CS-573 Fundamentals of Cyber Security
Final project - Authentication : Comparative Study
Group 2

- Michael Fang 10430930
- Chunli Liu 10430963
- Janiece Pandya 10431583
- Disha Sareen 10429086
- Siddhant Barua 10439929

DNA Biometrics in Authentication:

Michael Fang
CWID: 10430930

Introduction:
Biometrics add a layer of security to authentication that is, in general, difficult to fake, relatively unique to each user, and also generally accessible to the user. In the same vein as two-factor authentication, it adds difficulty for an outside third party to obtain or fake this form of data.

DNA authentication is perhaps the ultimate extension of this idea: for a variety of reasons, detailed below, DNA is the ultimate unique identifier of an individual. Its code literally defines every aspect of an individual. Without knowing the sequence, it is impossible to predict and fake, at the time of this writing, another individual's DNA sequence with any certainty. And the sequence itself remains accessible and stable for the duration of an individual's lifespan.

DNA authentication also faces a number of problems. It can be stolen, it is currently expensive and time-consuming to perform, and the high level of discrimination it provides is probably overkill for most applications in use in 2019.

In order to understand this relatively novel field, first let us briefly discuss what features of DNA are salient from a cybersecurity and computer science perspective.

Background:
Deoxyribose Nucleic Acid is a sequence of genetic instructions used in all aspects of life for all known organisms. This sequence is encoded as a string of characters known as *base pairs*: A, T, C, and G, for Adenine, Thymine, Cytosine, and Guanine. This code is

transcribed into RNA, which is the same sequence, except that T is replaced with U. The reasons for why this is done are complicated and outside the scope of our discussion, but it is important to note that most DNA sequences given are actually in their RNA form.

Each three-character grouping of RNA forms a codon, which *translates* to a particular protein; these proteins form the basis of most chemical catalysts of life.

For example, the DNA sequence ATGTTTACACATGGTTGGGTATAG is transcribed into the RNA sequence AUG-UUU-ACA-CAU-GGU- UGG-GUA-UAG, which is translated into the protein peptide chain START-Phenylalanine-Threonine-Histidine-Glycine-Tryptophan-Valine-STOP. [The code is not *necessarily* read from the 'beginning' of the chain, so it is possible to read the code from multiple *frames*.]

With 4 letters, 2 bits are required to encode each A, T, C, or G. As the human genome is roughly 3.088 billion base pairs long (human DNA is double-stranded, so there is a total of ~6 billion base pairs, but as the strands are *complementary*, it is only necessary to read one-strand), it takes approximately 750 MB of memory to encode this amount.

As it turns out, 750MB is not necessary to encode an individual's 'data'. The vast majority of our genomic data is the same, from person to person, or person to dog, or even person to fungus. Some sequences, such as those for ncRNA, ATP transporters, and Serine hydroxymethyltransferase, are so highly conserved between species that they are *identical* in *all* living organisms.

For Humans, all the variation (unique identifiers) occurs in just 0.5% of the genome (Abecasis), though that difference is spread out throughout the genome. Between any two individual humans, their genomes may only differ by 0.1% to 0.6% of their sequence (at most 20 million base pairs out of 3 billion) (Auton). 20 Mbp would require 4.76 Mb to encode; with some padding data to show start sequences and positions, it may bloat a bit more.

Roughly 5 Mb of space to encode the entirety of a human individual's identity is not a lot!

Discussion:

Methods for Authentication with DNA:

At the time of writing, even the fastest and most efficient methods of sequencing DNA are still too slow, expensive, and cumbersome for use in most every-day authentication situations. The following is a short summary of how this process, based on the Sanger method, is accomplished:

1) Clone DNA into bacteria.
2) Slice up the DNA into smaller chunks, because of degradation issues of long strands.
3) Purify the DNA.
4) Divide the sample into 4 separate sequencing reactions, one for each of the bases (A,T,C,G).
5) Extend the DNA strand and sometimes attach radioactive labels to ends.
6) Heat denature the strands.
7) Separate the fragments by size via Gel Electrophoresis.
8) Use X-ray or autoradiography to calculate weights and deduce sequence.

While much of the work can be automated, correct preparation of slides still takes some finesse. Steps 6 and 7 have, until fairly recently, typically taken place in completely separate table-sized machines, requiring sleep-deprived graduate students to manually prepare and load gels [which, by the way, is a pretty weak point-of-failure; what a grad student would do for a Big Mac…]. Furthermore, run-times are very slow. Even very fast systems (read: expensive setups) take 6 to 8 hours to read 700 bp. At this rate it would take 3,400 years to sequence an entire genome without some analytical shortcuts such as parallelization and template generation via genome fragmentation.

As of the time of writing in mid-2019, it takes between 24 and 48 hours to sequence enough of the human genome (satellite DNA is typically not sequenced, for technical reasons) to be able to template through the rest of it. Much of the actual sequencing is matching reading frames to known sequences, rather than actually, meticulously sequencing each and every base pair in the entire genome.

Computationally, distinguishing two string sequences is a trivial matter after the sequence is determined. The problem in terms of authentication is that so much of the human genome is shared between individuals. This is a needle-in-a-stack-of-needles problem, where discrimination of two individuals is reliant on a few differences in a sequence which are, in general, not predictable by any other method than sequencing the strand.

There are some promising future directions for faster and cheaper sequencing, though all are the providence of academia for the time being.

DNA Encryption:
The above discussion about DNA glosses over a critical fact: that the knowledge of an individual's DNA sequence is, very literally, the *entirety* of that individual's identity. While our current knowledge of epigenetics is still in its infancy, having possession of an individual's DNA data has significant consequences for our society.

Most factors that lead to pre-existing health issues can, eventually, be traced to a genetic factor, though it may not necessarily be the *only* factor. One individual's data can also elucidate data for many others, such as a family history of high cholesterol to susceptibility to a particular disease.

And there are society-level questions as well. If one were to know the particular sequence of certain deadly biological agents such as Ebola, Zika, or Smallpox, it may be possible to produce similar deadly vectors; perhaps such information should itself be quarantined. Alternatively, because alleles occur at different rates in different populations and ethnicities, DNA data must be treated carefully lest our own biases lead us to unsavory conclusions about different populations of people.

As such, it is prudent to at least imagine some schemes to encrypt DNA when passing the data over a network. Of particular concern is that because so much of the DNA sequence is shared between individuals, we must protect against educated 'guesses' to match certain allele placements.

If we need to encrypt a large sequence, perhaps an individual's genome, for transmission across an un-secure channel, compression along with header positions for pieces of DNA can be applied in conjunction with a normal encryption method such as AES in order to allow for relatively fast encryption/decryption and ordering of the data. This method is also relatively space-efficient, requiring only about 10-30 Mb of memory to store an entire human genome (Hosseini).

Alternatively, if one only wishes to send shorter sequences over a network, which is likely the case for near-future discrimination authorization schemes, then one could embed the target sequence within a non-self sequence. A primer sequence to determine where in the sequence to start retrieval via PCR, could then be encrypted (Bancroft). Without the primer sequence, it is not feasible to determine the start position of the sequence, as the majority of the DNA sequence is going to be similar but not distinguishable.

Future Directions:

As sequencing technology becomes ever cheaper and automated, the possibility of using DNA as a method of authenticating oneself seems to be an inevitable part of our future. The difficulties in sequencing DNA can be reduced by focusing on the particular variations between individuals, which will improve as our understanding of our shared DNA increases over time. Because this DNA data is inseparable from our identity, using DNA in this fashion also raises questions about security. More work, of course, needs to be done.

