Perception of Deception in Speech Using Homomorphic Technique

Sheela B.P[1], Sreepathi B[2], Shilpa patil[3], Priyanka Deshpande[4]
[1, 2] Dept of ISE
[1, 2] RYMEC/Ballari

Abstract- *This work studies the effect of the emotions that is experienced due to guilt situation on different vocal parameters in attempt to identify whether or not the suspect is lying. The homomorphic speech processing is applied to extract the vocal parameters related to the source excitation such as: pitch, pitch power and vowel duration and those related to the vocal tract such as: formant frequencies and its gain. Also the energy as a global vocal parameter is computed. The vocal parameters are extracted from normal speech utterances and from stressed utterances for the same suspect in order to determine the most significant vocal parameters that can be affected by emotional stress. Correlation coefficients were investigated between the pitch power and the digitized smoothed output of the Psychological Stress Evaluator (PSE).More than 0.8 correlation coefficient has been found.*

Keywords- Deception, , Psychological Stress Evaluator (PSE), Voice Stress Analyzer (VSA). (CVSA) Computerized Voice Stress Analyzer.

I. INTRODUCTION

Deception as a psychological process (from an evolutionary perspective), may be viewed as the action of human trying to hide information's from being showed up because the consequences, if these information's are shown up, may be against himself or some body that he is covering up. Such deceiving behavior could be called lying (The CIA's Secret manual on Coercive Questioning [l]). The tools for detection of deception, for purposes ofjustice, are being developed since in the middle of the century. Scientific techniques that can distinguish the changes in human physiological functions in case of deception or intention to deceive were initiated and continued to evolve. Law enforcement agencies are repeatedly exposed to offers of technological solutions that promise to reduce their officers' workload, improve their effectiveness, and/or save lives. While the thought of having the latest and greatest technology might seem enticing for many agencies, due to the constraints resulting from decreasing budgets, most agencies cannot afford the risk associated with the high cost of experimenting with different technology. The significance of VSA technology in military applications is extensive. During military field interrogations of potential informants, VSA can be applied in a manner similar to the application administered by law enforcement: to detect deception. Also, VSA can be used to study the impact stressed speech may have on the performance of speech processing technology, such as speaker identification and language identification. If stress can be detected in voice, speech recognition algorithms can then be modified to adapt to stress speech and improves the recognition performance.

Six cases of real time criminal suspects cases were investigated throughout this work. Traditional deception detection questioning techniques were used to develop questionnaires for these criminal cases. Moreover, a case for an actor simulating different emotional states (downloaded from the Internet) was investigated for the effect of different emotions on the vocal parameters. Speech vocal parameters and the PSE Hirch&Wiegele scoring method were investigated for stress (due to anxiety or guilt). Pitch contour exhibits the most significant sensitivity for speech-based stressed/unstressed classification.

II. RELATED RESEARCH

The United States judicial system places huge weights in the belief to juries are effectual and reliable in determining the credibility of the witness. However, behavioral and social research explains that humans are good at lying and quite poor at deception detection (Vrij, 2008).

For exemplar, an average individual's capability to detect deception in a face-to-face interaction through another individual is merely modestly enhanced than chance (Ekman & O'Sullivan, 1991). Therefore, the significant of truthful testimony in addition to the inadequacy of human deception detectors encompass prompted the perennial search intended for a technology-based objective method of deception detection or else truth verification; this search continues nowadays (Grubin, 2010) polygraph, measures activity of the peripheral nervous system towards gauging truthfulness, has been the primary technical method intended for deception detection through the last century. The scenario of a deception

