Supavit MUANGJAROEN[1], Sakol UDOMSIRI[1]

Pathumwan Institute of Technology, Thailand
ORCID: 0000-0002-2399-6619

doi:10.15199/48.2023.06.37

Continuous Speech Commands Recognition with Thai Language used Support Vector Machine Technique: A case study of speech commands control for mobile robots

***Abstract**. This article presents the application of an automatic speech recognition by continuous speech commands recognition with Thai language as a speaker verification model, this is a case study of speech commands control of mobile robots. The design of the automatic speech recognition system consisted of 3 steps: The first we analyzed the signal processing of the continuous speech commands and compared the accuracy of the speech recognition with a time frame adjustment and the overlapped period of signal filtered with the window function, The second we proceed to find the feature extraction of speech commands using format frequency techniques and configured the feature extraction with format frequencies of F1, F2, and F3,The last step was to design the recognition using Support Vector Machine technique to check the accuracy of an automatic speech recognition. These is support vector machine classification algorithm provides a comparison of the filtered function window and compares the accuracy of the time frame scaled and the overlapped time of the filtered, which gives different values of precision. From the experiment, the researcher found that are applied a Hanging function the test results of the test result of the "forward" speech commands has an accuracy of 81.92% but kind of Gaussian function the test results of the "backward" speech commands has an accuracy of 83.69%, the "turn left" speech commands had an accuracy of 82.81%, the "turn right" speech commands had an accuracy of 85.56% and the "Stop first" speech commands has an accuracy of 86.78% and speech recognition by continuous speech commands recognition with Thai language was applied an every function the test results of the overall performance of the speech commands has an accuracy of 83.88%.*

Streszczenie.** Artykuł przedstawia zastosowanie automatycznego rozpoznawania mowy poprzez ciągłe rozpoznawanie poleceń głosowych z językiem tajskim jako modelem weryfikacji mówiącego, jest to studium przypadku sterowania poleceniami głosowymi robotów mobilnych. Projekt systemu automatycznego rozpoznawania mowy składał się z 3 etapów: W pierwszym przeanalizowano przetwarzanie sygnału ciągłych poleceń głosowych i porównano dokładność rozpoznawania mowy z dopasowaniem przedziału czasowego i nakładającym się okresem sygnału filtrowanego funkcją okna, Następnie przystępujemy do znalezienia ekstrakcji funkcji poleceń głosowych przy użyciu technik formatowania częstotliwości i skonfigurowania ekstrakcji cech z częstotliwościami formatu F1, F2 i F3. Ostatnim krokiem było zaprojektowanie rozpoznawania przy użyciu techniki maszyny wektorów nośnych w celu sprawdzenia dokładności automatyczne rozpoznawanie mowy. Jest to algorytm klasyfikacji maszyny wektorów nośnych, który zapewnia porównanie przefiltrowanego okna funkcji i porównuje dokładność skalowanych ram czasowych oraz nakładających się czasów filtrowanych, co daje różne wartości precyzji. Na podstawie eksperymentu badacz odkrył, że po zastosowaniu funkcji wiszącej wyniki testu wyników poleceń głosowych „do przodu" mają dokładność 81,92%, ale rodzaj funkcji Gaussa wyniki testu poleceń głosowych „wstecz" mają dokładność 81,92% dokładność 83,69%, polecenia głosowe „skręć w lewo" miały dokładność 82,81%, polecenia głosowe „skręć w prawo" miały dokładność 85,56%, a polecenia głosowe „Najpierw zatrzymaj" mają dokładność 86,78%, a rozpoznawanie mowy przez zastosowano ciągłe rozpoznawanie poleceń głosowych w języku tajskim, a wyniki testu ogólnej wydajności poleceń głosowych mają dokładność 83,88%.* (Ciągłe rozpoznawanie poleceń głosowych w używanym języku tajskim Technika maszynowa wektorów pomocniczych: studium przypadku kontroli poleceń głosowych dla robotów mobilnych***)

Słowa kluczowe: proszę podać cztery terminy opisujące treść artykułu.
Keywords: Thai language, Continuous Speech Commands, Support Vector Machine**:** SVM**,** automatic speech recognition**:** ASR