Works Referenced:

- Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, et al. (October 2015). "A global reference for human genetic variation". *Nature*. **526** (7571): 68–74. Bibcode:2015Natur.526...68T. doi:10.1038/nature15393. PMC 4750478. PMID 26432245.
- Bancroft et. al. "Long-Term Storage of Information in DNA." *Science.* 293, pp 1763-1765 (2001*)*
- Bjornsson et al*. Journal of the American Medical Association.* 25 July 2008. Johns Hopkins Medical Institution.
- Chen, Yangyi; Peng, Bo; Wang, XiaoFeng; Tang, Haixu (2012). *"Large-Scale Privacy-Preserving Mapping of Human Genomic Sequences on Hybrid Clouds".* School of Informatics and Computing, Indiana University.
- Christley, Scott; Lu, Yiming; Li, Chen; Xie, Xiaohui (2009). *"Human genomes as email attachments". Bioinformatics.* ***25*** *(2): 274–275. doi:10.1093/bioinformatics/btn582. ISSN 1460-2059.*
- Drake JW, Charlesworth B, Charlesworth D, Crow JF (April 1998). "Rates of spontaneous mutation". *Genetics*. **148** (4): 1667–86. PMC 1460098. PMID 9560386.
- *Hosseini, Morteza; Pratas, Diogo; Pinho, Armando J. (2018). "Cryfa: a secure encryption tool for genomic data". Bioinformatics.* ***35*** *(1): 146-158. doi: 10.1093/bioinformatics/bty645.*
- *"Human Genome Project Completion: Frequently Asked Questions". National Human Genome Research Institute (NHGRI). Retrieved 2019-03-29.*
- Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, Rosenbaum AM, Wang MD, Zhang K, Mitra RD, Church GM (9 Sep 2005). "Accurate multiplex polony sequencing of an evolved bacterial genome". *Science*. **309** (5741): 1728–32. Bibcode:2005Sci...309.1728S. doi:10.1126/science.1117389. PMID 16081699.

- https://www.news-medical.net/news/20171130/New-software-could-make-real-time-DNA-authentication-a-reality.aspx

Blockchain based trust & authentication for decentralized sensor networks

Chunli Liu
CWID: 10430963

Abstract

This report is based on BATM(Blockchain based trust & authentication for decentralized sensor networks), a new security model and its protocol based on the blockchain technology to ensure validity and integrity of cryptographic authentication data and associate peer trust level, from the beginning to the end of the sensor network

Introduction

Security and privacy handling for Sensor Networks present new issues due to specific constraints. security and privacy for the data being sent over the network on one side, and node authentication and trust management on the other side. Both have been actively explored the last ten years but none of these works propose a complete model for both content access, security, privacy and trust management, so I will focus on addressing authentication and trust management issues based on BATM.

Report Main Body

1.1.1 Blockchain as a secured data structure

1. blockchain helps leveraging user control over data in the context of social networks and big data, it has been proved as a secure decentralized data structure for new applications, but none has been used to provide node authentication and trust management in Wireless Sensor Networks (WSN) and in the Internet of Things.

2. But now we propose a model based on blockchain data structure used to store decentralized authentication and node trust informations, it is BATM.

1.1.2 BATM authentication

1. Public Key Infrastructure is a major component to resolve authentication in networks which used by Pretty Good Privacy encryption program to provide Confidentiality and Authentication.

2. BATM proposes a new way to achieve these goals using the blockchain as the database to store public keys, digital signature and peer informations, allowing each component of the network to validate informations about every other node in the network.

3. BATM associates cryptographic keys with each NN andcAS in the network. a master key to identify a NN or AS among its lifespan. This key is only used to generate secondary keys for encryption and digital signature.

1.1.3 BATM trust management and trust evaluation

1. The BATM module includes a trust model called Humanlike Knowledge based Trust (HKT), based on human like behaviour to maintain a reputation level for each node.

2. Since we use the payloads contained in the blockchain as an indication of each node behaviour on the network over time that we ensure a node cannot fool others by tampering data or pretending to be someone else. Thus we ensure reliability of trust evaluation without the need of a trust center.

3. For each payload type, HKT defines events and associates them reputation factors. To make the NN reputation evolve naturally over time, each event reputation factor must be weighted by a function evolving in time since the event occurs During it first authentication, a NN has no passed action to compute a reliable trust value. Thus we choose to grant a base trust value to all nodes when a trusted node gives them access to the network by including their credentials in the blockchain.

4. We perform trust evaluation by comparing the current reputation level of a NN to trust him doing certain actions in the network.

1.1.4 BATM payload rules

1. To avoid abuse from NN which can overload the network with payloads to be validated, we introduce specific rules on the payload exchange protocol for BATM.

1.1.5 Future work

1. Trust model: in decentralized networks. More researches on HKT performance must be conducted, and the model itself may evolve to consider more parameters in trust and reputation evaluation.
2. Real world testing:If simulation results fulfill our expectations, BATM will be included in Multicast Services for Linux (MSL), an implementation our SOA network model.

Conclusions and Recommendations

1. We proposes a new application for the blockchain as a secured decentralized storage for cryptographic keys as well as trust informations in the context of autonomous
2. Wireless Sensor Networks.
3. The Blockchain Authentication and Trust Module and its Human-like Knowledge based Trust model shows how to use to immutability of the blockchain to provide solutions to high problematics in the field of decentralized ad-hoc networks.
4. More precisely, we show how it is possible to build a complete solution providing authentication mechanisms as well as trust evaluation in a self-organized and evolutive network.

Reference

- [1] C. M. Medaglia and A. Serbanati, *An Overview of Privacy and Security Issues in the Internet of Things,* D. Giusto, A. Iera, G. Morabito, and L. Atzori, Eds. New York, NY: Springer New York, 2010.
- [2] Y. L. Sun, Z. Han, W. Yu, and K. R. Liu, “A trust evaluation framework in distributed networks: Vulnerability analysis and defense against attacks.” vol. 2006, pp. 1–13, 2006.
- [3] G. Zyskind, O. Nathan, and A. S. Pentland, “Decentralizing privacy: Using blockchain to protect personal data,” *Proceedings - 2015 IEEE Security and Privacy Workshops*, SPW 2015, pp. 180–184,
- 2015.
- [4] N. Fotiou and G. C. Polyzos, “Decentralized name-based security for content distribution using blockchains,” *Proceedings - IEEE INFOCOM,* vol. 2016-September, pp. 415–420, 2016.
- [5] S. Nakamoto, “Bitcoin: A Peer-to-Peer Electronic Cash System,” *www.bitcoin.org, p. 9, 2008*. [Online]. Available: https://bitcoin.org/bitcoin.pdf
- [6] I. Eyal, A. E. Gencer, E. G. Sirer, and R. Van Renesse, *Bitcoin-ng: A scalable*

blockchain protocol, 2016.

- [7] B. Solhaug, D. Elgesem, and K. Stolen, “Why trust is not proportional to risk,” pp. 11–18, 2007.
- [8] T. Erl, *Service-oriented architecture: concepts, technology,* and design.Pearson Education India, 2005.
- [9] Z. Shelby, K. Hartke, and C. Bormann, “The constrained application protocol (coap),” *Standard Track, Internet Engineering Task Force (IETF), 2014.*
- [10] S. De, P. Barnaghi, M. Bauer, and S. Meissner, *Service modelling for the Internet of Things,*2011.
- [11] P. R. Zimmermann, *The official PGP user’s guide*. MIT press, 1995.
- [12] J. Callas, L. Donnerhacke, H. Finney, D. Shaw, and R. Thayer,“Rfc 4880 - openpgp message format,” *Proposed Standard, Internet Engineering Task Force (IETF),* 2007.
- [13] D. Gambetta, “Can we Trust Trust?” *Trust: Making and breaking coopeative relations, pp*.213–237, 1990.
- [14]International Journal of Network Security, Vol.19, No.5, PP.653-659, Sept. 2017 (DOI: 10.6633/IJNS.201709.19(5).01)

Passwordless Authentication:

Janiece Pandya
CWID: 10431583

Introduction

Passwordless Authentication is any method of verifying the identity of a user that does not require the user to provide a password. According to a survey, an individual has an average of almost 30 online accounts. It is almost difficult and tedious to create and maintain different passwords for all of them, so usually, the user reuses one or many of those passwords. In other words, it can be said that passwords are no longer enough. With this form of authentication, users are presented with the options of either logging in simply via a magic link, biometrics like fingerprint, facial recognition, iris scan, or using a token that is delivered via email or text message. The quote “Out with the Old - In with the new” definitely goes with this form of authentication.

