Scientific Journal of Impact Factor (SJIF): 4.14

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 3, Issue 2, February -2016

# System Propose For Be Acquainted With newborn Cry Emotion Using Linear Frequency Cepstral Coefficient

Sandhya S Jagtap [1], Premanand K Kadbe.[2]

[1] *Department of Electronics and Telecommunication, Vidya Pratishthan's College of Engineering,M.I.D.C ,Baram ati, Pune, India*

[2]*Department of Electronics and Telecommunication Vidya Pratishthan's College of Engineering, M.I.D.C ,Baram ati, Pune, India*

**Abstract** — *In this paper, we mainly paying attention on mechanization of Infant's Cry. For this implementation we use LFCC for feature extraction and VQ codebook for toning samples using LBG algorithm. The newborn crying samples composed from various crying baby having 0-6months age. There are 27 babie's sound as training data, each of which represents the 7 hungry infant cries, 4 sleepy infant cries, 10 in pain infant cries, and 6 uncomfortable infant cries. The testing data is one of the traning newborn crying sample. The discovery of infant cries based the least amount distance of Euclidean distance. The, classification of the cry in four classes neh for hunger owh for sleepy, heh for discomfort ,eair for lower gas.*

*Here for classification of the cry our system is alienated into two phases. First is training phase, in which LFCC is used for feature extraction, and then VQ codebooks are engender to compress the feature vectors. Second is the testing stage in which features extraction and codebook production of samples are recurring. Here, estimation of the codebook blueprint of samples to the all the existing patterns in the database are carried based on Euclidian distance between them. LFCC efficiently take into safekeeping the lower as well as higher frequency characteristics than MFCC, hence we will get high-quality results over MFCC.*

***Keywords***- *DBL, LFCC, Feature extraction, Euclidean distance, codebook, Matlab, VQ*

## I. INTRODUCTION

The first verbal communication of newborn baby with the world is baby's cry. Infant crying is a biological alarm system. An infant crying signal are the attention call for parents or caregivers and motivates them to alleviate the distress. Currently, there is a system that learns the meaning of a 0-3 month old infant cries which is called Dunstan Baby Language (DBL). DBL is introduced by Priscilla Dunstan, an Australian musician who has got talent to remember all kinds of sounds, known as sound photograph. According to DBL version, there are five baby languages: "neh" means hunger, "owh" means tired which indicates that the baby is getting sleepy, "eh" means the baby wants to burp, "eairh" means pain (wind) in the stomach, and "heh" means uncomfortable (could be due to a wet diaper, too hot or cold air, or anything else) Infant crying is characterized by its periodic nature, i.e. alternating cry utterances and inspirations. By using a rapid flow of air through the larynx burst sound is produced, because of that there is repeated opening and closing of the vocal folds, which in turn generates periodic excitation. This excitation is transferred through the vocal tract to produce the cry sound, which normally has a fundamental frequency (pitch) of 250-600 Hz. The acoustic signal of an infant's cry contains valuable information about their physical and physiological condition, such as health, weight, identity, gender and emotions. Here for cry detection, zero-crossing rate (ZCR) and fundamental frequency [1], Fast Fourier transforms coefficients, were determined and analyzed to detect the crying signals. In clinical settings one can assume noise-free conditions, and the research depends on finding subtle difference between cries that may be used for diagnostic purposes. In contrast to the clinical settings the detection problem does not assume noise-free conditions. In other words, the focus is on robustness in detecting crying signals in noisy and unpredictable environments.

In this paper, we explain an analysis of infants' cry and present an algorithm for cry detection, which is aimed to alert parents in potential physical danger situations. The proposed algorithm is based on two main stages. The first stage involves feature extraction using LFCC, in which pitch related parameters, are extracted from the signal. In the second stage, the signal is classified using Vector Quantization and later verified as a cry signal. LFCC is based on linear-frequency cepstral coefficients instead of MFCC as a short-time feature. LFCC effectively capture the lower as well as higher frequency characteristics than MFCC [2].Also, mel-frequency cepstral coefficients (MFCCs) and short-time energy were used to develop a noise-robust crying detection system [3] Motivated by this, we use LFCC for robust performances than MFCC. We hope that, by capturing more spectral details in the high frequency region, the linear scale in frequency may provide some advantages in speaker recognition over the mel scale.

