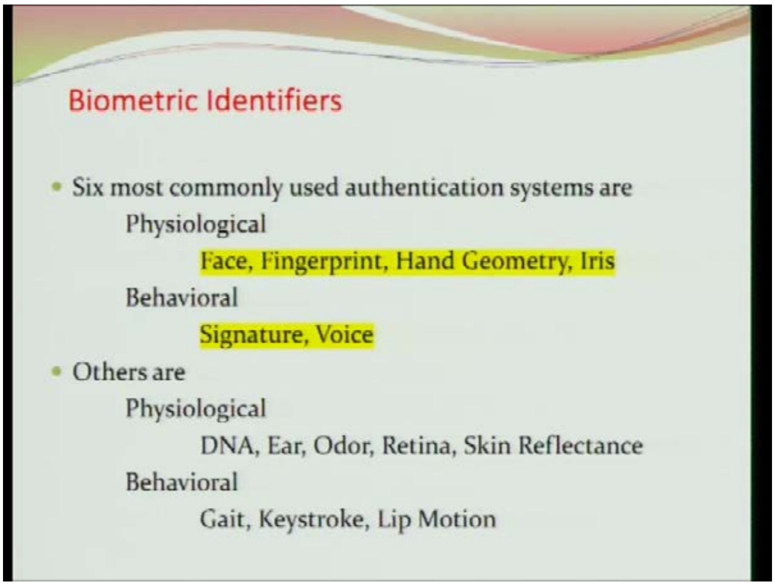
Biometrics
Prof. Phalguni Gupta
Department of Computer Science and Engineering
Indian Institute of Technology, Kanpur

Lecture No. # 10

(Refer Slide Time: 00:19)



Biometric Identifiers

- Six most commonly used authentication systems are
  - Physiological
    - Face, Fingerprint, Hand Geometry, Iris
  - Behavioral
    - Signature, Voice
- Others are
  - Physiological
    - DNA, Ear, Odor, Retina, Skin Reflectance
  - Behavioral
    - Gait, Keystroke, Lip Motion

We are discussing about the biometrics and now you can understand from here. What are the biometrics which are being commonly used. First one is that the face is one commonly used biometrics trait, even though we assume that face is a commonly used biometrics trait, but it is not acceptable too much by biometrics society reason behind it has lot of implications.

First one is that it is not age invariant, and if age go goes the face gets changed, then lot many you know at older age, you will find lot of wrinkles on the face then is coming that expression. It is not expression invariant, that you have these different types of mood at any instance of time. There what mood you have based on that your face will be there and may not be matched with the natural mood. It is not pose invariant, as your pose is changing your recognition rate also will be varying, say if it is a 30 degree rotated image, then it may you may not be able to match also.

So, even though we tell that face is a commonly used biometrics trait, but it is in biometric society we do not consider it is the best one of the best recognition system. Generally we observe that accuracy rate will be around 1885 while we will be discussing face recognition system we will realize this.

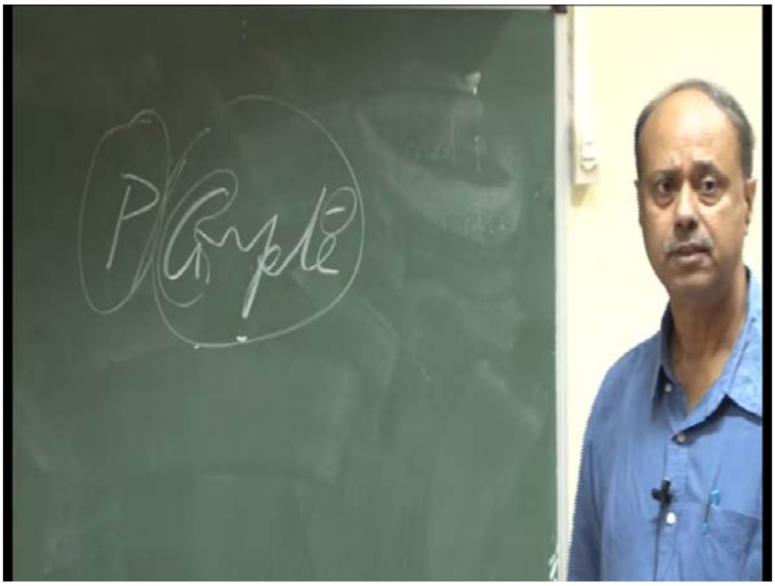
Fingerprint is also very commonly used biometrics system and it is not today. It is started in 1800s onwards. People are using fingerprints specially land record movements. You use fingerprints then money transaction. You use the fingerprints and so on but again it has been observed that it is almost it is very few mean very few cases. One has observed that there is a little similarity between the two fingerprints but those getting minutiae points is a challenge. Good minutiae points, because that once you take give the fingerprints then there is a chance of too many false minutiae points. How false minutiae points comes, because of scanner problem that this scanner due to that there may be some spot and it will be considered as a biometric because point and so on.

Then there is a possibility that people use the rubber stamp on similar to your fingerprints and he puts that finger to get the approval or recognition hand geometry, which I want to tell that it is a palm print and this is also well recognized system and same problem you will find in case of hand geometry. Here nobody keeps his data hand data for archival point of view. We cannot think that yes if I ask you, do you have your palm print data? You will tell no but when you need you have to give.

So, hand geometry is also another issue iris the getting data itself a challenge but it is a more accurate one and society now-a-day’s started accepting this one behavioral here. Signature is the well one well or commonly used biometric system but lot many forgery cases you will find on signatures and there are two types of signature verification say one is offline another one is online. Now offline is that based on your signature on a plane paper and we scan the image and then we want to do certain operation on it.

Now this operation involves while signing, we put the dot also. There are several ways you can use generally and we sign that one can sign this way.

(Refer Slide Time: 05:13)



So, these dots are important and you cannot ignore these dots. Sometimes see this is important, so this is one connected component. This is another connected component so these connected components are also important parameters for offline signature. Now in case of online signature, have you seen any online signature pad? This is a digital pad, where there is a digital pen is pen there you will be signing on it.

Now that the way you incline the pen is important, what angle, it is than how much pressure you are giving that is important for all those things parameters will be recorded. How much time you needed to sign the paper, that is important so online signature are for is a commonly used but it has also lot of issues.

Voice, it is a commonly used by listening my voice and you can tell he is mister x but there are several people who can copy my voice and if I am suffering from cold and cough then my voice will be changed. So, all those issues are lying with voice one beside that physiological biometric, there exist one is the DNA which is the most common most accurate one but it is the time consuming and it cannot be used as a near real-time system as then more over it is very costly.

