens when the system does not recognize the gesture within a series of gesture. If gestures are isolated gestures then the value of D and I will always be Zero otherwise the results occurs during continuous recognition. The calculation of the accuracy of the overall system can be found using the following formula:

$$\text{Accuracy} = \frac{N\text{-}S - D - I}{N} * 100 \qquad (12)$$

4.1 Features Effect on the Accuracy of Two Signs Recognition

Two signs which are ‘red’ and ‘black’ have been chosen in the first set of experiments because of their close similarity. Fig.8 shows that, these two signs are common in the position with respect to the face and the line shape motion.



(a) (b)

Figure 8. Sign (a) ‘black’ and (b) ‘red’ from MSL Database [12]

Two sets of experiments have been run based on the suggested features:

A. The first one using only geometric features from face and hand.
B. The second one is using motion and shape features.

The overall performance is recorded in the form of a confusion matrix as shown in Table 1 & Table 2.

Table 1: Overall Results for training and testing for gesture ‘red’ and ‘black’ using only geometric features from hand and face

```
------------------------ Overall Results --------------------------
SENT: %Correct=53.85 [H=7, S=6, N=13]
WORD: %Corr=53.85, Acc=53.85 [H=7, D=0, S=6, I=0, N=13]
------------------------ Confusion Matrix -------------------------
       b   r
       l   e
       a   d
       c
       k      Del [ %c / %e]
blac   3   1    0 [75.0/7.7]
 red   5   4    0 [44.4/38.5]
Ins    0   0
===================================================================
```

Table 2: Overall Results for training and testing for gesture ‘red’ and ‘black’ using motion and geometric features from hand and face

```
------------------------ Overall Results --------------------------
SENT: %Correct=85.71 [H=6, S=1, N=7]
WORD: %Corr=85.71, Acc=85.71 [H=6, D=0, S=1, I=0, N=7]
------------------------ Confusion Matrix -------------------------
       b   r
       l   e
       a   d
       c
       k      Del [ %c / %e]
blac   4   0    0
 red   1   2    0 [66.7/14.3]
Ins    0   0
===================================================================
```

In Table 1, the accuracy is 53.81%, when only geometric features have been used. In contrast, Table 2 shows an accuracy of 85.71% for the two signs ‘black’ and ‘red’ using both geometric and trajectory features. Both the black and red signs are located in the face with the hand shape for both signs is the same (see Fig.8). Therefore, geometric features cannot stand alone to distinguish between these two signs.

4.2 Vocabulary of Four Signs

Four signs ‘green’, ‘yellow’, ‘water’ and ‘rice’ have been chosen based on the similarities between each pair (green’ & ‘yellow’ , ‘water’ & ‘rice’) to run the second set of experiments. Green and yellow are very similar to each other (see Figs.9 a & b). It is difficult even for human being to differentiate between these two signs.

Table 3: Overall Results for training and testing for gesture ‘yellow’ and ‘green’ ‘water’ and ‘rice’ using motion features from hand and face

```
------------------------ Overall Results --------------------------
SENT: %Correct=70.00 [H=14, S=6, N=20]
WORD: %Corr=70.00, Acc=70.00 [H=14, D=0, S=6, I=0, N=20]
------------------------ Confusion Matrix -------------------------
       g   r   w   y
       r   i   a   e
       e   c   t   l
       e   e   e   l
       n       r   o  Del [ %c / %e]
gree   1   0   0   1    0 [50.0/5.0]
rice   0   6   0   0    0
wate   0   3   4   0    0 [57.1/15.0]
yell   2   0   0   3    0 [60.0/10.0]
Ins    0   0   0   0
```

Table 4: Overall Results for training and testing for gesture ‘yellow’ and ‘green’ ‘water’ and ‘rice’ using motion and geometric features from hand and face