Introduction

Currently, Thailand in the country's industrial work tends to use more robots and automation systems. We can clearly see that there will be adjustments to the infrastructure to foster transition for industries of all sizes. To be in line with the technological advancement of the current world after the COVID-19 situation. In order to solve the crisis situation that has arisen, the ratio of robots and automation to replace existing labor in the system is an important part. That is indicative of the country's industrial potential. Thailand must therefore be prepared in terms of human resources development with knowledge and skills in the robotics and automation industries to support the workforce that is diminishing day by day and being replaced by robots and automation systems, which educational institutions in Thailand have accepted the policy of the country. Each educational institution in Thailand therefore focuses on developing the potential of personnel in the country. To support the management of engineering teaching to suit the world's advanced technology in robotics and automation systems. And at the same time, educational institutions in Thailand have supported national and international skills competitions in robotics and automation. Every year, teachers and students are interested and present their works in robotics and automation contests. Researchers and students from various educational institutions are interested in designed and building robots to participate in such competitions as shown in Fig 1. [2] But to control, the robot must use the remote control to move or control the robotic arm or other devices that are installed on the robot to perform the mission. The limitation of the Joystick PlayStation PS2 wireless for Arduino remote control is popular among students studied robotics engineering. In which the use of the joystick requires many commands channels to be selected, for example: forward commands, reverse commands, left turn commands, right turn commands, commands to raise the front wheel arm up, commands to lift the front wheel arm down, commands to raise the rear wheel arm up, commands to lift the rear wheel arm down, commands to turn the robot arm to the right, commands to rotate the robot arm to the left, commands grabs the mechanical arm forward, commands grabs the mechanical arm backwards, commands to adjust the camera angle to the right, commands to adjust the camera angle to the left, commands to adjust the camera angle up and commands to adjust the camera angle down. There are a total of 18 orders as shown in Fig 2.[1]

From the problem of limitations of remote control robots, the researcher has an idea to control robots by speech commands to be more modern and comfortable. By application of automatic speech recognition: (ASR)[1-20] due to continuous speech commands, [3] therefore selecting speech recognition speaker verification. To help solve the problem of controlling the robot, to reduce the switch button to control the robot. In which the robot operator is not familiar with or skilled in controlling the robot, it can interact with the computer system. The use of speech recognition technology is a system that converts commands into data signals that the computer system can be used to process further results.[3] Therefore, to study the technique for the feature extraction of Thai speech.[4] The researcher chose a technique to find the mean of the magnitude value at the format frequency and among the samples used in the research were males aged about 18-25 years, but each person recorded 5 sounds. That is, forward, backward, turn right, turn left and stop first, which can design speech extraction techniques that are suitable for different sounds or accents in each dialect.[5]

Fig.1. A rescue robot created by students to study and research robotics.[2]

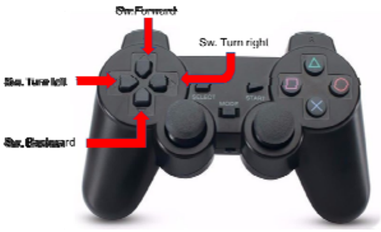


Fig.2. Application of speech commands to replace 4 push button switches.[2]

Theories and principles

A. Digital speech signal

The sound wave transforms that electrical wave used the microphone to receive the speech signal and the volume of the speech signal as the level of the electrical potential that the unit is in the speech. The changed electrical signals into sound waves bypass the electrical signal through the speech changed the speech wave to then the electrical wave is called the analog electrical signal. The analog speech system is the speech system that is used in the past and present as the speech system is changed to a digital system. Due to the analog sound system, there are disadvantages high noise there is a distortion of the speech signal from the original. The digital speech system has no noise, no distortion of the speech signal in the system. [6]

B. Interval division and window functions (Segmentation Windowing)

A speech signal is the time-varying signal. The speech signal analysis for one speech that is speech recognition. [7] Considering with size of the signal from the time, start and end of the speech also change with time. Analyzed the sound signal in the one speech, the signal must be divided into time frames, after which the signal is processed at each time. The time frames to find the feature values contained in each speech signal by filtering the signal at each time interval with a Finite Impulse Response:(FIR), known as a window function and important to attenuating the Gibbs Phenomenon effect [3] the results were taken to find the frequency component occurring at each time interval. The Fast Fourier Transform:(FFT), in which the key frequency components at each interval are used as the speech recognition characteristics with the support vector machine classification.

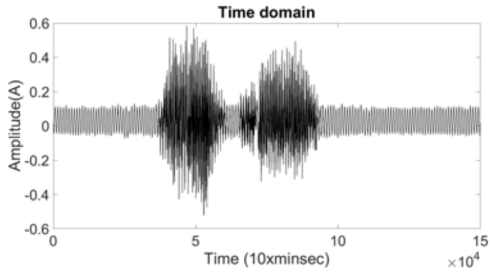
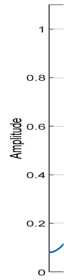


Fig.3. shows the speech signal 1 speech in the time domain at a Segmentation Windowing.



