Journal Homepage: - www.journalijar.com
INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR)
ISSN NO. 2320-5407
Article DOI: 10.21474/IJAR01/3460
DOI URL: http://dx.doi.org/10.21474/IJAR01/3460

RESEARCH ARTICLE

REVIEW ON BIOMETRICS TECHNOLOGIES.

Rakesh Dhull.
Chottu Ram Institutions of Law Rohtak, Haryana, India.

Manuscript Info

Manuscript History

Received: 19 December 2016
Final Accepted: 17 January 2017
Published: February 2017

Abstract

This paper presents the introduction of Biometric and multimodal biometric system. Then this paper addresses various advantages and applications of Biometric system. It also discusses about various existing algorithm for image processing and voice recognition.

Introduction:-

Security system can be used for Identity, Security and Time & Attendance. These systems can be used for all of your sites to accurately record time & attendance of all of your workers/solders including Home, Office and border.

Why do we use face and voice recognition?

1. Universal: Everyone has a face and voice, everyone can enroll.
2. Non-intrusive: A contact is not needed for verification process.
3. Incredibly Fast: Verification takes approximately 5 seconds.
4. Accurate: Better than humans at verifying identity, and able to work 24 hours a day.
5. Dependable: Successfully deployed in challenging real – life environments, overcoming usual biometric obstacles such as dust, dirt, grease, variable lighting and user co-operation.

Biometric system:-

Biometric consists of methods for uniquely recognizing humans based upon one or more intrinsic physical or behavioral traits. A biometric system is operates by acquiring biometric data from an individual. We are extracting a feature set from the acquired data, and comparing this feature set against the template set in the database. Depending on the application context, a biometric system may operate either in verification mode or identification mode. In the verification mode, the system validates a person's identity by comparing the captured biometric data with her own biometric template(s) stored system database. In the identification mode, the system recognizes an individual by searching the templates of all the users in the database for a match. Biometric system have now been deployed in various commercial, civilian and forensic applications as a means of establishing identity. These systems rely on the evidence of fingerprints, hand geometry, iris, retina, face, hand vein, facial thermo gram, signature, voice, etc. to either validate or determine an identity. (11) Most biometric systems deployed in real-world applications are unimodal, i.e., they rely on the evidence of a single source of information for authentication. These systems have to contend with a variety of problems such as: Noise in sensed data, Intra – class variations etc. (11)

Examples of some of the biometric traits associated with an individual Fig.1.1

(a) Fingerprint,(b) face,(c) hand geometry,(d) signature,(e) iris and (f) retina

Corresponding Author:- Rakesh Dhull.
Address:- Chottu Ram Institutions of Law Rohtak, Haryana, India.

Some of the limitations imposed by unimodal biometric systems can be overcome by including multiple sources of information for establishing identity. Such systems, known as multimodal biometric systems, are expected to be more reliable due to the presence of multiple, (fairly) independent pieces of evidence. These systems are able to meet the stringent performance requirements imposed by various applications.

Multimodal biometrics:-

A multimodal biometric system uses multiple applications to capture different type of biometrics. This allows the integration of two or more types of biometric recognition and verification systems in order to meet stringent performance requirements. A multimodal system could be, for instance, a combination of fingerprint verification, face recognition, voice verification and smart-card or any other combination of biometrics. This enhanced structure takes advantage of the proficiency of each individual biometric and can be used to overcome some of the limitations of a single biometric (14).

A multimodal system can combine any number of independent biometrics and overcome some of the limitations presented by using just one biometric as your verification tool. Multimodal are generally much more vital to fraudulent technologies, because it is more difficult to forge multiple biometric characterstics than to forge a single biometric characterstics. Multimodal biometric systems provide anti-spoofing measures by making it difficult for an intruder to simultaneously spoof the multiple biometric traits of a legitimate user. By asking the user to present a random subset of biometric traits (e.g., right index and right middle fingers, in that order), the system ensures that a "live" user is indeed present at the point of data acquisition. Thus, a challenge-response type of authentication can be facilitated using multimodal biometric systems.

