Journal of Theoretical and Applied Information Technology
20th February 2016. Vol.84. No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

AN INTRINSIC REAL-TIME MULTIMODAL RECOGNITION SYSTEM USING DEEP NEURAL NETWORKS

1 J.SEETHALAKSHMI, [2]C.JAYAKUMAR

[1] PhD Research Scholar, Bharathiar University, Coimbatore, India Assistant Professor., Department of Computer Application, S.S.K.V College of Arts and Science for Women, Kanchipuram,

[2] PhD Research Supervisors, Bharathiar University, Coimbatore, India, Professor, Department of CSE, RMK Engineering College, Kavaraipetai
E-mail: [1]seetha2207@gmail.com, [2]cjayakumar2007@gmail.com

ABSTRACT

In this paper, an embedded finger-vein and voice recognition system for authentication on ATM network is proposed. The system is implemented on an embedded platform and equipped with a novel finger-vein and voice recognition algorithm. The proposed system consists of four hardware modules: radio frequency identification system, image acquisition module, embedded main board, and human machine communication module. RFID module will start the very initial communication between the user and the device. The proposed system is designed for applications where the training database contains Finger-vein and voice for each individual. The final decision is made by multi model architecture. The multimodal system is developed through fusion of finger-vein and voice. These images of the user are acquired by the image acquisition module of the hardware. The finger-vein recognition algorithm is coded and targeted as firmware of the embedded main board. Human machine interface module communicates with the user through keyboard and displays the outcome of the match in human recognizable formats like LCD or LED displays as well as voice message. The RFID module enhances the security and credibility of the whole system by giving short message service (SMS) through GSM technology platform to the actual user. This system is tested on IITK database and overall accuracy of the system is found to be more than 99% accurate respectively.

Keywords: *Embedded Finger Vein Recognition, Voice Recognitions System, RFID Module, Embedded Platform, Communication Module, Image Acquisition Module, Deep Neural Network Etc.*

1. INTRODUCTION

Today, security is very much essential in all kind of activities. Illegal activities are happening in every place today. So government and corporate sections are concentrating mainly on the security levels with their every invention. This will bring privacy all over the world. So in a thought of bringing privacy through security level [2], this project has been developed. This FVIR system mainly uses three divisions which are image acquisition module, embedded main board, and human machine communication module. Each unit is having its own major role over the project. In this research, two major areas have been focused. Those are authentication [5] and identification. FVIR system performs the authentication function with the finger vein and Ear recognition. Every time when the user is going to use the system, the finger vein of will be scanned and comparison will be done.

2. LITERATURE SURVEY

The input to the ANN is the value of exponent of reactive power load-voltage characteristic (nq) and the output is the desired proportional gain (KP) and integral gain (KI) parameters of the SVC. Normalized values of nq are fed as the input to the ANN the normalized values of outputs are converted into the actual value. The process of Existing authentication systems have evolved over several decades . Each system has been very carefully scrutinized by security experts as well as the end users in terms of the level of security provided by the system as well as the inconvenience caused to the end user. Personal Identification Number [PIN] based authentication system[1] is found to have major flaws like

Journal of Theoretical and Applied Information Technology
20th February 2016. Vol.84. No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

complete breach of security once the PIN is revealed as well as the agony to the end user to memorize the PIN. Biometrics came as the potent solution to this problem of PIN based system. Biometric patterns [2] like iris, face, palm/fingerprints, voice etc have been developed to attempt the for ultimate security that cannot be breached. Notwithstanding the increasing complexity of implementation of the biometric based authentication system, no biometric system has yet been developed that provides a reasonably good convenience to the end user but at the same time provides impeccable defense against security breach. For example patterns are unique and are medically proven to be stable after a year or so after birth. But it suffers from the inconvenience to the end user especially those using contact lenses as well as elderly people. Fingerprints and palm prints based authentication systems suffer from being frayed and do not ensure that specimen being authenticated is from a live body. Face recognition system pose difficulty when the person has different expression or face-lift [3] from that in the pre stored data.