Background

Bill Gates quoted in the RSA Conference of 2004 that, “There is no doubt that over time, people are going to rely less and less on passwords. People use the same password on different systems, they write them down and they just don't meet the challenge for anything you really want to secure."

He had already predicted the death of passwords as a safe form of authentication. In today’s technocratic world identity protection is need of the hour. Millions and millions of accounts are compromised due to the data breaches that have happened in the last few decades. 81% of hacking-related breaches leveraged either stolen and/or weak passwords out of which 60-63% breaches were due to password leaks. So if people think that it did not happen to them, then it probably already has.

According to estimated cost analysis for Microsoft Inc. in the month of July 2017, almost 6,86,000 people forgot their passwords and asked for links to reset their passwords. The cost behind managing all this process of resetting the password to deleting the old passwords and updating the new ones rounded to more than $12 million.

Hence, it can be deduced that the economy and energy spent behind these databases which in turn can be breached, cracked or phished is just not worth it. Choosing the option of going passwordless is the wisest decision that anyone could take.

Ideology

A world without passwords should be delivered on two key-promises: User promise and Security promise. User promise guarantees that the end-users should never have to deal with passwords again in their day-to-day lives. Whereas, security promise guarantees that user credentials cannot be cracked, breached or phished.

This ideology can be broken into four steps:
1) **Develop password replacement offerings**: Replace the passwords with a new set of alternatives which would address the shortcomings of passwords while embracing the positive attributes.
2) **Reduce the user-visible password surface area**: Upgrade all your devices and systems with the passwordless settings so that the users experience these changes in a similar way they used to experience during the password oriented system.
3) **Simulate a password-less world**: Enable or convince the end users and IT admins to simulate and transition into this new world of authentication with confidence and ease.
4) **Eliminate passwords from the identity directory**: This step is the final frontier which means to delete passwords from the system's databases after transitioning into passwordless settings.

Working Mechanism

The ideal working environment for any passwordless authentication method would be when a user tries to login or access an application from his device using his biometrics, soft or hard tokens, risk-based authentication, push type methods or anything that does not involve a password. In order to authenticate his/her identity, an identity provider is required to validate the identity of this user by taking the user's passwordless credentials and matching it in its database.

- Let us consider a scenario saying that Alice wants to login into an application from her device using a bio-gesture which unlocks Trusted Platform Module (also known as TPM) which is holding her private key.

- When she enters her bio-gesture, the device would respond with “Hi Alice” signal along with key ID.
- The device asks the identity provider to validate her authentication for which the identity provider would send back a nonce and ask the device to sign it with its private key.
- The device uses the private key to sign nonce and returns to the identity provider along with key ID.
- Identity provider decrypts the encrypted nonce using the public key of the user by looking up in his database and returns a Primary Refresh Token along with encrypted session key protected in TPM.
- The device returns the signed PRT and derived a session key to identity provider to verify.
- Now Alice can access the applications without the need of authenticating again.

Hence, this kind of passwordless mechanism can be done by either using biometric-gesture or by sending the user a magic link or a one time password via email or SMS.

Benefits

- The first and foremost advantage is that the user would never have to deal with the passwords again which removes friction during digital experiences.
- It is very easy to use and convenient way of authentication. It is also definitely more secure than the password oriented method in lieu of all the data breaches that have happened so far.
- The IT department has to no longer worry about spending millions of dollars behind the databases that were required to store these passwords. So, the desktop cost is reduced to a higher amount making it easy to deploy during upgrades.
- Users have experienced a sudden improvement in their experience of accessing a website. At the end of the day, it is the users that the company has to please.

Drawbacks

- After researching a lot about the flaws of this method, it was concluded that this authentication method has only a few minor disadvantages.
- The entire working mechanism is dependent on the device and the identity provider. Therefore, if either of these systems is down then the authentication process would be unsuccessful.

- This technology is only prone to attacks with a compromised system. If the hacker is in possession of system and knows the passwordless credentials then and only then this method fails.
- As this is a developing technology, it is very hard to convince the users that are comfortable using password oriented applications to get compatible with and use the passwordless applications. This use case can be found similar with the concept of convincing a coffee lover to switch to green tea which he/she would not prefer at first even though it is more healthier but gradually after knowing the benefits the person might be ready to try.

Real-time Examples

Microsoft took the initiative of going passwordless with their latest applications like Windows Hello for Business and Microsoft Authenticator App. Other applications such as Duo Security, HYPR, Auth0 Universal Login, and many more applications are setting the trend of going passwordless. Stevens Institute of Technology use Duo Security for the authentication of the users logging in through Workday which is a career building application.

Conclusion

Going passwordless is a long-term approach for secure authentication, and it's still evolving and it can take time to transition. For users that can't go password-less, turn on MFA to validate users and minimize prompts based on the risk of the sign-in with conditional access capabilities.

The latest approach was made on April 10, 2018 when the World Wide Web Consortium (W3C) and FIDO Alliance announced the promotion of the WebAuthn spec to the "Candidate Recommendation" stage, the precursor to the final approval of a web standard and the W3C has invited online services and website application developers to implement WebAuthn. It can be foreseen that this passwordless strategy is the future of secure authentication.

References

- https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE2KEup
- http://jasonmun.com/wp-content/uploads/2013/03/6ws.jpg
- https://dzone.com/articles/how-passwordless-authentication-works
- https://www.tracesecurity.com/blog/articles/81-of-company-data-breaches-due-to-poor-passwords
- https://www.forbes.com/sites/forbestechcouncil/2018/06/04/the-future-of-authentication-is-here/#722d734d432e

- https://www.okta.com/blog/2018/04/its-a-new-world-with-webauthn-passwordless-authentication-goes-primetime/
- https://ieeexplore.ieee.org/document/8029677/references#references
- https://www.sitepoint.com/passwordless-authentication-works/
- https://dzone.com/articles/how-passwordless-authentication-works
- https://doubleoctopus.com/security-wiki/authentication/passwordless-authentication/
- https://www.pingidentity.com/en/company/blog/2017/10/03/achieving_secure_passwordless_authentication.html
- https://www.tracesecurity.com/blog/articles/81-of-company-data-breaches-due-to-poor-passwords
- https://www.theguardian.com/technology/2019/jan/17/breached-data-largest-collection-ever-seen-email-password-hacking
- https://nakedsecurity.sophos.com/2018/11/22/the-passwordless-web-explained/
- https://auth0.com/docs/getting-started/overview
- https://www.microsoft.com/security/blog/2018/05/01/building-a-world-without-passwords/
- https://www.w3.org/2018/04/pressrelease-webauthn-fido2.html.en
- https://channel9.msdn.com/Events/Ignite/Microsoft-Ignite-Orlando-2017/BRK2078