Modes of Operation:-

A multimodal biometric system can operate in one of three different modes: serial mode, parallel mode, or hierarchial mode. In the serial mode of operation, the output of one biometric trait is typically used to narrow down the number of possible identities before the next trait is used. This serves as an indexing scheme in an identification system. For example, a multimodal biometric system using face and fingerprints could first employ face information to retrieve the top few matches, and then use fingerprint information to converge onto a single identity. This is in contrast to a parallel mode of operation where information from multiple traits is used simultaneously to perform recognition. This difference is crucial. In the cascade operational mode, the various biometric characterstics do not have to be acquired simultaneously. Further, a decision could be arrived at without acquiring all the traits. (14)

Applications of biometric systems:-

The applications of biometrics can be divided into the following three main groups:

Government applications such as national ID card, correctional facility, driver's license, social security, welfare – disbursement, border control, passport control, etc.

Forensic applications such as corpse identification, criminal investigation, terrorist, identification, parenthood determination, missing children, etc. (14)

Still-image versus video:-

In the literature, two main forms of face recognition exist: Still image based face recognition and video face recognition. Still image face recognition relies on classifying an individual based on a single image obtained from a still shot camera. Conversely, video based face recognition relies on a sequence of frames to extract more information about the face of a subject.

An inherent advantage of using still-image-based face recognition over video based systems is that the images are of higher resolution. As a result, current face recognition algorithms are able to recognize a face more accurately. Further to this, still image based recognition is useful in controlled environments where pose and illumination are relatively fixed. One example of such an environment is while taking subjects photograph at the airport check in (8). The disadvantages of still-image-based face recognition occur when such a controlled environment is not easily attainable. An example of this scenario would be a security camera used to identify a subject in a public place. In this case, video-based recognition yields better results.

The clear advantage of video-based face recognition occurs in situations where the image resolution is low and the video feed is continous. Video-based algorithms capitalize on both spatial and temporal variations in a subjects face.

Nevertheless, natural disadvantage is the low resolution of the images being captured (10). Since an individual might be located at a distance, the pixels that represent this individual's face might not constitute a sufficient information base for the algorithm to operate correctly. Hence, the need for the two different approaches occurs in different situations.

Algorithms for face Recognition:-

Principle component analysis:-

PCA is an algorithm develop by Turk and Pentland that treats face recognition as a two dimensional recognition problem(1). The correctness of this algorithm relies on the fact that the faces are uniform in posture and illumination. PCA can handle minor variations in these two factors, but performance is maximized if such variations are limited. The algorithm basically involves projecting a face onto a face space, which captures the maximum variation among faces in a mathematical form. During the training phase, each face image is represented as a column vector, with each entry corresponding to an image pixel. These image vectors then normalized with respect to the average face. Next, the algorithm finds the eigenvectors of the covariance matrix of normalized faces by using a speed up technique that reduces the number of multiplications to be performed. This eigenvector matrix is then multiplied by each of the face vectors to obtain their corresponding face space projections. Lastly, the recognition threshold is computed by using the maximum distance between any two face projections (8). In the recognition phase, a subject face is normalized with respect to the average face and then projected onto face space using the eigenvector matrix. Next, the Euclidean distance is computed between this projection and all known projections. The minimum value of these comparisons is selected and compared with the threshold calculated during the training phase. Based on this, if the value is greater than the threshold, the face is new. Otherwise, it is a known face (8)