Hence it has been a demanding task for authentication system developers to set a high threshold of identity validation that is impregnable to the ever adapting deceptive practices [4]. A successful authentication system that provides this security at the same time which provides fairly good convenience to end users will have several wide ranging application from consumer electronics, banking, airport, business establishments to space and defense segments.

3. FINGER-VEIN AND VOICE RECOGNITION SYSTEM

3.1 Finger vein recognition

Finger vein recognition is very effective when compared with pattern recognition, pin number security the other type of Biometric[1] security methods like finger print security, palm print security, image scanning and some recognition techniques. FVVR system uses the vein scanning. As it is related to the biological factor, it is very difficult to change the vein information of a user. So, this system can provide more security[3] than any other security level. In this FVVR system, we are focusing on high security[4] with RFID technology. Initially each and every user will be given with one RFID secret card. This will make an effective initial communication between the user and the device. This technique will make the device to extract the user information from its memory.

There are four stages in finger vein authentication. These have analogues in most biometric techniques:
1. Capture of the finger vein image pattern
2. Normalization of the image
3. Feature pattern extraction from the image
4. Pattern matching and outcome decision

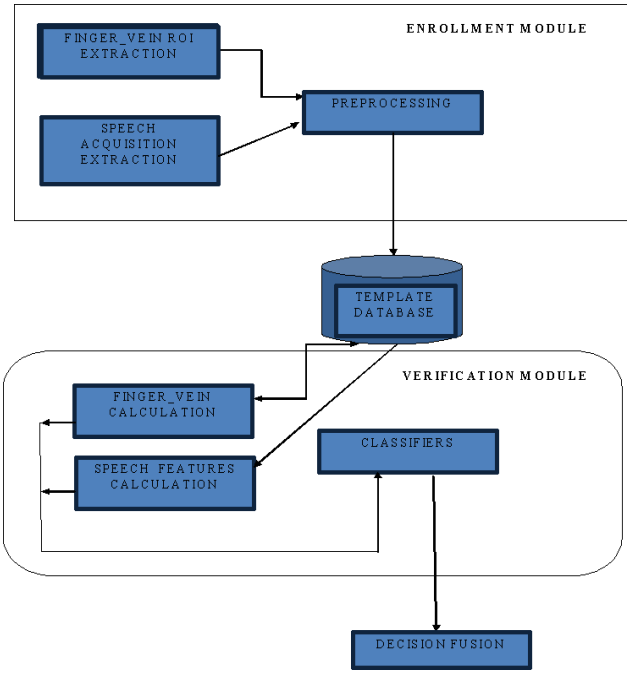


Fig 1:Proposed Multimodal Traits System

3.1.1 Image Acquisition

The CCD IR bullet camera comprises of IR LED source with an array of 36 IR LED's and required 12V DC power supply and has a 6mm lens which supports the IR rays with a maximum penetration distance of 30mm. The LED source is used to pass the IR rays in the finger, to get the finger vein images.

3.1.2 Image Segmentation and Alignment

Image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image (see edge detection). Each of the pixels in a region is similar with respect to some characteristic or computed property, such as color, intensity, or texture. The original image is captured with the black unwanted

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

background. Including the background reduced the accuracy of the original image, because the position of finger usually varies across different finger vein images, it is necessary to image segmentation in region of interest (ROI) before feature extraction and matching with database. The bone in the finger joint is articular cartilage. Unlike other bones, it can be easily penetrated by light source. When a finger is irradiating by the uniform light source, the image joint is brighter than the other part and is show in.

Fig 2: The ROI of the Finger-vein

3. 1.3.Image Enhancement

The segmented finger-vein image is then enhanced to improve
its contrast .The image is resized to 1/4 of the original size, and enlarged back to its original size. Next, the image is resized to 1/3 of the original size for recognition. Bicubic interpolation is used in this resizing procedure. Finally, histogram equalization is used for enhancing the gray level
contrast of the image.

