VIETNAMESE SOCIETY OF AUTOMOTIVE ENGINEERS

A member of the International Federation of Automotive Engineering Societies – FISITA

ICAT 2015

PROCEEDING

THE INTERNATIONAL CONFERENCE ON AUTOMOTIVE TECHNOLOGY FOR VIETNAM

HANOI, OCTOBER 9-11, 2015

Organizer	Patronage	Gold Sponsor	Main Sponsor

VIETNAMESE SOCIETY OF AUTOMOTIVE ENGINEERS

A Member of the International Federation of Automotive Engineering Societies - FISITA

ICAT 2015

PROCEEDING

THE INTERNATIONAL CONFERENCE ON
AUTOMOTIVE TECHNOLOGY FOR VIETNAM

HANOI, OCTOBER 9 - 11, 2015

BAN DO

CARTOGRAPHIC PUBLISHING HOUSE

THE INTERNATIONAL CONFERENCE ON AUTOMOTIVE TECHNOLOGY FOR VIETNAM

ICAT 2015

The International Conference on Automotive Technology for Vietnam
Hanoi, October 9 - 11, 2015

Gold Sponsor:

Toyota Vietnam

Main Sponsors:

Saigon Transportation Mechanical Corporation (SAMCO)

THE INTERNATIONAL CONFERENCE ON AUTOMOTIVE TECHNOLOGY FOR VIETNAM

ICAT 2015

Chairman

Dr. Do Huu Hao

Technical Sessions

Technical Session 1:

Chairman: Ph.D. Kirill Karpukhin

Federal State Unitary Enterprise Central Scientific Research Automobile and Automotive Institute "NAMI" (FSUE «NAMI»), Avtomotornaya Street, 2, Moscow, Russia, 125438

Secretary: PhD. Tran Thanh Tung

School of Transportation Engineering, Hanoi University of Science and Technology, No.1 Dai Co Viet, Hai Ba Trung, Hanoi, Vietnam

Technical Session 2:

Chairman: Prof. Yaojung Shiao

Department of Vehicle Engineering, National Taipei University of Technology, Taiwan

Secretary: MSc. Nguyen Anh Ngoc

Department of Vehicle Engineering, National Taipei University of Technology, Taiwan

Department of Automobile Technology, Hanoi University of Industry, Vietnam

Congress Secretary

Ass. Prof. Du Quoc Thinh

PhD. Trinh Minh Hoang

CONTENTS

Ord. No.	Paper No.	Title of Technical Paper
1	ICAT2015-A1	**Toyota Vietnam presentation**
2	ICAT2015-A2	**SAMCO presentation**
3	ICAT2015-001	**Comprehensive life cycle analysis of different types of energy storage for electric or hybrid vehicles** Kirill Karpukhin, Andrey Kozlov, Sergey Bakhmutov, Alexey Terenchenko
4	ICAT2015-002	**Load Control Algorithm for Efficient Energy Management in Electric Vehicles** Ilya Kavalchuk, M Seyedmahmoudian, Ben Horan, Aman Maung Than Oo, Alex Stojcevski
5	ICAT2015-003	**Semi-active Magnetorheological Mount for Vibration Suppression of Camcorder in Bicycle** Yaojung Shiao, Nguyen Anh Ngoc, Thai-Son Nguyen
6	ICAT2015-004	**Regenerative Braking Strategy for Power-Split Hybrid Electric Vehicle with Fuzzy Logic** Jia-Shiun Chen, Quoc-Viet Huynh, Ngoc-Tham Ngo Thi
7	ICAT2015-005	**Directional Stability Control of Electric Vehicles with Four In-Wheel Motors** Bo Chiuan Chen, Trung Hieu Nguyen, Dejun Yin
8	ICAT2015-006	**Research Algorithm of On-Board Mass (OBM) Monitoring based on Air-Spring Suspension Characters** Nguyen Quang Anh, Nguyen Cong Tuan
9	ICAT2015-007	**Research and evaluation simulation on the working process of the suspension system on KIA MORNING car manufactured in Vietnam** Cao Hung Phi, Tran Hoang Thai
10	ICAT2015-008	**Experimental Research on Coefficient of Slip at Brake Time of Tractor Semi-Trailer** Nguyen Thanh Tung, Vo Van Huong
11	ICAT2015-009	**Applying SAE J1266 Standard to Test the Efficiency of Active Rear Axle on Light Trucks** Nguyen Thanh Quang, Le Hong Ky, Le Hoang Anh
12	ICAT2015-010	**Using Inventor Software to Analyze the Kinetics of Active Rear Axle on Light Trucks with locked differential** Nguyen Thanh Quang, Le Hoang Anh, Hoang Viet
13	ICAT2015-011	**Application of ANSYS Software to simulate durability calculation cardan shaft on Light Trucks** Nguyen Thanh Quang, Tran Huu Danh
14	ICAT2015-012	**A study on a fundamental methodology to solve vibration, acoustical problems of automobile structures** Duong Ngoc Khanh
15	ICAT2015-013	**Study on rollover of tractor semi-trailer in several situations** Ta Tuan Hung, Duong Ngoc Khanh, Vo Van Huong
16	ICAT2015-014	**Speech emotion recognition of Vietnamese driver** Pham Thi Nhung, Pham Viet Anh, Nguyen Hong Quang, Cao Hung Phi
17	ICAT2015-015	**A Study on Effects of Dynamic Properties of Automobile Structures on Precision of Estimated Modal Parameters** Duong Ngoc Khanh, Takuya Yoshimura
18	ICAT2015-016	**Crosswind Effects on Stability of Tractor Semi-Trailers** Vu Thanh Niem, Luu Van Tuan, Dang Viet Ha
19	ICAT2015-017	**Study on detection of jackknifing problems for tractor semi-trailer in turning maneuvers** Ta Tuan Hung, Duong Ngoc Khanh, Vo Van Huong
20	ICAT2015-018	**A Study on Precision Improvement Modal Parameter Estimation of Automobile Structures** Duong Ngoc Khanh, Takuya Yoshimura
21	ICAT2015-019	**Experimental research determine dynamic type load of articulated vehicle** Phan Tuan Kiet, Vo Van Huong, Nguyen Phu Hung, Nguyen Thanh Tung, Le Thanh Dat
22	ICAT2015-020	**Comparison Simulation Study on Performance characteristics of Single-stage and Two-stage Turbocharging Diesel Engine** Nguyen Tien Han
23	ICAT2015-021	**Simulation of ignition timing control with maximum pressure timing in Suzuki engine LTR 450/Z** Nguyen Xuan Khoa, Le Hong Quan, Nguyen Tien Han, Nguyen Thanh Vinh

