Thai Voice-Controlled Analysis for Car Parking Assistance in System-on-Chip Architecture

Sethakarn Prongnuch[1,*], Suchada Sitjongsataporn[2]

[1]Department of Computer Engineering, Suan Sunandha Rajabhat University, Bangkok, Thailand

[2]Department of Electronic Engineering, Mahanakorn University of Technology, Bangkok, Thailand

Abstract

This paper introduces an analysis of Thai speech recognition for controlled car parking assist in the system-on-chip architecture. The objective is to investigate the male and female voice command signals, including Thai and English words, issued by the native Thai users. Hardware and software co-design by the Xilinx VIVADO are designed on an ARM multicore processor and a reconfigurable system on a ZYBO board. The experiments for Thai and English word recognition are conducted by using the Mel-frequency cepstral coefficient approach and presented in the form of spectrograms. The comparison of a voice command via Bluetooth and a reference command stored on an SD card and the ZYBO embedded board on a miniature electric vehicle is verified with the Pearson's correlation coefficient (PCC). The experimental results show the accuracies of the received Thai/English, male/female, and indoor/outdoor voice commands as compared with the reference voice commands in the noisy surroundings. Hence, our system can support Thai/English and male/female voice commands to perform a set of actions for maneuvering a car by the PCC.

Keywords: speech recognition, voice-controlled parking assist, reconfigurable embedded systems, system-on-chip

1. Introduction

With the increasing number of automobiles in crowded cities, the number of parking lot accidents that drivers troubles while parking on the rooftop under possible space restrictions have also increased. Traffic accidents during parking include hitting, denting, and being scratched. Drivers often encounter difficulties while maneuvering a vehicle adjacent to a narrow space and the roof of a building. For instance, a parking deck located on a roof may pose some dangers because of the short rooftop barrier as shown in Fig. 1(a). In 2019, car crash news [1-3] from Thailand described the case of a vehicle plunging from a parking building in Thailand as the driver applied the wrong gear [2]. The Indianapolis Star news also reported incidents of cars plunging from parking garages and that 46 accidents across the United States have been recorded in the last two decades[1, 3]. Many devastating incidents have been identified to be cased by elderly drivers who accidentally stepped on the accelerator instead of the brakes. Hence, to ensure safety, using the parking assist outside a car may be preferable. Fig. 1(b) shows a driver worrys about maneuvering a vehicle because of the narrow space restrictions. Using a voice-controlled assist outside the car would be the safer option because the driver can see obstacles on the road.

A parking assist is a voice-controlled vehicle maneuvering system functioning at low velocities. The system can detect the parking space and ensure safety while parking. However, the existing parking assist systems are mostly used inside a car,

---

* Corresponding author. E-mail address: sethakarn.pr@ssru.ac.th

Tel.: +662-1-601435; Fax: +662-1-601440

thereby limiting effectivity. Hence, to improve the parking skills of the driver, Yukawa [4] proposed using the reverse parking direction through the auditory assistance approach. Several studies have also proposed parking assist systems, such as an autonomous path planning in the narrow environments [5], a surround-view parking system [6], and a voice-controlled embedded system for parking a prototype car [7-8].

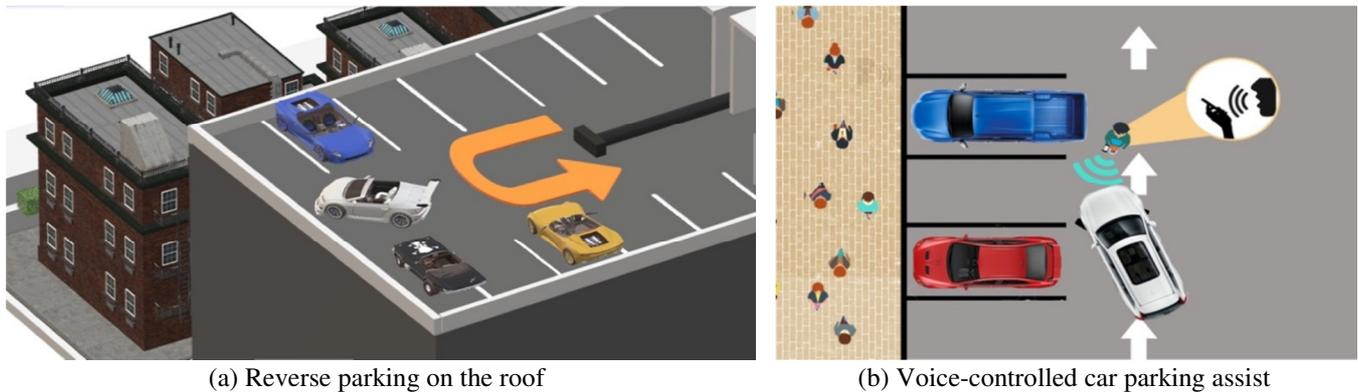(a) Reverse parking on the roof (b) Voice-controlled car parking assist

Fig. 1 Voice-controlled car parking assist

The need to minimize drive disruption in the operation of the car led by Honda [9] use analog control to assist the driver in operating the car more confidently. This usability will be adopted for aging owners. A voice-controlled interface will be built-in to encourage the interaction of humans and smart applications that understand voice commands boosted by the use of speech recognition.

To ensure their safety and well-being, drivers can employ the car parking assist while maneuvering the car from the outside so that they would have an unobstructed view around the car. The idea of the parking assist is an exterior usage from the safety control. In this paper, we propose a voice-controlled parking system for an automatic car parking assist that comprises a communication system adapted to receive a set of voice commands. We also evaluate Thai speech recognition for voice-controlled car parking assist in system-on-chip architecture on a prototype electric vehicle.

Previous studies have proposed various approaches regarding driver assistance. In [10], the authors proposed a parking-space detection system for occupancy by using Bluetooth communication and Arduino-based sensor devices. Other studies have also presented advances in voice-based recognition. In [11], a bias analysis of recognition phrases that utilized the accent of the Southeast region for smartphone interaction with a voice personal assistant was performed. The preliminary results were found to be skewed towards the voices of individuals. In [12], a Raspberry Pi-based voice-controlled car using the Baidu speech recognition technology and tested with remote control interaction was presented. In [13], the interoperability of speech recognition in Romanian for voice-controlled appliances in smart buildings was explored. In [14], a voice-based unmanned aerial vehicle approach using the MFCC analysis of semantic identification was introduced. The experiments in previous studies have shown that the MFCC was capable of voice command recognition.

The paper is organized as follows. Section 2 describes speech recognition shortly, and Section 3 outlines briefly the Pearson's correlation coefficient. Section 4 introduces the proposed voice-controlled car parking assist. Section 5 presents the experimental results and performance. Section 6 concludes the paper.