## II. REVIEW OF PRIOR CRY DETECTION TECHNIQUES

Researchers did research on infant cries, such as: cries classification of ordinary and irregular infant by means of a neural network which generates 85% accuracy[4], the cataloging of healthy infants and infants who had pain like brain damage, lip cleft palate, hydrocephalus, and sudden infant death syndrome by using Hidden Markov Model (HMM) which makes 91% accuracy. [5]. Previous classification of infant cries have learned HMM as the classifier. The research to recognize the DBL infant cries used codebook as a pattern identifier which is achieved from the k-means clustering and MFCC as feature taking out. The selection of the method is based on the high exactness outcomes from the researches, such as: automatic recognition of birdsongs making use of MFCC and Vector Quantization codebook and creates 85% accuracy [6]. Moreover, the research on the subject of speaker recognition system also successfully created by using MFCC and VQ [7]. The equivalent research that Singh and Rajan did, acquired 98. 57% [8] accuracy by accomplishing speaker recognition research using MFCC and VQ. The research about speech recognition and verification by means of MFCC and VQ which Patel and Prasad (2013) did can perform recognition with training error rate about 13% [9]. The formulating of this codebook using k-means clustering.

## III. ANALYSIS OF INFANT CRYING

Infant crying signals have the diverse temporal and spectral characteristics, which are vital nods in distinguishing them from all-purpose sounds such as speech. Fig.1 demonstrates the classic spectrograms of speech and infant crying signal.

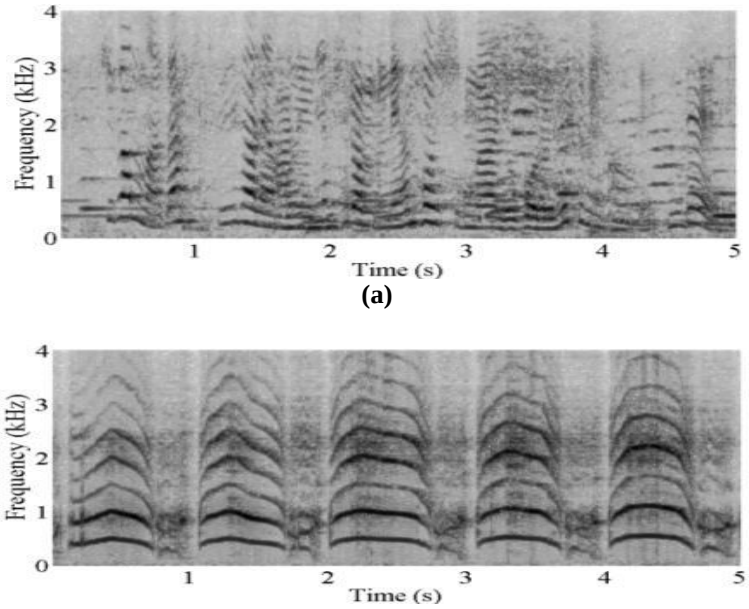



**Fig. 1.** Typical examples of spectrograms of (a) speech signal from conversation, (b) infant crying signal.

In common, the speech signals shown in Fig. 1(a) have stumpy pitch in the range of 110-310 Hz and corresponding plain harmonics in the lower frequency area below 2 kHz, whereas the harmonic formation becomes radically weaker as the frequency increases. Means, the largest part of energy tends to be focused on the lower frequency region, and the intermediary patterns of speech are diverse in that region. This is the motive why mel-scale frequency warping is talented for speech recognition. Fig. 1(b) show diverse time-frequency patterns. In general, it has far above the ground pitch of in relation to 500 Hz and has duration of in relation to 500- 700 ms. supplementary, infant crying is characterized by its sporadic nature, alternating crying and encouragement. Also, there are patent harmonic structures and inimitable melody patterns contained by the target province. Therefore, characterizing these idiosyncratic and customary patterns is important in effectively become attentive of the infant crying sounds. In this paper, we provide work for LFCC to capture the global time-frequency characteristics of newborn crying sounds. In toting up, LFCC is put forward to be a sign of local transitions. Crying signals relatively have easygoing variations with time. Thus, it is expected that LFCC can make available a enhanced results for infant's cry.