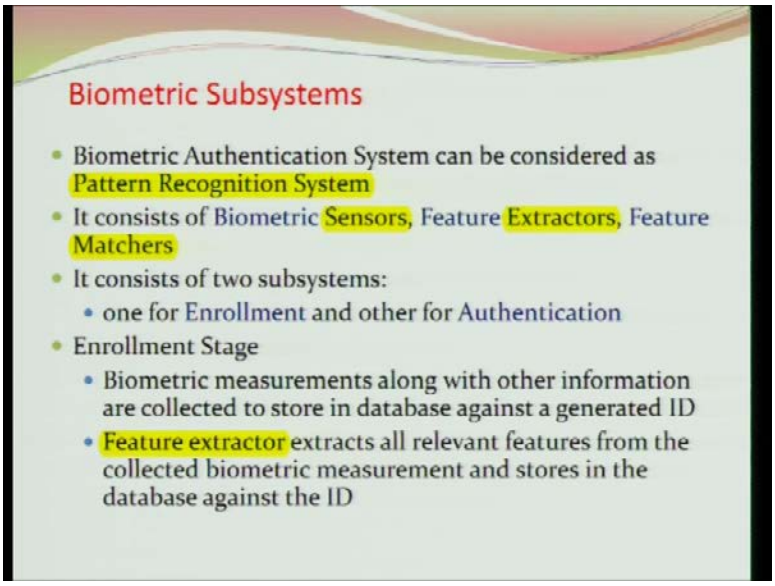
Ear biometrics, just it has been started few years back and what we have shown that it is giving us very good result but it is not yet well accepted. We have not yet able to prove that it is the one of the best biometric system body. Odor is also another one by getting the smell of body people can understand who is he especially, I can give you an example that dog they believe on body odor. They detect determine who is you or who are you based on your body odor.

Retina is one like iris but getting data itself is a challenge but it is also giving us very good results then skin reflectance parameters. That is the light source, you put on the skin then the skin gives you certain indications. Those are also good physiological measure then behavior one is gait well being studied but not yet well accepted, because y one is that walking style the way you walk but which can be copied.

So, beside that you know the way, you give the pressure on road. That is also an important parameter. There are some people they give the pressure on hill. There are some people who give the pressure on the front part of the foot.

So, these are the gait parameters and there is a need a footpath where the pressure will be measured. So this is being studied now a day's key strokes the way or the amount of pressure you give on keys of keyboard. This can be considered as a behavioral biometrics. Biometrics lip motion is another one the way you speak that lip moves that also being considered as a biometric based system.

(Refer Slide Time: 09:15)



Biometric Subsystems

- Biometric Authentication System can be considered as Pattern Recognition System
- It consists of Biometric Sensors, Feature Extractors, Feature Matchers
- It consists of two subsystems:
  - one for Enrollment and other for Authentication
- Enrollment Stage
  - Biometric measurements along with other information are collected to store in database against a generated ID
  - Feature extractor extracts all relevant features from the collected biometric measurement and stores in the database against the ID

Now any biometric based system is based on pattern recognition problem, because you cannot match the two images or two scanned images to draw a conclusion. There are several reasons, first one is image to image matching to do that you need to see the both the images are registered. The term register one day that there will be some feature points or some key points. They should be matched and then you have to adjust it. So to match two images you need to have the scenario like that they are registered and second thing is that if it is a frontal face then another one also must be frontal face. It cannot be rotated or it cannot be expression different expression.

Or simple is the case of finger prints say I have the full finger print and another case. I may have the half finger print may not be exactly same size so those are the issues. There image cross image versus image matching is not possible. It takes time also, if the database is very large then you are in problem solution that why not takes out certain patterns from the system or from the image and those patterns you use for matching.

Now this patterns means that you'll pattern means what here are the certain peculiar characteristics what you expect or you feel is lying in the image so those peculiar characteristics you take out and then you use them as features for mismatching so the what are the things involve in that case I need one very good biometric sensor now by the word sensor we mean that we need a camera we need a scanner or digitizer something like that.

So, sensor is that biometric sensor we need next one is that I have to know very good feature extractors by what it means that given the image you must be able to extract the important parameters or important characteristics from the image and then feature matchers and once you have the features than you have to match them so matching algorithm is required.

Now you observed that once I tell that it is a matcher. The matcher means what it is basically you will be finding the distance, how much he is away from the original one so basically you need to obtain a distance matrix. Now distance matrix may give you one of the two one is the similarity matrix another one is dissimilarity matrix and also in some case you will find that matching algorithm matching matcher gives you the result lying between 0 and 100 in some cases. You may find that no it is giving you 0 to 1 so it need not be normalized one.

So, matcher gives return some value matching value which is similarity or dissimilarity and it is not need to be a normalized value so it is may not be possible always by knowing the matching score, which one is good or which one bad you cannot draw the conclusion you need to do normalization before draw the conclusion. Now in any biometric system it has the two sub system, one is the enrolment another one is the authentications by enrolment. We mean what that you have come and you want to put your name in the database.

So, enrolment stage you have come, you are not only giving your biometrics data but also you need to give certain other information's say for example: your name, date of birth, gender, maybe father's name house number all those information, you may have to give. So, those information once you give along with the biometrics data that we will be putting in the database to generate an ID.

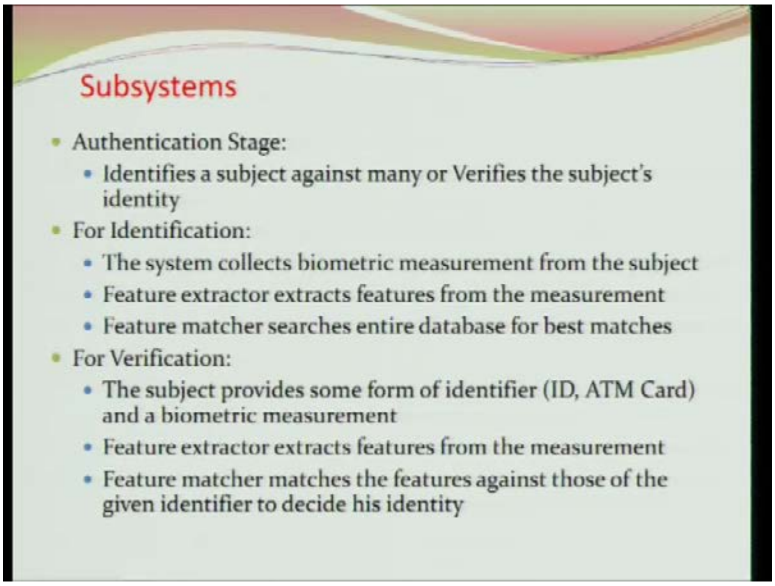
Now once you have got an ID then your image of biometric data will be taken and process to extract the features. Now by the word, it is very easy to tell the word that feature extraction technique involves lot many things, first one is that your image quality may be poor.

So, you need to enhance the image, so the features can be extracted. Your image may contain lot many noises, you may have to remove those noises after that you may have to do the feature extractions.

Or your image may find that it is a smudged. One says for example, if there is a face or fingerprint, if I use the inkpad and give the pressure. If you put more pressure, you will find that it is a smudged. One so you have to extract only the thin image out of it so the feature extraction is required to design to get the features from the biometric image and you want to store them in ID.

Now that means what your database contains in one database contains the ID against all features against all images and biometric demographic information. Another one contains the ID and the feature vectors remember that the second database that where it contains the feature vectors an ID. That will be used for our purpose, because we are not using the images, we are not using the demographic data the second database is important for us.