The speech signals are classified as time-varied signals in the addition, the division of each word in the sentence signals analysis of one continuous speech word for this speech recognition. Considered the size of the speech signal from the time the speech starts and ends, there is also a change in time.[7]

In order to analyze one continuous speech signal, it is necessary to divide the signal into time frames and then process the speech signal in each time frame to find feature values, which are included in the speech signal in each word. Known as a window function and is important in Gibbs attenuation Phenomenon [8] was then applied to find the frequency components occurring at each time interval by means of Fast Fourier Transform: (FFT) [9] transformation, herein which the significant frequency component at each interval was used as the idiosyncratic value of discriminant speech recognition the support vector machine: (SVM).[10]

To prove the effectiveness of each window filter affecting the speech discrimination fidelity with SVM and scaled the overlap length during window function shift to determine the convolution between the window function and the speech signal. Each time frame has the different overlap length value. How will this be affected the accuracy of speech classification with SVM.[11] In this study, with a time frame adjustment and the overlapped period of signal filtered with the window function were used in this study. To differentiate the fidelity of the speech recognition system with SVM, here four window function filters were selected as follows [12].

The followed equation generates the coefficients of a Gaussian window as shown in Fig. 4.

(1) $$w(n) = e^{-\frac{1}{2}\left(\alpha \frac{n}{(L-1)/2}\right)^2}$$

(2) $$= e^{-n^2/2\alpha^2}$$

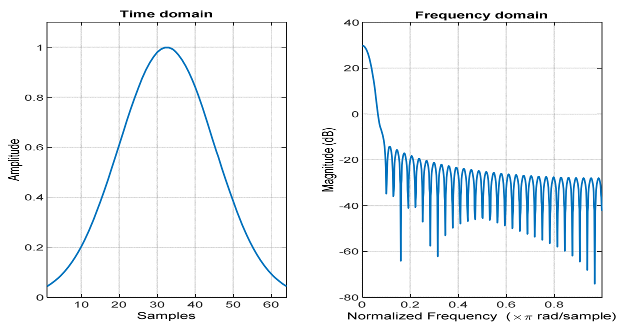


Fig.4. Gaussian window.

The followed equation generates the coefficients of a Hamming window as shown in Fig. 5.

$$w(n)=0.54-0.46\cos\left(2\pi\frac{n}{N}\right), 0\leq n\leq N \tag{3}$$

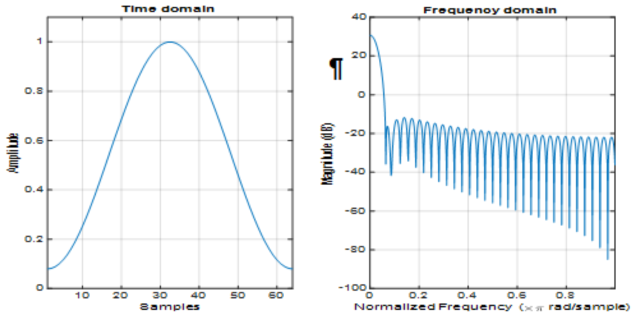


Fig.5. Hamming window function.

The followed equation generates the coefficients of a Hanning window as shown in Fig. 6.

$$w(n)=0.5\left(1-\cos\left(2\pi\frac{n}{N}\right)\right), 0\leq n\leq N \tag{4}$$

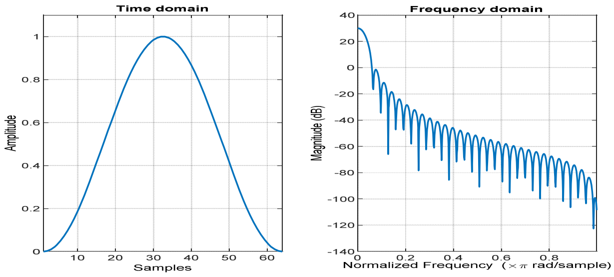


Fig.6. Hanning window function.

The followed equation generates the coefficients of a Kaiser window as shown in Fig. 7.

$$w(n)=\frac{I_0\left(\beta\sqrt{1-\left(\frac{n-N/2}{N/2}\right)^2}\right)}{I_0(\beta)}, 0\leq n\leq N \tag{5}$$

Where N is the length of the window function and β is the sidelobe attenuation of the window's Fourier transform.

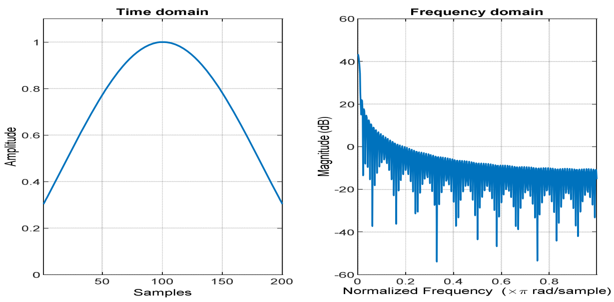


Fig.7. Kaiser window function.

C. Feature Values

From executed the speech transcoded with window functions, the next step is to find the feature extraction of speech commands by format frequency techniques. It is the mean of the magnitude values at the F1, F2 and F3 format frequencies, respectively.