2. Speech Recognition

Speech signal [15] refers to a sequence of encoding message symbols as shown in Fig. 2. The speech recognition system begins by transforming the speech waveform into the sequence of a discrete vector. Then, a recognizer is used as a mapping tool for the discrete speech vector, which is underlined by a sequence of symbols. Feature extraction is the method of reducing the number of features from the measured input signal, which leads to the identification of the speaker.
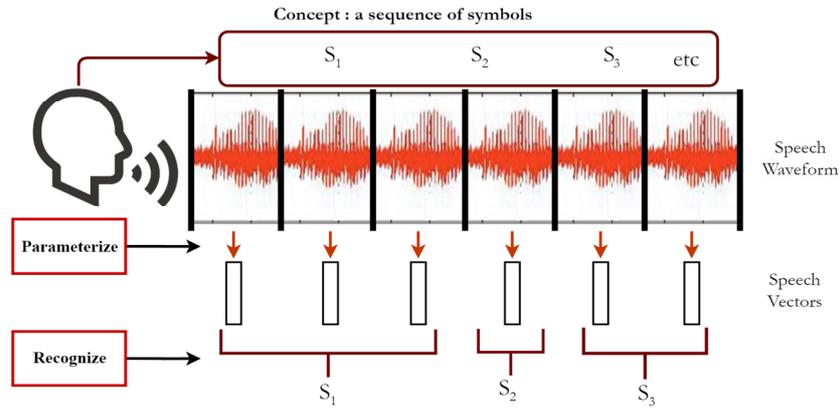
Fig. 2 Speech recognition system

2.1.   Feature extraction

Fig. 3 illustrates the feature extraction process. An acoustic signal is converted to digitize the features. Then, the pre-emphasis procedure is implemented to enhance the signal to noise ratio in the higher frequencies. The window framing and fast Fourier transform (FFT) are performed to avoid discontinuities. The Mel-frequency cepstrum coefficients (MFCC) technique [16] is used to extract the features as an envelope of the spectrum in the cepstral domain. The transformed acoustic signals are obtained based on their features in the time and frequency domains. In relation to establishing the feature vector [17], the spectrogram is modified through a short-time Fourier transform (STFT) to demonstrate the features of the acoustic signal in the time-frequency domain applied through the MFCC approach.
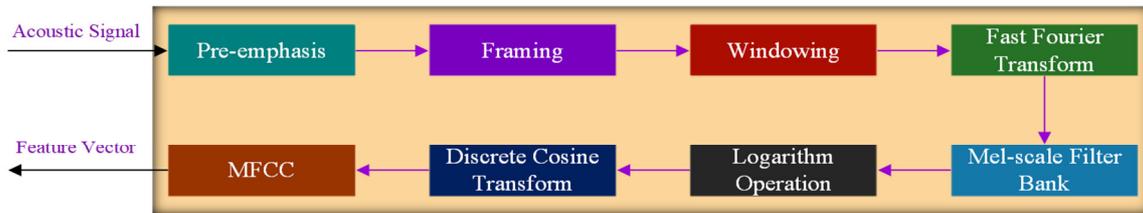

Fig. 3 Feature extraction process using the Mel-frequency cepstral coefficient approach

2.2.   Short-time fourier transform and spectrogram

The STFT of a signal is represented as [18]:

$$STFT\{x(k)\}(m,\phi) \equiv X(m,\phi) \equiv \sum_{n=-\infty}^{\infty} x(k)w(k-w)e^{-j\phi n} \tag{1}$$

where $x(k)$ and $w(k)$ are the input signal and window, respectively. The spectrogram is the square of the STFT magnitude and is calculated as:

$$Spectogram\{x(k)\}(m,\phi) \equiv \left| X(m,\phi) \right|^2 \tag{2}$$

2.3.   Mel-frequency cepstral coefficients

The MFCC is the real cepstrum of a windowed short-time signal developed from the FFT [17]. Its difference from the real cepstrum is assumed to be the characteristic of the auditory system by using a nonlinear frequency scale.

For the pre-emphasis process, the first order high-pass filter is performed as in [18]:

$$y(k) = x(k) - \gamma x(n-1), \ 0.9 \le \gamma \le 1.0 \tag{3}$$

where $x(k)$, $w(k)$, and γ are the input signal, output signal, and filter coefficient, respectively. The discrete Fourier transform input signal can be computed as [16]:

$$X_{DFT}(k) = \sum_{k=0}^{K-1} x(k)e^{\frac{-j2\pi nk}{N}}, \ 0 \le n \le N \tag{4}$$

The *N*-point Hamming window function can be expressed as:

$$w(k) = \alpha - \beta\cos(\frac{2\pi n}{N-1}), \ 0 \le n \le N-1 \tag{5}$$

where $\alpha = 0.54$ and $\beta = 0.46$. Mel-scale filter bank is modified by binding the frequencies onto the Mel-scale in the frequency domain as [17]:

$$Mel(\omega) = 1127\ln(1+\frac{\omega}{700}) \tag{6}$$

We define a filter bank with *M* filters as:

$$H_m(n) = \begin{cases} 0, \ n < \omega(m-1) \\ \dfrac{n-\omega(m-1)}{\omega(m)-\omega(m-1)}, \ \omega(m-1) < n < \omega(m) \\ \dfrac{\omega(m-1)-n}{\omega(m+1)-\omega(m)}, \ \omega(m) \le n \le \omega(m+1) \\ 0, \ n > \omega(m+1) \end{cases} \tag{7}$$

where $1 \le m \le M$, *M* is the total number of filters in the frequency band, and $\omega(\cdot)$ is the Mel-scale frequency. The log-energy at the output of the filter bank may be expressed as:

$$s(m) = \ln(\sum_{k=0}^{N-1}\left|X_{DFT}(k)\right|^2), \ 0 \le n < M \tag{8}$$

The discrete cosine transform (DCT) is operated to achieve the nth cepstral coefficients c_r as:

$$c_r = \sum_{m=0}^{M-1} s(m)\cos(\frac{\pi r(m-0.5)}{M}), \ r = 1, \ 2, \ldots, \ R \tag{9}$$

where $s(m)$ is the logarithm of the Mel spectrum, and *R* is the total number of cepstral coefficients. The speech recognition procedure typically uses the 13-cepstrum coefficients.