Linear Deiscriminant Analysis (lda):-

Another popular algorithm used in face recognition is LDA. Although this algorithm was initially develop for data classification, it has been adapted to face recognition. Whereas PCA focuses on finding the maximum variation within a pool of images. LDA distinguishes between the differences within an individual and those among individuals. That is, the face space created in LDA gives higher weight to the variations between individuals than those of the same individual. As a result, LDA is less sensitive to lighting, pose, and expression variations (22). The drawback is that this algorithm is significantly more complicated than PCA. As an input, LDA takes in a set of faces with multiple images for each individual. These images are labeled and divided into within-classes. The former captures variations within the image of the same individual while the latter captures variations among classes of individuals. LDA thus calculates the within-class scatter matrix and the between-class scatter matrix, defined by two respective mathematical formulas. Next, the optimal projection is chosen such that it "maximizes" the ratio of the determinant of the between –class scatter matrix of the projected samples to the determinant of the within-class scatter matrix of the projected samples"(7). This ensures that the between-class variations are assigned higher weight than the within-class variations. To prevent the within-class scatter matrix from being singular, PCA is usually applied to initial image set. Finally, a well known mathematical formula is used to determine the class to which the target face belongs. Since we have reduced the weight of inter-class variation, the result will be relatively insensitive to variations.

Independent component analysis (ica):-

ICA is the third mathematically-based algorithm for face recognition. Whereas PCA depends on the "pair wise relationships between pixels in the image database," ICA strives to exploit "higher-order relationships among pixels." (8) That is, PCA can only represent second-order inter-pixel relationships, or relationships that capture the amplitude spectrum of an image but not its phase spectrum. On the other hand, ICA algorithms sue higher order relationships between the pixels and are capable of capturing the phase spectrum. Indeed, it is the phase spectrum that contains information which humans use to identify faces (3). The ICA implementation of face recognition relies on the infomax algorithm and represents the input as an n-dimensional random vector. This random vector is then reduced using PCA, without losing the higher order statistics. Then, the ICA algorithm finds the covariance matrix of the result and obtains its factorized form. Finally, whitening, rotation, and normalization are performed to obtain the independent components that constitute the face space of the individuals. Since the higher order relationships between pixels are used, ICA is robust in the presence of noise. Thus, recognition is less sensitive to "lighting conditions, changes in hair, make-up, and facial expression"(4).

Neutral networks:-

Unlike the above three algorithms, the neutral networks algorithm for face recognition is biologically inspired and based on the functionality of neurons. The perceptron is the neural network equivalent of a neuron. Just like a neuron sums the strengths of all its electric inputs, a perceptron performs a weighted sum on its numerical inputs. Using these perceptions as a basic unit, a neural network is formed for each person in the database. The neural network usually consists of three or more layers (8). An input layer takes in a dimensionally reduced (using PCA) image from the database. An output layer produces a numerical value between 1 and -1. In between these two layers, there usually exist one or more hidden layers. For the purposes of face recognition, one hidden layer usually provides a good balance between complexity and accuracy. Including more than one such layer exponentially increases the training time, while not including any results in poor recognition rates. Once the neural network is formed for each person, it must be trained to recognize that person. The most common training method is the back propagation algorithm (8). This algorithm sets the weights of the connections between neurons such that the neural networks exhibits high activity for inputs that belong to the person its represents and low activity for others. During the recognition phase, a reduced image is placed at the input of each of these networks, and the network with the highest numerical output would represent the correct match. The main problem with neural networks is that there is no clear method to find the initial network topologies. Since training takes a long time, experimenting with such topologies become a difficult task (8). Another issue that arises when neural network are used for face recognition is that of online training. Unlike PCA, where an individual may be added by computing a projection, a neural network must be trained to recognize an individual. This is a time consuming task not well suited for real – time applications.

Genetic algorithms:-

Another biologically inspired algorithm that is commonly used for face recognition is the Genetic Algorithm (GA). While neural networks mimic the function of a neuron, genetic algorithms mimic the function of chromosomes. Like neural networks, genetic algorithms are only well suited for the recognition of a limited number of individuals and are generally not too scalable. To start with, the images are divided into two classes: those that belong to the target person and those that belongs to other people. Each of these images is transformed into a binary coded truth table. Within each of the above mentioned classes, the image 13 are further subdivided into F-tables and T-tables, where each image occupies a row in the table. Initially, the rows in F-tables and T-tables do not match. However, by gradually changing some of the F-table value to don't –cares, some rows end up matching with each other. Hence, the F-table obtains the generalization ability. The evolution process ensures that the modified F-table includes as many rows in the T-table as possible. Once evolution is complete, the modifications that the result in the best fitness are chosen for each category (target person and unknown people) and applied to the F-table (15). During the recognition phase, the input image is passed through the tables that correspond to both categories. Two counters keep track of the number of pixel matches in each of the categories and the counter with the highest value classifies the input face as belonging to the corresponding category (15). The obvious drawback of this algorithm is that entire tables have to be created whenever a new individual is to be detected. As in neural networks, the scalability of this algorithm is hindered by the exponential complexity involved when training for multiple target faces.