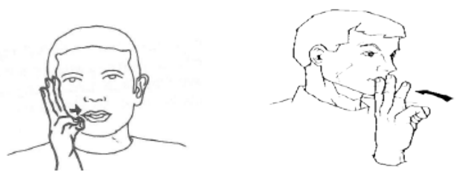
```
------------------------ Overall Results --------------------------
SENT: %Correct=58.33 [H=7, S=5, N=12]
WORD: %Corr=58.33, Acc=58.33 [H=7, D=0, S=5, I=0, N=12]
------------------------ Confusion Matrix -------------------------
       g   r   w   y
       r   i   a   e
       e   c   t   l
       e   e   e   l
       n       r   o  Del [ %c / %e]
gree   1   0   0   0    0
rice   1   2   0   1    0 [50.0/16.7]
wate   0   0   1   0    0
yell   2   1   0   3    0 [50.0/25.0]
Ins    0   0   0   0
===================================================================
```

In Table 3, it is clear that the confusion mainly happened between signs that have similarities such as ‘green’ and ‘yellow’ because the motion trajectory is the same. In such case, one can estimate that geometric features could play a role in the accuracy enhancement. In fact, Table 4 shows that, the existence of geometric features could drop the system accuracy. We could verify that as follows.

a. Signs ‘water’ and ‘rice’ have better accuracy in Table 4 rather than the confusion that happened between the two signs in Table 3. This due to nature of the two signs; i.e. slow dynamic motion (see Fig.10). This leads to low accuracy in Table 3.
b. Sign ‘green’ and ‘yellow’ are very similar even for human vision (see Fig.9 a & b). Therefore, none of the features could enhance the system accuracy in that case.

(a) (b)

Figure 9: Sign (a) ‘green’ and (b) ‘yellow’ from MSL Database [12]



(a) (b)

Figure 10: Sign (a) ‘rice’ and (b) ‘water’ from MSL Database [12]

4.3 Study the Effect of Motion Features Vector Dimensions on the Recognition Stage

Deeper analysis of the extracted features from the hand shape and motion trajectory has given us another dimension of choosing the best features that can describe the system characteristics

As mentioned in Section 3.3.1, there are eight motion features which have been considered in this research. Mainly in gesture recognition systems, researchers based their experiments on motion features [5][8] rather than other types of features because of the nature of the gesture trajectory formulation. Therefore, in this Section, the experiments will be conducted using the proposed eight motion features vectors to prove that features are very crucial for getting better recognition accuracy. However, feature vector dimension should be chosen carefully. Hence, two types of experiments have been conducted, i.e. using two hands and one hand with their corresponding datasets. Experiments conducted 100 times using 8 motion trajectory features and for the two signs dataset (i.e. two hands and one hand) in total 38 combinations of feature vectors have been tested.

Table 6: Six Signs using Two Hands

No	Image From MSL Database [12]	Name	Category
1		Car	Transportation
2		School	Place
3		Chicken	Food/ Bird
4		Noodles	Food
5		Dog	Animal

4.3.1 Features Vector Effect using Two Hands Signs

Five signs 'car', 'school' ,' chicken', 'noodle' and 'dog' have been chosen from MSL database to cover three MSL categories such as nouns, food and animals. Also for the two hands signs, it is very crucial to know the term "active and semi active" hand. When looking at Table 5 , it is clear for sign 'car', 'school' and 'noodles' that the two hands act equally to form the motion trajectory of the signs. Meanwhile for sign 'chicken' it is very obvious that one hand, the lower one, is not as active as the upper hand. Furthermore the motion trajectory shape of signs ('school' ,'noodles') and ('car', 'dog') have similarities in the motion trajectory direction and the trajectory shape. But because of the two hands provide the system with more information, we can find high recognition accuracy as shown in Table 6 for two hands' signs.

From the definition above for the active and semi active hand and the nature of the trajectory shape using two hands, a test has been conducted using different features vector combinations and the results are shown in Table 6.