task refers in the direction of the hypothetical setting in experimental deception occurs. For illustration, a few experiments engage participants in a mock crime situation in addition question them about it (Kozel, 2005). Others probe participants concerning autobiographical information of dissimilar levels of intimacy (Abe, 2009). Ultimately, experiments to treated emotion, embarrassment in addition to autobiographical memory since confounds rather than variables of interest, used comparatively "neutral" scenarios to necessitated concealing possession of a playing card intended for a monetary reward. The task scenario as well determines the risk or benefit ratio of the deception experiment. The investigational deception model refers towards the method used to generate deceptive responses in addition to the appropriate controls. The two fundamental deception-generating models are the CQT (Comparison Question Test) and the GKT (Guilty Knowledge Task), as well referred to since the CIT (Concealed Information Test). These models are not unique to MRI research along with have been developed intended for forensic investigative use (Stern, 2003) through the polygraph along with later through EEG (Rosenfeld, 1988). An additional parameter of significance to the experimental deception-generating models is whether responding deceptively is being endorsed through the experimenter (Miller, 1993). Whereas in the real world, a person's deception would usually be undesirable to its target (a feature recognized to the deceiver, through definition), in most deception experiments, subjects are provided explicit instructions (i.e. endorsement) to deception to a few of the questions (Spence, 2001). Such endorsement severely limits the ecological validity of the experiment. A few deception experiments encompass attempted towards enhancing ecological validity to commence intent through allowing the subjects to choose while to deception during the task (Lee, 2002). Others have eliminated the appearance of endorsement of deception through separating the research team member who instructs participants to deception as of the rest of the team, therefore creating a "co-conspirator" (Langleben, 2005). A rather new fraud approach that is becoming more common is the use of anonymous telephone hotlines (Holtfreter, 2004). It is an extremely cost effectual for detecting occupational fraud along with abuse. A hotline permits employees to give confidential, inside information without the fear of reprisal to accompany being a whistleblower (Pergola and Sprung, 2005). Common Term Better results are provided in big data and analysis of the functions according to the volume of information. To find out the better results, in testing parameters and constraints are increased on that time. These will be accurate in results, however in some time that might be unnecessary or else irrelevant. The generation of the report is compared with the causes and matching. At present, different technologies are used to determine an individual's speech has truth percentage or else that speech have truth or deceptions. Fraudulent financial reporting and asset misappropriation have turn out to be major costs for most of the organizations. To reduce the direct and indirect costs associated with all forms of, fraud numerous fraud prevention and detection techniques are now utilized. These different techniques comprise however they are not limited to: fraud policies, telephone hot lines, employee reference checks, fraud vulnerability reviews, vendor contract reviews and sanctions, analytical reviews (financial ratio analysis), password protection, firewalls, digital analysis and other forms of software technology, and discovery sampling (Carpenter and Mahoney, 2001; Thomas and Gibson, 2003). Organizations so as to have not been fraud victims depend more on intangible prevention tools similar to code of conduct or else fraud reporting policies whereas those to have suffered fraud have implemented additional tangible measures, for instance whistle-blowing policies and fraud prevention and detection training (Price Waterhouse Coopers (PWC), 2003). Liars are so smart at present to interact through the police and investigation team. Therefore individuals require being extremely sharp compare with liars.

Polygraph Test Polygraph machines are designed to detect the motion or else the waves in an asynchronous or simultaneous process in addition to this by using at health care or medical segments headed to record changes in physiological unexpected changes on characteristics. That might be an individual's pulse rate, breathing rate per minute along with these are used at a same time for detect liars (lie detection). In the deception detection system individual has to consider a few of the most important characteristics that might be controllable with the assistance of human otherwise uncontrollable characteristics.

The sources are listed beneath,

- Direct speech by means of eye sight
- Answering with no relevant
- Answer indirectly
- Irregular speech vs. flow of regular speech
- Questionnaire preparation
- Answering speed
- Polygram verification (Variation)
- Report of EEG /ECG

Little Self-control and Sensitivity to Reward An additional interesting dissimilarity was to discover between non-criminal individuals with multiple psychopathic traits and criminal people through psychopathic traits. There is dissimilarity in the communication among the reward centre as well an area in the middle of the forebrain. Excellent

communication among these areas might appear to be a condition meant for self-control. Here the outcomes seem to indicate so as to the tendency towards committing an offence arises as of a combination of a strong focus on reward along with lack of self control. This is said to be the initial research project where convicted criminals were actually examined. Here individual can find a few of the physiological waves through the speech. A few of the similar wave files are declared in figures. The reports are generated from the Speech Analyser Software. The fuzzy logic is applicable in these terms as well.Since every question needs to finalize the result no one is not able to give percentage of truth or the percentage of deception. By receiving the answer alone one can compare with another question and its answer. In the same time through ANN concept the entire combination of question will calculate the final outcomes.