Issue with face Recognition:-

Although face recognition systems have advanced remarkable over the past few years, there still exist some major obstacles that need to be overcome. In general, still image face recognition accuracy fades away as image variations are increased. The main image variations are illumination levels, pose variation, and changes in facial expression. Moreover, the problem of face detection, or the extraction of a face from an image, is a required first step for face recognition. The illumination problem occurs in an uncontrolled environment where "the same face appears different due to a change in lighting" (5). The problem is emphasized when the variations in lighting are greater than the variations between people. One solution to this problem involves preprocessing the images and introducing contrast normalization and compensation. Another approach attempts to reconstruct all possible lighting variations from a selection of training images for each individual. A third method relies on creating a separate linear illumination subspace. This is similar to the space created to capture face variations, except that it captures lighting variations (8). Pose variation also impairs the face recognition process. Pose variation becomes especially pronounced when it is combined with illumination changes. One solution to the pose variation problem involves obtaining images with multiple views of an individual. In this case, multiple process are available during both training and recognition. During the recognition process, each pose is aligned with a similar pose in the database to achieve correct classification. The obvious drawbacks are that multiple views of an individual are not always available. A more popular solution involves using multiple poses during training but only a single pose during recognition. One such implementation creates and Eigen spaces for each pose to achieve pose-invariant recognition

(8). The problem of facial expression variation is also common in the literature. If only one image of an individual is available, recognition accuracy drops considerably. However, if many images are available, algorithms like PCA can absorb these changes. It is important to note that during expression changes, parts of the face remain largely unchanged. As a result, algorithms that segment the face are more robust to these variations (8). Many databases available today contain training images with multiple expressions, and face recognition systems have been capable of making accurate image classifications despite expression variations. Lastly, it is important to discuss face detection in the context of the face recognition problem. The need for face detection arises when one or more faces must be extracted from an image. Furthermore, face detection and extraction is essential to reduce external factors that might hinder the recognition process. One common method of face detection relies on the use of Haar classifiers. These classifiers sweep through the image and apply several filters to detect the presence of a face. Another method, mentioned earlier, relies on skin color to detect a face. As such, face recognition is a growing field with potential applications in security, entertainment, and personal identification. The recognition algorithms can be grouped in to mathematical/statistical (PCA, ICA, LDA) algorithms and biological (NN, GA) algorithms.(22) Many of these algorithms have been implemented by several researchers with high recognition rates and recognition times within the margin of real time applications. However, long training times and the scalability of face recognition has been a recurring concern in all of these implementations. Finally, common face recognition problems include illumination changes, pose variations, and the issue of face detection and extraction.

Speech Recognition:-

Speaker recognition is basically divided in to two-classification: speaker recognition and speaker identification and it is the method of automatically identifying who is speaking on the basis of individual information integrated in speech waves (6). Speaker recognition is widely applicable in use of speaker's voice to verify their identity and control access to services such as banking by telephone, database access services, voice dialing telephone shopping, information services, voice mail, security control for secret information areas, and remote access to computer AT and T and TI with Sprint have started field tests and actual application of speaker recognition technology; many customers are already being used by Sprint's Voice Phone Card (7). Speaker recognition technology is the most potential technology to create new services that will make our everyday lives more secured. Another important application of speaker recognition technology is for forensic purposes. Speaker recognition has been seen an appealing research field for the last decades which still yields a number of unsolved problems.