3. Pearson's Correlation Coefficient

Following [18], we consider the signal in the time domain as:

$$\mathbf{y}(t) = \mathbf{x}(t) + \mathbf{\eta}(t) \tag{10}$$

where $\mathbf{y}(t)$, $\mathbf{x}(t)$, and $\mathbf{\eta}(t)$ are the signal from the microphone, reference, and noise signals, respectively. Thus, the covariance matrix of the signal from the microphone can be expressed as:

$$\mathbb{R}_y \simeq E\left\{\mathbf{y}^T(t)\bullet\mathbf{y}(t)\right\} \simeq E\left\{\mathbf{x}^T(t)\bullet\mathbf{x}(t)\right\} + E\left\{\mathbf{\eta}^T(t)\bullet\mathbf{\eta}(t)\right\} \simeq \mathbb{R}_x + \mathbb{R}_\eta \tag{11}$$

where $\mathbb{R}_x$ and $\mathbb{R}_\eta$ are the covariance matrices of $\mathbf{x}(t)$ and $\mathbf{\eta}(t)$, respectively, the operator $E\{\cdot\}$ is defined as an expectation operator. The input signal to noise ratio can be defined as:

$$iSNR = \frac{\sigma_x^2}{\sigma_\eta^2} \tag{12}$$

where σ_x^2 and σ_n^2 are the variances of $\mathbf{x}(t)$ and $\mathbf{\eta}(t)$ as $\sigma_x^2 = E\{x^2(t)\}$ and $\sigma_x^2 = E\{\eta^2(t)\}$, where $x(t)$ and $\eta(t)$ are the reference and noise signals in the time domain respectively. We focus on how to estimate $x(t)$ from the observations by using PCC on the most interesting signals.

The PCC (p_{xy}) is a measurement of the linear correlation between two random variables, $x(t)$ and $y(t)$, which is defined as in [16]:

$$p_{xy} = \frac{E\{y^T(t) \cdot y(t)\}}{\sigma_x \cdot \sigma_y} \tag{13}$$

where $E\{y^T(t) \cdot y(t)\}$ is the crosscorrelation between $x(t)$ and $y(t)$. The square of Pearson's correlation coefficient (SPCC) is computed as:

$$p_{xy}^2 = \frac{\left[E\{y^T(t) \cdot y(t)\}\right]^2}{\sigma_x^2 \cdot \sigma_y^2} \tag{14}$$

where σ_x^2 and σ_y^2 are the signal variance of $x(t)$ and $y(t)$, respectively, and can be presented as:

$$\sigma_x^2 = E\{x^T(t) \cdot x(t)\} \tag{15}$$

$$\sigma_y^2 = E\{y^T(t) \cdot y(t)\} \tag{16}$$

The important property of SPCC is $0 \leq p_{xy}^2 \leq 1$. The SPCC indicates the linear relationship between two random variables, $x(t)$ and $y(t)$. If p_{xy}^2 is equal to zero, $x(t)$ and $y(t)$ are uncorrelated. The value of p_{xy}^2 is closer to 1, which means $x(t)$ and $y(t)$ have strong correlations.

The PCC (p_{xy}) is reformulated by substituting the estimates of the covariance and variance with the method of a sum over the samples, which can be obtained by:

$$p_{xy} = \frac{n\sum x_i y_i - \sum x_i y_i}{\sqrt{n\sum x_i^2 - \left(\sum x_i\right)^2 \cdot \left[n\sum y_i^2 - \left(\sum y_i\right)^2\right]}} \tag{17}$$

where n is the size of the sample, x_i and y_i are the individual sample points at index i.

4. Proposed Voice-Controlled Car Parking Assist

In this section, we present the voice-controlled car parking assist system. The system consists of hardware and software co-design with a systems-on-chip architecture [19]. The system design is composed of four parts: a ZYBO board, an Arduino UNO board, a Bluetooth module, and the sensors.

4.1. Hardware co-design system

Fig. 4 is the overview of the hardware co-design system, which comprises six hardware installations for exterior usage: a Bluetooth device, a motor device, a buzzer, a servo motor managed through an Arduino UNO, and two ultrasonic sensors. The Bluetooth device collects the input voice command data and then transmits it to other devices.
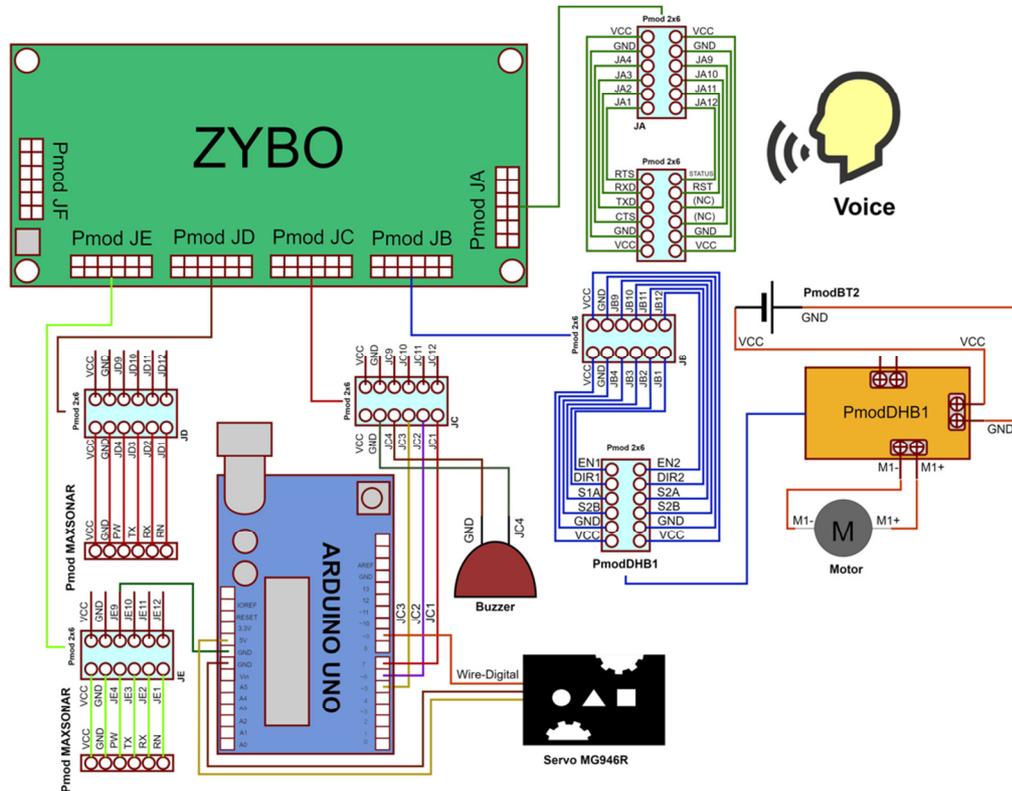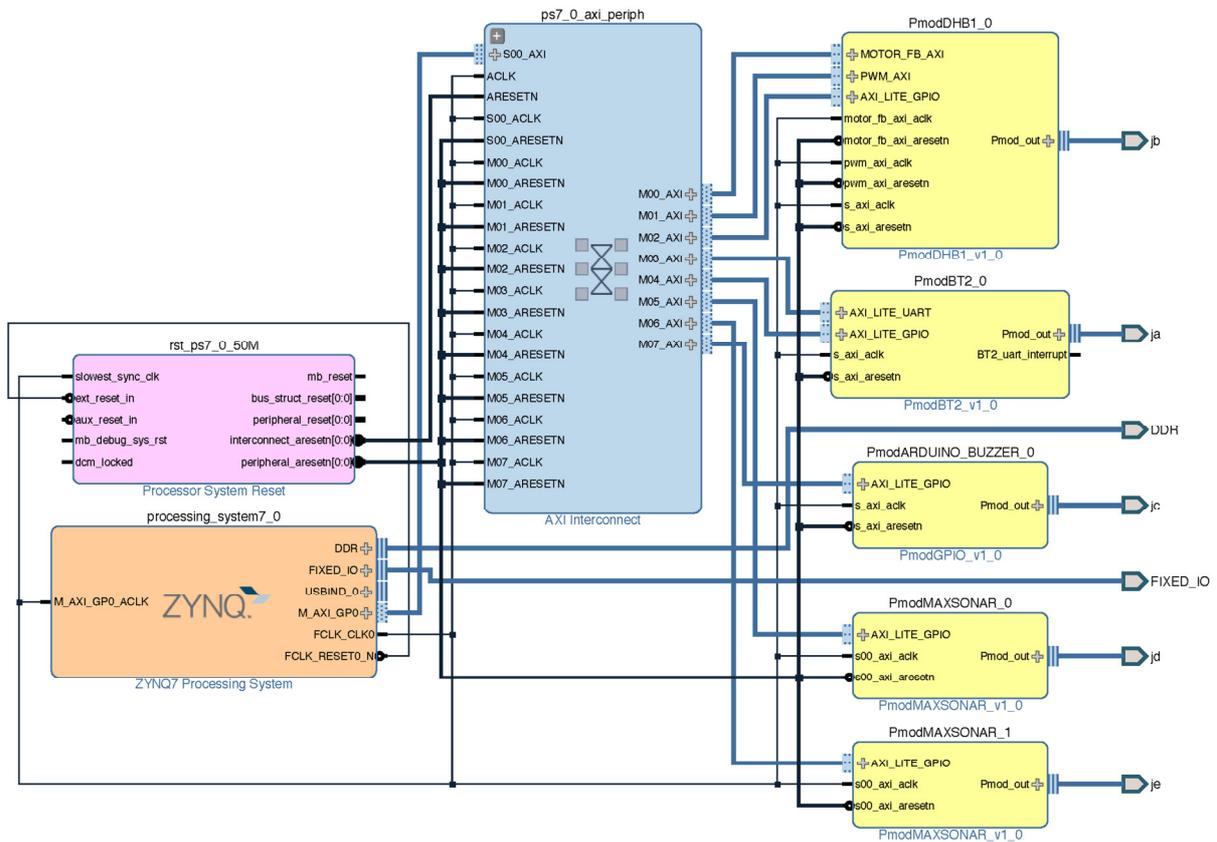
Fig. 4 Overview of the system