## IV MATERIALS AND METHODS

### *A. Data Collection*

The cry signals bring into played in this paper were recorded data commencing the Neonatal Intensive Care Unit (NICU) of Department of Neonatology, K.E.M. Hospital ,Pune. This cry signals of babies ranging in era sandwiched between 0 – 8 months. In order to appraise the performance of the put forward algorithm in a earsplitting environment,

we make use of quite a lot of types of noise, including engines, casual observer, motorcycles and speech signals, obtained from more than a few databases. The data is carve up into two, training data and testing data. There are 27 training data, each of which represents the 7 hungry infant cries, 4 sleepy infant cries, 10 in pain infant cries, and 6 uncomfortable infant cries.The testing data is of the 27 traning data.

### *B. METHODOLOGY*

The methodology of this research consists of more than a few phase of process: data collected works, preprocessing, codebook representation of newborn cries, testing and analysis, and interface manufacturing. The methodology of the infant cries is made known in Figure 2.

The methodology consists of subsequent steps
Step1: By making use of LFCC extracting the features from the input speech samples.
Step2: Training the features and producing the code book of them by making use of Vector Quantization technique.
Step3: Extract the features of trial sample and come to a decision on the Euclidian distance/aloofness to each and every one trained samples in code book.
Step4: Come again the sample as matched sample which is in the vicinity of the test sample.

.

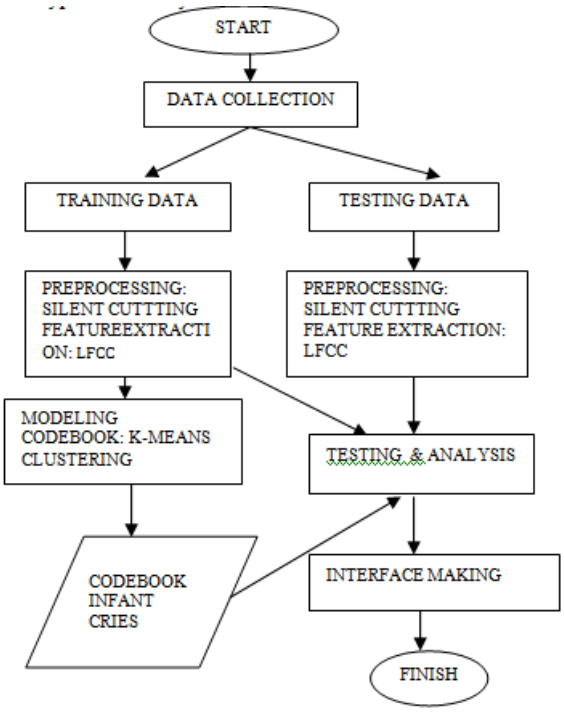



Figure 2: The methodology of identifying the meaning of a crying baby.

### *C. PREPROCESSING*

In preprocessing we do away with noises due to engines, casual observer, motorcycles and speech in attendance in the crying sound. LFCC has the narrow banded linear filter-bank.Near the beginning reflection in a room impulse response is more often than not less than 25 ms and it can be take into custody by the narrow-banded linear filter-bank in the far above the ground frequency section and do away with the cepstral mean subtraction, whereas the mel filter bank in the high frequency section is broad-banded and does not have this property. LFCC is strong in babble noise, but not in the white noise. The energy in the high frequency section of speech is as a rule pathetic and it is supplementary susceptible to noise corruption. LFCC has additional filter banks in the high frequency section and this is why it is a lesser amount of robust in the white noise than MFCC.Underneath function used in matlab for noise removal.

**function output=denoise(signal,fs,IS)**

S is the noisy signal, FS is the sampling frequency and IS is the preliminary silence length in seconds(0 .25 sec)

**Y=OverlapAdd(X,A,W,S)**

Y is the signal reconstructed signal from its spectrogram. X is a matrix with every one column being the FFT of a subdivision of signal. A is the phase angle of the spectrum which should have the same dimension as X. W is the window length.. S is the shift length of the segmentation process. Y is the reconstructed time domain signal.