III. EXISTING SYSTEM AND PROPOSED SYSTEM

The inventors have introduced adeception detector machine that can measure the human physiological changes (such as: blood pressure and pulse rate) during an interrogation, this device is known as the Polygraph, which depends on sensors attached to the suspect during an investigation to measure the physiological variance through it. These attachments to the suspects were the reason for a set of constrains which make the polygraph test complex. Later trials from the inventors to overcome the problem of attaching equipments to the suspect were directed to the speech, since the speech conveys mental balance and the general state of functioning of the entire organism, thus a speech baseddeception detectors were introduced to the market in attempt to replace the polygraph. These speech baseddeception detectors namely were: the Psychological Stress Evaluator (PSE) and the Voice Stress Analyzer (VSA). The PSE was developed by DEKTOR Corporation of Springfield, Virginia (1970) [21. The PSE response is defined by Smith [31 as a rise of the pen from its zero base-line, followed by a number of pulses of varying amplitude at the fundamental frequency. B.F.Fuller [4] claimed that stress arousal causes a less amplitude variations as shown in Fig. 1:

Fig 1: Stressed utterance output from the PSE

Hirsch & Wiegele [6] has developed a Scoring method to enhance the poor reliability (which was reported as low as 0.38 Nachshon, et. [5]) of the PSE. Fig. 2 illustrates Hirsch & Wiegele scoring technique at which they count the number of adjacent vocal pulsations that do not differ in height and divides this number by the total number of pulsations in the voice sample. Hirsch & Wiegele have reported a reliability of 074 for their scoring technique In Fig.2 the scoring result for the left pattern is 6/4 = 1.5, the scoring for the right pattern is 1/3. = 0.33, i.e. the right pattern seems to be a stressed utterance.

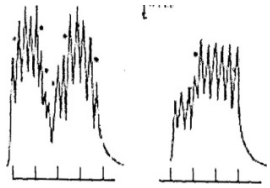

Fig. 2 Hirch&Wiegele scoring method

IV. DECEPTION DETECTION METHOD

PSYCOLOGICAL STRESS EVALUATOR AND COMPUTERISED VOICE STRESS ANALYZER:

The Dektor Psychological Stress Evaluator (PSE-) detects, measures and graphically displays certain specific, stress-related components of the human voice: superimposed on the audible voice frequencies are inaudible frequency modulations (FM) whose strength and pattern relate inversely to the degree of psychological stress in the speaker at the moment of utterance.

The VSA process of detecting deception is similar to that used in polygraph testing. VSA examiners use pre-test interviews to prepare the subject for the forthcoming test. The purpose of this process is to establish the credibility of the test procedures, develop a rapport with the subject, make observations of the subject's behavior, develop · questions for testing and breakdown any internalized barriers that may exist in the subject making admissions. CVSA is an energy-based detection system that produces a filtered waveform for evaluation and is computer based. Testing is live and multiple waveforms can be displayed on a single page. The audio is passed through the sound card and is automatically directed to the CVSA system.

V. SPEECH AND EMOTIONS

Carl Williams & Kenneth Stevens [8] reported that respiration is frequently a sensitive indicator in certain; emotional situations, such as, conscious attempts at deception, and conflict. The respiratory pattern is frequently disturbed in