Fig. 5 Xilinx VIVADO design architecture

We apply the Xilinx VIVADO [19] shown in Fig. 5, which includes the X Software Development Kit (XDSX) software design tool on the reconfigurable areas inside the ZYBO board, ZYNQ7 system for the interior usage combined with an ARM multicore and a reset system, Advanced eXtensible Interface (AXI) Interconnect, fixed I/O, DDR RAM, and six hardware devices.

4.2. Software co-design system

The flow chart of the proposed system is shown in Fig. 6. Initially, the headset Bluetooth microphone receives English or Thai voice command data, which are then sent to an application on a smartphone for pre-processing. Then, the ZYBO FPGA embedded board checks the received voice command and compares the command with the reference stored in the micro SD card. The ARM main processor on the ZYBO board reads the voice input and compares it with the original voice by using the PCC function defined in Eq. (17). Finally, the results are compared to match the control commands to regulate the peripherals, which include the Arduino board, the DC motor, and the buzzer. The final step involves the car moving by the voice-controlled assist.
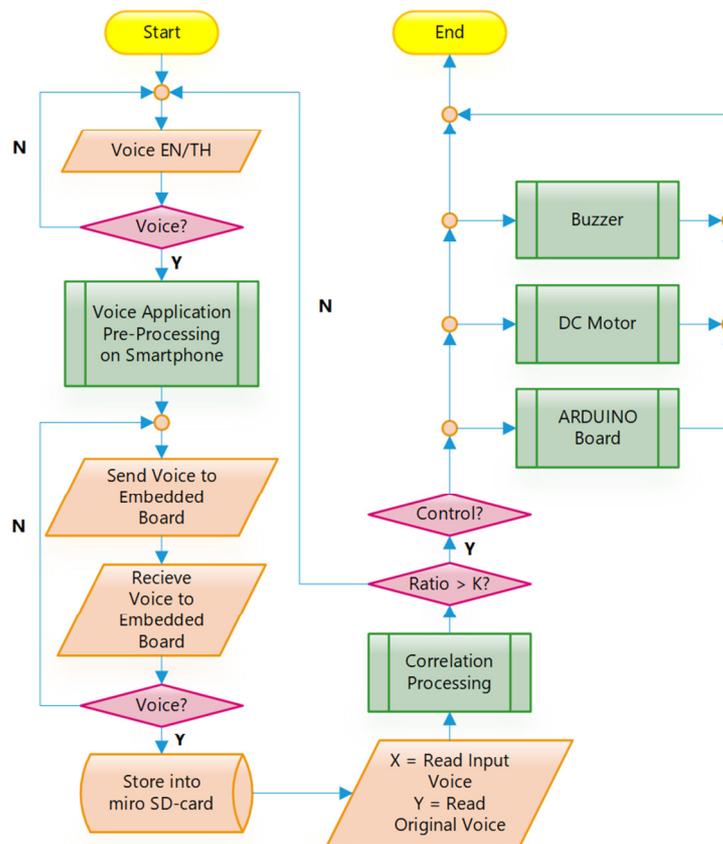

Fig. 6 Flow chart of the proposed system

5. Experimental Results and Performance

This section expounds on the experimental results and the performance evaluation for the voice-controlled car parking assist.

5.1. Simulation results

This section provides the analysis results of Thai female and male voice signals in the form of the spectrograms by using the MFCC approach. The voice signals features are normally operated in the time and frequency domains. Hence, to form the feature vector [20], a spectrogram using STFT to form the acoustic features in the time-frequency domain is provided, and the MFCC approach is utilized following [21].