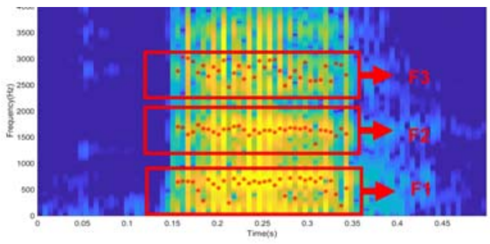


Fig.8. F1,F2 and F3 (frame=250,overlap=0%).

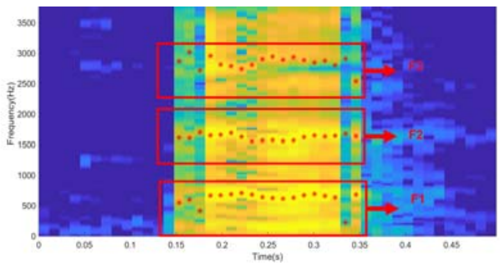


Fig.9. F1,F2 and F3 (frame=500,overlap=0%).

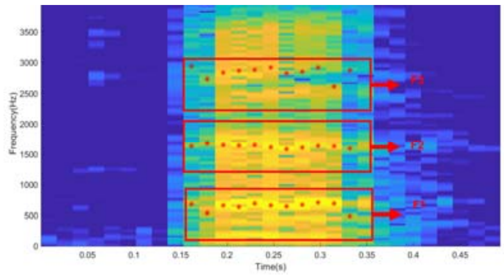


Fig.10. F1,F2 and F3 used function Gaussian window (frame=750,overlap=0%).

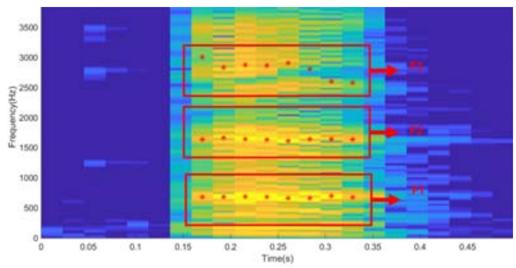


Fig.11. F1,F2 and F3 (frame=750,overlap=0%).

From Fig. 8-11 that has been presented, it is only the Gaussian window function that are divides the time frame adjustment and the overlapped period. The to find optimum value for feature extraction. The speech commands signal was analyzed to find the mean magnitude at the F1, F2, and F3 of format frequencies, replacing the mean of the three frequencies with and respectively.

D. Classification with Support Vector Machine:(SVM)

Support Vector Machine data classification is one of the key algorithms in Machine Learning. [13] In the Supervised Learning model, data types can be distinguished by specified a margin. The maximum margin between the boundary points of two or more different data samples. The SVM's mapping is shown in Fig. 12.[4]

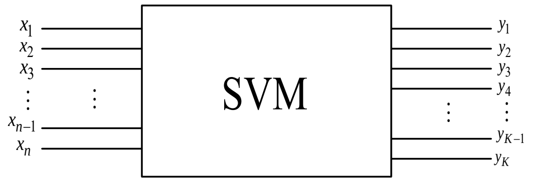


Fig.12. SVM for classification of multiple data sets.[4]

Where are the characteristic values of classification and what is the type of data resulting from classification. For example, the classification of two object samples with SVM can be specified.[11] Any is one of the two sample data as shown in Fig. 13.

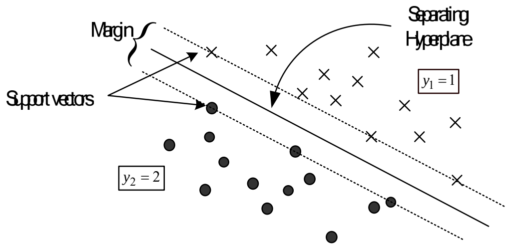


Fig.13. Classification of two data types with SVM [4].

The support vectors are vectors that are drawn through critical data points near the separated hyperplane. For the classification of clearly differentiated datasets, a linear separation plane can be used. can line, however, if the dataset is not very distinct, It is necessary to create a non-linear separation plane. By converted into the original data set to space new by the mapped method $\Phi : \mathrm{R}^n \rightarrow \mathrm{R}^N$ with kernel functions, it measures the distance or similarity between pairs in a data set, where N is the number of original eigenvalues and N is the number of eigenvalues in the new space. The kernel functions come in many forms and have different properties.