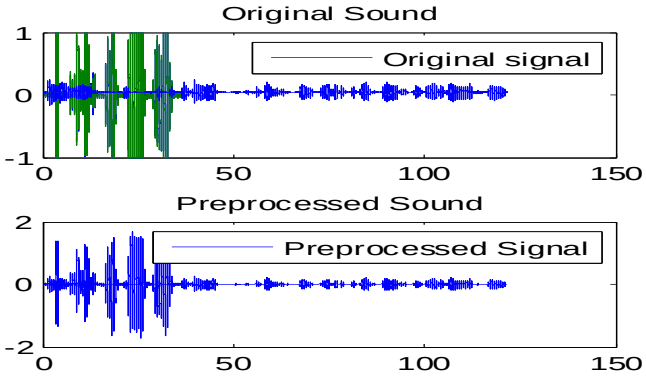



Figure 3: Preprocessed Sound and Original Sound

### *D. LFCC ALGORITHM*

The foremost step in the implementation of in the least speech recognition system is extracting the features. There are a lot of well recognized algorithms for this intention. LFCC is an well-organized feature extracting algorithm.

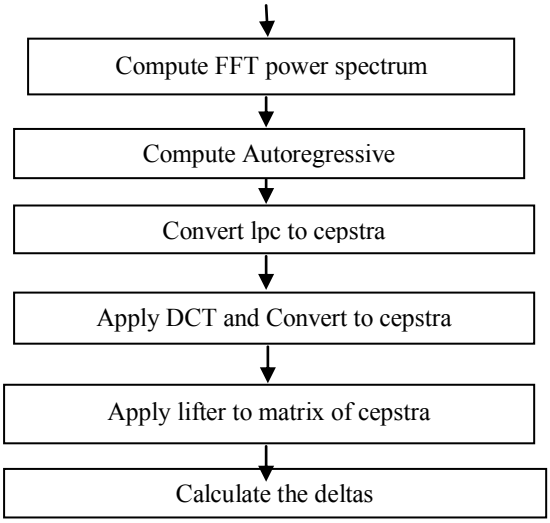



Figure 4: Flow-chart of LFCC feature Extraction.

To dig out LFCC, every audio clip is first break up into several segments composed of a fixed number of frames. The segment volume is 500 ms, which are typical period of crying sounds, and the frame size is 25 ms by means of 50% overlap to reflect time evolutions contained by the segment. Second, LFCCs are computed in every one frame. Next, LFCC can be obtained by applying a DCT to a sequence of T consecutive LFCCs along the time axis contained by the segment.

### *E. VECTOR QUANTIZATION*

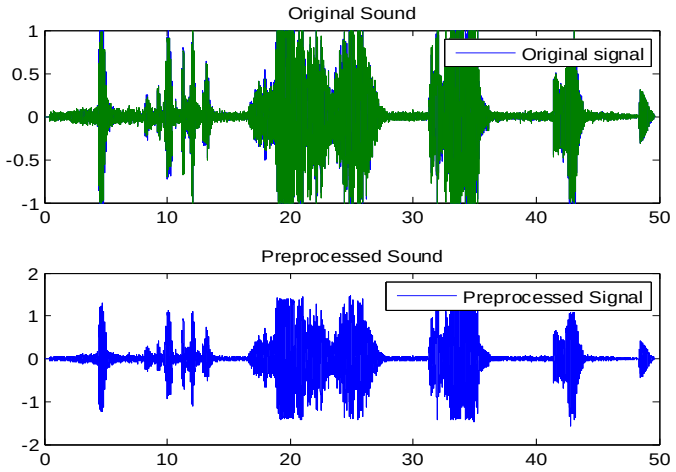
Here for Speaker recognition we are put side by side an unknown baby sound clip by means of a set of known sound clip in a database and discovering the most excellent matching speaker. Vector quantization (VQ) is a lossy data compression technique based on the principle of block coding. Vector quantization is a process of redundancy taking away that makes the effective make use of nonlinear dependency and dimensionality by compression of speech spectral parameters. In general, the use of VQ results in a lower distortion than the use of scalar quantization at the same rate. VQ is one of the preferred methods to map gigantic amount of vectors from a space to a predefined number of clusters each of

which is defined by its central vectors/centroids. In VQ a large set of feature vectors are taken and a smaller set of measure vectors is produced which correspond to the centroids of the distribution.