24	ICAT2015-022	**Construction dynamics model amphibious excavator** Bui Van Hai, Le Hong Quan, Tran Quang Hung
25	ICAT2015-023	**Application of Takagi-Sugeno fuzzy system for automotive fault detection and isolation** Tran Quang Ha, Pham Van Thoan, Pham Huu Nam
26	ICAT2015-024	**Estimate effects of load condition on replacement proportion of syngas on dual fuel diesel/syngas** Bui Van Chinh, Khong Vu Quang, Pham Van The
27	ICAT2015-025	**Simulation study on the effect of the ratio of biodiesel on performance and emissions of diesel engines** Nguyen Tuan Nghia
28	ICAT2015-026	**Building a model studying transverse fluctuation of the hang-up mass of automobiles that use semi-dependent suspension system** Nguyen Xuan Tuan, Nguyen Van Bang, Nguyen Thanh Cong
29	ICAT2015-027	**Design optimally the structure of automobile clutch diaphragm spring** Nguyen Thanh Cong, Nguyen Quang Cuong
30	ICAT2015-028	**Optimization Design of the Automobile Gearbox Structural Parameter based on Matlab** Nguyen Thanh Cong, Le Hong Quan
31	ICAT2015-029	**Simulate dynamic of hydraulic braking model with pneumatic booster** Trinh Duc Phong, Nguyen Tuan Nghia, Trinh Ngoc Long, Nguyen Phuc Dinh , Nguyen Si Dinh
32	ICAT2015-030	**The method of vibration calculation of multi axle vehicles and parameter optimization of the elastic element in the suspension system** Nguyen Huy Truong, Le Ngoc Tu, Đao Đinh Nam, Le Van Anh
33	ICAT2015-031	**Studying taking full advantage of inertia force of Motorcycle to save fuel** Le Van Anh, Le Hong Quan, Nguyen Xuan Khoa, Ngo Quang Tao
34	ICAT2015-032	**Matlab software applications to diagnosis and evaluation technical state of the system of hydraulic power steering cars** Nguyen Can, Le Van Anh, Vu Anh Quang, Le Duy Long
35	ICAT2015-033	**Modeling research the selective catalytic reduction of NOx by urea from light truck diesel exhaust gases** Nguyen Tung Lam, Nguyen Manh Dung, Pham Minh Tuan, Tran Quang Vinh
36	ICAT2015-034	**Researched applied technology for plasma nitriding to improve the planetary gear transmission quality** Pham Van Dong
37	ICAT2015-035	**Estimation of the wearing amount of plasma nitriding planetary gear transmission** Pham Van Dong
38	ICAT2015-036	**Study on establishing flexible hydraulic vibratory hammer control in pile driving in unconsolidated foundation** Tran Quang Hung, Nguyen Van Thang, Le Duy Long
39	ICAT2015-037	**Research and evaluation mechanic-hydraulic automatic controller method on combinatory leveling, compacting and finishing cement concrete road surface** Tran Quang Hung, Vu Van Trung
40	ICAT2015-038	**Study on the change of pressure in hydraulic driving motor in rotary press-in tubular** Tran Huu Ly, Nguyen Van Thang
41	ICAT2015-039	**Fuel Consumption and Operating Characteristics of a Gasohol Fuel Injection Motorcycle Using Japan 10-15 and Constant Velocities** Huynh Thanh Cong, Nguyen Doan Hong
42	ICAT2015-040	**Research of Developing Vietnam Car Design Evaluation Software (VCDES)** Pham Viet Anh, Trinh Minh Hoang, Nguyen Trong Hoan
43	ICAT2015-041	**Study on the reliability of active axle of light trucks produced and assembled in Vietnam** Tran Phuc Hoa, Trinh Minh Hoang, Le Hong Quan, Du Quoc Thinh
44	ICAT2015-042	**Using multi-elements generalized polynomial chaos to study on the stability analysis in the presence of uncertainties of clutch system on automobile** Trinh Minh Hoang
45	ICAT2015-043	**Research Dynamics of a Compacting and Flattening-Ciment Concrete Self-Moving Equipment, Vibrated-Steel Frame Type** Tran Quang Hung, Bui Van Tuong, Le Thanh Dat
46	ICAT2015-044	**Studying the influence of vibration on rollover** Nguyen Van Viet, Tran Minh Son, Nguyen Van Hung