(6) $$K_{gaussian}\left(x^{(i)}, x^{(j)}\right) = \exp\left(-\frac{\left\|x^{(i)} - x^{(j)}\right\|^2}{2\alpha^2}\right)$$

In order to find the maximum margin between different datasets, the most appropriate parameter θ must be calculated from the evaluation function [3] as follows:

(7) $$J(\theta) = C\sum_{i=1}^{m} y^{(i)} \cos t_1\left(\theta^T x^{(i)}\right) + \left(1 - y^{(i)}\right)\cos t_0\left(\theta^T x^{(i)}\right) + \frac{1}{2}\sum_{j=1}^{n}\Theta_j^2$$

(8) $$\cos t_0\left(\theta^T x^{(i)}\right) = \max\left(0, k\left(1 + \theta^T x^{(i)}\right)\right)$$
$$\cos t_1\left(\theta^T x^{(i)}\right) = \max\left(0, k\left(1 - \theta^T x^{(i)}\right)\right)$$

By finding the minimum error value from the evaluation function $j(\theta)$ obtained by adjusting the parameters θ each value by calculating the value min with $j(\theta)$ computer program to find the evaluation function $j(\theta)$ the minimum value of the parameter θ most suitable for this research application SVM [4-17]to distinguish Thai language speech signals in the sample group with a limited number of classifications of into different types therefore, several classifications must be used (Multi-Class Classification) set to y represented the result set of the Thai language speech commands $y = \{1, 2, 3, \ldots, K\}$ here, in this set of Thai language speech commands, there are equal numbers of that are interested in classified them. Where K speech and used a sequence of numbers to represent the type of speech command for each speech therefore, it is necessary to provide a set of data for teaching SVM to recognize the characteristics of each of those. Here we use a method to classified many types of objects called one approach compares to all (1-vs-ALL) therefore provides a tutorial for all the SVM. Where K speech signal set, algorithm SVM will give the result of classification as $y = i$ the i this is the possible value $i = 1, 2, 3, \ldots, K$ which is the sequence of numbers representing each group of , including all K type, so the SVM classification parameter for (1-vs-ALL) classification can be found as.

Speech recognition algorithm

This research applied the speech recognition algorithm with SVM algorithm to compare the performance of various window functions. The speech signal collected as a sample dataset as a speech in Thai Includes training set data is given to classifiers obtained from SVM algorithm[12] and one test data set each. The speech sounds in Thai can be divided according to their pronunciation into 3 types, namely, single. Compound and excess [13] in this paper presents a test with a single speech can also be divided into two sub-, namely in Thai language. As shown in details along with international symbols as shown in Table 1.

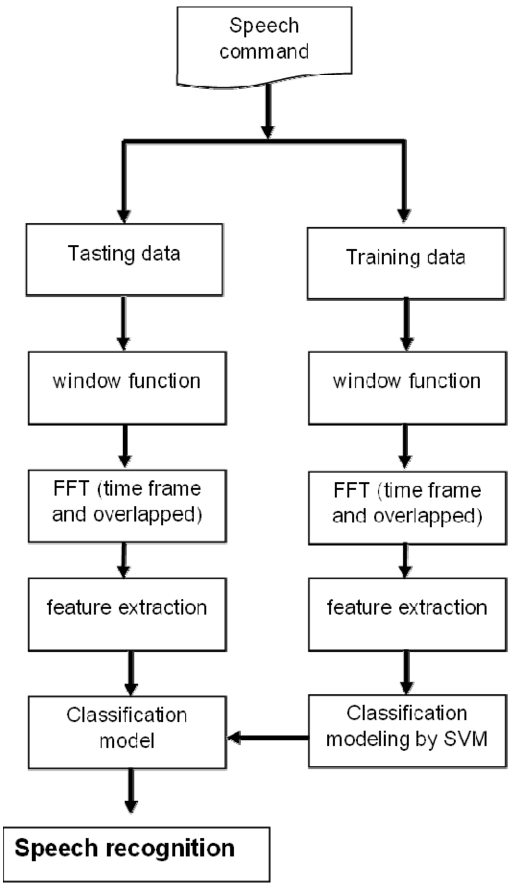


Fig.14. Speech Recognition with SVM.

Table 1. It is a comparison of the meaning in Thai language with international languages.

Speech commands (Thai language)	Speech commands (universal language)
เดินหน้า	Forward
ถอยหลัง	Backward
เลียวซ้าย	Turn left
เลียวขวา	Turn right
หยุดก่อน	Stop first

There are 5 continuous speech commands with Thai languages in this study. The continuous speech commands for each speech is recorded in Table 1. The undisturbed speech recorded with male speech is a speech file (.wav) wave frequency values. A sampled frequency of 44,100 Hz, 16-bit size, was recorded as a data file of 100 individual Thai languages speech and divided into trained data sets for 50 speech and divided into tested data sets for 50 speech. The labeled information will be available in the speech recognition test.[18-20]

Experimental results

The form ASR by continuous speech commands recognition with Thai language as a speaker verification model as shown in Fig. 14. And the test by inputted a continuous speech commands signal 1 speech at a time. Then, each time frames are filtered with a FIR filter, and four window functions are selected. The results obtained by filtered the signal with the four window filters were fed into frequency processed with the FFT transformation, and the characteristic vector of each word was determined as the average of the format frequencies at F1, F2, and F3. Respectively, this is an SVM classification method using eigenvalues for classification as the key frequencies of each sound word. Is the mean value of the magnitude at the format frequency, showing the result of the sampled word identification on the characteristic vector space, as shown in Fig. 15-18.