anxiety states. An increase in respiration rate would presumably result in an increased subglottal pressure during speech. This height ended subglottal pressure would give rise to a higher pitch frequency Fo during voiced sounds in speech. The increased respiration rate could also lead to shorter durations of speech between breaths, with a consequent effect on the basic temporal pattern of speech. They also claimed that other relevant physiological effects of certain emotions are dryness of the mouth often observed undér conditions of emotional excitement, anticipation, fear, and anger, and tremor and disorganization of motor response, observed under conditions of emotional conflict. These effects can have an influence on various components of the speech system, including the larynx, which is directly involved in the control of Muscle activity in the larynx and the condition ofthe vocal cords are likely to have a more direct influence on the sound output and, in particular, on the fundamental frequency, than changes in muscle activity in other parts of the speech generating system, such as the tongue, lips, and jaw. The reason is that the vibrating vocal cords have a direct effect on the volume velocity through the glottis, whereas the other muscles and vocal fract components simply shape the resonant cavities for sound that is generated at the vocal cords. Such physiological changes as increased subglottal pressure, excessive dryness or salivation, and decreased smoothness of motor control can have an influence on the waveform of the pulses from the vocal cords, as well as on their frequency.

For example, increased subglottal pressure generally gives rise to a narrowing of individual glottal pulses, and hence to a change in the specmlm of the pulses. Under some circumstances, such as excessive salivation, there may be irregularities in the waveform of the glottal output from one pulse to the next (this was noticed during the data collection for this work that for the excessive salivation pitch contour is down). It has been assumed by Bonner, With regard to the consonant, the terms easy and hard, applied to the attack and release of the hyphen, refer in a very general and non-mathematical way to compare the length of the consonants beginning and end Bonner and Jones supposed that most individuals under emotional stress would speak more jerkily than is normal, i.e. would attack and release the hyphen more abruptly. They reported that since speech is a physiological process, based on the functioning of the organism at large, it is reasonable that they should find no definite onedirectional trends in any of the attributes of speech studied under emotional tension. They also reported that there is nearly always a fairly marked change in the attributes of speech under emotive strain, but the nature of the change shows wide individual differences, as has also been found in the physiological factors of pulse rate and respiration, The same results has been obtained from the studies of the physiological factors of pulse rate, respiration, etc. Darrow says that strong emotion may cause a fall rather than a rise of blood pressure. It is reported in [9] that some researchers concluded that rate of breathing under fear increases in some people and decreases in others.

Since phonetic events play an important role in the transmission of the emotional modes, Lieberman [101 wondered whether pitch frequency plays a secondary role in the presence of phonetic information and whether pitch information is immaterial or negligible in the presence of a correct and complete phonetic description of the speech material. Lieberman concluded that phonetic content, gross changes in fundamental frequency, the fine structure of the fundamental frequency, and the speech envelope amplitude, are all contributing to the transmission of the emotional modes. The different emotional modes did not all depend to the same degree on all the acoustic parameters.

VI. MODEL DESCRIPTOR

Criminal suspects

The inventors have introduced adeception detector machine that can measure the human physiological changes (such as: blood pressure and pulse rate) during an interrogation, this device is known as the Polygraph, which depends on sensors attached to the suspect during an investigation to measure the physiological variance through it. These attachments to the suspects were the reason for a set of constrains which make the polygraph test complex. Later trials from the inventors to overcome the problem of attaching equipments to the suspect were directed to the speech, since the speech conveys mental balance and the general state of functioning of the entire organism, thus a speech baseddeception detectors were introduced to the market in attempt to replace the polygraph.

Homomorphism speech

The respiratory pattern is frequently disturbed in anxiety states. An increase in respiration rate would presumably result in an increased sub glottal pressure during speech, This height ended sub glottal pressure would give rise to a higher pitch frequency during voiced sounds in speech. The increased respiration rate could also lead to shorter durations of speech between breaths, with a consequent effect on the basic temporal pattern of speech. They also claimed that other relevant physiological effects of certain emotions are dryness of the mouth often observed under conditions of emotional excitement, anticipation, fear, and anger, and tremor and disorganization of motor.

Normal speech

They reported that since speech is a physiological process, based on the functioning of the organism at large, it is reasonable that they should find no definite one directional trends in any of the attributes of speech studied under emotional tension. They also reported that there is nearly always a fairly marked change in the attributes of speech under emotive strain, but the nature of the change shows wide individual differences, as has also been found in the physiological factors of pulse rate and respiration. The same results have been obtained from the studies of the physiological factors of pulse rate, respiration, etc.