3.1.4.Feature Extraction

The fractal model developed by Mandelbrot [7] provides an excellent method for representing the ruggedness of natural surfaces and it has serve data successful image analysis tool for image compression and classification. Since different fractal sets with obviously different textures may share the same fractal dimension[8] Fractal Dimension allows us to measure the degree of complexity by evaluating how fast our measurements increase or decrease as our scale becomes larger or smaller . Two types of fractal dimension: self similarity dimension and box-counting dimension .Essentially, data behave with a power law relationship if they fit the following equation:

y=c*x^d where c is a constant.

One way to determine if data fit a power law relationship is to plot the log(y) versus the log(x). If the plot is a straight line, then it is a power law relationship.

3.2 Voice recognition

3.2.1. Deep Neural Network HMM

A deep neural network (DNN) is a conventional multi-layer perceptron (MLP, [8]) with many hidden layers, optionally initialized using the DBN pre-training algorithm. In the following, we want to recap the DNN from a statistical viewpoint and describe its integration with HMMs for speech recognition. For a more detailed description, please refer to [6].

3.2.2.. DBN Pre-Training

The deep belief network (DBN), proposed by Hinton [11], provides a new way to train deep generative models. The layerwise greedy pre-training algorithm developed in DBN was later found to be also effective in training DNNs. The DBN pre-training procedure treats each consecutive pair of layers in the MLP as a *restricted Boltzmann machine* (RBM) [11] whose joint probability is defined as

$$P_{h,v}(h,v) = 1 . \frac{e^{V^T W h + V^T b + a^T h}}{Z_{h,v}}$$

2.3. Integrating DNNs with HMMs

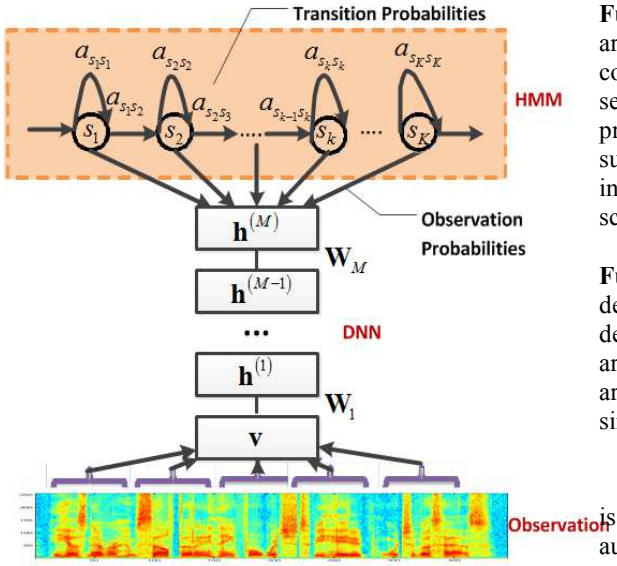
Following the traditional ANN-HMMs of the 90's [1], we replace the acoustic model's Gaussian mixtures with an MLP and compute the HMM's state emission likelihoods $P_{o|s}(o|s)$ by converting state posteriors obtained from the MLP to likelihoods:

$$P_{o|s}(o|s) = P_{s|o}(s|0) . \text{const}(s). \quad (2)$$
$$\frac{}{P_s(s)}$$

Here, classes *s* correspond to HMM states, and observation
vectors *o* are regular acoustic feature vectors augmented with neighbor frames (5 on each side in our case). $P_s(s)$ is the prior probability of state*s*. This is a critical factor in achieving the unusual accuracy improvements in this paper.

Journal of Theoretical and Applied Information Technology
20th February 2016. Vol.84. No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195



3. TRAINING DNN-HMMS

In this section, we will describe the process and some practical
considerations in training DNN-HMMs.

3.2.3. Basic Training Process

DNN model learning begins with the DBN pre-training (section 2.2), using one full sweep through the 309 hours of training data for all hidden layers but the first, where we use two full sweeps. Slight gains may be obtained if the pre-training procedure sweeps the data additional times. However, this seems to be not critical [5]. RBMs are not scale-invariant, so the training
corpus is normalized to zero mean and unit variance [13]. The alignment is updated once during training.