Speech emotion recognition of Vietnamese driver

Pham Thi Nhung[1], Pham Viet Anh[1], Nguyen Hong Quang[1], Cao Hung Phi[2]

1) School of Information and Communication Technology, Hanoi University of Science and Technology
2) Vinh Long University of Technology Education

Abstract

Emotion recognition of driver is needed to help prevent traffic accidents. This should be an intelligent driver support system to accurately monitor the driver's state and give the alert. The most linguistically resources for emotions recognition is speech. However, the most results of these researches in the world focused for the common languages like English or French. So this paper presents the first study for recognizing emotions of Vietnamese speaker. The system is trained and tested with Vietnamese Emotion Speech Database which include four emotions: happiness, sadness, anger and normal. The emotion recognition system was studied by using the MFCC features and GMM model. We have integrated the gender recognition into emotions recognition system and the results accurately emotion recognition have improved from 62.46% to 65.75%.

Keywords: speech emotion recognition, emotional states, driver emotion, Vietnamese emotion speech database

1. INTRODUCTION

Driving is a job that requires the concentration of the driver. In addition to biological status as well as the emotions of the driver also significantly affect driving performance. In recent years, modern technologies are increasingly being integrated into cars. These technologies allow the drivers in the easier and more favorable automotive operation such as: satellite navigation, audio/video systems, etc. Compared with video technology, voice inquiry technology will help the drivers operate the car safer. Based on the driver's recorded voice, emotion recognition module can pinpoint emotions of the driver, and then give a warning when necessary.

Voice is a sound sequence emitted from the vocal tract of human, used to exchange information, ideas, and feelings between people, particularly among members of society. Voice is one of the most fundamental and most important means of communication of human. Currently, studies and applications related to speech processing still focus on developing and improving, such as voice synthesis, speech recognition, voice storage, voice signal processing... One of these areas is gender identity, feelings of the speaker. Gender identity of the speaker is a system that defines whether the speaker is male or female by using characteristics derived from their voices. While speaker's emotion recognition system will determine the emotional state of the speaker: sad, happy, angry, scared, or normal.

In recognition of the speaker's feelings, emotions expressed by male and female voices are different because the information in the voice of men and women is different, such as the fundamental frequency of the male voice is 80-200 Hz and fundamental frequency of female voice is 150-450 Hz [10]. Therefore, when building the identity system of the speaker's feelings, we consider and incorporate gender recognition system into the emotion identification system of the speaker [1] [7] [8] [14].

The identity of Gender and feeling of the speaker is a field of voice recognition. In the identity of gender, feelings of the speaker, the sample recognition method is often used. In this method, the database used for training must have the full versions of the forms that need to identify, the training process can accurately identify the characteristics of the samples. This method consists of two phases: training and recognition.

In the training step, first, we need to record the emotions to be trained: with each recording, encode the speaker, identify the gender and label the emotion. Each speaker will be asked certain questions and expresses different emotions with the same sentence. Depending on the number of emotions that the system needs to identify, the speaker would express such number of emotions. After obtaining database of emotional voices, select characteristics: is the change of real voice signal into a specific parameter for identification and test. The selection of characteristics is extremely important, directly affects recognition results. One characteristic widely used in the identification system is Mel Frequency Ceptral Coefficients (MFCC) [2]. Next, train a statistical model for each gender and emotion: Gaussian Mixture model (GMM) [2][3], hidden Markov model (HMM), artificial neural network model (ANN)...

Recognition phase is the process of determining the gender and the emotion of the speaker. The process is based on gender model, emotion model that has been chosen. In this step, at first, record the voice to check gender, emotional state. Then select the characteristic. Finally, make the identification decision: the system will compare the characteristics of the voices entered to identify with the characteristics of the gender model, emotion model that has been chosen. The model with the highest similarity or the biggest scores will be selected as the result model.