Table 7: Signs with two hand system accuracy using various motion features vectors with 5 states HMM

No	Features in Use	System Accuracy using HMM '5' States (%)
1	v	59.4677
2	α	47.6798
3	x, y	74.0659
4	d_x, d_y	78.4619
5	ra_x, ra_y	73.7413
6	v , α	81.6428
7	x, y, v	82.5752
8	x, y, α	75.1584
9	x, y, d_x, d_y	87.5521
10	x, y, ra_x, ra_y	70.0199
11	v, α, x, y	82.7633
12	$x, y, d_x, d_y v$	85.6113
13	x, y, d_x, d_y, α	84.8595
14	x, y, ra_x ,ra_y ,v	83.0598
15	x, y, ra_x, ra_y, α	79.7853
16	$x, y, d_x, d_y ,ra_x, ra_y$	87.9845
17	$x, y ,d_x ,d_y, ra_x ,ra_y, \alpha$	86.3602
18	$x, y ,d_x ,d_y, ra_x ,ra_y, v$	90.5809
19	$x, y ,d_x ,d_y, ra_x ,ra_y,v, \alpha$	88.4964

4.3.2 Features Vectors Effect using One Hand Signs

A random set of six signs 'like', 'white' beat', 'bag', 'green', 'throw' using one hand has been chosen from the MSL database and their definition is shown in Table 7.

Table 8: Six Signs using One Hand

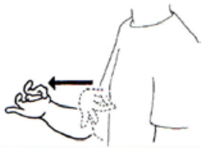
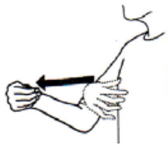
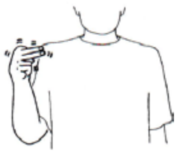
No	Image From MSL Database [12]	Name	Category
1		Like	Feeling
2		White	Colour
3		Beat	Verb
4		Bag	Name
5		Green	Colour
6		Throw	Verb

Table 8 shows the results obtained after training 6 gestures using one hand and 5 states HMM. From Table 6, the accuracy obtained has been analyzed based on the feature combination and features vector dimension chosen as follows:

(a) The higher the vector dimension the better the results are. When 8 feature vector dimension, (i.e. $x, y, d_x, d_y, ra_x, ra_y, v, \alpha$) was used, the system accuracy reached up to 88.4964%. Meanwhile, the two lowest accuracy 47.6798% and 59.4677% are obtained, when feature vector 'α' and 'v' (with vector dimension of 1) are used, respectively.

(b) From the features vector (x, y), (dx, dy), and (rax, ray), one can get an accuracy of 74.0659%, 78.4619% and 73.7413%, respectively. However, when these six features ($x, y, d_x, d_y, ra_x, ra_y$) are combined together, an accuracy of 87.9845% is obtained.

(c) Further observation found that if the feature 'α' is omitted, the accuracy increased up to 90.5809%. Whereas if all 8 features is used (i.e. α is included) the accuracy dropped to 88.4964%.

(d) The absence of 'v' reduces the system recognition accuracy from 88.4964% for ($x, y, d_x, d_y, ra_x, ra_y, v, \alpha$) features to 86.7258%.
Observation (c) and (d) can be verified as follows:
- For signs involved two hands, whether the two hands are both active or one is less active, the two hands could provide rich information of the velocity 'v' vector.
- The angle 'α' obtained from both hands could cause a conflict between the two hands because they might have different direction during the motion duration.