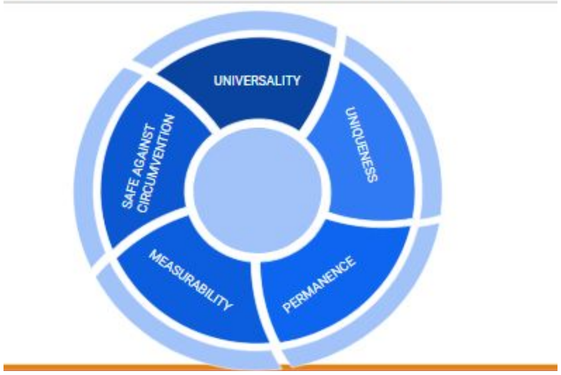
Biometric Authentication
Disha Sareen
10429086

Introduction

Authentication by Biometrics is a way to uniquely verify a human being’s identity by their physiological characteristics, such as the eyes & ears or behavioral attributes such as the unique way one solves a security-authentication puzzle.
To be useful, biometric data must be unique, permanent and collectible. Once measured, the information is compared and matched in a database.[1]

Checklist for Biometric features

In order for us to successfully use a human trait for authentication, it must satisfy the following attributes:



1. Universality: All people possess the feature.
2. Uniqueness: The feature is different for people so the system can distinguish between them.
3. Permanence: The feature only varies slightly over time.
4. Measurability: The system can acquire and process the feature in an efficient way.
5. Safe against circumvention: The system can distinguish between the real feature and a dummy.[2]

Physical components for a Biometric Authentication System

A typical biometric authentication system comprises of the following 3 components:
1. Sensor: This is what records your information, as well as reads it when your biometric information needs to be recognized.
2. Biometric Application: There has to be a computer system to compute & compare

features.

3. Database: The database is required to store the biometric templates.[1]

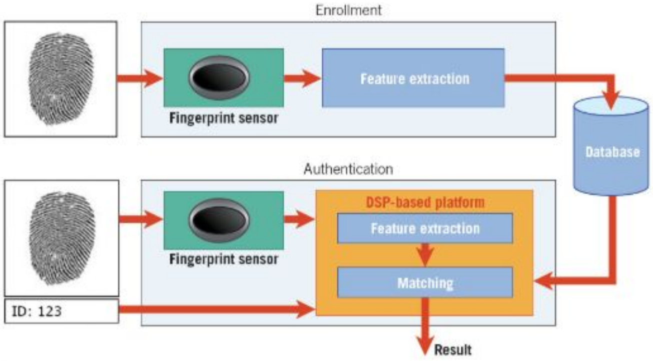
Stages in Biometric Authentication

There are 2 stages in a Biometric Authentication which are the following:

1. Enrollment: : During enrollment, each person's fingerprints/Face's are scanned, analyzed, and then stored in a coded form on a secure database.

2. Verification: The scanner takes their biometric, checks it against all the templates in the database stored during enrollment, and decides whether the person is entitled to gain access or not. [3]

The following diagram describes the system relationship between the two.



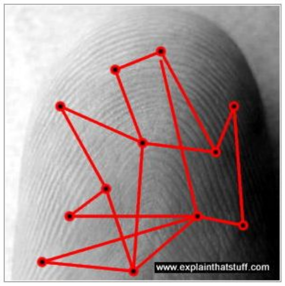
[3]

Fingerprint as a Biometric Authentication Trait

Fingerprint authentication is the process of matching fingers based on the structure of the upper skin. During the prenatal development of a human being, the skin of hands and feet fold in a random process leading to ridges and valleys. Those end or bifurcate at certain points called minutiae. Although some genetic influence, the position, orientation and type of these minutiae points are unique for each human and even each finger. [2]

The applications of Fingerprint authentication can be found in Civilian and commercial applications like military, law enforcement, medicine, education, civil service, forensics, driver license registration, cellular phone access, computer log-in and so on. More specifically we can find the applications in Apple's Touch ID.

Reading the fingerprint



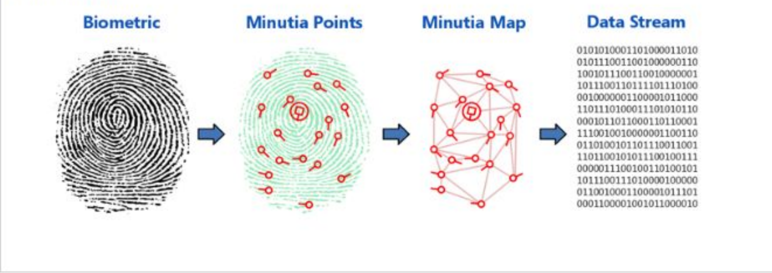
During enrollment or verification, each print is analyzed for very specific features called minutiae, where the lines in your fingerprint terminate or split in two.

The computer measures the distances and angles between these features—a bit like drawing lines between them—and then uses an algorithm to turn this information into a unique numeric code.

Comparing fingerprints is then simply a matter of comparing their unique codes. If the codes match, the prints match, and the person gains access.

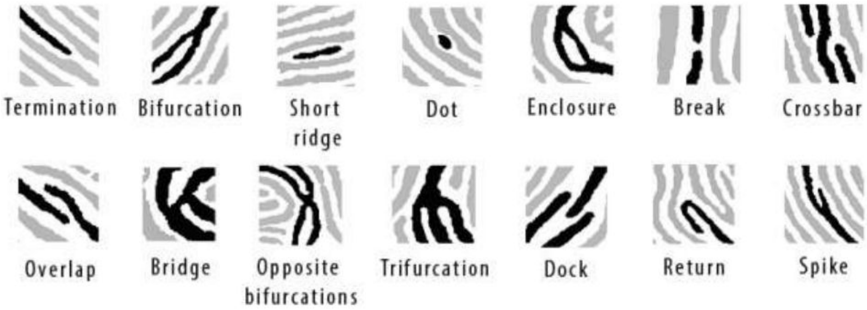
The Data Process

A pattern or “map” of the minutiae is stored in a database as a representation of the fingerprint. In essence, what’s stored in the memory is not the fingerprint itself but a set of key minutiae, as shown below. [4]



Possible Minutiae Detection Patterns

The following are some of possible minutia features which are usually appeared in a fingerprint. [5]



The Pros and Cons of Biometrics for Authentication

Authentication is a critical process. Hence its pros and cons must be carefully analysed before it is put into practise. The following must be taken into account:

PROS:

- Can't Fake it (Highly difficult)
- Convenience is Key
- Role in 2 Factor Authentication
- Reliable in Authentication

CONS:

- Privacy at Risk
- Feels Invasive
- Injury & Illness
- Common across multiple systems
- Cannot reset or revoke

Conclusion

We can see the intricacies of an authentication system based on Biometrics. One must carefully weigh the advantages & disadvantages of such a system before putting it into use as Identity & System Security are crucial areas in the modern day.It is also important to consider biometric information related to an individual as a high value asset. Exfiltration of such information can be devastating to the entire system and other systems who rely on biometrics as an authentication process.

References

[1] https://us.norton.com/internetsecurity-iot-biometrics-how-do-they-work-are-they-safe.html
[2] https://www.usenix.org/system/files/conference/woot18/woot18-paper-fietkau.pdf
[3] https://www.explainthatstuff.com/fingerprintscanners.html
[4] https://semiengineering.com/fingerprint-senor-technology-and-security-requirements/
[5] https://ijerat.com/uploads/2/3788_pdf.pdf
[6]https://www.bayometric.com/12-reasons-consider-fingerprint-authentication/
[7] https://patentimages.storage.googleapis.com/83/2b/58/ba2cf29229a75f/US20190034606A1.pdf

[8]https://www.androidauthority.com/how-fingerprint-scanners-work-670934/
[9]https://it.toolbox.com/blogs/carmashoemaker/pros-and-cons-of-biometrics-for-security-053118
[10]https://www.ieee-security.org/TC/SP2011/PAPERS/2011/paper003.pdf

Application of Computer Vision in Authentication:

Siddhant Barua

CWID: 10439929

1.Introduction

Computer vision is defined as a field of study that seeks to develop techniques to help computers "see" and understand the content of digital images such as photographs and videos. Typically, this involves developing methods that attempt to reproduce the capability of human vision.Computer Vision includes methods for acquiring, processing, analysing, and understanding images or image sequences from the real world in order to produce information.