Algorithms for Speech Recognition:-

There are a lot of approaches to speech recognition. Algorithms like zero crossing, energy, measurement and feature extraction are based on the acoustic-phonetic approach. Algorithms such as template matching come under the pattern recognition approach, while algorithms that depend on knowledge sources, stochasticity of speech signals and neural networks are based on the artificial intelligence approach. However, an important approach to speech recognition is stochastic modeling, in particular stochastic using Hidden Markov Models (25) in conjunction with the Viterbi algorithm. Among these the most popular and accurate algorithm is the template based Dynamic Time Warping (25).

Zero-crossing and energy-based speech Recognition:-

As compared to other approaches, zero-crossing and energy-based recognition systems require far fewer computations and fewer sets of parameters (30). For the purpose of recognition, sets of parameters are obtained from several frames of speech data. Parameters such as average zero-crossing rate, density of zero-crossings within frames, excess threshold duration, standard deviation of the zero-crossing within frames, mean zero-crossing within frames, and energy estimates for each frame have all been used. These parameters are then compared with fixed thresholds to determine the spoken word. For example, the zero crossing rate at the beginning of words starting with strong fricatives is higher than for words starting with weak consonants. This classifies a set of words into two groups and makes the process of recognition easier. A system develop by Chok-ki Chan (20) is based on the parameters such as zero crossing and energy. It is a speaker dependant, isolated Cantonese digit and words, limited vocabulary SR system, developed and implemented on a PC-386, with a recognition accuracy of 97.2%. This system works reasonably well for isolated digits but when tested to isolated words, the accuracy dropped to 76% 920). One more example of such a system is the one developed by Chan Y.T. (29) on an MC68000 microprocessor based system. It was a speaker independent, isolated-word, limited vocabulary SR system. The system when operated in speaker-dependent mode had an accuracy of 87%, but when operated in speaker independent mode, the accuracy dropped down to 78%. The system when developed for a 10 word vocabulary took around 6 minutes (19) to recognize a word but with inclusion of every 3 words in the vocabulary the recognition speed dropped by 40%.

Feature-dependent speech recognition:-

Feature-dependent speech systems are based on principles of human speech perception. These systems try to mimic human performance in recognizing speech. Some of the features that are used in such system are: frequency location of the first few formants, the maximum and minimum frequency of the first few formants, the duration of a periodic energy, formant transitions and the ratio of high frequency energy to low frequency energy. To obtain the above sets of features, parameters such as spectrum, pitch, zero crossings, total energy, energy in low,mid, and high frequency bands are produced using signal processing routines. The decision structure is arrived at by first selecting a set of measures that are likely to provide a clean grouping of letters into a set of classes. A recursive clustering algorithm is used to determine the optimal grouping of letters into clusters until the errors are minimized. An example of such a system is, FEATURE, a speaker independent isolated letter recognition system (21). This system performs recognition on an 80 letter vocabulary generated by 20 different speakers. The recognition accuracy of 85% was obtained across 2 speakers for a speaker-independent mode while an accuracy of 91% was observed when operated in dynamic adaptation mode] . The different algorithms for the system were implemented on a Fast Processor Array, Motorola 68000 microprocessor, and VAXI 1/750.