(Refer Slide Time: 15:44)



Subsystems

- Authentication Stage:
  - Identifies a subject against many or Verifies the subject's identity
- For Identification:
  - The system collects biometric measurement from the subject
  - Feature extractor extracts features from the measurement
  - Feature matcher searches entire database for best matches
- For Verification:
  - The subject provides some form of identifier (ID, ATM Card) and a biometric measurement
  - Feature extractor extracts features from the measurement
  - Feature matcher matches the features against those of the given identifier to decide his identity

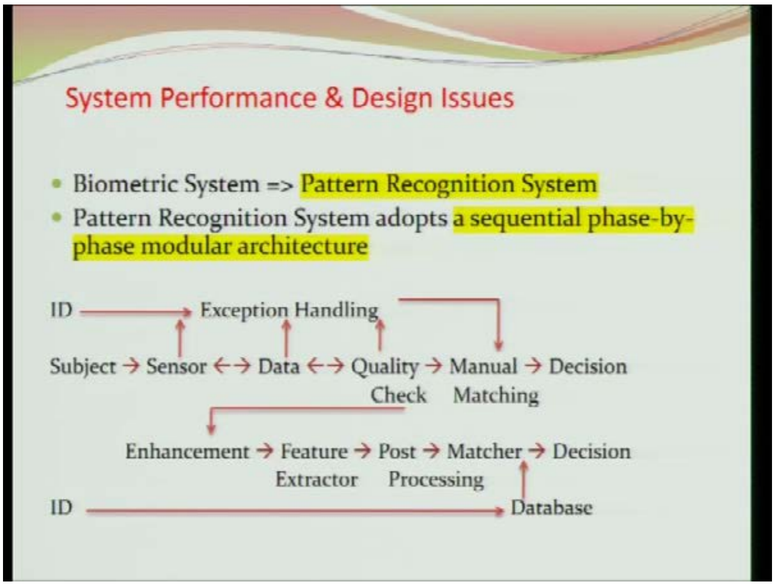
Now authentication stage here you have two component, one is the identification another one is verification, in the case of identifications what you have to do given. One biometric data you have to find out the list of persons the list of subjects available in the database who are similar to this database similar to this feature vectors that means you have the list of subjects in the database and given a subjects image. You will be extracting the features, these features will be matched in the database and find out who are the people or who are the subject having the similar type of features. Those names will be or those ID's will be broadcasted.

So, it needs also use it the feature extractor but he needs also a good matcher algorithms to extract the list of people, In the case of verification here what is given, everybody is having some ID and also feature vectors in the database. Once I give my ID and I give my features biometrics data. The feature vectors will be extracted and against my ID that features will be matched and you will be taking the decision whether he is the man or not.

So, there is a difference between this here you remember that the feature vectors whatever features or whatever data you are getting or biometrics data you will be getting that data may not exist in your database, because you got some biometric data and you want to extract that x number of people who are nearly matched with this biometrics data so that data may exist and may not exist in the database that mean database may not be closed.

But here in this case since I know my ID is this, so in that ID my biometric data is there. So, it is expected in the case of verification my database is closed.

(Refer Slide Time: 18:10)



System Performance & Design Issues

- Biometric System => Pattern Recognition System
- Pattern Recognition System adopts a sequential phase-by-phase modular architecture

ID → Exception Handling

Subject → Sensor ↔ Data ↔ Quality Check → Manual Matching → Decision

Enhancement → Feature Extractor → Post Processing → Matcher → Decision

ID → Database

So, we told that biometric system is based on the pattern recognition system and pattern recognition system adopts a sequential phase-by-phase modular architecture. What it means that I have given my ID and in the case of verification I have given my ID and its subject once a person has come and he has given his ID and the sensor tries to extract the image from the subject. Now it may so happen that sensor is not working, you do not expect because these are hardware component and it may fail at any instance of time. If sensor does not working correctly. What will be doing? You will be taking the help of manual verifications.

So, this comes under exception handling, if sensor does not work then manual thing you have rule. So exception handling will come in between, suppose sensor works so sensor gives you the data. Now you may find that even you are trying to give the data but data you are not capturing data. System is not able to capture the data. So, you will be sending back data, I have not got the data you send again but if it fails after a certain number of times it will be going again exception handling.

Now you may get the data but it is poor quality, because you do not know the quality. How much quality is good for you, so once you see the in the screen that the quality is very poor. You are telling no, which is not good. I need again, so you will be sending back and I need again data. The sensor will give you another data after some time and say after 5 shot you find that no. The quality is coming very poor then you will be taking the help of exception handling.

If it is the case then you go for the manual matching and take the decision but if it passes quality is checked and quality is good then you enhance the algorithm image. Now enhancement means, here it involves too many things you first remove the noise and then you enhance then you extract the features and once you extract the features then you have to do the post processing and matchers.

In order to match it you need the data from the database, so you ask the data and then you compare and take the decision. So, this is the generic structure. Now if you observe that this is most critical one say for example: in the field where scheme is there. What happen that government of India has certain schemes where poor people get something at the end of the month, two hundred or three hundred rupees. Now they distributed through banks. Now bank in the village, you cannot have several branches and something what they do, they appoint business correspondent.

Now business correspondent every month they distribute the money and most of the time you know they the villager come and they give the money. Sometimes they may not like to verify whether he is the man or not, because in the village may be fifty people everybody knows each other. Now suppose I want know that I do not want only one village to be controlled by the b c one, b c will be appointed for few villages and it is very difficult to remember who is who.

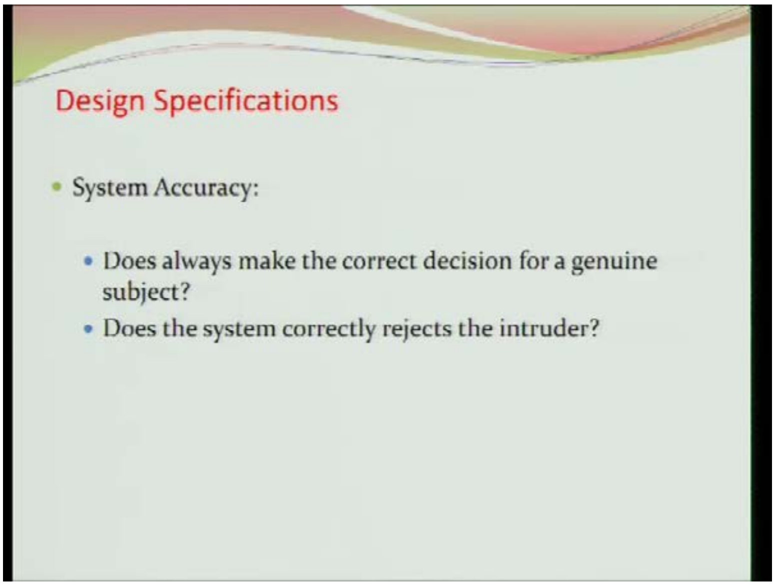
So, suppose we introduce the biometric system, now the villager’s finger print quality will be very poor and there is a possibility that it will fake not only possibility. There is a high possibility, it will fail or after sometime or may be in the beginning itself. If it is the case, what business correspondent will do? It is very difficult for him to tell no. I will not pay, because the finger print is not matched this this poor person may be waiting for this 200 rupees to get his rice and daal.