Biometrics deals with the recognition of persons based on physiological characteristics, such as face, fingerprint, vascular pattern or iris, and behavioural traits, such as gait or speech.

We can hence make a strong correlation between Computer Vision and Biometrics. Described in this paper, are a few Biometric authentication methods and how Computer Vision relates to each of those methods.

2.Retinal Scan

The human retina can be described as a thin tissue composed of neural cells, and is located in the posterior portion of the eye.

The retina contains a complex network of blood capillaries, that are unique to each individual. The network of blood capillaries are so unique, that even identical twins don't share the same network.

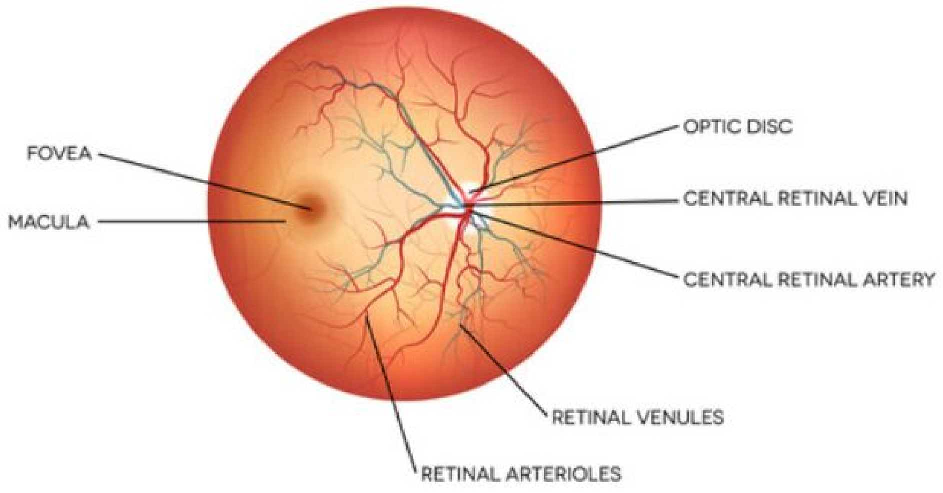
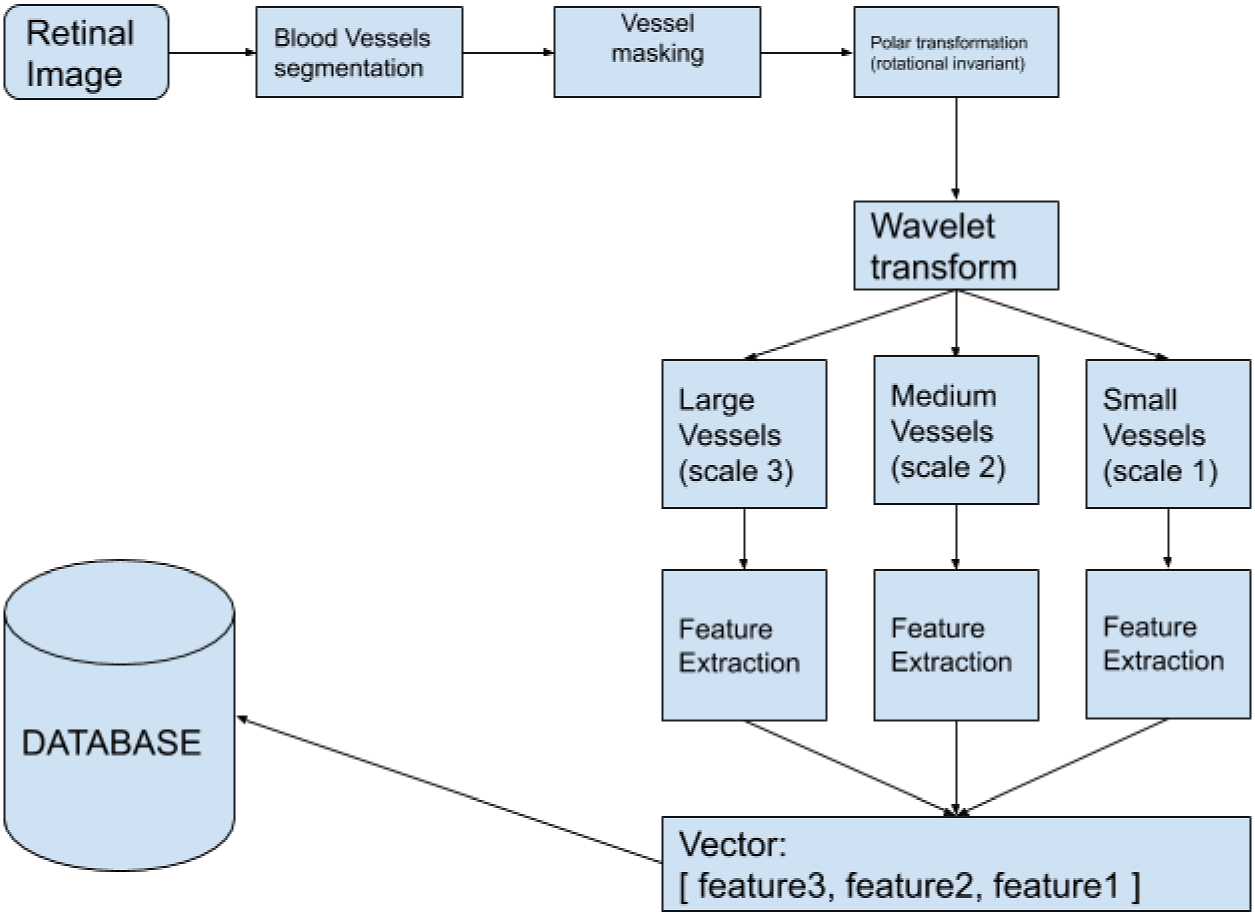


Figure 1

2.1.How Retinal Scans work?

A Biometric identifier known as a retinal scan is used to map out the unique patterns of a person's retina. The person places his eye in close proximity to the retinal scanner. A low energy infrared light is hence passed onto a person’s eye as they look through the scanners eyepiece.

The blood vessels and capillaries absorb light more readily than the surrounding auxiliary tissues. This leads to a variation in the amount of infrared light reflected. This pattern of variation is converted to computer code and hence stored in the database for future retrieval.



[2] Figure 2: Retinal identification system. (Feature extraction (F.E.)).

Figure 2 describes a novel Retinal Identification System [2]. After the image of the retina is obtained,

- The blood vessels are then segmented,
- We then perform vessel masking ([3] *In computer science, a **mask** or **bitmask** is data that is used for bitwise operations*) occurs in the vicinity of the Optical Disc of the eye.
- We then perform polar transformations ([4] ***Polar coordinates** provide a method of rendering graphs and indicating the positions of points on a two-dimensional (2D) surface. ... The **polar** plane consists of a reference axis, or ray, that emanates from a point called the origin.*), this is to obtain a rotational invariant binary image, which highlights the major blood vessels and capillaries.
- We then use wavelet transforms in order to separate the vessels according to their diameter sizes.
- The blood vessels are then classified into 3 scales (*scale 1 for smaller blood vessels, scale 2 for medium sized blood vessels and scale 3 for larger blood vessels*).
- We then push these classified blood vessels into a feature vector. Which is then stored in the database with an ID tag to identify which user it belongs to.