The next section of the article includes the following parts:

- Part 2 describes the identification method GMM and the parameter set MFCC used in emotion identification.

- Part 3 presents the tests and emotion identification results.
- Part 4 summarizes and gives direction to subsequent studies.

2. EMOTIONAL RECOGNITION METHOD

In this paper, we modeled each emotion and gender with a GMM model. Each sample of an emotion and gender is represented by a series of MFCC vectors.

2.1. Gaussian Mixture Model

A mixed model Multivariate Gaussian is the weighted total of M density component Gauss as the expression (1):

$$p(X|\lambda) = \sum_{i=1}^{M} \pi_i \, g_i(X|\mu_i, \Sigma_i) \qquad (1)$$

In (1), X is the data vector (containing the parameters of the objects to be represented), πi, i=1, M is the weights of the mixture and $g_i(X|\mu_i, \Sigma_i)$ is the component Gaussian density functions (equation (2)) with average vector μi of D-dimensional vector and covariance matrix Σi dimension DxD.

$$g_i(X|\mu_i, \Sigma_i) = \frac{1}{(2\pi)^{D/2}|\Sigma_i|^{1/2}} \exp\left\{-\frac{1}{2}(X-\mu_i)'\Sigma_i^{-1}(X-\mu_i)\right\} \qquad (2)$$

The mixed weights must meet the condition $\sum_{i=1}^{M} \pi_i = 1$.

A full GMM is parameterized by average vectors, covariance matrix and the mixed weights from all Gauss components. These parameters can be represented neatly under (3).

$$\lambda = \{\pi_i, \mu_i, \Sigma_i\}, i = 1,2, \ldots, M \qquad (3)$$

Assuming T is number of eigenvectors or also the entire amount of the frames (frame) of voice, M is the number of Gauss components: $X = \{x_1, x_2, \ldots, x_T\}$. Then the probability of GMM model with sample X is represented by Equation 4.

$$p(X|\lambda) = \prod_{t=1}^{T} p(x_t|\lambda) \qquad (4)$$

Model λ is determined by the algorithm expectation-maximization (EM). The idea of the EM algorithm is started with prototype λ, assess new model $\bar{\lambda}$ by Equation 5.

$$p(X|\bar{\lambda}) \geq p(X|\lambda) \qquad (5)$$

The new model is the starting model for the next iteration step and repetitive process until achieving the threshold of convergence.

2.2. Mel Frequency Cepstral Coefficients

In a study published in [8], GMM model is used effectively with parameter set MFCC (Mel Frequency Cepstral Coefficients: the spectral coefficients by mel-frequency scale). Algorithm perform the calculation of parameters is described in Figure 1.

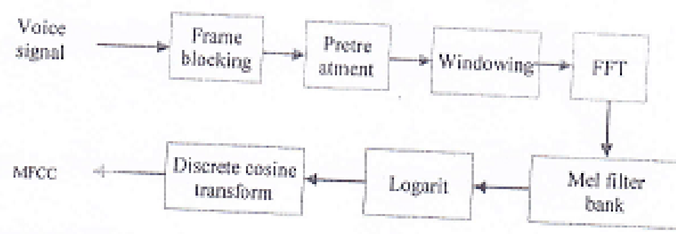

Fig. 1. Block diagram of the algorithm for MFCC parameter calculation

In Figure 2, at first, voice signal is divided into frames with a length of 0.1 seconds and the shift of a frame is of 0.01 seconds. Then each voice signal frame will be preprocessed according to equation (6).

$$y(n) = x(n) - 0{,}96x(n-1) \qquad (6)$$

The signal after being pre-processing will be taken through Hamming (Equation 7), where N is the number of samples of a voice signal frame).

$$w(n) = 0{,}54 - 0{,}46\cos(2\pi n/(N-1)) \text{ với } 0 \leq n < N \qquad (7)$$

Next, calculate the Fast Fourier Transform (FFT: Fast Fourier Transform) on the signal given through hamming. Spectrum of received signal will pass the filter by Mel scale. The number of filters is the number of MFCC coefficients to be calculated. Finally calculate logarithms on the output values of filter and carry out discrete cosine transform to obtain MFCC coefficients.