Table 1 The parameters of feature extraction

Parameter	Value in unit
Length of Thai female and male voice data per instance	1 second
Sampling frequency (f_s)	48 kHz/44 kHz
FFT size per frame	512
Frame length (T_f)	25 milliseconds
Frame shift (ΔT_f)	10 milliseconds
Number of filter banks	20
Number of cepstral coefficients	13

The experiments for Thai female and male speech recognition are studied to establish the investigation process. Table 1 shows the parameters used for the feature extraction that were obtained by using computer simulations. Figs. 7-12 depict the spectrograms of the various Thai female and male voices tested and compared with the reference voice data. The words consist of six Thai and English words of the basic commands for the proposed voice-controlled car parking assist, including "𝒫𝒾"/"GO", "𝒥𝒽𝒶𝓎"/"BACK", "𝒦𝓀𝓌𝒶"/"RIGHT", "𝒮𝒶𝓎"/"LEFT", "𝒫𝒽𝓋𝓂"/"READY", and "𝐻𝓎𝓊𝒹"/"STOP".

(a) Thai male outdoor voice

(b) Thai male original voice

(c) Thai female outdoor voice

(d) Thai female original voice

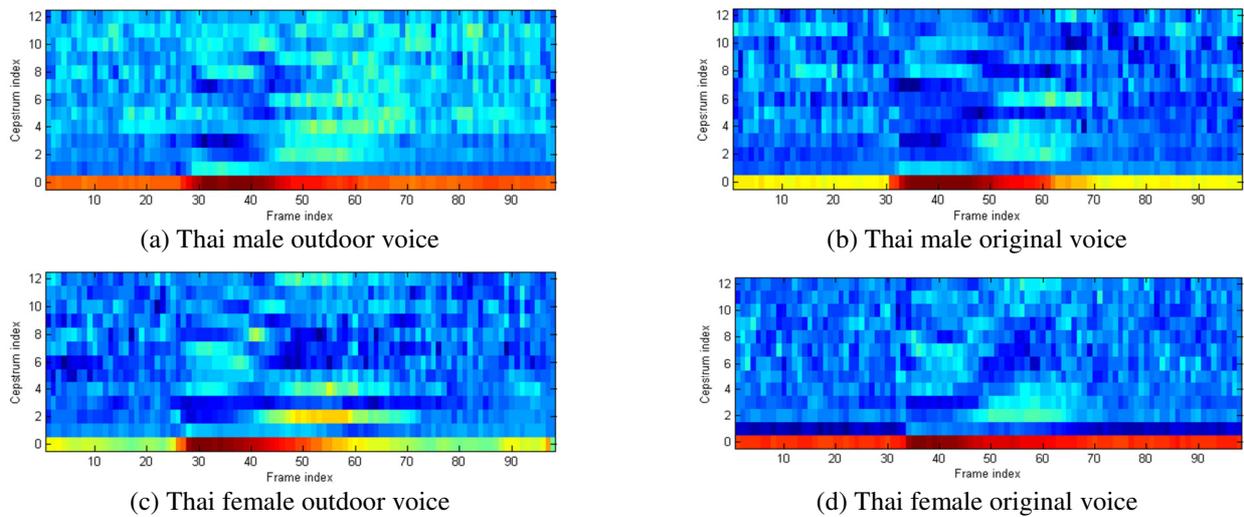Fig. 7 Mel-Frequency Cepstral Coefficients computed from male and female speech signals shown in the form of spectrograms of "𝒫𝒾" or "GO" in Thai

(a) Thai male outdoor voice

(b) Thai male original voice

(c) Thai female outdoor voice

(d) Thai female original voice

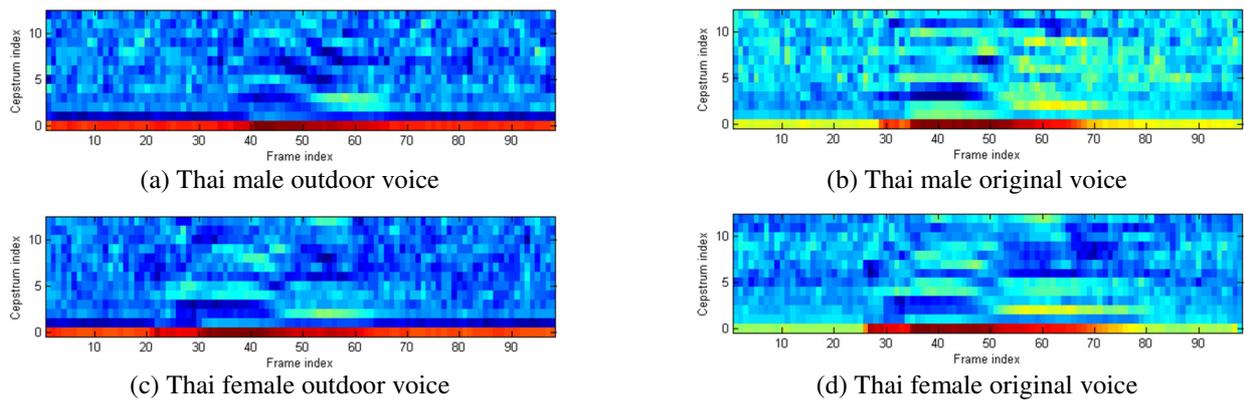Fig. 8 Mel-Frequency Cepstral Coefficients computed from male and female speech signals shown in the form of spectrograms of "𝒥𝒽𝒶𝓎" or "BACK" in Thai

(a) Thai male outdoor voice

(b) Thai male original voice

Fig. 9 Mel-Frequency Cepstral Coefficients computed from male and female speech signals shown in the form of spectrograms of "Kkwa" or "RIGHT" in Thai

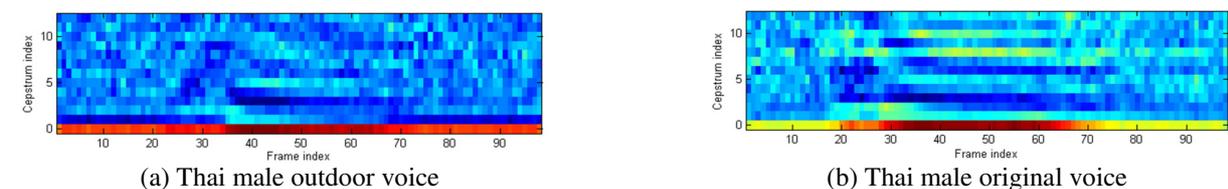(c) Thai female outdoor voice
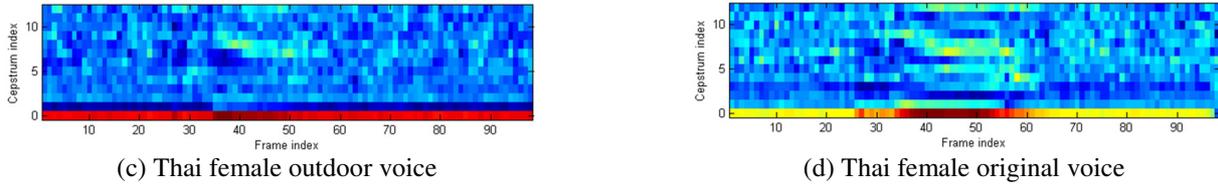