4. MULTIMODAL BIOMETRICS SYSTEM

Multimodal Biometrics are looked to as a means of reducing false match and false match rates. The levels fusions proposed [2] for multimodal system are broadly categorized into three system architectures which are according to the strategies used for information fusion as shown in figure.

- Fusion at the Feature extraction level
- Fusion at the Matching Score level
- Fusion at the Decision level

In Fusion at the Feature extraction level, information extracted from the different sensors is encoded into joint feature vector, which is then compared to an enrollment template and assigned score as in a single biometric system.

Fusion at the Matching Score level, feature vectors are created independently for each sensor and are then compared to the enrollment templates which are stored separately for each biometric trait. Based on the proximity of feature vector and template, each subsystem computes its own matching score. These individual scores are finally combined into a total score which is passed to the decision module.

Fusion at the Decision level, a spate authentication decision is made for each biometric trait. These decisions are the combined into final vote. This architecture is rather loosely coupled system architecture, with each subsystem performing like a single biometric system.

Finger vein and Voice recognition (FVVR) is very effective when compared pin number based authentication and other types of Biometric security methods like finger print security, palm print security, image scanning and some recognition techniques. Finger-vein being hidden inside the human body,
is difficult to be duplicated. Also, it is very convenient for the end user due to non-intrusive nature of acquisition. FVVR system uses the vein scanning. As it is related to the biological factor, it is very difficult to change the vein information of a user. So, this system can provide more security than any other security level. In this FVVR system, we are focusing on high security with RFID technology. Initially each and every user will be given with one RFID secret card. This will make an effective initial communication between the user and the device. This technique will make the device to extract the user information from its memory. The transmitter section of the FVVR system is shown in Figure 5 and the receiver section in Figure 6.

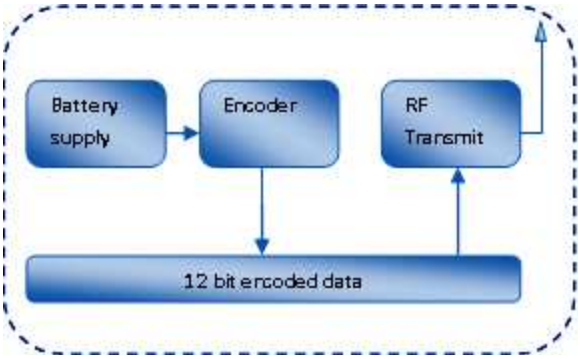


Fig 5: FVVS – Transmitter Section

In FVVR system, the RFID module is used to collect the user database. With this system, a unique code will be generated for each and every user for storing the finger vein details in the server. Here an active RFID technology is used for creating the secret signal. The

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

encoded signal will be continuously transmitted by the card if it is in on state. This RFID will reduce the complexity of the image acquisition module. As the RFID have an unique signal it can store only one vein information. So authentication and identification will become soon. Because of these features the FVVRS will be a faster recognition system.