So, he has to distribute the money moreover, if he fails to distribute the money. The bank will feel in different way that you are not doing your duty and you have not gone to the village to distribute the money. So he has to clear his transactions and so that he himself can claim some money from the bank as commission. Sooner is the better for him but if it is not matched in the biometric data. It is not matched with his data then what is he going to do.

So, what generally they do three times four times check face that he take the risk. He tells that I am over righting the system. He pays the money so the customer is happy, bank is also happy. He is also happy, because he is getting his commission. Bank will feel that I have distributed the money but if one finds that a business correspondent is giving the money always by giving the through over righting power then it is a questionable.

There are people that every time he is over righting. Now say out of 100 people one of them. One case he has over righted there is no problem but out of 100 people if he over rights 100 of them or 90 of them then the government will feel that means, he is not doing his job or government may feel that he does not know whether this money has been distributed to the concern people or not.

(Refer Slide Time: 24:53)

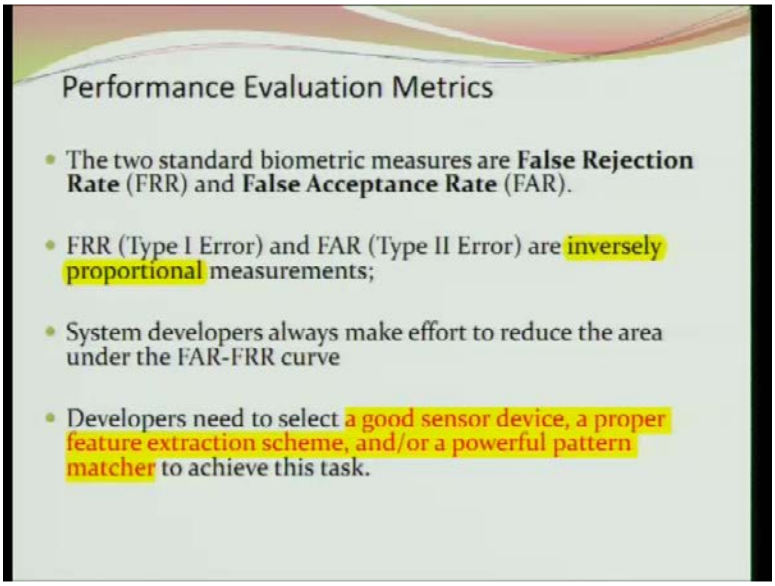


Design Specifications

- System Accuracy:
  - Does always make the correct decision for a genuine subject?
  - Does the system correctly rejects the intruder?

So, this is the very weak point and this weak point has to be minimizing as far as possible but one thing you remember this weak point will be there. You cannot have a system without exception handling reason behind it that suppose I want to enter into the my class room and I have put a business biometric system outside the room and every time I am giving them a finger prints. It is not allowing me to enter but I am a genuine person what I am going to do at this stage. So there might be some exceptional handling through which you will be allowed to enter this is the case with your attendance system whatever you have in the department. If you fail you go there. There is a register you sign there and so on.

(Refer Slide Time: 25:44)



Performance Evaluation Metrics

- The two standard biometric measures are **False Rejection Rate** (FRR) and **False Acceptance Rate** (FAR).
- FRR (Type I Error) and FAR (Type II Error) are inversely proportional measurements;
- System developers always make effort to reduce the area under the FAR-FRR curve
- Developers need to select a good sensor device, a proper feature extraction scheme, and/or a powerful pattern matcher to achieve this task.

But in reality in will not happen, because you will be matching pattern against pattern. Some pattern may not occur in one instance. Some pattern may come out in one instance or there may be some false pattern also on your biometrics data.

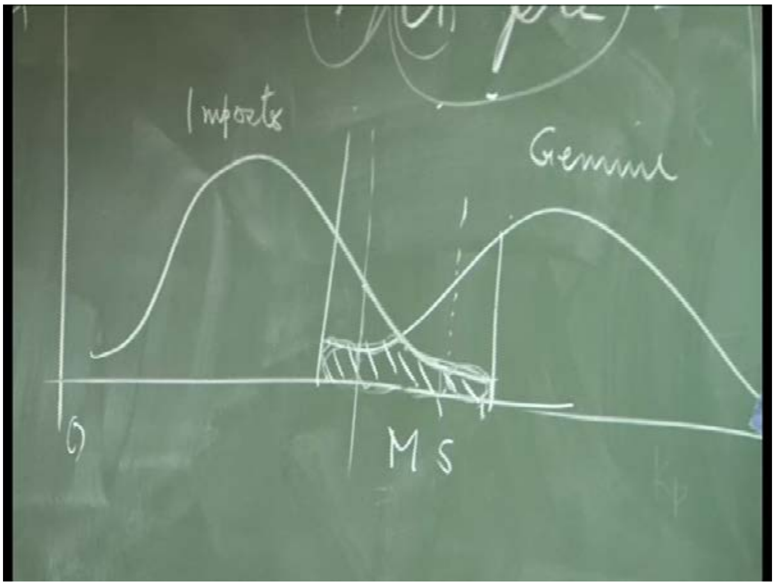
(Refer Slide Time: 26:12)

So, as a result what happens at any instance of time, you may have some pattern here? This is one second; one is that there may be additional patterns which are also true pattern. It may so happen that there some false patterns also there, because of this there is a possibility that whatever system accuracy or ideal system accuracy environment that may not be achievable.

That means what we are telling that even if you are a genuine user, sometimes you will be thrown out Even if there is a forger sometimes, you will be accepted by the system. So, your problem is to minimize these two factors. These two factors are known as gives you the two errors one is known as false rejection errors and another one is false acceptance rate and they are inversely proportionate. If false rejection you are increasing means what you are making very tight system where all genuine users are entering into the system.

So, if it is the case then what will happen some intruder also? They will be entering into the system and I will give you one example later on and similarly, is the case that if I feel that no. I will not allow any false acceptance then you will find some genuine persons also will be rejected.

(Refer Slide Time: 28:08)



So, what happens that in ideal situation, let us assume the similarity matching algorithms. I have or my distance method is based on the higher the matching the matching score probability that I am the genuine person is high. This is a similarity measures

So, in reality you will find this is a matching score and say let us assume that it is a normalize one 0 to 1. This frequency so the number of people, they are having the matching score. This is very close so generally most of them will have the average people. Who will have the higher frequency and again very large matching score? You will be getting very less number of people, because probability that all genuine people will have the high matching score is very less because of those parameters. Some of the matching, some of the feature point will be missing or some addition matching point feature points will be coming or there may be some false minutiae false feature points.

So, that is why very less people will have very large minutiae or large matching score. Now in ideal scenario you will find that if I am imposter, my matching score will be less, because I am in imposter. So, you will find that matching score will be like this. So, this is imposters and this is the graph of genuine. So, matching score will be maximum matching score may be here for imposter.