(d) Thai female original voice

Fig. 9 Mel-Frequency Cepstral Coefficients computed from male and female speech signals shown in the form of spectrograms of "Kkwa" or "RIGHT" in Thai (continued)
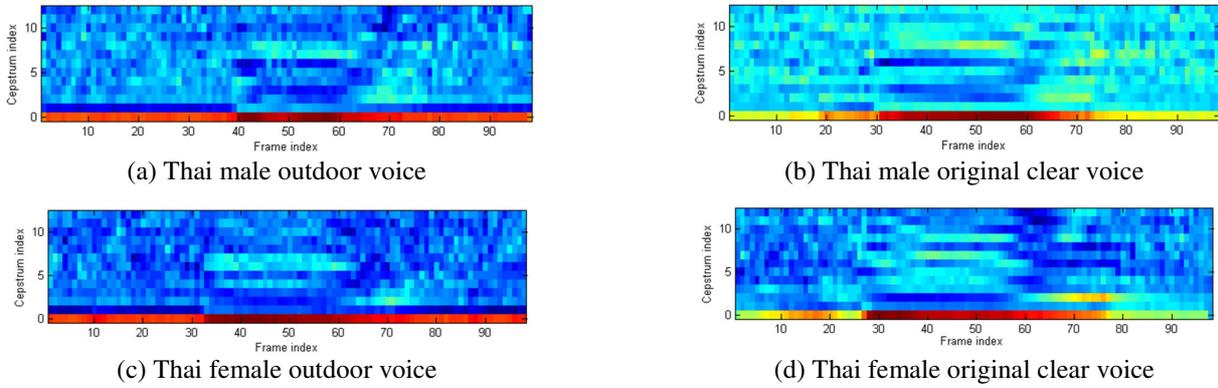

(a) Thai male outdoor voice

(b) Thai male original clear voice

(c) Thai female outdoor voice

(d) Thai female original clear voice

Fig. 10 Mel-Frequency Cepstral Coefficients computed from male and female speech signals shown in the form of spectrograms of "𝓢𝓪𝔂" or "LEFT" in Thai

(a) Thai male outdoor voice

(b) Thai male original clear voice

(c) Thai female outdoor voice

(d) Thai female original clear voice

Fig. 11 Mel-Frequency Cepstral Coefficients computed from male and female speech signals shown in the form of spectrograms of "𝓟𝓱𝓿𝓶" or "READY" in Thai

(a) Thai male outdoor voice

(b) Thai male original clear voice

(c) Thai female outdoor voice

(d) Thai female original clear voice

Fig. 12 Mel-Frequency Cepstral Coefficients computed from male and female speech signals shown in the form of spectrograms of "𝓗𝔂𝓾𝓭" or "STOP" in Thai
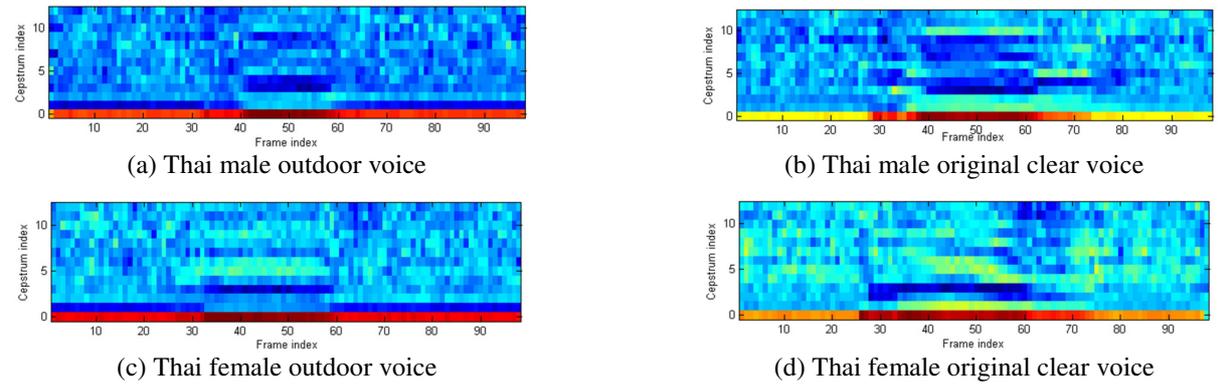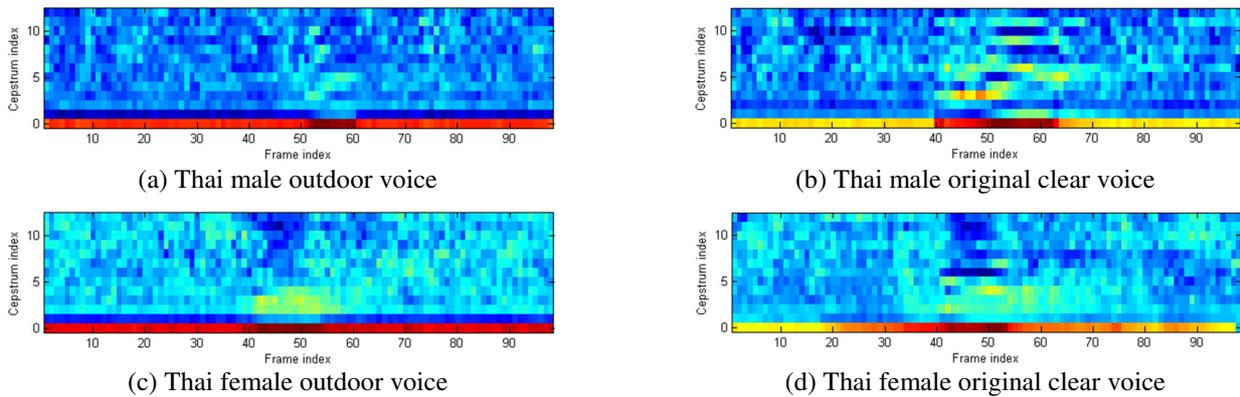
5.2. Hardware implementations

The proposed hardware co-design system was implemented with ZYBO and Arduino boards installed in a 234×434 mm scale electric vehicle as shown in Fig. 13. The properties of the scale electric vehicle are as follows: the motor type is a DC

motor gear, the maximum gross vehicle weight is 1.75 kg, the Kerb weight is 1.38 kg, the wheelbase is 257 mm, the radius of the turning wheel is 22 mm, the tire type is rubber, and the size of the front/rear wheel is 163/161 mm.