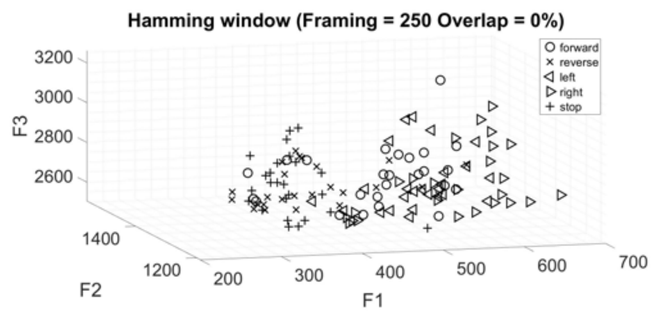


Fig.15. A single speech signal data set on a characteristic vector space.

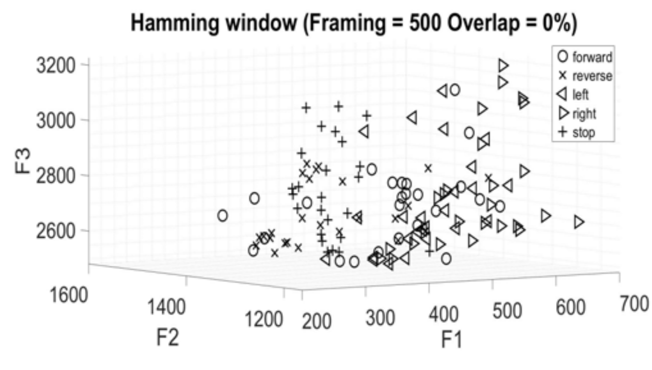


Fig.16. A single speech signal data set on a characteristic vector space.

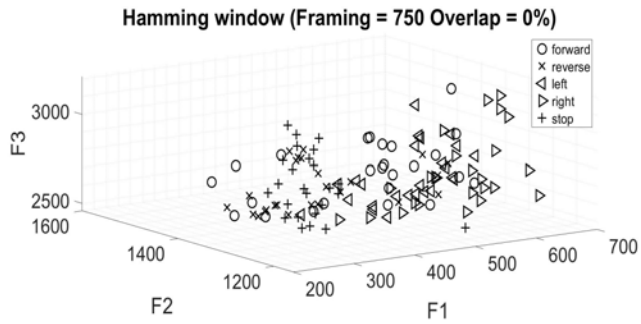


Fig.17. A single speech signal data set on a characteristic vector space.

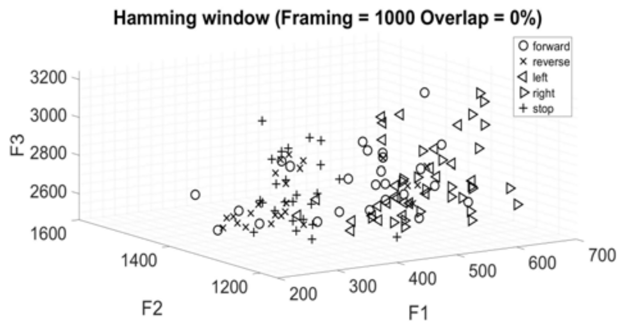


Fig.18. A single speech signal data set on a characteristic vector space.

The next table compares the accuracy of speech recognition by continuous speech commands recognition with Thai language by applyed Gaussian window function by tested each time, setted time frames equal to 250,500,750 and 1000 points, and adjusted the percentage of overlaps in each time frames equal to 0%, 25%, 50% and 75% shown in Table 2 shows the comparison of the accuracy overall of a ASR by continuous speech commands recognition with Thai language model.

Table 2. Comparison of Gaussian window function performance for Thai language recognition.

Speech Recognition Accuracy.				
Gaussian window function	time frame 250	time frame 500	time frame 750	time frame 1000
Forward	80.395%	81.36%	81.33%	81.86%
Backward	83.69%	83.67%	83.58%	83.81%
Turn left	84.00%	83.14%	82.14%	81.94%
Turn right	85.72%	85.61%	85.50%	85.39%
Stop first	86.78%	87.06%	86.81%	86.47%
Overall performance	84.12%	84.17%	83.87%	83.89%

From Table 2, the accuracy of speech recognition by continuous speech commands recognition with Thai language was compared with the percentage of time frame equal to 500 points the test result of the "Stop first" speech commands has an accuracy maximum of 87.06%.

The next table compares the accuracy of speech recognition by continuous speech commands recognition with Thai language by applied Hamming window function by tested each time, settled time frames equal to 250,500,750 and 1000 points, and adjusted the percentage of overlaps in each time frames equal to 0%, 25%, 50% and 75% shown in Table 3 shows the comparison of the accuracy overall of an ASR by continuous speech commands recognition with Thai language model.