Now this is the region and is very difficult for us. This is known as critical region. Suppose I decide if my matching score is here or all those elements all those persons whose matching score is larger than this value then you can assume that they are matched now. If you assume that they are matched that means this many people they will be also falsely matched.

And this many people will be falsely rejected, because they are also genuine people but your threshold you have made this one so these are genuine people but you are rejecting. These are the genuine forger and you are accepting them so that the reason is that if suppose I increase do this one. Now what happens that you are accepting more number of people but also very less number of genuine people? You are rejecting or larger number of people who are forger and you are accepting them.

So, if one you are increasing, another will be decreasing. So, that is the reason why they are inversely related. So your aim is to reduce this critical region and if it is an ideal scenario then this is the best one that two independent distributer enough. So, you are happy, you put the threshold. This is the result but you will not get this one but this should be as minimum as possible that is your idea.

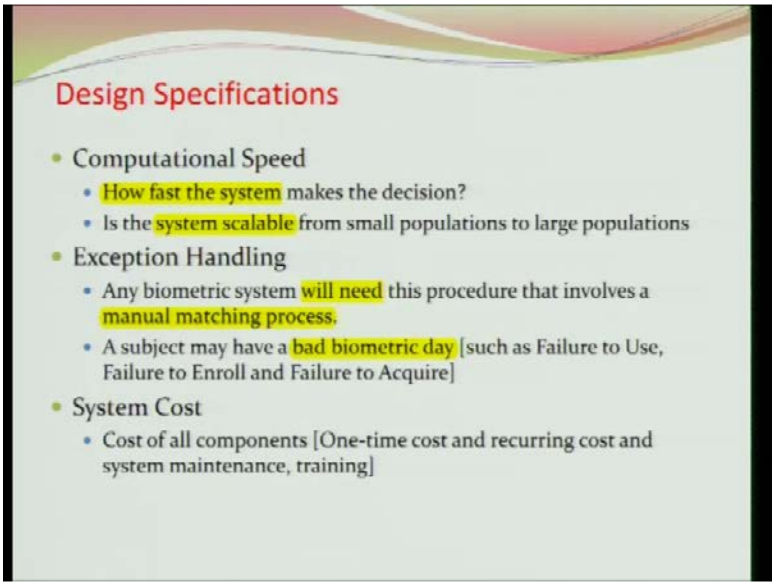
Now in order to get the good system, First thing that, it is dependent on patterns. So, pattern will be good if I have very good sensor if I have a very good condition. Suppose I take the data under uniform condition then say under this light source. Under this temperature if I start taking your data and if I take the same data outside this room. They will be difference if I take the high resolution cameral based on the high resolution camera and the low resolution camera data will be different.

So, we need an environment where you have very good Sensor device and the control environment as far as possible then you need a proper feature extraction algorithm. You have a sensor and you are extracting certain features that feature action technique may not good for that sensor data of that sensor it may be different. So you have to understand the sensor device parameter based on that you can think that whether feature extraction can be used or not. I have camera to take your photograph and now face photograph and I have a feature extraction algorithm for that.

Now I have a finger print and now can use that whatever feature extraction algorithm I have used for face. Can I use on my finger prints, because it is not possible the face data? It is very large and it feature points are different finger points. Features are different so you have to handle different way. So, same feature extraction algorithm will not work and then you need once you get the features you have to match them with another feature vectors.

So the matching algorithm which is true for face may not be good for iris. We will see this one. So, feature vector itself is different.

(Refer Slide Time: 34:45)



Design Specifications

- Computational Speed
  - How fast the system makes the decision?
  - Is the system scalable from small populations to large populations
- Exception Handling
  - Any biometric system will need this procedure that involves a manual matching process.
  - A subject may have a bad biometric day [such as Failure to Use, Failure to Enroll and Failure to Acquire]
- System Cost
  - Cost of all components [One-time cost and recurring cost and system maintenance, training]

Now once you have these three parameters based in this, if you design a system what are the things you should look into first. One is that the speed of your system and once you develop a system. People will tell how much time it will take? Is it near real-time say DNA and DNA it takes a few days now will it be useful as a real-time biometric system, is not?

But if it is a fingerprint then it will be useful for real-time system. Answer is yes so how fast system take make the decision, this is important parameter. You need second one is that scalability what it means that whatever we develop that is on small database size 11500 and so on can I use it or can I deploy it. For one million people can I deploy it? For billion people so that domain you have that has to be done properly or studied properly so that it is scalable.

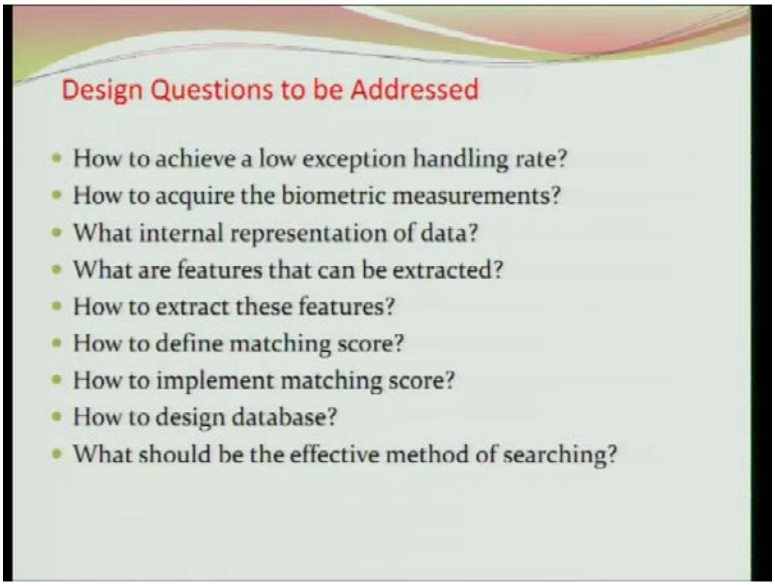
Now exception handling as I mentioned earlier that any biometric system must have the concept of acceptance handling otherwise it will be very difficult for you to use or implement the system, because there are some days you may find that your data is not able to give your data, because of many reason. May be, you have come from the field and your hand is very dirty and you have been asked to give your finger print data. The system is not accepting that or because of weather see every sensor has certain property as sensor must work under temperature of this between this and this.

But what happens if it exceeds say forty nine degree will your sensor work so even if the sensor is giving some signal but because of temperature you are not able to give the data correctly similarly, you have come from the field work and your finger print quality may be poor because you have come from the field and you will not be able to provide the data so that biometric data will be there.

So, your aim is that your data should be acceptable but at the same time the exceptional handling also will be as minimum as possible. System cost, if it is a very costly then nobody will use your system. So, system should be as cost as minimum as possible. Now the system cost involves, what one time cost? I am ready to spend one time money but recurring if you ask that no you have to pay every month x amount of money. That may not be acceptable or reverse way that you take only one time. First time you take very less amount and every month.

