Enhancing Voice Print using Computer Vision through Optical Character Recognition with Subscriber Identity Module

[1]SADO – ALIMIKHENA ELIZABETH, [2]UMAR ADAM ISAH, [3]ADOGA EMMANUEL ALEONOLU,
[4] OLOTU YAHAYA, [5]GARUBA ISAH ANIRU
SCHOOL OF ENGINEERING TECHNOLOGY, AUCHI POLYTECHNIC AUCHI, AUCHI, EDO STATE, NIGERIA.
[1, 2]MECHANICAL ENGINEERING TECHNOLOGY
[3]DEPARTMENT OF CIVIL ENGINEERING TECHNOLOGY
[4]AGRICULTURAL ENGINEERING TECHNOLOGY
[5]ELECTRICAL / ELECTRONIC ENGINEERING TECHNOLOGY

Abstract:- This research focuses on voiceprint collection of fully registered subscriber identity module using Smartphone to generate computer coded language inform of identification audio image signal text. It is propagated at a speed of 4.21m/s, frequency $0.048H_Z$ and wavelength of 7083.3 λ through a vacuum to form constructive wave signal with light source from the Smartphone screen. At this point, a voice pattern with word syllable stress with be captured and transmitted into an articulated text binary value number with digitizer to convert analog calls into digital barcode text. The subscriber identity module card will be seal with the user signature for further sequence and chart with fire wall through an audio image signal feedback. This process was made possible through smart phone location notification street map summary. To generate accurate biometric of recorded speech into an image for crime scrutiny. The optical character recognition efficacy was merged with fingerprint call answering mode which act a synchronous biometric. So that, deep learning and convolutional neural network work out solutions with the algorithm to differentiate true tone from fake tone by air flow, air filter and audible pulses to know the exact caller for identity through vocal tract to reduce crime rate with the aid of internet cloud.

Keywords:- Voiceprint, Computer Vision, Optical Character Recognition, Subscriber Identity Module.

I. INTRODUCTION

Voiceprint is a process of transmitting voice signal into constructed sequence of word character. By using adaptive algorithm that is friendly to interact with human being preferably in computer language called code. The ability of computer to think deals with artificial intelligence which is called computer vision. Computer vision is a field of an artificial intelligence which permit computers and systems to collect useful information from [1] digital image, videos and other inputs to give an assessment with network service provider experts assistance such as mobile telephone network (MTN), affectionate interested respectful tolerant energetic and loving (Airtel) and global communications limited (Glo) companies based on the given command. In this process, the subscriber identity module (SIM) must complete a full registration process through fingerprint [2]

biometric and voice random sampling routine. Voiceprint is a unique behavioral identity of human being produced inform of sound propagation as an energy wave. Voiceprint has certain element which differentiate one person to another by the rate of opening and closing of the vocal tract. Due to the high crime rate within Africa and other foreign countries this research is very significant for scientist and engineering to seek empathy. Concurrent engineering material should be initiated with an integration to seek smart devices prototype [3] installation on building wall and street map signal. In cases of insecurity, fraud and cyber crime, swift synchronous web data is necessary to be installed in cloud [4] by fire walls or sacrificial shield with an encryption barcode.

- **Significance of voiceprint**
  - ➢ It is used for data sharing and approval workflows
  - ➢ It is used for password setting and validation process
  - ➢ It is used for telephonic payment processes
  - ➢ It is used for an interactive voice response system
  - ➢ It assist disable people to enroll into bank and other biometric process

When making voiceprint three things should be considered:
- Frequency
- duration
- amplitude.

However, language and voiceprint has many advantages compared to other biometric technology. People might pronounce word with same sound but vary with slight syllables[stress alphabets. It means voiceprint has more secrecy[5] than other means of biometric open biometric process. That is, why this research combine voiceprint into text image for coding into computer vision for users privacy. Voiceprint isolator was inserted in to the Smartphone mouth piece to avoid poor sensing of sound when making phone calls to minimize density and temperature effect which is a major problem in sound wave propagation. Sound speed from the vocal tract changes slightly in [6] space and travel less in air compared to liquid and solid media. The voiceprint convert sound into electrical signal through an acoustic vacuum in the smart phone mouth piece into digital print for further sequence activation. Voiceprint is a biometric technology in terms of

handwriting [7-8] on a screen touch, stamping by fingerprint on paper (bank chequebook) and iris showing to camera .Areas where computer vision are applicable are in [9] quality assurance, image detection, surveillance, cinema graphics, vital organ study such as x ray, safety in airline and automobiles.