Table 3. Comparison of Hamming window function performance for Thai language recognition.

Speech Recognition Accuracy.				
Hamming window function	time frame 250	time frame 500	time frame 750	time frame 1000
Forward	81.08%	81.50%	81.42%	81.75%
Backward	83.67%	83.83%	83.44%	83.64%
Turn left	84.17%	82.69%	82.11%	81.94%
Turn right	85.78%	85.39%	85.36%	85.53%
Stop first	86.92%	86.92%	86.61%	86.56%
Overall performance	84.32%	84.07%	83.79%	83.88%

From Table 3, the accuracy of speech recognition by continuous speech commands recognition with Thai language was compared with the percentage of time frame equal to 250 points the test result of the "Stop first" speech commands has an accuracy maximum of 86.92%.

The next table compares the accuracy of speech recognition by continuous speech commands recognition with Thai language by applied Handing window function by tested each time, settled time frames equal to 250,500,750 and 1000 points, and adjusted the percentage of overlaps in each time frames equal to 0%, 25%, 50% and 75% shown in Table 2 shows the comparison of the accuracy overall of an ASR by continuous speech commands recognition with Thai language model.

Table 4. Comparison of Hanning window function performance for Thai language recognition.

Speech Recognition Accuracy.				
Hanning Window function	time frame 250	time frame 500	time frame 750	time frame 1000
Forward	82.19%	81.97%	81.58%	81.92%
Backward	83.81%	83.72%	83.42%	83.67%
Turn left	84.06%	83.14%	82.19%	81.67%
Turn right	85.69%	85.56%	85.28%	85.36%
Stop first	86.36%	87.03%	86.89%	86.56%
Overall performance	84.42%	84.28%	83.87%	83.84%

From Table 4, the accuracy of speech recognition by continuous speech commands recognition with Thai language was compared with the percentage of time frame equal to 250 points the test result of the "Stop first" speech commands has an accuracy maximum of 86.36%.

The next table compares the accuracy of speech recognition by continuous speech commands recognition with Thai language by applied Kaiser window function by tested each time, settled time frames equal to 250,500,750 and 1000 points, and adjusted the percentage of overlaps in each time frames equal to 0%, 25%, 50% and 75% shown in Table 2 shows the comparison of the accuracy overall of an ASR by continuous speech commands recognition with Thai language model.

Table 5. Comparison of Kaiser window function performance for Thai languagei recognition.

Speech Recognition Accuracy.				
Kaiser Window function	time frame 250	time frame 500	time frame 750	time frame 1000
Forward	80.33%	81.00%	81.78%	80.11%
Backward	82.78%	82.67%	83.00%	83.56%
Turn left	84.00%	82.44%	82.00%	81.67%
Turn right	85.67%	85.33%	85.33%	85.33%
Stop first	86.00%	84.22%	84.44%	86.11%
Overall performance	83.76%	83.13%	83.31%	83.36%

From Table 5, the accuracy of speech recognition by continuous speech commands recognition with Thai language was compared with the percentage of time frame equal to 250 points the test result of the "Stop first" speech commands has an accuracy maximum of 86%.

Table 6. Comparison of window function performance for Thai language recognition.

Speech Recognition Accuracy.				
Window function	Gaussian window function	Hamming window function	Hanning Window function	Kaiser Window function
Forward	81.24%	81.44%	81.92%	80.81%
Backward	83.69%	83.65%	83.65%	83.00%
Turn left	82.81%	82.73%	82.76%	82.53%
Turn right	85.56%	85.51%	85.47%	85.42%
Stop first	86.78%	86.75%	86.71%	85.19%
Overall performance	84.02%	84.02%	84.10%	83.39%

From Table 6, the accuracy of speech recognition by continuous speech commands recognition with Thai language was compared with the percentage of overall performance the test result of the speech commands has an accuracy maximum of 84.1% of the Hanning window function.

Conclusions

This paper aims to applied automatic speech recognition with continuous speech command recognition for Thai language. When analyzing the signal processing of continuous speech commands and comparing the accuracy of the "forward", "backward", "left turn", "right turn" and "stop first" speech commands with page functions, it was found that the Hanning window function the most effective was 84.10%, and when comparing the methods, the Gaussian window function and the Hamming window function were equal with 84.02%. [1] and this is the Kaiser window function were equal with 83.39%. In this paper, the Hanning window function has a more accurate result than the others by Thai language may be different from other languages, therefore is the accuracy of the Hanning window function test result is the appropriate to use this method in the future research to control other parts of the robot by voice command in conjunction with the original is the two systems in the next step.

***Authors**: Mr. supavit Muangjaroen, E-mail: supavitjack@gmail.com;*
Assoc.Prof.Dr. Sakol Udomsiri, E-mail: sakol.udm@gmail.com, Faculty of Engineering, Department of Electrical Engineering, Pathumwan Institute of Technology, 833 Rama1 Wangmai District, Bangkok, Thailand.