2.3. Emotional Recognition Method

Figure 2 describes the steps incorporate gender identity system into emotion recognition system. This system consists of two steps:

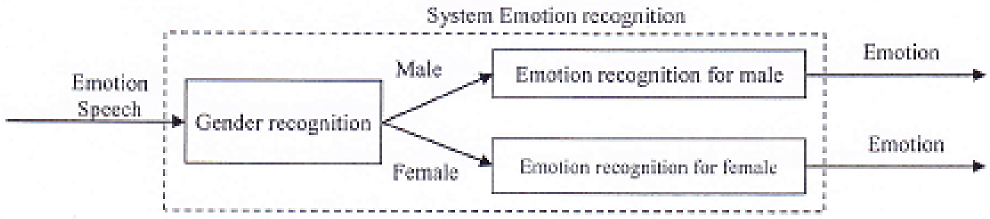

Fig. 2. Chart of incorporating gender identity system into emotion recognition system

- Step 1: implement gender recognition: the emotional voice given into gender identity system, the system will identify whether the emotional voice belongs to male or female and then the system will switch to the next steps.
- Step 2: implement emotion identity: in step 1, if the gender recognition system results male, emotion identification system will identify emotion with emotion model has been chosen for male voice, if gender identity system results female, emotion identification system will identify emotion with emotion model has been chosen for female voice.

3. EMOTIONAL RECOGNITION EXPERIMENTS

3.1. Constructing emotional speech corpus for Vietnamese

Is the database built and recorded by MA. Le Xuan Thanh (lecturer, graduate students in the Department of Computer Engineering) at the VTV, speakers expressing emotions are actors. General information about the emotion database:

- Language used: Vietnamese
- The emotions expressed are: angry, happy, sad and normal.
- Number of sentences: 55 Vietnamese sentences without emotion identity.
- Number of actors: a total of 55 actors, but in the thesis only uses recordings of feelings from 50 actors (25 male and 25 female) to test data sets with the equal gender as well as the number of emotion data files.
- Recording: the sound files recorded as 1 mono channel, 16-bit, 16KHz sampling frequency. Each actor would say 55 sentences, each sentence shown with 4 different emotions and recorded 4 times.

This database is named Vemo-DB with many actors (50 people) so each 5 actors are gathered into one group. Each group is denoted as shown in Table 1.

Table 1: Table of Vemo-DB database

Emotion \ Speaker		Male					Female					Total file
	Symbol	A	B	C	D	E	F	G	H	I	J	
Angry	W	1095	1100	1100	1098	1098	1100	1099	1100	1100	1099	10989
Happy	F	1082	1100	1100	1100	1099	1099	1096	1100	1099	1100	10975
Sad	T	1095	1099	1099	1100	1100	1100	1099	1100	1100	1099	10991
Normal	N	1092	1100	1100	1100	1100	1098	1100	1100	1100	1099	10989
Total file		4364	4399	4399	4398	4397	4397	4394	4400	4399	4397	43944

3.2. Experiments of gender and emotional identification of speakers

The experiments have conducted by two criteria:

- Criterion 1: The speaker who used to identify is not trained in the system.
- Criterion 2: The speaker who used to identify is trained in the system.

In this study, we conducted 6 experiments and subject to each experiment, the experiments will be conducted with the two criteria (Table 2).

Table 2: Experiment of gender and emotional identification of speakers

Experiment	Contents of the experiment	Symbol
Test 1	Experiment of gender identification	G (Gender)
Test 2	Experiment emotional identification with the speakers including male and female voice. Conduct the training including male and female voice.	E (Emotion)
Test 3	Experiment emotional identification as a result gender identification is male. Conduct training including only male voice.	GME (Gender Male Emotion)
Test 4	Experiment emotional identification as a result gender identification is female. Conduct training including only female voice.	GFE (Gender Female Emotion)
Test 5	Experiment emotional identification with male speakers. Conduct training including only male voice.	EGM (Emotion Gender Male)
Test 6	Experiment emotional identification with female speakers. Conduct training including only female voice.	EGF (Emotion Gender Female)

Experiment of gender identification

For the experiment of gender identification, we conduct the experiment in turn, each experiment will use 4 groups of speakers (2 men and 2 women) to experiment the remaining groups of speakers used for training. Summarize the experiments presented in Table 3.

Table 3: Summarize the experiment of gender identification (experiment 1) and experiment of emotion identification (experiment 2)

No.	Test 1		Test 2	
	Criteria 1	Criteria 2	Criteria 1	Criteria 2
1	G01_ABFG	G01_ABFG	E01_AEFG	E01_AEFG
2	G01_CDHI	G01_CDHI	E01_CDGH	E01_CDGH
3	G01_AEFJ	G01_AEFJ	E01_AEFJ	E01_AEFJ
4	G01_BCGH	G01_BCGH	E01_BCGH	E01_BCGH

Symbols in Table 3 are interpreted as follows:

- G is gender identification;
- 01 is the experiment by the criteria of untrained speaker system;
- 02 is the experiment by the criteria of trained speaker system;
- ABFG is four groups of speakers A, B, F and G used for the experiment (similar to CDHI, AEFJ, BCGH).

Experiment of emotional identification

Experiment emotional identification includes two steps: training and identification. In training step, 4 emotional models will be trained. For getting 4 emotional models, it should have 4 sound sets of data corresponding to 4 emotions as the input to the training process. Each experiment will use 4 speakers and the remaining groups will be used for training. After the training step, we will have GMM model for four emotions.