Another thing we do not take into account, it is training. It needs lot of training for a like any other system for biometric also. You need training otherwise what will happen that they will pass everything and they will enroll everybody without understanding what has happened. What in my feeling in some of the database collection of biometric for biometrics system.

(Refer Slide Time: 38:30)



Design Questions to be Addressed

- How to achieve a low exception handling rate?
- How to acquire the biometric measurements?
- What internal representation of data?
- What are features that can be extracted?
- How to extract these features?
- How to define matching score?
- How to implement matching score?
- How to design database?
- What should be the effective method of searching?

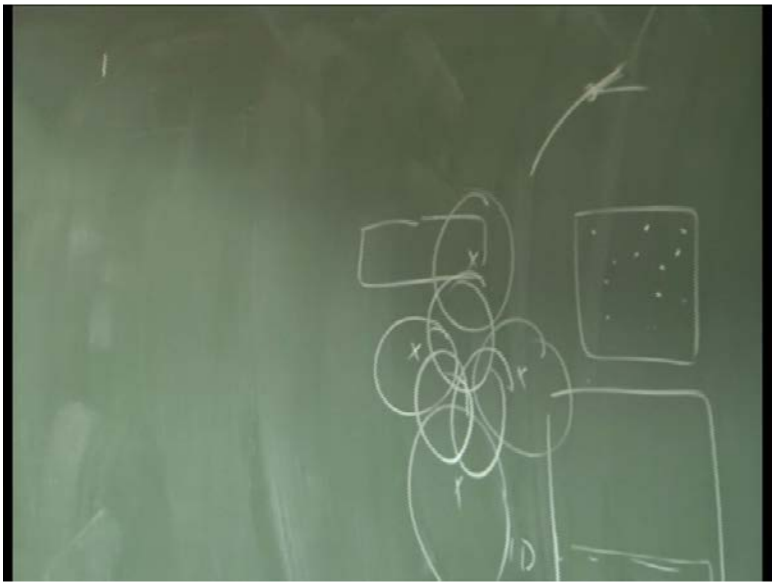
Now while you have to answer these entire questions, this is the most important parameters. You have to check, how to achieve a low exceptional handling rate, because this is the most important thing and corruption will come from there. So, you have to find out the way that your system should be searching that exceptional rate handling. Rate has been very less that means the two graphs are separated as far as possible. They are independent in nature.

Then how to get the biometric data is it flat based scanner? Is it no there? is a small scanner design for it or so on but it should be easily deployable. You should not design a system which takes a huge system for deployment. Then again cost will go up and you have to ensure that the biometric measurement or biometric data will be acquiring. The sensor will be low cost one but it should give you moderately good features or pattern patterns.

What internal representation of data, so once you get the data. How are you going to get this data? or How are you going to represent the data? There is a standard or ISO standards are there. You have to follow that standards otherwise nobody will be an issue and nobody will accept your system and there it is written that first bite is this purpose second bite is reserve everything is mentioned so you have to follow that pattern.

Then next one is what the features you like to extract. For a example: in the case of fingerprint, you will be taking out the bifurcation or end the minutiae points regenerating minutiae points and that is represented by coordinate x y theta and the type is x y theta means, what x is the coordinate and y is the coordinate and theta is the angle. If it is this then what is the angle and the type whether it is a ridge ending or the bifurcation so that has to be decided.

(Refer Slide Time: 41:00)



Similar is the case of iris. Iris has also patterns and those patterns to be represented in the binary form. So, it will be to an iris code which consist of 0 1 and so on then how to extract the features now you need mathematics extract the features.

So, this technique is your technique and this is important. These are the things you like to obtain but how that you have to design then matching score see depends upon that. Your distance matrix is you going to use the equidistance or you are just using or you are using the D and there are several distance matrix available, which one you are using and you have to decide whether it is giving the similarity measure or the dissimilarity measure. Whether it is normalized or all those things you have to find out at this stage.

What is implication you get some matching score? Now you have to base on the matching score. You have to tell whether it is accepted or rejected. Whether he is the authenticated person or not so that threshold will come from this matching score and now you have the data where the feature vectors are there or the patterns are there. Now store the database, how are you going to store the database storing? The database is it the only the ID and patterns or ID plus demographic data and patterns and so on.

Then you need a database, large database you have to search this database. If it is a case of verification, there is no problem. The ID and the feature vectors so you go to ID and you get this ID and the feature vector is try to match it and how do you think? how do you think that this will be easy for you suppose, I have 125 crore people in your database and you got an ID. How are you going to search it?

You must ensure that face fault is not, easily you can get that ID. There are several way you can think that one is that. Last in first out or based on, everybody has certain parameters. Suppose I ask you that how often you go to the bank to withdraw money. Withdraw your, now there are people they go on the first day of the month and he withdraws whole month's money that I know that my monthly expenditure is this much and I withdraw this money. This is one type of person.

Another people is there, they go one week every weeks that one week my expenditure is this and I will withdraw this much and again next week budget and so on. There are some people, they do not withdraw at all then only after he finishes. He goes there irregularly and he withdraws there are people there are different ways.

So, if this knowledge you keep then you can design your database. I go every month first day and one day is coming my priority and will be higher. If I go every week, so my priority will be higher week wise so that knowledge can be taken into account in creating your database so the search area is search zone is very less.

Another issue you have to keep in mind that whatever I am talking. This is for the verification but what happens for the identification. I do not have ID, but I have feature vectors. I only have one feature vectors this database have to be matched and all similar feature vectors you have to bring them out and tell that these are the people their features are almost alike of this. You have to rank them that is the problem.

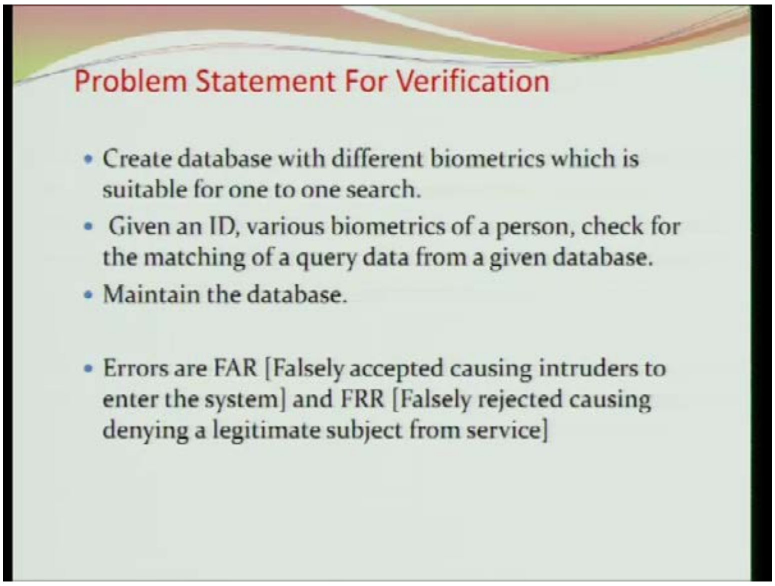
In that case, if you start checking 125 crore people, you are in problem. Suppose it takes one or what some micro second, micro nano second or vey very less amount then you also multiply 125 crore people. You will find that it is a huge number. I am not considering any face fault or any other information here it is huge.