Vocal parameters

Seven case studies were collected and investigated for stress indications through relevant vocal parameters. The first case represents an actor simulating different emotional utterances was downloaded from the Internet. Cases two to seven were recorded from real crime situations. The efficiency of the different features fordeception detection has been evaluated for all the cases considered in this work. This accuracy is computed by initially determining the ensured utterances that represents deception, and then we calculate the ratio between the answers to irrelevant questions that contains significant deviations the vocal parameters and the total number of the ensured deception utterances.

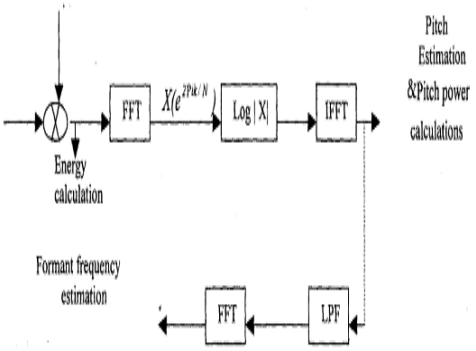

Fig.3 vocal parameters estimation

An algorithm for vocal parameters extraction is developed in this work and it is based upon the homomorphic technique. We choosed this technique because it is accurate in pitch calculations as reported by M.Nool [l ll. Fig. 3 shows the block diagram of the vocal parameters extraction via homomorphic technique. The speech signal is classified into voiced and unvoiced intervals. The voiced intervals are segmented into frames of 256 samples (23 ms) and the following steps are applied to these frames:

Step l:

1. Multiply the frame (frame i) by a Hamming window:

Y, (j) = s(j). w(j) where w(j) = $(0.54 + 0.46\cos(2\pi j/255))$

2. Compute the energy of the frame:

$$= \sum_{j=0}^{256} y_{ij}^2 \quad ,i = 1 \ldots, N$$

Where N is the total number of the frames in an utterance .

3. Compute the FFT for the frame: $Y_i(f) = \text{Real}\{\text{FT}(y_i[j])\}$

4. Compute the logarithm of the magnitude spectrum of the frame .

5. Yl(f) = 20 logl Yl(f) I .

5. Compute the Inverse FFT for the frame to get the output of the homomorphic operation (Cepstrum).

$$y_i(j) = \text{IFT}\{Y_i(f)\}$$

6. Beginning from a high time sample location (40 for male, 20 for women and children) seek for the peak in the remainder of the frame.

$$X_i = \max_{\substack{j=40 \\ \text{or } 20}}^{128} y_i(j).$$

$$P_i = \text{index } \{Xi\}.$$

$$F\,P_i = i.$$

$$Ai = \{Xi\}.$$

Where P, is the pitch period at frame i, F P, is the frame number whose pitch period is PI and Ai is the amplitude of the signal at Pi .

● Step 2: For the whole voiced interval ao the following .

1. Compute the mean energy for the utterance.

$\overline{E} = \cdot(\sum_{i=1}^{N} E_i) / N$ N = total number of frames.

2. Compute the maximum pitch Amplitude for the utterance.

Amax - MAX(AI) .

3. Compute the vowel duration in

$$Vd = 0.23 \sum_{i=1}^{N} (V/2) \quad ,V = \begin{cases} 1 & \text{for v} \geq s \\ 0 & \text{elsewhere} \end{cases}$$

Where s is a percentage threshold of Amax

4. Compute the mean of the pitch amplit

$$\overline{AP} = (\sum_{i=0}^{an} A_i)/an \ , A_i = \begin{cases} A_i & \text{for } P_i(A_{max})+5 > P_i(A_i) > P_i(A_{max})-5 \\ 0 & \text{elsewhere} \end{cases}$$

, an .... the number of P_i's that that lie in the range[11] :

$(P_i(A_{max})+5) > P_i(A_i) > (P_i(A_{max})-5)$.