Template-based speech Recognition:-

Template-base speech recognition systems have a database of prototype speech patterns (templates) that define the vocabulary. The generation of this database is performed during the training mode. During recognition, the incoming speech is compared to the templates in the database, and the template that represents the best match is selected. Since the rate of human speech production varies considerably, it is necessary to stress or compress the time axes between the incoming speech and the reference template. This can be done efficiently using Dynamic Time Warping (DTW). In a few algorithms, like Vector Quantization (VQ), it is not necessary to vary the time axis for each word, even if any two words have different utterance length. This is performed by splitting the utterance into several different sections and coding each of the sections separately to generate a template for the word. Each word has its own template, and therefore this method becomes impractical as the vocabulary size is increased (500 words). Itakura of NTT (31) developed an SR system with the concept of TW for non-linear alignment of speech. The system performed a speaker-dependant isolated-word recognition task with a 97% accuracy on a 200-word vocabulary. Performance of this system degraded to 83%), when the vocabulary size was increased to 1000 words. A template-based system for connected speech was implemented by H. Sakoe (17). It was a speaker-dependent connected-digit large vocabulary speech recognition system that had an accuracy of 99.6%. It was implemented on a NEAC-3100 computer, with a memory requirement of 10 Kbytes per word, and had a vocabulary of 600-2500 words depending on the training set (17). This system performed well for speaker-dependent speech recognition, but one of the shortcomings of the system was that it took more than 150 minutes (25) to recognize a word with a vocabulary size of 600. For the same set of space. This represents a 130:1 reduction for the VQ recognizer would require around 250 Kbytes to 1.2 Mbytes of memory space. This represents a 130:1 reduction for the VQ recognizer over the DTW model. In some cases of template matching, no time alignment is needed to perform the recognition, as in the case of the system developed by J.E. Shore et.al. (21). It was a speaker independent, isolated word, limited vocabulary SR system, implemented using a VQ approach. It had an accuracy of 98% for speaker dependent recognition while for speaker independent recognition the accuracy reduced to 85%. The system was implemented on a DEC VAXI 1/750 with floating point accelerator. Template generation required 20-22 minutes, while classification of a single utterance with these templates took about 1-1.5 minutes.

Knowledge –based speech Recognition:-

Knowledge-based speech recognition systems incorporate expert knowledge that is, for example, derived from spectrograms, linguistic, or phonetics. The goal of a knowledge-based system is to include the knowledge using rules or procedures. The drawback of these systems is the difficulty of quantifying expert knowledge and integrating the multitude of knowledge sources (16). This becomes increasingly difficult if the speech in continuous, and the vocabulary size increases (16). A good example of a knowledge-based system is HEARSAY, developed on a PDP-11 microcomputer, at Carnegie-Mellon University. It was a speaker-dependent continuous-speech recognition system with a vocabulary of 1011 words. Using a very restrictive syntax (perplexity4.5), it achieved a recognition accuracy of 87% (32). It took 13 hours to compile the algorithm developed for HEARSAY. Fifty knowledge sources were used to develop this system and the amount of storage required for each one was about 80 Kbytes. Comparing this storage requirement with that needed for the Vector Quantization algorithm shows a reduction (22) of the order of 30 or more. This would mean far less memory storage would be required for codebooks generated by VQ approach than by the knowledge sources in the above approach.

Stochastic speech Recognition Systems:-

Speech recognition based on stochastic modeling is another approach for Speech Recognition systems. Probalistic models of speech are used in this approach to deal with incomplete information or uncertainty. The most widely used model is the Hidden Markov model (HMM) (11). The HMM uses states that model generic speech sounds and transitions between states with associated transition probabilities to model the temporal behavior of speech. This model assumes that speech was produced by a hidden markov process. To derive the transition probabilities, an efficient estimate-maximize algorithm, the forward backward algorithm, is often used (33). Though the HMM approach can give substantially accurate results, if the time factor is taken into consideration, then algorithms based on template matching using Vector Quantization or DTW prove to be much faster than the ones based on HMM (11). An example of a system based on the HMM approach is the one developed by Rabiner et al. at Bell Labs (12). The system used the techniques of Vector Quantization in conjunction with HMM in a speaker independent, isolated-word recognition system. It achieved an accuracy of 96.3% word accuracy for a 10 digit vocabulary. The system was implemented on a SUN-workstation. Despite the fact that the system could recognize words from any speaker, the system had a drawback, which was the large amount of time it took to train the system to a given set of words. For the given case, 10 digits took more than 15 hours to train the system (12). Another example of a system based on HMM and statistical methods (14) is SRI's DECYPHER system, which has a 1000 word vocabulary, is speaker independent, accepts connected speech input, and has a recognition accuracy of 95.6%. Since DECYPHER is a connected speech recognition system, the recognition algorithm involves more computation than a discrete recognition system (by a factor of 3 (23)), and real time performances is more difficult to achieve. In this system, a statistical bigram grammar is used that reduces perplexity to 60 (without grammar, any word can follow a given word). Using this grammar, the recognition accuracy for DECYPHER improves from 75.5% (no grammar) to 95.6%. This system was implemented on a microcomputer and required more than 16 Mbytes of workspace to perform calculations (14).