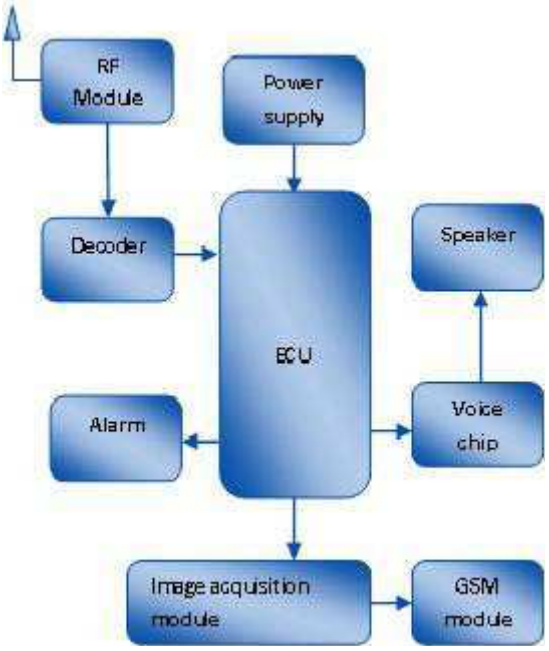


Fig 6: FVVRS –receiver section

In this section, RF receiver module is attached with the embedded control unit. This unit receives the secret digital data from the card and it will be given to the controller. In the FVVRS recognition unit, vein images will be stored in the image acquisition module. If RF receiver receives any digital code, then automatically code verification will be done inside the embedded control unit. If the code is matched then an asynchronous command will be given to the image acquisition module. Then the vein image comparison will be done inside the processor. If the image is matched then automatically the device will go to its working state. To this ECU further a GSM module is interfaced. With this module a password system can be developed. During every successful access, automatically intimation will be given to the controlling authority. The unit will send a password with this intimation. It will make a very effective security to the user. This password will be working for one time. It will play an effective authentication process. This mobile GSM communication module will not only send the intimation for authorize but also for unauthorized.

5. SYSTEM HARDWARE ARM PROCESSOR

The ARM7 family includes the ARM7TDMI, ARM7TDMI-S, ARM720T, and ARM7EJ-S processors. The ARM7TDMI is the industry's most widely used 32-bit embedded RISC microprocessor solution. Optimized for cost and power-sensitive applications, the ARM7TDMI solution provides the low power consumption, small size, and high performance needed in portable, embedded applications. The ARM7TDMI core uses a three-stage pipeline to increase the flow of instructions to the processor. allows multiple simultaneous operations to take place and continuous operation of the processing and memory systems. As the processor is having a high speed it is easy to make the communication between the RF module and the Image acquisition module Operating modes The ARM7TDMI core has seven modes of operation:

- User mode is the usual program execution state
- Interrupt (IRQ) mode is used for general purpose interrupt handling
- Supervisor mode is a protected mode for the operating system
- Abort mode is entered after a data or instruction pre fetch abort.

The interrupt setting of ARM supports the DHLS to response to the interrupt coming from the server section. Interrupt controller The Vectored Interrupt Controller (VIC) accepts all of the interrupt request inputs from the home server section and categorizes them as Fast Interrupt Request (FIQ), vectored Interrupt Request (IRQ), and non-vectored IRQ as defined by programmable settings. These interrupt settings will give a quick response to the RF decoder. So that address verification will be very faster and signal for image processing will be given to the image acquisition module.

6. WIRELESS COMMUNICATION:

RF communication:
Radio Frequency, any frequency within the electromagnetic spectrum associated with radio wave propagation. When an RF current is supplied to an antenna, an electromagnetic field is created that then is able to propagate through space. Many wireless technologies are based on RF field propagation

Journal of Theoretical and Applied Information Technology
20th February 2016. Vol.84. No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Transmitter:

The TWS-434 extremely small, and are excellent for applications requiring short-range RF remote controls. The TWS-434 modules do not incorporate internal encoding. If simple control or status signals such as button presses or switch closures want to send, consider using an encoder and decoder IC set that takes care of all encoding, error checking, and decoding functions The transmitter output is up to 8mW at 433.92MHz with a range of approximately 400 foot (open area) outdoors. Indoors, the range is approximately 200 foot, and will go through most walls.

Figure 3: RF Transmitter

RF receiver:

RWS-434: The receiver also operates at 433.92MHz, and has a sensitivity of 3uV. The WS-434 receiver operates from 4.5 to 5.5 volts-DC, and has both linear and digital outputs. A 0 volt to Vcc data output is available on pins. This output is normally used to drive a digital decoder IC or a microprocessor which is performing the data decoding. The receiver's output will only transition when valid data is present. In instances, when no carrier is present the output will remain low. The RWS-434 modules do not incorporate internal decoding. If you want to receive Simple control or status signals such as button presses or switch closes, you can use the encoder and decoder IC set described above. Decoders with momentary and latched outputs are available