So, what is it possible based on the features? You make the cluster so you have the database of cluster. These are cluster database and there exist a center this is known as cluster center. You got a features you find out in which center in which cluster it is lying and then that cluster only you search and give the your x number of list of people.

Now here the problem you will be getting and the cluster is a challenge, because this will not be like this. You will find like this, because while we will be discussing, you will find that the patterns are very closely associated making the cluster itself is a challenge.

For example; I decide that based on the ear structure, I cluster them now there are four types of ear. One is oval and another one is like that triangle. Four type of oval, one is this part little lengthy and now you will find that by doing that only 5 percent or 10 percent different type and most of us 80 percent will be oval shape. So, the database size will not be reduced further. That is not a good classification technique.

(Refer Slide Time: 48:28)



Problem Statement For Verification

- Create database with different biometrics which is suitable for one to one search.
- Given an ID, various biometrics of a person, check for the matching of a query data from a given database.
- Maintain the database.
- Errors are FAR [Falsely accepted causing intruders to enter the system] and FRR [Falsely rejected causing denying a legitimate subject from service]

So, these are the certain issues that you have to keep in mind. Let us do some more things, because I am little running late. Now in the case of verification what you have create database of different biometric. My ultimate goal is that to design a multi model biometric system that is ultimate thing. So, I am not thinking that I have only one biometric. I have different biometric so I have to create the database having the different type of biometrics characteristics.

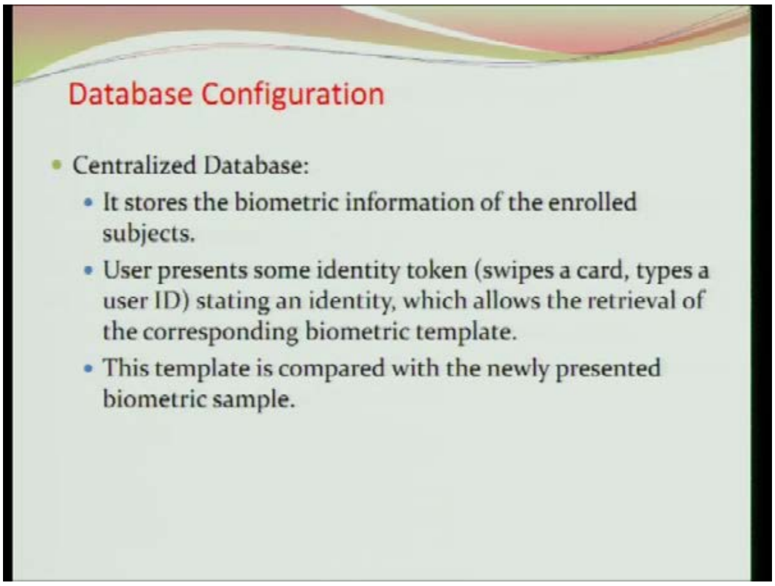
Now remember here, there are several problems. One problem is that some of us may not have one type of biometric. I got the same finger print iris. Suppose these are the three biometric data and I want take some of us or all of us and having face. Photograph is there, some of us may not have finger print, and some of us are not able open eyes in front of camera. This exists or why because IR light will be passing you may not like to open that much and so on.

So as a result even though we are taking the multiple biometric but you have to think in your mind that some biometric is missing. For somebody, some of them may have the all the biometric and so on given an ID and this various biometric traits. You will be extracting the features and now from using the features you will be taking the decision or using matcher will be getting the matching score and finally, you will be drawing certain decision.

If you find that element and does not exist in your database, he is telling that sir I do not have the database. I do not have my enrolment number then what you are going to do. You will be enrolling first so you add in your database sometimes. I come with an ID and you know that he is a genuine person but he has come after a long period or long number of years. Obviously his face has changed or there is some cut in the fingerprint. So, what you need and what will you be updating his biometric data and it will be putting.

So, this is the maintaining the database either creation or update your database. Now there are two types of error, you will be getting. One is the false acceptance rate and another one is false rejection rate in this case.

(Refer Slide Time: 51:14)



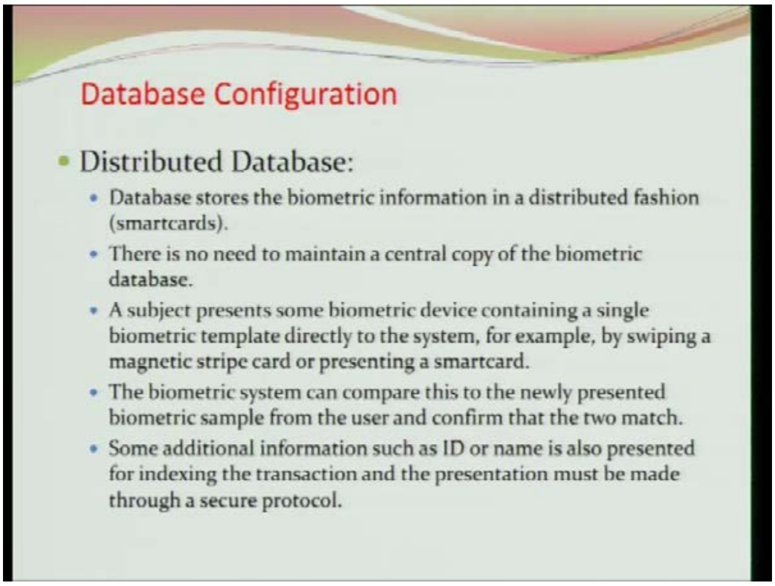
Database Configuration

- Centralized Database:
  - It stores the biometric information of the enrolled subjects.
  - User presents some identity token (swipes a card, types a user ID) stating an identity, which allows the retrieval of the corresponding biometric template.
  - This template is compared with the newly presented biometric sample.

Now generally there are two types of database and you will be finding. One is known as centralize database. Centralize database means that there is a system where all the data will be kept along with their features.

Now once you ask for something verifications, the system will send the data to the central area and then they will match and they will give you the authentication part. Whether it is accepted or rejected, this is happening for that adhar thing. The bank wants to use the adhar card things. So, what do they do? They collect the data and they send to the adhar server and adhar gives you the indication yes or no matched or not matched.