(a) 234×434 mm Scale electric vehicle
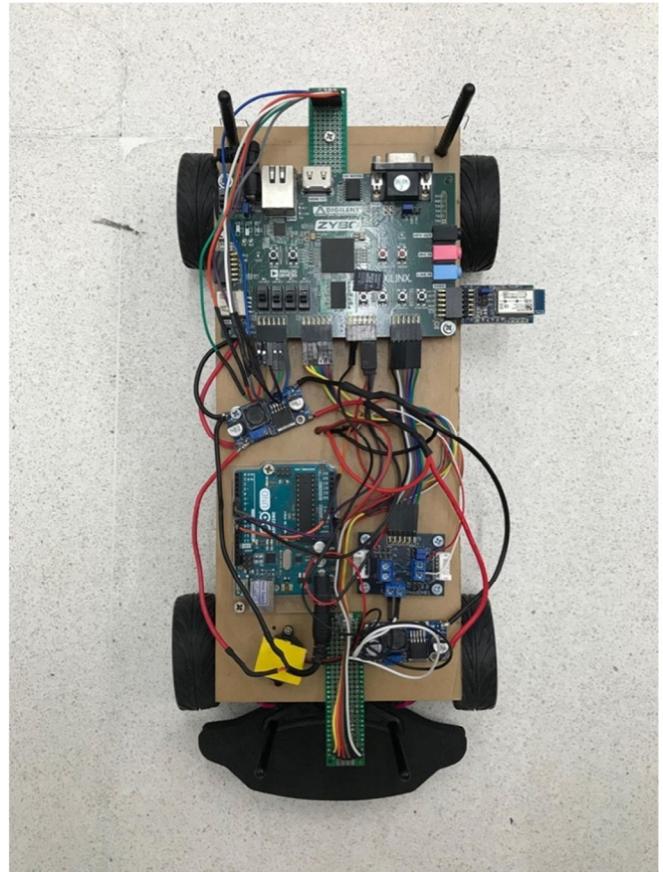
(b) ZYBO and Arduino boards implementation inside

Fig. 13 ZYBO and Arduino boards for installation on a small-scale electric vehicle
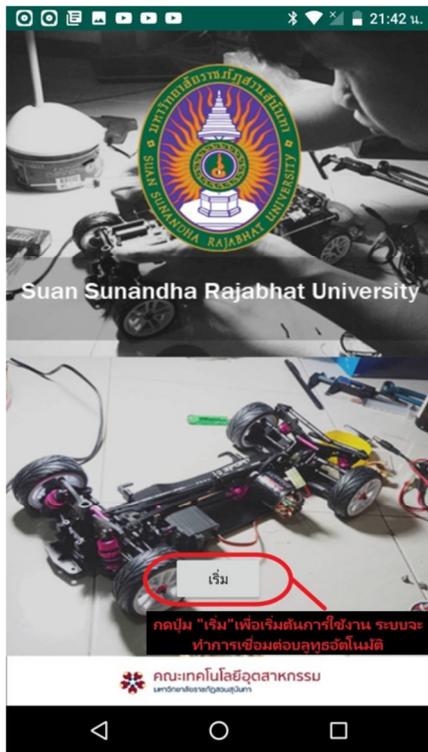
(a) First page of the application

(b) Tap the microphone icon to record the voice command

(c) Send the voice command data via bluetooth to start the systems

Fig. 14 Application design of the proposed voice-controlled car parking assist interface on the smartphone

A prototype application was developed. The application design of the proposed voice-controlled car parking assist interface on the smartphone is provided in Fig. 14. The first page of the voice-user interface is shown in Fig. 14(a); the driver taps the microphone icon to record the voice command depicted in Fig. 14(b) to control car parking; the recorded voice command data are sent through Bluetooth, which is indicated in Fig. 14(c), to start the voice-controlled car parking assist.

The reference command patterns are stored on an SD card in the ZYBO FPGA-embedded board and loaded into the memory while running the software. A comparator using PCC shown in (16) checks the reference command patterns and compares the patterns with the voice command data received via Bluetooth device. All results are transmitted to control the motor drive and the servo motor. The prototype electric vehicle in Fig. 13 will perform a set of actions that involve moving forward/backward, turning right/left, ready, and stop as shown in Table 2. Table 2 demonstrates the comparison of Thai and English voice commands used for maneuvering an automobile.

Table 2 Set of Thai and English word command used for maneuvering a miniature electric vehicle

A set of actions	Thai word command	English word command
Moving Forward	"*Pi*"	"GO"
Moving Backward	"*Jhay*"	"BACK"
Turning Right	"*Kkwa*"	"RIGHT"
Turning Left	"*Say*"	"LEFT"
Ready	"*Phvm*"	"READY"
Stop	"*Hyud*"	"STOP"

For the experimental operations, six voice commands in Thai and English were used, including "*Pi*"/"GO", "*Jhay*"/"BACK", "*Kkwa*"/"RIGHT", "Say"/"LEFT", "*Phvm*"/"READY", and "*Hyud*"/"STOP". We concentrated on the experimental testing of the proposed voice-controlled car parking assist while the driver issuing the commands through the smartphone is outside the car. Tables 3 and 4 show the PCC (p_{xy}) and SPCC (p_{xy}^2) of the received Thai/English, male/female, and outdoor/indoor voice commands compared with the reference Thai and English original clear voice commands stored on the SD card in the ZYBO board. Table 5 illustrates the percentage accuracies of Thai and English voice commands in a noisy environment.

Table 3 PCC (p_{xy}) and SPCC (p_{xy}^2) of the received Thai/English male/female **outdoor** commands
compared with the reference voice commands stored on the SD card in the ZYBO board

	"*Pi*"	"GO"	"*Jhay*"	"BACK"	"*Kkwa*"	"RIGHT"	"Say"	"LEFT"	"*Phvm*"	"READY"	"*Hyud*"	"STOP"
Male:												
p_{xy}	0.981	0.904	0.900	0.865	0.923	0.889	0.902	0.877	0.909	0.847	0.915	0.916
p_{xy}^2	0.963	0.817	0.808	0.749	0.852	0.791	0.814	0.769	0.826	0.718	0.836	0.839
Female:												
p_{xy}	0.932	0.934	0.866	0.978	0.859	0.924	0.904	0.910	0.874	0.913	0.957	0.920
p_{xy}^2	0.869	0.872	0.786	0.956	0.738	0.853	0.818	0.829	0.765	0.833	0.916	0.847

Table 4 PCC (p_{xy}) and SPCC (p_{xy}^2) of the received Thai/English male/female **indoor** commands
compared with the reference voice commands stored on the SD card in the ZYBO board