## V. RESULTS AND DISCUSSION

Considered two samples
1.Sample one




After LFCC

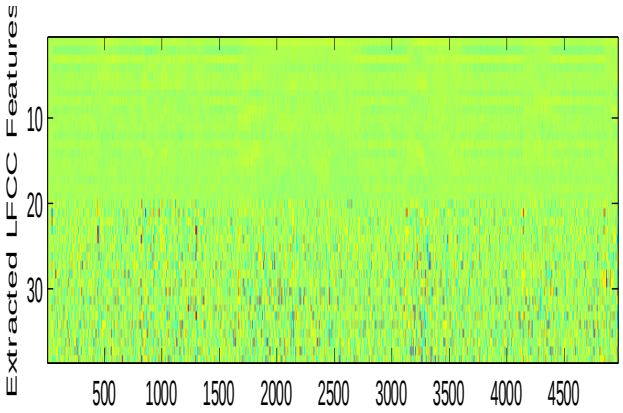



Figure 6: Representation of LFCC feature of sample one

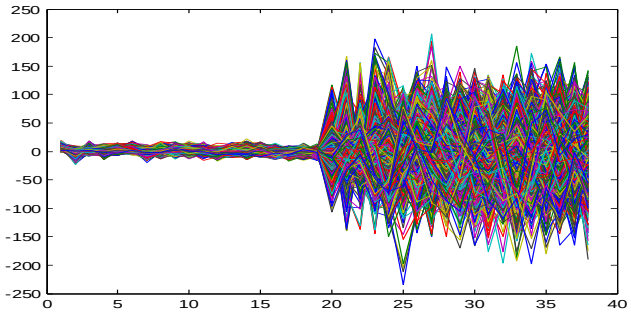
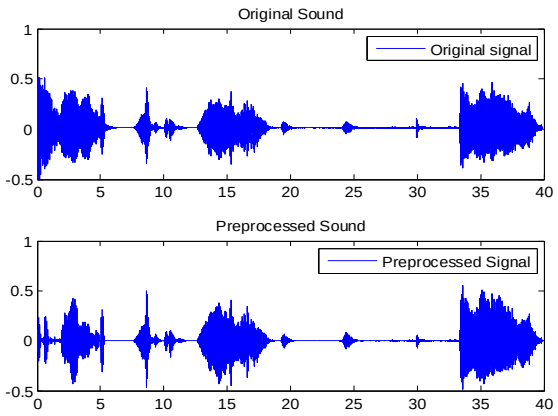


Figure 7: Codebook illustration of testing data of sample one

Speaker is Recognized and Baby Cry is Detected as Infants!

2.Sample Two




After LFCC

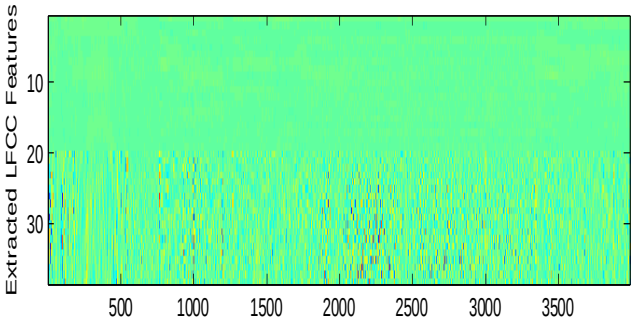



Figure 8: Representation of LFCC feature of sample two

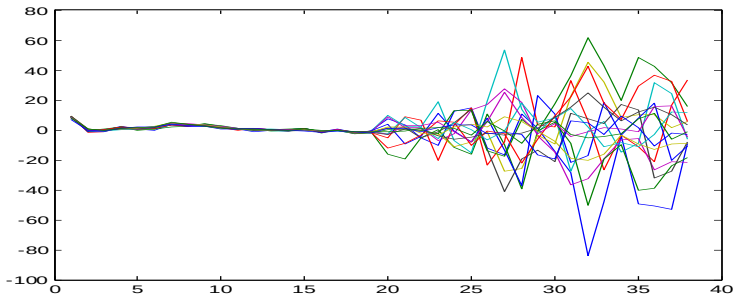


Figure 9: Codebook illustration of testing data of sample two

Speaker is Recognized and Baby Cry is Detected as Neonatal!