Connectionist speech recognition systems:-

Connectionist speech recognition is based on artificial neural networks that use learning strategies to organize and optimize a network of processing elements (neurons). These networks are used as classifiers or mapping functions to recognize the incoming speech. Thus the speech knowledge or constraints used for speech recognition are distributed among many, but simple processing elements (13). The concepts and ideas of applying neural networks to speech recognition are relatively new, and researchers are investigating a number of approaches. An example of connectionist system developed for keyword spotting (23) is the one implemented using a time-delay neural network (22). This speech recognition system, implemented on a PC-486, was compared to an HMM recognizer in the task to recognize the phones “B”, “D”, “G” out of a database of 5240 Japanese words. For different speakers, the HMM had a recognition accuracy of 90.9% to 97.2%, while the neural net achieved an accuracy of 97.5% - 99.1% (24). One of the advantages of neural network based speech recognition systems is that they can achieve the capability of handling a really large vocabulary (24). One of the disadvantages of the system, which is the same as that of an HMM-based system, is that a large time may be required for training the system.

Speech Recognition using c#:-

Speech recognition in C# we use Microsoft speech SDK tool. The tools use to speech enable would be the speech SDK 5.1. Speech SDK 5.1 is the latest release in the speech product line from Microsoft. Microsoft speech SDK is one of the many tools that enable a developer to add speech capability into applications. Speech SDK can be used in either C#, C++, VB or any COM compliant language. Speech recognition can be of two types based on the grammar that the recognition is based on. (Grammar is in other words the list of possible recognition outputs that can be generated.) An application can limit the possible combination of the words spoken by choosing proper grammar. In a command and control scenario a developer provides a limited set of possible word combinations, and the speech recognition engine matches the words spoken by the user to the limited list. In command and control the accuracy of recognition is very high. It is always better for applications to implement command and control as the higher accuracy of recognition makes the application respond better. In Dictation mode the recognition engine compares the input speech to the whole list of the dictionary words. For the dictation mode to have a high accuracy of recognition is it important that the user has prior trained the recognition engine by speaking in to it. The training or creating of a profile can be done by using the speech properties in the control panel.

Hardware/software co design methodology:-

G.F. Zaki, R.A. Girgis, W.W. Moussa and W.R. Gabran proposed that an embedded face recognition / verification systems has been implemented on an FPGA. Its small size allows us to apply these systems wide range of applications. The system is based on FPGA which offers high configurability in the design phase. The hardware/software co design methodology was used which enabled optimizing the system to meet different design constraints including size, cost, and power dissipation. A number of DSP algorithms were used or created to detect the face from a background, enhance the image and recognize the person. The principal component analysis algorithm was used for feature extraction (5).

Che Ming, Chang Yisong proposed that the implementations of a face detection algorithm on FPGA for eye mouse control system. An improved algorithm of skin color module and binary image projection is used to ensure real-time detection. The system is based on a hardware/software co-design, which consist of a dedicated hardware accelerator that solves the parts of the algorithm with higher computational cost and an embedded microprocessor that manages the control process and executes the rest of the algorithm. Several optimization methods have been accomplished to enhance performance. The system has been implemented on an Altera Cyclone 2 FPGA using a Nios embedded soft-core processor and it is benchmarked against a soft ware implementation.

Objective:-

Aims and objectives of this dissertation work are summarized as follow:

1. We have implemented a security system that based on both hardware and software along with wireless connectivity.
2. It's based on face and voice biometrics for recognition.
3. Soul of the system is based on PCA algorithm for face recognition and for voice recognition using Vb.net.
4. This security system based on microcontroller to reduce the cost and complexity of the system.
5. Comparing the result with FPGA based system.

Conclusion:-

This paper described the introduction of Biometric and multimodal biometric system their mode of operation, advantages and applications of Biometric system. The future, we will discuss about various existing algorithm for image processing and voice recognition. The last section of chapter presents the aim and objective of the dissertation work.