	"*Pi*"	"GO"	"*Jhay*"	"BACK"	"*Kkwa*"	"RIGHT"	"Say"	"LEFT"	"*Phvm*"	"READY"	"*Hyud*"	"STOP"
Male:												
p_{xy}	0.929	0.951	0.939	0.973	0.924	0.952	0.953	0.972	0.913	0.969	0.902	0.983
p_{xy}^2	0.862	0.905	0.881	0.947	0.854	0.906	0.908	0.945	0.834	0.939	0.813	0.966
Female:												
p_{xy}	0.927	0.934	0.969	0.969	0.927	0.912	0.920	0.910	0.965	0.934	0.998	0.967
p_{xy}^2	0.859	0.872	0.939	0.939	0.860	0.832	0.846	0.827	0.931	0.872	0.995	0.935

Table 5 Percentage accuracies of Thai and English voice commands in a noisy environment

	"*Pi*"	"GO"	"*Jhay*"	"BACK"	"*Kkwa*"	"RIGHT"	"Say"	"LEFT"	"*Phvm*"	"READY"	"*Hyud*"	"STOP"
%Accuracy	89	90	90	87	89	89	89	88	90	85	91	92

6. Conclusions

In this study, voice interfaces are used to control a car parking assist based on the word command patterns of Thai/English and males/females. Gender analysis was used as the basis for an MFCC technique to extract the features represented in the spectrogram of Thai/English and male/female word recognition through STFT. The simulation results indicate the better accuracies of the received Thai/English, male/female, and indoor/outdoor commands as compared with the reference voice commands. The results of experimental tests with Thai and English voice commands for maneuvering a miniature electric vehicle using the PCC and the SPCC of Thai/English, male/female, and indoor/outdoor voice commands were also shown.

Our research contributes to the literature on Thai speech recognition and the performance of the proposed voice-controlled parking assist in noisy environments. This study is also expected to provide parking assistance to ensure drivers' safety and well-being through testing and implementation in commercial eco-cars in the future.

Acknowledgments

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1] D. Johnson, "2 people dead after car falls 4 stories from downtown Indianapolis parking garage," https://fox59.com/news/2-people-dead-after-car-falls-4-stories-from-downtown-indianapolis-parking-garage/, October 23, 2019.

[2] "Another driver narrowly escapes FDA car park plunge," https://www.bangkokpost.com/thailand/general/1804184, November 28, 2019.

[3] T. Cook, R. Martin, E. Hopkins, and T. Evans, "Simple mistakes can lead to cars falling from parking garages in freak accidents," https://www.indystar.com/story/news/2019/10/30/city-market-parking-garage-crash-among-46-last-20-years/2457201001, October 30, 2019.

[4] M. Yukawa, K. Sonoda, and T. Wada, "Auditory assist method to indicate steering start timing in reverse parking for improvement of driver performance," IEEE Transactions on Intelligent Vehicles, vol. 5, no. 1, pp. 32-40, November 2019.

[5] D. Kiss and G. Tevesz, "Autonomous path planning for road vehicles in narrow environments: An efficient continuous curvature approach," Journal of Advanced Transportation, vol. 2017, pp. 1-27, October 2017.

[6] D. Buljeta, M. Vranjes, Z. Marceta, and J. Kovacevic, "Surround view algorithm for parking assist system," Zooming Innovation in Consumer Technologies Conference, IEEE Press, May 2019, pp. 21-26.

[7] S. Prongnuch, M. Ardwichai, N. Pong-ngam and A. Chana, A. Thitinaruemit, P. Uttaphut, and N. Areerachakul, "Embedded system design for car park prototype by voice controlled," Proc. Suan Sunandha Rajabhat University National Conference, March 2016, pp. 3034-3047 (in Thai).

[8] S. Prongnuch and S. Sitjongsataporn, "Exterior car parking assistance algorithm in reconfigurable system-on-chip," unpublished.

[9] R. Burgess, "Honda bucks industry trend by removing touchscreen controls," https://www.autocar.co.uk/car-news/motor-shows-geneva-motor-show/honda-bucks-industry-trend-removing-touchscreen-controls, March 30, 2020.

[10] K. Marso and D. Macko, "A new parking-space detection system using prototype devices and bluetooth low energy communication," International Journal of Engineering and Technology Innovation, vol. 9, no. 2, pp. 108-118, August 2018.

[11] L. Lima, V. Furtado, E. Furtado, and V. Almeida, "Empirical analysis of bias in voice-based personal assistants," The 2019 World Wide Web Conference, May 2019, pp. 533-538.

[12] L. Liu, S. Zhu, D. He, Y. Ma, X. Zhang, J. Huang, and J. Li, "Design and realization of intelligent voice-control car based on Raspberry Pi," International Conference on Smart Vehicular Technology, Transportation, Communication and Applications, Springer press, October 2018, pp. 87-95.

[13] A. Caranica, H. Cucu, C. Burileanu, F. Portet, and M. Vacher, "Speech recognition results for voice-controlled assistive applications," 2017 International Conference on Speech Technology and Human-Computer Dialogue, IEEE Press, July 2017, pp. 1-8.

[14] O. Lavrynenko, G. Konakhovych, and D. Bakhtiiarov, "Method of voice control functions of the UAV," 4th International Conference on Methods and Systems of Navigation and Motion Control, IEEE Press, October 2016, pp. 47-50.

[15] S. Young, G. Evermann, M. Gales, T. Hain, D. Kershaw, and X. Liu, et al., The HTK book, vol. 3, no. 175, Cambridge university engineering department, 2009.

[16] X. Huang, A. Acero, A. Acero, and H. W. Hon, Spoken Language Processing: A Guide To Theory, Algorithm, and System Development, New Jersey: Prentice Hall, 2001.

[17] L. Shi, I. Ahmad, Y. He, and K. Chang, "Hidden Markov model based drone sound recognition using MFCC technique in practical noisy environments," Journal of Communications and Networks, vol. 20, no. 5, pp. 509-518, November 2018.

[18] J. Benesty, J. Chen, Y. Huang, and I. Cohen, Noise Reduction in Speech Processing, New York: Springer Science & Business Media, 2009.

[19] S. Prongnuch, S. Sitjongsataporn, and T. Wiangtong, "A heuristic approach for scheduling in heterogeneous distributed embedded systems," International Journal of Intelligent Engineering and Systems, vol. 13, no. 1, pp. 135-145, November. 2019.

[20] J. Benesty, Fundamentals of Speech Enhancement, Springer Briefs in Electrical and Computer Engineering, Springer, 2018.

[21] Siri team, "Personalized Hey Siri," Machine Learning Journal, vol. 1, no. 9, April 2018.