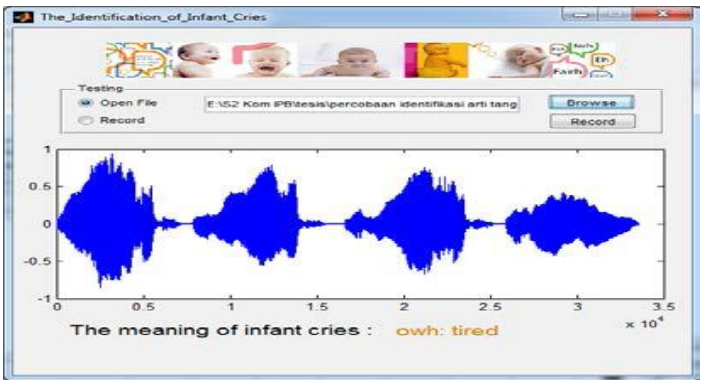



Figure 10: The Interface of Identification Of Infant Cries

***Accuracy Computation***

We have computer recognition rate for MFCC and LFCC with proposed architecture given above.

Recognition_rate = number_of_speech_recognized
/number_of_speech_presented;

| Test Cases | LFCC | MFCC |
|---|---|---|
| 1 | 92.30 % | 81.2 % |
| 2 | 89.1 % | 84.33 % |
| 3 | 93.33 % | 80.89 % |
| Average Recognition Rate | **91.58 %** | **78.8 %** |

Table 1: Recognition Rate Performance Analysis

From above table 1, it is clearly showing the performance of projected approach is improved significantly as judge against to existing MFCC method.

## VI. CONCLUSION

With the lend a hand of Codebook model and LFCC without difficulty recognize infant's baby cry and verified babies emotions by using KNN with the superior correctness. Codebook model and LFCC with the superior correctness is: frame length = 440, overlap frame = 0.4, k = 18. The distance using which create the higher accuracy is euclidean distance. That model can produce correctness recognition of infant cries with the higher about 93%. The research is just slash the silent at the beginning and at the end of speech signal. With a bit of luck, in the next research, the silent can be slash in the central point of sound so that it can produce more specific sound. It has impact on the superior accuracy as well. LFCC resulting higher formant frequencies in speech. LFCC is as robust as MFCC.

## VI. REFERENCE

[1] A. M. Prathibha, R. Putta, H. Srinivas, and S. B. Satish, "An eclectic approach for detection of infant cry and wireless monitoring of swinging device as an alternative warning system for physically impaired and better neonatal growth," *World Journal of Science and Technology*, vol. 2 no. 5, pp. 62-65, 2012.

[2] X. Zhou, D. Garcia-Romero, R. Duraiswami, C. Espy-Wilson, and S. Shamma, "Linear versus mel frequency cepstral coefficients for speaker recognition," in *Proc. Autom, Speech Recogn. Understand.*, 2011, pp. 559-564.

[3] R. Cohen, Y. Lavner, "Infant cry analysis and detection," *IEEE 27th Convention of Electrical and Electronics Engineers in Israel*, 2012, pp. 1-5.

[4] Poel M, Ekkel T. Analyzing Infant Cries Using a Committee of Neural Networks in order to Detect Hypoxia Related Disorder. International Journal on Artificial Intelligence Tools (IJAIT) Vol. 15, No. 3, 2006, pp. 397-410.

[5] Lederman D, Zmora E, Hauschildt S, Stellzig-Eisenhauer A, Wermke K. 2008. Classification of cries of infants with cleft-palate using parallel hidden Markov models. International Federation for Medical and Biological Engineeering, Vol. 46, 2008, pp. 965-975.

[6] Lien C, Huang R. Automatic Recognition of Birdsongs Using Mel-frequency Cepstral Coefficients and Vector Quantization. International MultiConference.

[7] R. M. Gray, ``Vector Quantization," IEEE ASSP Magazine, pp. 4--29, April 1984.

[8] Y. Linde, A. Buzo& R. Gray, "An algorithm for vector quantizer design", IEEE Transactions on Communications, Vol. 28, pp.84-95, 1980.

[9] Mahmoud I. Abdalla and Hanaa S. Ali "Wavelet-Based Mel- Frequency Cepstral Coefficients for Speaker Identification usingHidden Markov Models".

[10] X. Zhou, D. Garcia-Romero, R. Duraiswami, C. Espy-Wilson, and S. Shamma, "Linear versus mel frequency cepstral coefficients for speaker recognition," in *Proc. Autom, Speech Recogn. Understand.*, 2011, pp.