II. METHODOLOGY

A. Optical character recognition And Computer Vision Using Subscriber Identity Module

A qualitative methodology of phone call was used to detect identity of voice wave samples. To get an accurate result a finger print was used to [10] software called audio image signal (AIS) which is compatible with all internet phone. The apparatus used is 360º camera reader, smart phone touch screen to enable the full collection of fingerprint image and analytic thinking to initialed by computer vision. The validation process are orderly recorded through an interface at early open page [11-12] of the software at the telecommunication service outlet centre to create database record in cloud.

www.Voiceprint biometric.com

- On the modify toolbar click fingerprint thumb and index finger for image code processing, Click session to define voiceprint active status.
- Press enter to finish fingerprint and voiceprint programming and authentication[13-16] with encryption
- Enter signature for each typing speed pulses
- On the command session make a call for five minutes to fully generate a secret question and answer incase when your phone get loss or stolen for easy recovery process.
- Enter your landmark address for street map[17-19] activation and agreement into management information system protection protocol
- Identity information executed press ok with the result given in the table below

Feedback loop	Word syllable stress captured	Image error signal failed	Instant location traced user
Unregistered subscriber identity module (sim)	10	80	N/A
registered subscriber identity module	60	7	97

Table 1: Evaluation Of Audio Image Signal Software

III. RESULTS AND DISCUSSION

Given voiceprint as word with binary code: 76205

A block chain will open for biodata cloud file[21-24] then follow the command prompt

Input full Name

↓

Code the full name into binary number value

↓

Create sequence of barcode with the above details into microfluidic chip file

↓

Generate token identity image and text strings with encryption

↓

Decision feedback loop

↓ Transport layer Security (TLS)

Reset biodata

↓

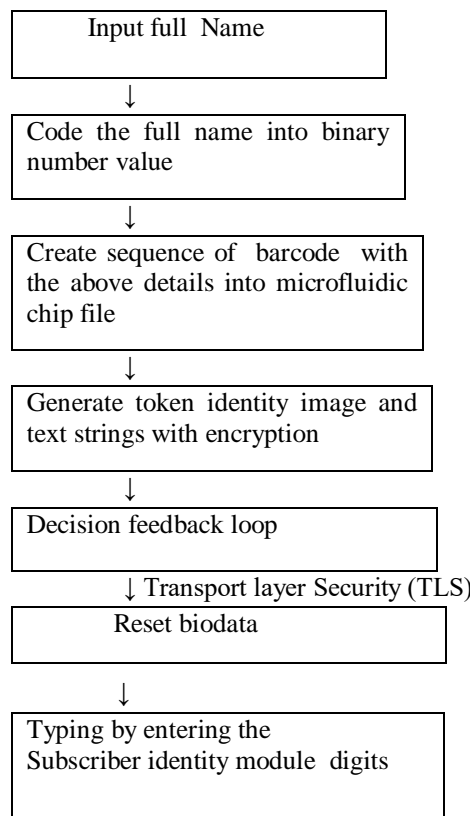Typing by entering the Subscriber identity module digits

Fig 1: Block Flowchart Diagram Validation Process And Outcome

Each process work with audio image signal (AIS) for easy recognition. The reason behind this concept is preventing the attacker to fool and [25-27]gain access into the owner subscriber identity module without been traced with fingerprint shield then okay the button of the smart phone or internet phone at the time of calls making This idea is a synchronous forensic control block chain that protect victims in time of emergency.

A. Factors Affecting Call Detection Intelligence

There are many constraint the basic one is temperature and pressure. Sound moves in a vacuum so it frequency remain constant and wavelength. The lower the temperature the slower the speed and the higher temperature the faster the speed. As the speed increases the wavelength[28-32] increases .When the speed decreases the wavelength decreases. A simple experiment was conducted for call detection intelligence quadrant sampling within different locations with subscriber telecommunication for identification using automatic scanner[33-34] with three local area network (LAN) the results are given below:

IV. DISCUSSION OF RESULTS

Distance (m) (d_1)	Distance (m) (d_2)	Distance(m) $(d_{1)}+(d_{2)/2}$	Time (sec) t_1	Time (sec) t_2	Time (sec)$t_{1}+$(sec) $t_{2/2}$
10.00	8.12	9.10	1.20	1.10	1.15
20.00	6.30	13.15	2.03	2.50	2.27
30.00	4.10	17.05	4.47	3.30	4.00
40.00	3.40	21.70	6.33	5.05	6.00
50.00	1.00	25.50	4.06	10.06	7.10
total=150	22.90	86.50	18.09	22.01	20.52