2.1.1.Authentication

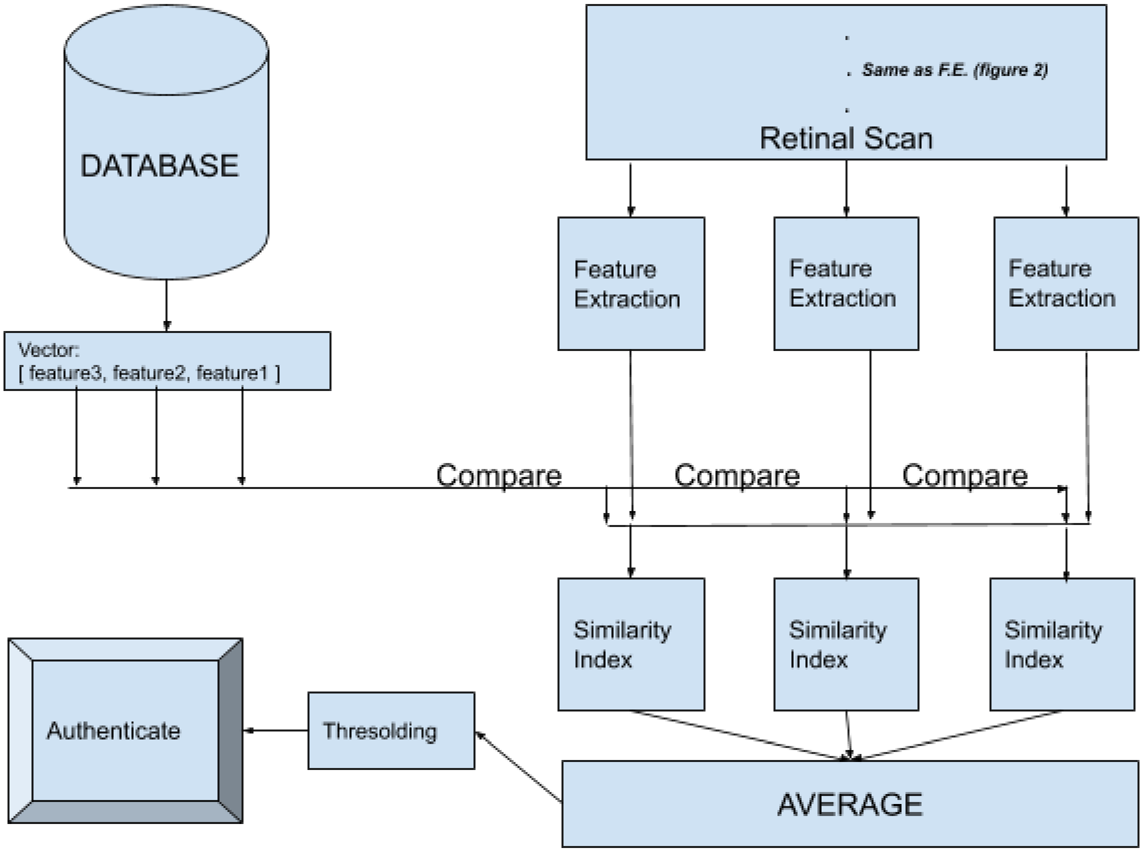


Figure 3: Retinal Authentication system

- We check for the degree of similarity between the features in the database and the received features, this is done between the 3 aforementioned scales. (*scale 1 for smaller blood vessels, scale 2 for medium sized blood vessels and scale 3 for larger blood vessels*)
- We then take the average and check against a set threshold value.
- If the computed average is within the threshold range $[0\ to\ t]$ the user is authenticated.

2.2.Advantages

- High level of accuracy
- Low occurrence of false positives
- Most accurate method of biometrics
- Reliable due to the fact that retina capillary structures doesn't change over a lifetime.
- No two individuals share the same retinal pattern

2.3.Disadvantages

- Diseases, and genetic birth defects can affect accuracy.
- High costs
- Need of very specific equipment, requiring the person's eye to be in very close proximity to the equipment in order to get maximum coverage of the retinal pattern .
- Perception of invasiveness.

3.IRIS Scan

The **iris** (plural: irides or **irises**) is a thin, circular structure in the eye, responsible for controlling the diameter and size of the pupil and thus the amount of light reaching the retina. Eye color is defined by that of the **iris**.[5].

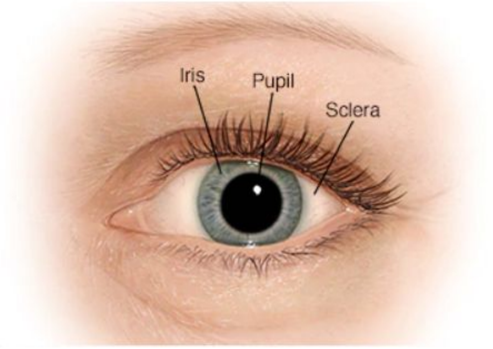


Figure 4: Structure of the eye. [6]

3.1.How Iris Scans work?

Unlike retina scanning, iris recognition uses camera technology with subtle infrared illumination to acquire images of the detail-rich, intricate structures of the iris. Uses a camera similar to any digital camera, to capture an image of the Iris.
Digital templates encoded from these patterns by mathematical and statistical algorithms allow unambiguous positive identification of an individual. Databases of enrolled templates are searched by matcher engines at speeds measured in the millions of templates per second per (single-core) CPU, and with infinitesimally small False match rates.
Mostly used for mobile authentication, passport-free automated border-crossings, and some national ID systems based on this technology are being deployed.

3.2.Simple Iris Scan Model.