Figure 4: RF receiver GSM

7. GSM MODEM

A GSM modem is a wireless modem that works with a GSM wireless network. Global system for mobile communication (GSM) is a globally accepted standard for digital cellular communication. GSM is the name of a standardization group established in 1982 to create a common European mobile telephone standard that would formulate specifications for a pan-European mobile cellular radio system operating at 900 MHz GSM provides recommendations, not requirements. The GSM specifications define the functions and interface requirements in detail but do not address the hardware. The reason for this is to limit the designers as little as possible but still to make it possible for the operators to buy equipment from different suppliers. The GSM network is divided into three major systems: the switching system (SS), the base station system (BSS), and the operation and support system (OSS). The basic GSM network elements are shown in below *figure.*

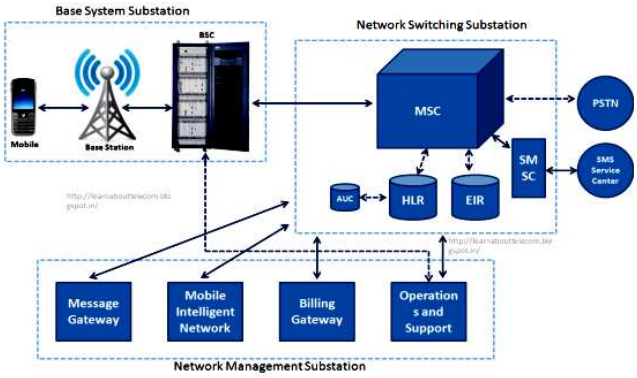


Figure 5: GSM network Topology

GSM modems support an extended set of AT commands. These extended AT commands are defined in the GSM standards. With the extended AT commands, you can do things like:

- ❖ Reading, writing and deleting SMS messages.
- ❖ Sending SMS messages.
- ❖ Monitoring the signal strength.
- ❖ Monitoring the charging status and charge level of the battery.
- ❖ Reading, writing and searching phone book entries.

Sending the message : To send the SMS message, type the following command: AT+CMGS="+31638740161" <ENTER> Replace the above phone number with your own cell phone number. The modem will respond with: >

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

(Response from the modem) You can now type the message text and send the message using the <CTRL>-<Z> key combination: Hello World ! <CTRL-Z> Here CTRL-Z is keyword for sending an sms through the mobile device.After some seconds the modem will respond with the message ID of the message, indicating that the message was sent correctly. CMGS:

8. IMAGE AQUISTION

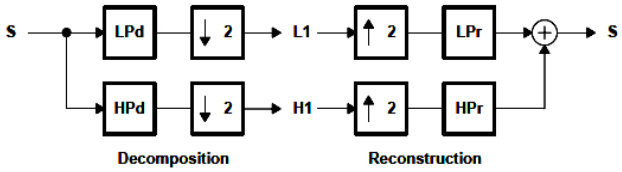
A color model is an abstract mathematical model describing the way colors can be represented as tuples of numbers, typically as three or four values *color components* (e.g. RGB and CMYK are color models). However, a color model with no associated mapping function to an absolute color spaceis a more or less arbitrary color system with no connection to any globally understood system of color interpretation.
Adding a certain mapping function between the color model and a certain reference color space results in a definite "footprint" within the reference color space. This "footprint" is known as a gamut, and, in combination with the color model, defines a new **color space**. For example, Adobe RGB and RGB are two different absolute color spaces, both based on the RGB model.
In the most generic sense of the definition above, color spaces can be defined without the use of a color model. These spaces, such as Pantone, are in effect a given set of names or numbers which are defined by the existence of a corresponding set of physical color swatches. This article focuses on the mathematical model concept.