In emotional identification step: for emotional identification experiment on Vemo-DB database, we will perform the experiment in turn, each experiment will use 4 groups of speakers (2 men and 2 women) to identify, the remaining groups of speakers are used for training. Summarize the experiments presented in Table 4.

Symbols in Table 4 are constructed as follows:

- E is emotional identification;
- 01 is the experiment by the criteria of untrained speaker system;
- 02 is the experiment by the criteria of trained speaker system;
- ABFG is four groups of speakers A, B, F and G used for the experiment (similar to CDHI, AEFJ, BCGH).

Experiment emotional identification combination with gender identification system in emotional identification system

After the result of gender identification result is a list of files of male and female identification, the files of the list will be used as experimental files of emotional identification on training database of male and female voice. Summarize the experiments presented in Table 4.

Symbols in Table 4 are constructed as follows:

- GME: Experiment emotional identification in case of the results of male gender identification on a database of training emotion of male voice.
- GFE: Experiment emotional identification in case of the results of female gender identification on a database of training emotion of female voice.
- 02: is the experiment by the criteria of trained speaker system.
- ABFG: is four groups of speakers A, B, F and G used for the experiment (similar to CDHI, AEFJ, BCGH).
- Experiment emotional identification in case of knowing gender.

Table 4: Summary emotional identification combination with gender identification system in emotional identification system (experiment 3, 4) and emotional identification experiments in case of knowing gender (experiment 5, 6)

No.	Criteria 3	Criteria 4	Criteria 5	Criteria 6
1	GME02_AEFG	GFE02_AEFG	EGM02_AB	GFE02_FG
2	GME02_CDGH	GFE02_CDGH	EGM02_CD	GFE02_HI
3	GME02_AEFJ	GFE02_AEFJ	EGM02_AF	GFE02_FJ
4	GME02_BCGH	GFE02_BCGH	EGM02_BC	GFE02_GH

Use two groups of speakers in turn to experiment emotional identification on trained emotional database of male voice and 2 female speakers to experiment emotional identification on trained emotional database of female voice. Summarize the experiments presented in Table 5.

Symbols in Table 5 are constructed as follows:

- EGM: Experiment emotional identification with male speaker on trained emotional database of male voice.
- EGF: Experiment emotional identification with female speaker on trained emotional database of female voice.
- 02: is the experiment by the criteria of trained speaker system.
- AB: two groups of speakers A and B are used for identification (similar to CD, AF, BC, FG, HI, FJ, GH).

3.3. Experimental Results

The gender identification results by criterion 1 and 2 as described in Table 5 and Table 6.

Table 5: Percentage of correct gender identification (%) of the experiments; criterion 1 (untrained speakers in system).

Test / Gender	G01_ABFG	G01_CDHI	G01_AEFJ	G01_BCGH	Rate average
Male	88,41	91,13	95,30	88,13	90,74
Female	93,35	93,65	92,52	92,84	93,09
Rate average	**90,89**	**92,39**	**93,91**	**90,48**	**91,92**

In Table 5, we see that gender identification system reached 91.92% of average rate. The percentage of male gender identification is 90.74%. The percentage of female gender identification is 93.09%. The identification percentage between experiments is similar because the identification and training files of experiments are similar.

Table 6: Percentage of correct gender identification (%) of the experiment; criterion 2 (the trained speakers in system)

Test / Gender	G02_ABFG	G02_CDHI	G02_AEFJ	G02_BCGH	Rate average
Male	90,22	91,22	95,16	88,74	91,34
Female	93,48	93,11	92,88	92,34	92,95
Rate average	**91,85**	**92,16**	**94,01**	**90,54**	**92,14**

Table 6 shows the percentage of gender identification by criterion 2 improved from 91.92% (by criterion 1) to 92.14% by trained speakers used for identification in system.

Table 7: Percentage of correct emotional identification (%) of the experiment; emotional identification on emotional training database on both male and female voice; criterion 1 (untrained speakers)

Emotion \ Test	E01_AEFG	E01_CDGH	E01_AEFJ	E01_BCGH	Rate average
Angry	59,99	58,36	57,03	62,83	59,55
Happy	54,67	52,64	51,46	53,11	52,97
Sad	68,38	75,01	70,93	73,20	71,88
Normal	42,34	51,86	43,85	49,43	46,87
Rate average	**56,35**	**59,47**	**55,83**	**59,64**	**57,82**

Table 7 shows that the emotional identification system for emotional identification percentage for angry and sad feelings. Percentage of the correct emotional identification is among similar experiments and about 56.35% to 59.64%.