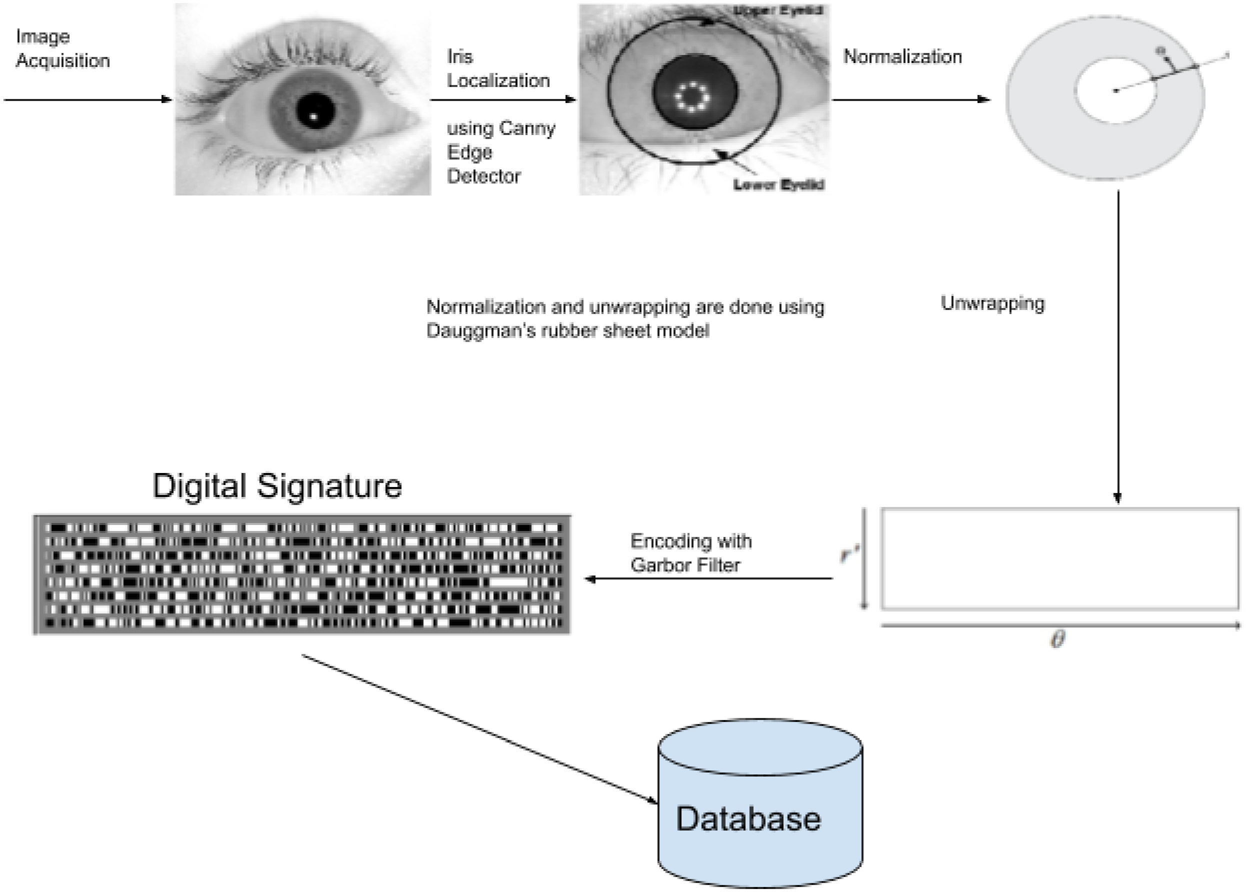


Image Acquisition
Iris Localization
using Canny Edge Detector
Upper Eyelid
Lower Eyelid
Normalization
Normalization and unwrapping are done using Dauggman's rubber sheet model
Unwrapping
r'
θ
Digital Signature
Encoding with Garbor Filter
Database

Figure 5: Iris Digital Signature Extraction.

- We first capture the image of the person.
- We then locate the eye in the captured image.
- After which, the Localization of the Iris (***Iris localization means segmentation*** *of* ***iris*** *from the other parts of eye like* ***pupil****, sclera, eyelids and eyelashes in the image of an eye*) takes place using a Canny Edge Detector.[7]
- Then the data is normalized and unwrapped using the Dauggman's rubber sheet model.[8]
- The resulting data is encoded with a Garbor Filter.[9]
- After encoding we Obtain the Digital Signature as the output, which is unique to each iris, *(including your left and right iris)* .
- This Digital Signature is stored in the database for further retrieval.

3.2.1.Authentication

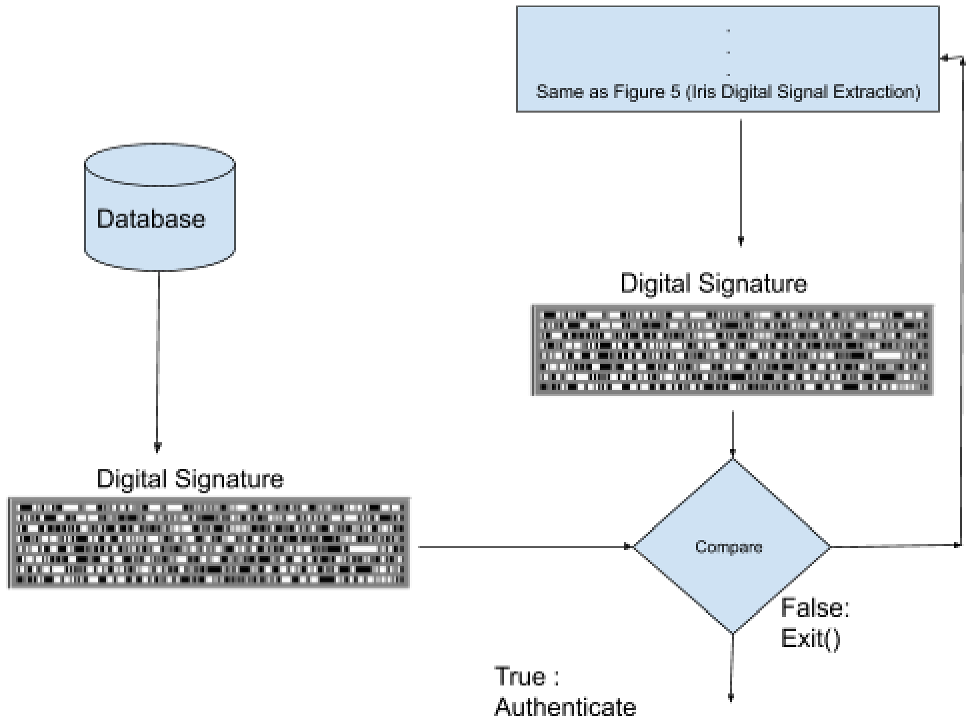


Figure 6: Iris Digital Signature Authentication.

- We perform the iris scan just like in section 3.2 (Iris Digital Signal Extraction)
- Then we take the digital signature obtained from the scan, compare it with the digital signatures stored in the database.
- Upon a successful match, the user is Authenticated and upon failure a count variable is updated, and then we start the scan again.
- If the count fails $n\ (variable)$ times then access is blocked.

2.3.Advantages

- High Accuracy.
- Can be scaled to larger applications.
- Distance isn't as big a factor as compared to Retinal Scans.
- Perceived Non Intrusive nature.
- No two individuals share the same Iris encoded pattern, no set of eyes share the same iris encoded pattern.

2.4.Disadvantages

- Subject needs to be steady.
- Glossy eyes, leads to reflection.
- A lot of memory is required to detect the iris pattern and store it.
- High cost.

4.Ear Scan

The ear is the organ of hearing and balance. The ear is usually described as having three parts - the outer ear, the middle ear and the inner ear.

Each human is said to have a unique set of ears , this was tested by Iannerelli - who conducted an ear identification study among 10,000 individuals, each of whom had a unique set of ears.

Ear biometrics is a unique class of biometrics. It is a type of passive biometrics, similar to iris or facial biometrics.

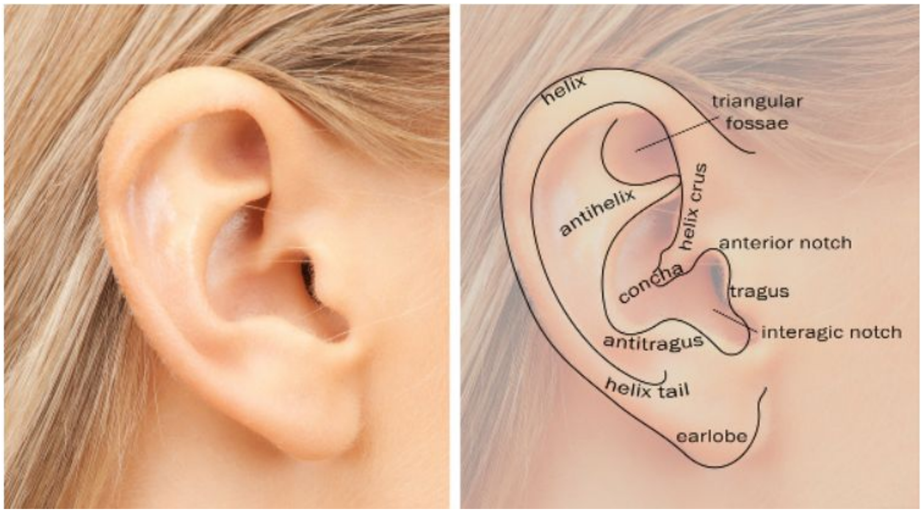


Figure 7: Ear anatomy.