Compute the H&W scores The chart of the PSE height is 40 mm (the distance allowed for the heat pen to vary the peak from its base zero-level to the maximum) and the counting score for H&W method is computed if the difference of 4 mm or more between the peak and its successor peak occurred . It is assumed that a 0.1 of the maximum pitch amplitude to be the threshold difference between every pitch amplitude peak and its successor peak.

$$H\&W = 2(\sum_{i=0}^{v} k)/v \ , \ k = \begin{cases} 1 & \text{for } |(A_i - A_{i+1})| > 0.1\,A_{max} \\ 0 & \text{elsewhere} \end{cases}$$

VII. ADVANTAGES AND DISADVANTAGES

In this work we have studied the effect of stress on the five features mentioned above. Seven case studies were collected and investigated for stress indications through relevant vocal parameters.

The first case represents an actor simulating different emotional utterances was downloaded from the Internet. Cases two to seven were recorded from real crime situations.

They reported that since speech is a physiological process, based on the functioning of the organism at large, it is reasonable that they should find no definite one directional trends in any of the attributes of speech studied under emotional tension.

They also reported that there is nearly always a fairly marked change in the attributes of speech under emotive strain, but the nature of the change shows wide individual differences, as has also been found in the physiological factors of pulse rate and respiration.

VIII. CONCLUSION

Homologue speech based stress evaluators: PSE and CVSA are actually classified by some researchers as stress arousal detectors. The CVSA may have a better chance since the operation is computerized and consequently less sensitive to the analyst experience compared to the PSE. It is noteworthy that the polygraph doesn't distinguish anxiety or indignation from guilt and suffers from the deception complegties and Consequently, the success of the results depends on the skills of convenience of analyst and the questionnaire. The addressed problem of speech-based deception detection has been investigated. We have relied upon the claim that deception does not cause a known distinctive physiological reaction in response to the relevant questions. Therefore, what we have implemented actually is a stress detector. Several case studies have been elaborated and automatic stress detector has been proposed on the basis of a pitch contour classification criterion. Classification accuracy scored higher than 90% for the implemented automatic stress detector.

REFERENCES

[1] N. Abe, T. Fujii, K. Hirayama, A. Takeda, Y. Hosokai, T. Ishioka and H. Fukuda, "Do parkinsonian patients have trouble telling lies?", The neurobiological basis of deceptive behaviour, Brain, Vol. 132, No. 5, Pp.1386-1395, 2009.

[2] B.W. Carpenter and D.P. Mahoney, "Analyzing organizational fraud", Internal Auditor, Pp. 33-38, 2001.

[3] P. Ekman and M. O'Sullivan, "Who can catch a liar?", American Psychologist, Vol. 46, Pp. 913-920, 1991.

[4] D. Grubin, "The polygraph and forensic psychiatry", Journal of the American Academy of Psychiatry & the Law, Vol. 38, Pp. 446-451, 2010.

[5] K. Holtfreter, "Fraud in US organisations: an examination of control mechanisms", Journal of Financial Crime, Vol. 12, No. 1, Pp. 88-95, 2004.

[6] J.L. Bierstaker, R.G. Brody and C. Pacini, "Accountants' perceptions regarding fraud detection and prevention methods", Managerial Auditing Journal, Vol. 21, No. 5, Pp. 520-535, 2006.

[7] F.A. Kozel, K.A. Johnson, Q. Mu, E.L. Grenesko, S.J. Laken and M.S. George, "Detecting deception using functional magnetic resonance imaging", Biological Psychiatry, 2005.

[8] D.D. Langleben, J.W. Loughead, W.B. Bilker, K. Ruparel, A.R. Childress, S.I. Busch and R.C. Gur, "Telling truth from lie in individual subjects with fast event-related fMRI", Human Brain Mapping, Vol. 26, No. 4, Pp. 262-272, 2005.

[9] T.M. Lee, H.L. Liu, L.H. Tan, C.C. Chan, S. Mahankali, C.M. Feng and J.H. Gao, "Lie detection by functional magnetic resonance imaging", Human Brain Mapping, Vol. 15, No. 3, Pp. 157-164, 2002.

[10] G.R. Miller and J.B. Stiff, "Deceptive Communication", Deceptive Communication, Vol.14, Pp. 32–39, 1993.

[11] C.W. Pergola and P.C. Sprung, "Developing a genuine anti-fraud environment", Risk Management, Vol. 52