Table 9: Signs with one hand system accuracy using various features dimension with 5 states HMM

No	Features in Use	System Accuracy with 5 States HMM (%)
1	v	29.1502
2	α	36.82963
3	x, y	66.0561
4	d_x, d_y	33.6767
5	ra_x, ra_y	61.1086
6	v, α	31.3251
7	x, y, v	57.7896
8	x, y, α	61.2003
9	x, y, d_x, d_y	55.8674
10	x, y, ra_x, ra_y	64.1225
11	v, α, x, y	55.8515
12	$x, y, d_x, d_y v$	51.0751
13	x, y, d_x, d_y, α	65.7133
14	x, y, ra_x, ra_y, v	71.4987
15	x, y, ra_x, ra_y, α	75.9883
16	$x, y, d_x, d_y, ra_x, ra_y$	71.6515
17	$x, y, d_x, d_y, ra_x, ra_y, \alpha$	67.3555
18	$x, y, d_x, d_y, ra_x, ra_y, v$	64.7272
19	$x, y, d_x, d_y, ra_x, ra_y, v, \alpha$	60.3255

From Table 8, the system accuracy has been analyzed based on the feature combination and features vector dimension is stated below:

(a) The higher the vector dimension the better the results are. When 8 feature vector dimension (x, y ,d_x ,d_y, ra_x ,ra_y,v, α) are used, the system accuracy reached up to 60.3255 %. Meanwhile, the two lowest accuracy 29.1502% and 36.82963% are obtained, when feature vector 'v' and 'α' (with vector dimension of 1) are used respectively.

(b) In the case of one hand, the *(x, y)* vector could be the most descriptive vector for the system as has been observed in many other research [19][20]. But when more analysis is done, we found that the system characteristics require more than random data analysis. It requires deep investigation on the combination of the observed system image to find the best descriptive information. This is the case because of the following observations:

- Feature vector (x, y, d_x, d_y, ra_x, ra_y) with 6 features vector dimension gives an accuracy of 71.6515% and feature vector (x, y, ra_x, ra_y, α) with 5 features vector dimension gives much better accuracy that reached up to 75.9883%.
- Another test has been conducted to confirm the above point has proven that the feature vector with vector dimension of 3 (x, y, v) and (x, y, α) have much higher accuracy 57.7896% and 61.2003%, respectively whereas with four dimensional feature vector (x, y, v, α), an accuracy of 55.85% is obtained.

  Hence, these observations contradict the hypothesis that having more vector dimension may help to improve accuracy level.

(c) From the features vector (x, y), (dx, dy), and (rax, ray) we can get an accuracy of 66.0561%, 33.6767% and 61.1086%, respectively. But the most interesting result has been obtained when combining these six features together (x, y, d_x, d_y, ra_x, ra_y), an accuracy of 71.6515% is obtained. This gives us another dimension that:

- Having one feature vector or two feature vector dimension which drop the system accuracy as the step (c) , could be repaired by adding more features in the vector dimension that have proven their ability to enhance the system accuracy

(d) A case of study was tested carefully when 'v' and 'a' are absent from the features vectors. Table 8 shows that, by omitting feature vector 'v', a higher accuracy of 67.3555% has been obtained comparable to (x, y ,d_x ,d_y, ra_x ,ra_y,v, α) features vector accuracy of 60.3255%. Meanwhile, the absence of 'α' from the features vector could bring the accuracy down to 64.7272%.

The verification of that can be taken as the opposite of two hands results because here one hand involve then there are less info

5 Conclusion

In this work, it is observed that signs have variations in terms of hand shape, trajectory motion and starting and ending position of the sign. As a result, features' effect on the automatic SL translator has been verified. HMM is used to train and test the system using one hand and two hands isolated signs, from the MSL database. Features are extracted from the blobs motion trajectory and blobs shape. The results show that, the system accuracy depends mainly on the features' vector. Different feature vectors were generated and experiments were conducted 100 times for 38 features vectors, making it in total 3800 experiments. It was also observed that, the feature vector combination is much important than the feature vector dimension.

The aforementioned result affirms that the understanding of the system's nature is crucial and vital at the same time we must find the best features that could describe the systems' characteristics in an augmented mode. For future work, more feature vector such as hue features, polynomial interpolation and other relative feature vector should be studied in order to select the best feature vector combination.