9. RECOGNITION

Applications range from tasks such as industrial machine vision systems which, say, inspect bottles speeding by on a production line, to research into artificial intelligence and computers or robots that can comprehend the world around them. The computer vision and machine vision fields have significant overlap. Computer vision covers the core technology of automated image analysis which is used in many fields. Machine vision usually refers to a process of combining automated image analysis with other methods[12] and technologies to provide automated inspection and robot guidance in industrial applications.
As a scientific discipline, computer vision is concerned with the theory behind artificial systems that extract information from images. The image data can take many forms, such as video sequences, views from multiple cameras[11], or multi-dimensional

10. WAVELET ANALYSIS

The discrete wavelet transform (DWT) was developed to apply the wavelet transform to the digital world. Filter banks are used to approximate the behavior of the continuous wavelet transform. The signal is decomposed with a high-pass filter and a low-pass filter. The coefficients of these filters are computed using mathematical analysis and made available to you.



Where

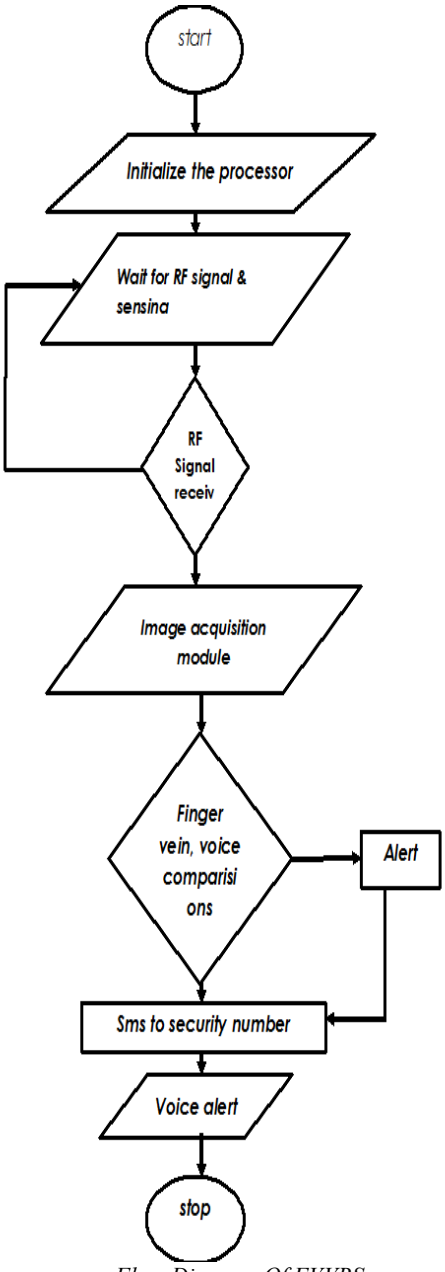
- LPd: Low Pass Decomposition Filter
- HPd: High Pass Decomposition Filter
- LPr: Low Pass Reconstruction Filter
- HPr: High Pass Reconstruction Filter

11. HAAR WAVELET ANALYSIS

In mathematics, the Haar wavelet is a sequence of rescaled "square-shaped" functions which together form a wavelet family or basis. Wavelet analysis is similar to Fourier analysis in that it allows a target function over an interval to be represented in terms of an ortho normal function basis. The Haar sequence is now recognized as the first known wavelet basis and extensively used as a teaching example. DESIGN FLOW The flow diagram of FVVRS- Mobile unit is given below. It shows all the step by step function of finger vein recognition system. Initially the device will wait for an RF signal from the user to activate the communication between the embedded control unit and the image acquisition unit. Then the finger vein image of the user will be compared with the unique data base image. Then the authentication result will be send to the security number of the user. Firstly initialize the image from the data base through the matlab then the image is resize to1/3 size for low noise image and go for histogram for the enhancement of the image and compare the image with user and database, if image is not compared then security

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

number is get to the mobile, if image is compared then security number is get to the mobile through voice alert. According to the security number The transaction of ATM is opened according to Bank name &Pin Number. After entering the pin number the transaction is being processed.



Flow Diagram Of FVVRS

10. CONCLUSION

Security is becoming essential in all kind of application. This research is implemented in a way to improve the security level. As the finger-vein and Voice is a promising biometric pattern for personal identification in terms of its security and convenience. The non-invasive and contactless capture of finger-veins and voice ensures both convenience and hygiene for the user, and is thus more acceptable. It has been shown both qualitatively and quantitatively that finger vein and Voice recognition The proposed multimodal biometric identification and authentication system is considered a robust combination of finger vein and voice that have matching score level with accuracy of 99.5%. So this system is more hopeful in improving the security level.