Table 2: Variation Of Distance And Time Of Voiceprint Under Variable Constraint

Voiceprint were taken on various pitch vibration rate, location and design composite wall materials to check the distance and time behavior. As the distances increases at first[35-36] noise vibration, the distance for next decreases with less noise vibration this phenomenon disagrees with the law of superposition in physics which that when two or more waves cross at a point, the displacement at that point is equal to the sum of the displacement of the individual wave vector. Normal scene provides good results than where noise density is higher From the table 2 there are both constructive and destructive wave pattern. It was found that the speed of the voiceprint (sound) was 4.21m/s lesser than the constant speed of gas 340 m/s at sea level 20ºc due to smoke generated pollutant in the air by Brownian motion.

V. CONCLUSION

Voiceprint has a significant role in biometric process to verify individual and crime. This research merged physiological and behavior method using voiceprint with call answering mode alongside fingerprint press button for swift match of images into text code in cloud to defined full details of location call. Using encrypted secret signature of data recovery process even the third party regulating the process would be panic about identity theft by internet grooming or phishing. With this cyber crime, fraud and forensic swift detection is scale up to 97 percent positive output.

VI. RECOMMENDATION

For future research, sophisticated hardware and software should be developed to accurately detect and communicate with live human being without contact with any biometric device such as barriers like composite wall with poor wifi would hinder the efficacy performance.

REFERENCES

[1.] Abraham A., Dote Y., Furuhashi T., Köppen, M., Ohuchi A. and Ohsawa Y. (Editors.), Soft Computing as Transdisciplinary Science and Technology (series: Advances in Soft Computing), Springer-Verlag Berlin Heidelberg, Germany, pp. 192-201, 2005.

[2.] Abdulla, W. H., Chow, D., & Sin, G, Cross-words reference template for DTW-based speech recognition systems.In TENCON Conference on Convergent Technologies for the Asia-Pacific Region (Vol. 4, pp. 1576-1579). IEEE, October, 2003.

[3.] Al-Qatab, B. A., & Ainon, R. N. Arabic speech recognition using hidden Markov model toolkit (HTK). In Information (June)Technology (ITSim), International Symposium in 2010 (Vol.2, pp. 557-562). IEEE.

[4.] Campbell, W., Sturim, D., & Reynolds, D. Support vector machines using GMM supervectors for speaker verification. IEEE Signal Processing Letters, pp.13, 308-311, 2006.

[5.] Cai, S., Jin, X., et al. Subband energy regularized perceptual linear prediction coefficient for noise robust speech recognition. Acta acoustics Sinica, (6): 667-672, 2012.

[6.] Dumitru, C. O., & Gavat, I. A Comparative Study of Feature Extraction Methods Applied to Continuous Speech Recognition in Romanian Language. In Multimedia Signal Processing and Communications,

48th International Symposium ELMAR-2006 focused on June,2006. (pp. 115-118). IEEE.

[7.] D. Yu and L. Deng. "Automatic Speech Recognition- A deep learning approach", Springer- Verlag London, 2015

[8.] Da, Z., Li, Q., et al. The influence of different recording systems on voiceprint detection. Journal of Nanjing University (NATURAL SCIENCE), 2011, (02): 201-207.

[9.] Dehak, N., Kenny, P., et al. Front-End Factor Analysis for Speaker Verification. IEEE Transactions on Audio, Speech, and Language Processing, . 2011 Vol.19,pp. 788-798

[10.] Furui, S., Ichiba, T., Shinozaki, T., Whittaker, E. W., & Iwano, K. Cluster-based modeling for ubiquitous speech recognition. Inter speech 2005, 2865-2868.

[11.] Hinton et.al . "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups", IEEE signal process. Magazine, vol. 21, no. 6, pp. 82-97, 2012. GIN-DER WU AND YING LEI, "A register array based low power FFT processor for speech recognition. "JOURNAL OF INFORMATION SCIENCE AND ENGINEERING, 2008.

[12.] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition", IEEE Transactions on Acoustic Speech and Signal Processing, vol. 26, no. 1, pp. 43-49, 1978.

[13.] Heigold, G., Moreno, I., et al. End-to-end text-dependent speaker verification. 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2016 ,pp. 5115-5119.

[14.] Juang, B. H., & Rabiner, L. R. Automatic speech recognition–A brief history of the technology development. Encyclopedia of Language and Linguistics, 2005.

[15.] King, S., Frankel, J., Livescu, K., McDermott, E., Richmond, K., & Wester, M Speech production knowledge in automatic speech recognition. The Journal of the Acoustical Society of America,2007, 121(2), 723-742.