References

[1] Intitute, G. R. (2001). Regions regional and national summary report of data from the 1999-2000 annual survery of deaf and hard of hearing children and youth. Washington, D. C: GRI, Gallaudet University.

[2] Derpanis, K. G. (2004). A review of vision-based hand gestures (D. o. C. Science, Trans.): York University.

[3] Mitra, S., & Acharya, T. (2007). Gesture recognition: a survey, systems, man, and cybernetics , part C: applications and reviews. *IEEE Trans, 37*(3), 311–324.

[4] Bilal, S., Akmeliawati, R., Salami, M. J. E., & Shafie, A. A. (2011). *Vision-based hand posture detection and recognition for Sign Language — A study* Paper presented at the International Conference On Mechatronics (ICOM).

[5] Bhuyan, M. K., Bora, P. K., & Ghosh, D. (2008). Trajectory Guided Recognition of Hand Gestures having only Global Motions. *International Journal of Electrical and Computer Engineering, 3*(4).

[6] Suk, H.-I., Sin, B.-K., & Lee, S.-W. (2008). *Recognizing hand gestures using dynamic Bayesian network* Paper presented at the Conference on Automatic Face & Gesture Recognition. FG '08. 8th IEEE International, Amsterdam

[7] Hunter, E., Schlenzig, J., & Jain, R. (1995). *Posture estimation in reduced-model gesture input systems*. Paper presented at the Proceedings of the International Workshop on Automatic Face-and Gesture-Recognition.

[8] Starner, T., Pentland, A., & Weaver, J. (1998). Real-Time American Sign Language Recognition Using Desk and Wearable Computer Based Video. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 20*(12), 1371-1375.

[9] Segan, J. (1993). *Controlling computer with gloveless gesture.* Paper presented at the Virtual Reality System '93.

[10] Jung, H.-S. Y., Soh, J., Min, B.-w., & Yang, H. S. (1999). Recognition of Alphabetical Hand Gestures Using Hidden Markov Model. *IEICE TRANS. FUNDAMENTALS E82-A*(7).

[11] Yoon, H.-S., Soh, J., Bae, Y. J., & Yang, H. S. (2001). Hand Gesture Recognition Using Combined Features of Location, Angle and Velocity. *Pattern Recognition Lettters, 34*, 1491-1501.

[12] Society, M. (2011). Malaysian Federation Deaf fromhttp://www.mfd.org.my/public/edu_eSign.asp

[13] Bilal, S., Akmelawati, R., Salami, M., Shafie, A., & Bouhabba, E. (2010). *A hybrid method using haar-like and skin-color algorithm for hand posture detection, recognition and tracking.* Paper presented at the International conference on mechatronics and automation (ICMA), Xi'an, China.

[14] Imagawa, K., Lu, S., & Igi, S. (1998). Color-based Hands Tracking System for Sign Language Recognition. *Proc. Third IEEE International Conference on Automatic Face and Gesture Recognition*, 462– 467.

[15] Lowe, D. G. (1991). Fitting parameterized three-dimensional models to images. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 13*(5), 441–450.

[16] Aran, O., Keskin, C., & Akarun, L. (1998). *Sign Language Tutoring Tool*. Paper presented at the International Conference on Computer Vision (ICCV'98), Mumbai, India.

[17] Kilian, J. (2001). Simple Image Analysis By Moments.

[18](GT2k), T. G. T. G. T. (2011). The Georgia Tech Gesture Toolkit (GT2k) Retrieved June, 2011

[19] Chen, F.-S., Fu, C.-M., & Huang, C.-L. (2003). Hand gesture recognition using a real-time tracking method and hidden Markov models. *Image and Vision Computing, 21*(8), 45–758.

[20] Zahedi, M., & Manashty, A. R. (2011). Robust Sign Language Recognition System Using ToF Depth Cameras. *World of Computer Science and Information Technology Journal (WCSIT), 1*(3), 50-55.