Table 8: Percentage of correct emotional identification (%) of the experiment; emotion identification on emotional training database on both male and female voice; creation 2 (trained speakers in system);

Emotion \ Test	E02_AEFG	E02_CDGH	E02_AEFJ	E02_BCGH	Rate average
Angry	61,62	61,55	59,58	64,60	61,84
Happy	59,94	60,46	56,62	60,75	59,44
Sad	71,79	80,49	74,55	76,71	75,89
Normal	49,70	57,34	50,05	53,47	52,64
Rate average	**60,77**	**64,96**	**60,21**	**63,88**	**62,46**

Table 8 shows the percentage of correct identification of similar experiments and correct identification percentage of no more difference emotions due to a few of emotional identification in system (4 emotions). Compared to emotional identification by creation 1, the emotional identification by creation 2 improved from 57.82% to 62.46%.

Table 9: Percentage of correct emotional identification (%) of the experiments; in case of the results of male gender identification on emotional training database of male voice by criterion 2 (trained speaker in system)

Emotion \ Test	GME02_ AEFG	GME02_ CDGH	GME02_ AEFJ	GME02_ BCGH	Rate average
Angry	60,38	60,39	62,29	59,25	60,58
Happy	62,24	69,34	58,36	66,79	64,18
Sad	68,36	77,71	69,80	74,54	72,60
Normal	57,15	57,30	55,78	55,69	56,48
Rate average	**62,01**	**66,00**	**61,61**	**63,79**	**63,35**

Table 10: Percentage of correct emotional identification (%) of the experiments; in case of the results of female gender identification on emotional training database of female voice by criterion 2 (trained speaker in system)

Emotion \ Test	GFE02_ AEFG	GFE02_ CDGH	GFE02_ AEFJ	GFE02_ BCGH	Rate average
Angry	68,26	63,00	63,02	73,81	67,02
Happy	64,82	69,28	69,85	61,99	66,49
Sad	76,46	88,50	82,42	80,95	82,08
Normal	53,12	62,21	58,08	55,27	57,17
Rate average	**65,58**	**70,84**	**68,42**	**67,71**	**68,14**

Results from Table 10 and Table 11 shows that the integration of gender identity module with emotion recognition system gave better recognition results. Percentage of correct emotional identification with male voice was 65.75% (63.35% and 68.14% female voice) versus emotion recognition results when gender is unknown 62.72% (Table 8).

Table 11 and Table 12 describe the emotion recognition results where known gender. These tests were performed to assess the ability to improve the performance (accuracy of identification) maximum of emotion recognition system with the aid of gender identity module.

Table 11: Percentage of correct emotional identification (%) of the experiments; the emotional identification of male speakers on emotional training database of male voice by criterion 2 (trained speaker in system)

Emotion \ Test	EGM02_ AE	EGM02_ CD	EGM02_ AE	EGM02_ BC	Rate average
Angry	61,82	62,46	66,30	61,68	63,07
Happy	66,26	73,09	59,19	70,81	67,34
Sad	67,63	78,58	71,48	74,29	73,00
Normal	54,92	57,00	56,56	53,50	55,50
Rate average	**62,66**	**67,78**	**63,39**	**65,07**	**64,73**

Table 12: Percentage of correct emotional identification (%) of the experiments; the emotional identification of female speakers on emotional training database of female voice by criterion 2 (trained speaker in system)

Emotion \ Test	EGF02_FG	EGF02_HI	EGF02_FJ	EGF02_GH	Rate average
Angry	69,89	63,68	64,98	75,35	68,48
Happy	71,93	76,71	73,12	70,08	72,96
Sad	76,48	88,27	82,40	80,44	81,90
Normal	56,09	63,45	58,35	58,04	58,98
Rate average	**68,60**	**73,03**	**69,71**	**70,98**	**70,58**

Experimental results show that the hypothesis has to know the gender of the speaker, the emotional identification rate increases: with the male voice from 63.35% (Table 9) to 64.73% (Table 11), with the female voice from 68.14% (table 10) to 70.58% (table 11).

In table 13, summarize the following experiments:

- Experiment 1: Experiment emotional identification in case of unknown gender.
- Experiment 2: Experiment emotional identification combination with gender identification system in emotional identification system.
- Experiment 3: Experiment emotional identification in case of knowing gender.

Table 13 shows that if the gender identification system combines with emotional identification system, the emotional identification system could not been achieved the percentage of correct emotional identification (67.65% in case of unknown gender) but the emotional identification system has been improved from 62.46% to 65.75%. When the gender identification system combines with emotional identification system, the emotional identification system has not reached the emotional identification percentage as the emotional identification system in case of knowing gender because gender identification system only reached 92.14% of gender identification and will have 7.86% of gender identification confusion.

Table 13: Summary the percentage of the correct emotional identification (%) of the experiments

Emotion \ Test	Test1	Test2	Test3
Angry	61,84	63,80	65,77
Happy	59,44	65,33	70,15
Sad	75,89	77,34	77,45
Normal	52,64	56,83	57,24
Rate average	**62,46**	**65,75**	**67,65**

With the experimental results on two sets of emotional database in German and Vietnamese when the gender identification system combines with the emotional identification system, the emotional identification system has been improved.