[16.] K. A. Qazi, T. Nawaz et.al " A hybrid technique for speech segregation and classification using a sophisticated deep neural network", PLoS ONE 13(3):e0194151, , March, 2018, Vol.20, pp. 1-15.

[17.] L. Deng and X. Li. "Machine learning paradigms for speech recognition: An overview", IEEE transactions on Audio, Speech and language processing, 2013 vol. 21, no. 5,pp. 1060-1089.

[18.] K. Kumar and R. K. Aggarwal "A Hindi speech recognition system for connected words using HTK", Int. J. Computational systems Engineering, 2012, vol. 1, no.1.

[19.] Lu, X.Q., Guan, S.X. Real time voiceprint recognition based on VQ and GMM. Computer system application,2014. 23(09): 6-12. Morales, N., Hansen, J. H., & Toledano, D. T.. MFCC Compensation for Improved Recognition of Filtered and Band-Limited Speech. In ICASSP . 2014, (1) (pp. 521-524).

[20.] N. Morgan et.al " Pushing the envelope- Aside [Speech Recognition]", IEEE Signal Processing Mag, 2005 vol. 22, no. 5, pp. 81-88.

[21.] N. Smith and M. J. F. Gales, "Using SVM's and discriminative models for speech recognition", Proc. ICASSP, 2002, vol. 1, pp. 77-80.

[22.] P. Kannadaguli, P. Bhat "A Comparison of Gaussian mixture modeling (GMM) and Hidden Markov modeling (HMM) based approaches for automatic phoneme recognition in Kannada",IEEE, 2015 pp. 257-260.

[23.] Qu, F.M., Liu, J.H., et al A Multi-fault Detection Method with Improved Triplet Loss Based on Hard Sample Mining. IEEE Transactions on Sustainable Energy, 2020, 99: 1.

[24.] [24] R. K. Aggarwal and M. Dave Using Gaussian Mixtures for Hindi Speech Recognition System", International Journal of Signal Processing and Pattern recognition, 2011 vol.4, no. 4 , December [25] S. Katagiri, " Speech pattern recognition using neural networks", W. Chou and B. H. Juang (Eds.) Pattern recognition in Speech and language Processing, CRC Press, 2003 pp. 115-147.

[25.] Shanthi, T. S., Lingam, C.. Isolated word speech recognition system using htk. International Journal of Computer Science Engineering and Information Technology Research, 2014 4(2), 81-86.

[26.] Saeed K. Minimal-Eigenvalue-Based Face Feature Descriptor. In: Dramiński M., Grzegorzewski P., Trojanowski K., Zadrożny S. (Eds), Issues in Intelligent Systems Models and Techniques. Institute of System Research, Polish Academy of Sciences, Akademicka Oficyna Wydawnicza EXIT, Warsaw, Poland, 2005, pp. 185-196.

[27.] Saeed K., Nammous M., "A Simple Speech-and-Speaker Identification System," Accepted for publication in IEEE Trans. on Industrial Electronics, IEEE Computer Society, 2006.

[28.] Saeed K., Nammous M., "Heuristic Method of Arabic Speech Recognition," Proc. 7th Intern. Conf. on Digital Signal Processing and its Applications – IEEE-DSPA'05, Moscow, Russia, 2005 pp. 528-530

[29.] Saeed K., Kozłowski M "An Image-Based System for Spoken-Letter Recognition. In: Petkov N. and Westenberg M. (Eds), Lecture Notes in Computer Science – LNCS 2756, Springer-Verlag Heidelberg, Germany,2003 pp. 494-502006.

[30.] W. M. Campbell, D. E. Sturim et.al. 2006 "The MIT-LL/IBM speaker recognition system using high performance reduced complexity recognition" MIT Lincoln Laboratory IBM

[31.] Xu, W.Y. A fast discrete wavelet transform algorithm and its application in speech signal.Computer engineering and application, 2011, 47(35): 143-146.

[32.] Xin, W., Zhang, H.R Robust i-vector speaker recognition method based on DNN processing. Computer Engineering & Applications, 2018.

[33.] Xie, T, Zheng, X.D., et al. Analysis of seismic facies based on cepstrum coefficient of linear prediction. Acta geophysica Sinica, (11).ICAACE 2021 Journal of Physics: Conference Series .

[34.] Zhu, S.X. Research on voiceprint recognition system and pattern matching algorithm.Dalian University of technology, 2006.

[35.] Z. Yu et.al " The Teaching Experiment of Speech Recognition based on HMM", IEEE,2014 pp. 2416-2420.