(Refer Slide Time: 52:10)



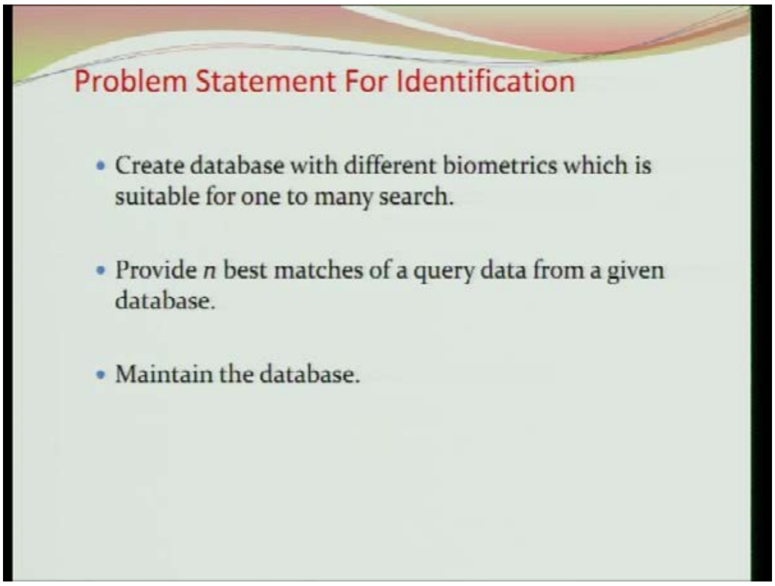
Database Configuration

- Distributed Database:
  - Database stores the biometric information in a distributed fashion (smartcards).
  - There is no need to maintain a central copy of the biometric database.
  - A subject presents some biometric device containing a single biometric template directly to the system, for example, by swiping a magnetic stripe card or presenting a smartcard.
  - The biometric system can compare this to the newly presented biometric sample from the user and confirm that the two match.
  - Some additional information such as ID or name is also presented for indexing the transaction and the presentation must be made through a secure protocol.

So, this is your centralize database. Another one is that distributed database here the concept of server is not coming. You have a smart card and in the smart card itself I have kept the templates or patterns and once you have come to withdraw money. You give your biometric data. The local system gets the pattern and match with the patterns you kept in your smart card and take the decision.

So, there is no concept of central database. It is a locally available and this is distributed and then is a very simple one but I lose the card then I am in problem. First problem is that I cannot prove myself by giving my biometric data. This is the card I do not have the card with whom I will be matching second point is that. Suppose, I lose that card then how safe will be my card. Is it possible that by I read your finger print patterns? Can I have an artificial finger print with that patterns in that case I will put that artificial finger print on my finger and I will give the data that is possible?

(Refer Slide Time: 53:44)

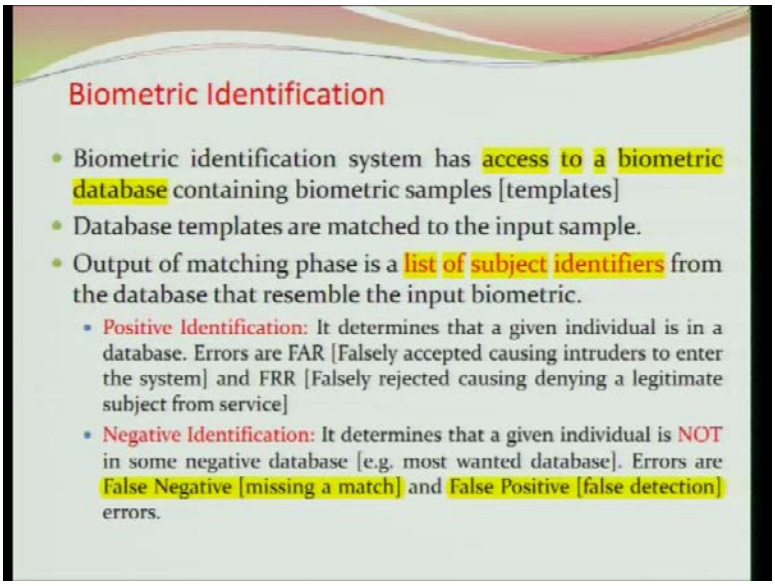


Problem Statement For Identification

- Create database with different biometrics which is suitable for one to many search.
- Provide *n* best matches of a query data from a given database.
- Maintain the database.

So, those issues are there and generally what we think that there will be distributed centralized database and also the distributor one.

(Refer Slide Time: 53:54)

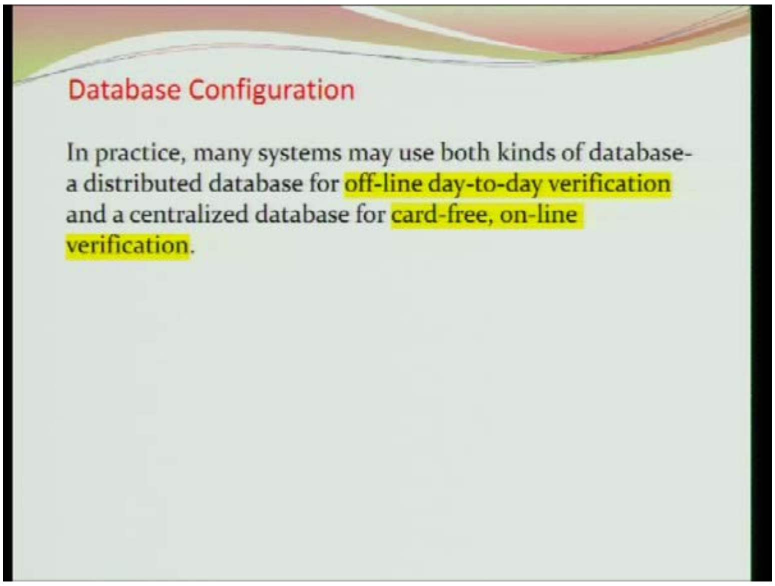


Biometric Identification

- Biometric identification system has access to a biometric database containing biometric samples [templates]
- Database templates are matched to the input sample.
- Output of matching phase is a list of subject identifiers from the database that resemble the input biometric.
  - Positive Identification: It determines that a given individual is in a database. Errors are FAR [Falsely accepted causing intruders to enter the system] and FRR [Falsely rejected causing denying a legitimate subject from service]
  - Negative Identification: It determines that a given individual is NOT in some negative database [e.g. most wanted database]. Errors are False Negative [missing a match] and False Positive [false detection] errors.

So, distributor will be taking day to day transaction that I want 200 rupees and I give my smart card and I give the finger print I get the data but when I need to upgrade my database.

(Refer Slide Time: 54:04)



Database Configuration

In practice, many systems may use both kinds of database- a distributed database for off-line day-to-day verification and a centralized database for card-free, on-line verification.

So, you go there and give it and that will be centralizing database. We will take care that one this type of problem occurs; for examples in the smart card or ATM machine. Suppose I decide that give 2 fingers and I do not tell which 2 fingers. I have taken your 10 finger data and I will display the point finger and middle finger. You will be giving point finger and middle finger matched it locally and tells you that you got the money.

Now suppose you lost your ATM card, because ATM card contains your data that has been matched. Now you lost your ATM card or because of some reason then the ATM card is spoiled. What are you going to do? You will be going to the bank and tell him that I lost my ATM card and I want to get back another one. Now bank cannot issue the same ATM card, there ATM card number will be different. They will give you the different pin and you have to regenerate your own pin, because once you accept it the responsibility goes to you.

No, he will not give you the same 2 finger prints. He will give another and randomly two more finger prints, you have 10. We have in the database and we have 10 finger prints out of 10, we will select any two randomly. So, those two will be shown and you will be giving the data. So, the centralize database will be useful for this type of thing that it takes care your future need or your online you want to update the data, that is possible through centralize one, which is not the case in the smart card.