REFERENCES

[1] A. Vishwakarma and M. Soni, "Speaker Recognition System – A Review,"International Journal of Emerging Technologies in Engineering Research (IJETER), 85-89.

[2] J. Nuchrak, K. Fathanom, C. Prakhom, C. Mueantaeng, S. Muangjaroen, "Designed of rescue robot's controller with a board Arduino Mega 2560, "*Proceedings of the 10thConference of Electrical Engineering Network 2018* (EENET 2018),455-458.

[3] S.Muangjaroen and S. Udomsiri,"The study of continuous speech recognition with Thai language using Mel Frequency Cepstral coefficient extraction method,*10th International Science, Social Science, Engineering and Energy Conference, (2019) , pp. 1-12.*

[4] S.Muangjaroen, W.Chaemla, S.Duangsungnern and S.Udomsiri ,The study of speech recognition for the Thai language by using audio extraction techniques with the method Mel frequency cepstral coefficients, *The 42nd Electrical Engineering Conference eecon42*(2019),-301-304.

[5] D. Anggraeni, W.S.M. Sanjaya, M. Munawwaroh, M. Y.Nurasyidiek,and I. P. Santika,"Control of Robot Arm based on Speech Recognition using Mel-Frequency Cepstrum Coefficients (MFCC) and K-Nearest Neighbors (KNN) Method ," *Intelligent Manufacture, and Industrial Automation (ICAMIMIA*),217-222.
[6] S. K. Shevade, S. S. Keerthi, C. Bhattacharyya, and K. R. K. Murthy, "Improvements to the SMO Algorithm for SVM Regression," *IEEE TRANSACTIONS ON NEURAL NETWORKS*, VOL. 11, NO. 5, (2000),1188-1193.
[7] Y. Li, K. Bontcheva and H. Cunningham, "SVM Based Learning System For Information Extraction," *International Workshop on Deterministic and Statistical Methods in Machine Learning*, (2005), 319–339.
[8] M. Zhang, X. Ai and Y. Hu , "Chinese Text Classification System on Regulatory Information Based on SVM,"*Earth and Environmental Science*, (2019), 1-10.
[9] P. TAO , ZHE. SUN, AND ZHI. SUN, "An Improved Intrusion Detection Algorithm Based on GA and SVM,"*Special section on human-centered smart systems and technologies*, (2018), 13624-13631.
[10] Y Zhang, M. Ni, C Zhang, S Liang, Sg Fang, R Li1 and Z. Tan, "*Research and Application of AdaBoost Algorithm Based on SVM," IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC 2019)*,662-666.
[11] L. Auria and R. A. Moro, "Support Vector Machines (SVM) as a Technique for Solvency Analysis," *DIW Discussion Papers*, (2008), No. 811, 1-16.
[12] A. Mathur and G. M. Foody, "Multiclass and Binary SVM Classification: Implications for Training and Classification Users," *In IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 2, 241-245.
[13] M. E. Mavroforakis and S. Theodoridis, "A geometric approach to Support Vector Machine (SVM) classification," *In IEEE Transactions on Neural Networks,* vol. 17, no. 3, 671-682.
[14] Y. Bazi and F. Melgani, "Toward an Optimal SVM Classification System for Hyperspectral Remote Sensing Images," *In IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 11, 3374-3385.
[15] A. Patle and D. S. Chouhan, "SVM kernel functions for classification," *2013 International Conference on Advances in Technology and Engineering (ICATE)*, 2013, pp. 1-9
[16] M. Pal and G. M. Foody, "Feature Selection for Classification of Hyperspectral Data by SVM," in IEEE Transactions on Geoscience and Remote Sensing, vol. 48, no. 5, 2297-2307.
[17] J. Augustine, B. Bellermin, J. K Martin, Deepthy J, "Emotion Recognition in Speech Using MFCC with SVM, DSVM and Auto-encoder" *International Journal for Research in Applied Science & Engineering Technology (IJRASET*), (2021),1021-1026.
[18] B. Kollmeier and B.T. Meyer, "Exploring Auditory-Inspired Acoustic Features for Room Acoustic Parameter Estimation from Monaural Speech," *IEEE/ACM Transactions on AUDIO, Speech, and Language Processing*, VOL. 26, NO. 10, 2018,1809-1820.
[19] M. Walid, S. Bousselmi, K. Dabbabi and A. Cherif, "Real-Time Implementation of Isolated-Word Speech Recognition System on Raspberry Pi 3 Using WAT-MFCC," *IJCSNS International Journal of Computer Science and Network Security,* VOL.19 No.3, (2019), 42-49.
[20] M. AUGUSTYN, "Recognition of handwritten digits on the basis of signals from a 3-axis accelerometer using the DTW method taking into account various detection criteria," Przeglad Elektrotechniczny,VOL. 11, (2022), 177-180.