3. CONCLUSIONS

This article has presented the first studies of the emotional identification of Vietnamese speakers. Our system is trained and experimented with emotional database in Vietnamese for 4 emotions: happiness, sadness, anger and normality. The system uses MFCC features and GMM model to identify emotions. With emotional database in Vietnamese, the results of the emotional identification by criterion 1 (untrained speakers in system) reached 57.82% of average percentage and the results of the emotional identification by criterion 2 (trained speakers in system) reached 62.46% of average percentage. When the gender identification system combines with the emotional identification system, the emotional identification system has been improved from 62.46% to 65.75%.

In the future, we will develop the gender and emotional identification system of the speaker using the second type of database to apply to the actual systems. In addition, we will combine applications of emotional identification of speakers to emotional identification via facial movements.

References

1. Ashish Tawari and Mohan Trivedi, "Speech Based Emotion Classification Framework for Driver Assistance System", 2010 IEEE Intelligent Vehicles Symposium University of California, San Diego, CA, USA, June 21-24, 2010.
2. Krajewski, J., Nöth, E., & Batliner, A.. Speech emotion recognition. Measuring driver states from voice characteristics. In A. Lichtenstein, C. Stößel, & C. Clemens (Eds.), Der Mensch im Mittelpunkt technischer Systeme (S. 98-103). Fortschritt-Berichte VDI, Reihe 22, Mensch-Maschine-Systeme. Düsseldorf: VDI-Verlag, 2009
3. N. Kamaruddin and A. Wahab, "Driver Behavior Analysis through Speech Emotion Understanding", 2010 IEEE Intelligent Vehicles Symposium University of California, San Diego, CA, USA, June 21-24, 2010.
4. Norhaslinda Kamaruddin, Abdul Wahab, and Hüseyin Abut, "Chapter 2 Driver Emotion Profiling from Speech", Digital Signal Processing for In-Vehicle Systems and Safety, Springer, 2012.
6. Elif Bozkurt, Engin Erzin, Çiğdem Eroğlu Erdem, A. Tanju Erdem, "Improving Automatic Emotion Recognition from Speech Signals", Proceeding of 10th Annual Conference of the International Speech Communication Association (INTERSPEECH 2009), pp. 324-327, 2009
7. Felix Burkhardt, Markus van Ballegooy, Klaus-Peter Engelbrecht, Tim Polzehl, Joachim Stegmann, "Emotion Detection in Dialog Systems: Applications, Strategies and Challenges", ACII 2009: Affective Computing and Intelligent Interaction and Workshops, 2009
8. Igor Bisio, Alessandro Delfino, Fabio Lavagetto, Mario Marchese, And Andrea Sciarrone, "Gender-Driven Emotion Recognition Through Speech Signals for Ambient Intelligence Applications", Proceeding IEEE, Vol. 1, No. 2, pp. 244-257, 2013
9. Kalyana Kumar Inakollu, Sreenath Kocharla, "Gender Dependent and Independent Emotion Recognition System for Telugu Speeches Using Gaussian Mixture Models", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, Issue 11, pp. 4172-4175, 2013
10. Laurence Vidrascu, Laurence Devillers, "Detection of real-life emotions in call centers", Proceeding of 9th European Conference on Speech Communication and Technology (INTERSPEECH' 2005), pp. 1841-1844, 2005
11. Prasad Reddy P.V.G.D, Prasad A, Srinivas, Brahmaiah P, "Gender Based Emotion Recognition System for Telugu Rural Dialects Using Hidden Markov Models", Journal of Computing, Vol2, Issue 6, pp. 94-98, 2010
12. Slobodan T. Jovičić, Zorka Kašić, Miodrag Đorđević, Mirjana Rajković, "Serbian emotional speech database: design, processing and evaluation", SPECOM'2004: 9th Conference Speech and Computer, 2004
13. Thurid Vogt, Elisabeth André, "Improving Automatic Emotion Recognition from Speech via Gender Differentiation", Proceedings of Language Resources and Evaluation Conference (LREC), 2006
14. Hoang Anh Viet, Ngo Van Manh, Ban Ha Bang, Huynh Quyet Thang, " A Real-Time Model based Support Vector Machine for Emotion Recognition through EEG", International Conference on Control, Automation and Information Sciences ICCAIS 2012, Ho Chi Minh City 26-28 November 2012, pp.191 – 196, Digital Object Identifier: 10.1109/ICCAIS.2012.6466585, 2012

Bộ Văn hóa, Thể thao và Du lịch
Cục xuất bản

cuốn, khổ 297 mm x 210 mm tại Nhà máy in Bản đồ - CN Công ty TNHH MTV NXB Tài nguyên môi trường và bản đồ Việt Nam - 85 Nguyễn Chí Thanh - Hà Nội
Nộp lưu chiểu tháng 10 năm 2015