REFRENCES:

[1] A. K. Jain, S. Pankanti, S. Prabhakar, H. Lin, and A. Ross, ―Biometrics:a grand challenge‖, Proceedings of the 17th International Conference onPattern Recognition (ICPR), vol. 2, pp. 935-942, 2004.

[2] P. Corcoran and A. Cucos, ―Techniques for securing multimedia content inconsumer electronic appliances using biometric signa tures,‖ IEEE Transactionson Consumer Electronics, vol 51, no. 2, pp. 545-551, May 2005

[3] P. J. Phillips, A. Martin C. L. Wilson and M. Przybocki,―An Introduction to Evaluating Biometric Systems,‖ IEEEComputer, Vol.33, No.2, Feb. 2000, pp. 56-63.

[4] S. Pankanti, R. M. Bolle and A. Jain, ―Biometrics: TheFuture of Identification,‖ IEEE Computer, Vol.33, No.2,Feb. 2000, pp. 46-49.

[5] H. Lee, S. Lee, T. Kim, and Hyokyung Bahn, ―Secure user identificationfor consumer electronics devices,‖ IEEE Transactions on ConsumerElectronics, vol.54, no.4, pp.1798-1802, Nov. 2008.

[6] D. Wang , J. Li, and G. Memik, ―User identification based on finger vein patterns for consumer electronics devices‖, IEEE Transactions on Consumer Electronics, vol. 56, no. 2, pp. 799-804, 2010.

[7] D. Mulyono and S. J. Horng, ―A study of finger vein biometric for personal identification‖, Proceedings of the International SymposiumBiometrics and Security Technologies, pp. 134-141, 2008.

[8] Y. G. Dai and B. N. Huang, ―A method for capturing the finger-veinimage using nonuniform intensity infrared light‖, Image and SignalProcessing, vol.4, pp.27-30, 2008

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[9] X. Sun, C. Lin, M. Li, H. Lin, and Q. Chen, ―A DSP-based finger veinauthentication system‖, Proceedings of the Fourth InternationalConference on Intelligent Computation Technology and Automation,pp.333-336, 2011.

[10] D. D. Hwang and I. Verbauwhede, ―Design of portable biometricauthenticators - energy, performance, and security tradeoffs,‖ IEEETransactions on Consumer Electronics, vol. 50, no. 4, pp. 1222-1231,Nov.2004.

[11] N. Miura, A. Nagasaka, and T. Miyatake, ―Feature extraction of fingerveinpatterns based on repeated line tracking and its application topersonal identification, Machine Vision Application, vol. 15, no.4,pp.194–203, 2004.

[12] W. Song, T. Kim, H. C. Kim, J. H. Choi, H. Kong and S. Lee, ―Afinger-vein verification system using mean curvature‖, PatternRecognition Letters, vol. 32, no.11, pp. 1541-1547, 2011.

[13]. L. Fallah, H. Shahhosseini, and F. Setoudeh , "voice Recognition Using Neural Network",Proceedings of the international Multi-conference of Engineers and Computer Scientists (IMECS) , 2010.

[14] A. Azizi and H. R. Pourreza , " Efficient voice Recognition Through Improvement of Feature Extraction and subset Selection" , (IJCSIS) International Journal of Computer Science and Information Security, Vol. 2, No.1, June 2009.

[15]. Melin, and O. Castillo, "A New Method for Adaptive Control of Non-Linear Plants Using Type-2 Fuzzy Logic and Neural Networks", International Journal of General Systems, Taylor and Francis, Vol. 33, 2004, pp. 289-304

[16] G. Hinton, S. Osindero, and Y. Teh, "A Fast Learning Algorithmfor Deep Belief Nets", Neural Computation, vol. 18, pp. 1527– 1554, 2006.