4.1.How Ear Scans work?

Similar to iris scanning, the person's image is taken where the ear is located.

Then the curvature of the ear is taken into account, and graphs are built based on these curvatures.These graphs, being different for each human being is stored in a database for further retrieval.

4.2.A simple Ear Scan model

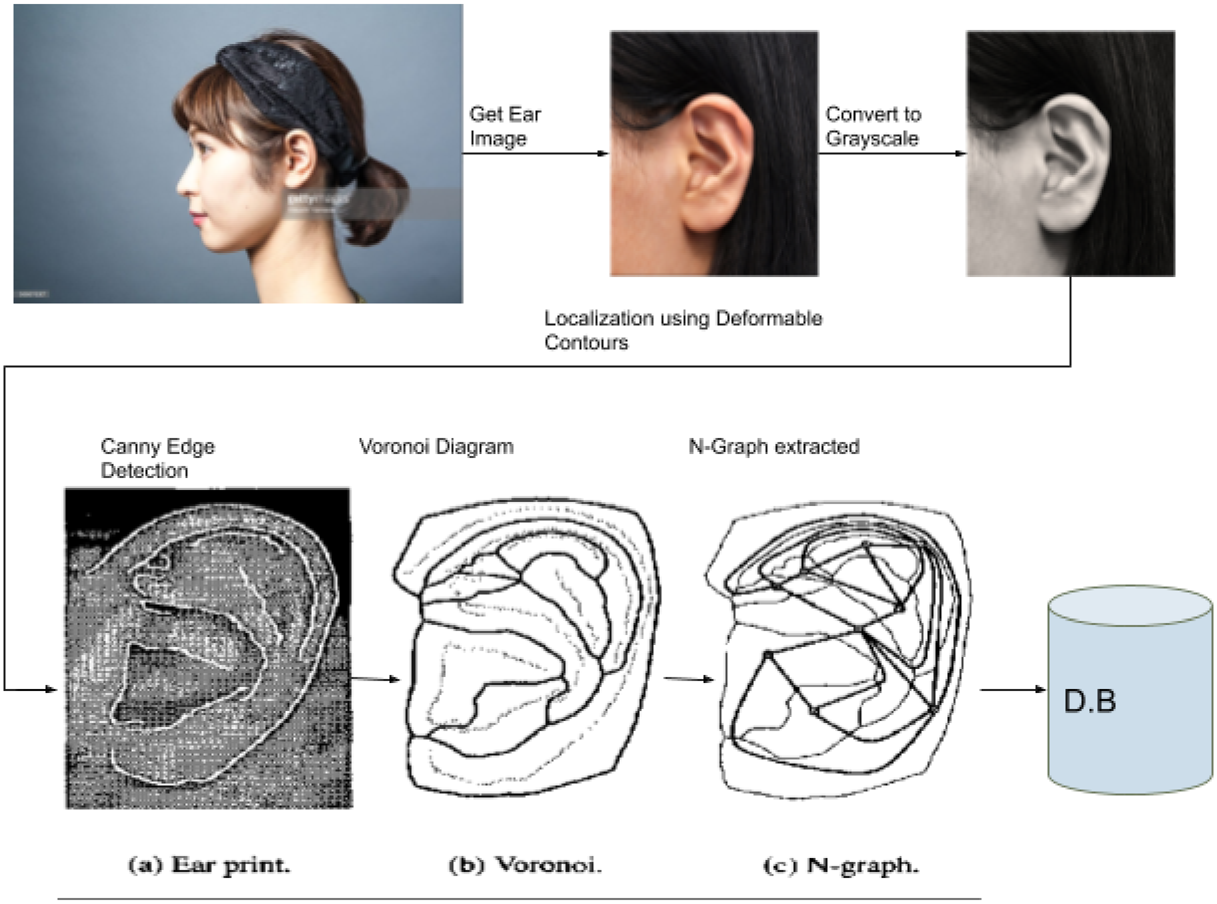


Figure 7: Ear Graph Extraction.

- We first take the image of our subject.
- Then we find the ear in the image.
- The image is then converted to greyscale.
- We then localize the ear using Deformable contours [10], on a Gaussian Pyramid representation of the Image gradient. [11].
- We then run the Canny Edge Detection algorithm, to extract the edges of the ear. [12]
- After this we relax the edges, *(edge relaxation)* to get only the major curve segments.
- We then create a voronoi diagram (*In mathematics, a **Voronoi diagram** is a partitioning of a plane into regions based on distance to points in a specific subset of the plane.*) from the edges and then extract the corresponding Neighbourhood graph or N-Graph. [13]
- This is then stored in the database for further Retrieval.

4.2.1.Authentication

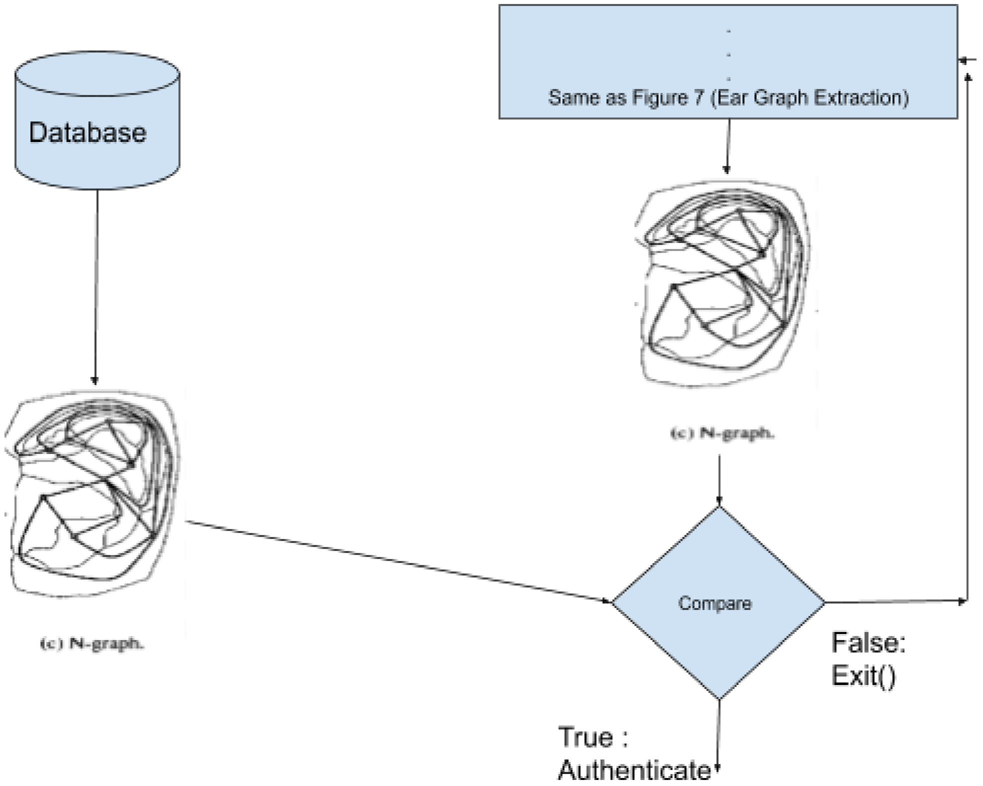


Figure 8: Ear Graph Authentication

- The persons N-Graph is obtained by the same process as Section 4.2 *(Ear Graph Extraction.)*
- We then test the N-Graph against the N-Graph(s) in the database.
- If the Graphs match then the user is authenticated else the scanning loop starts again till the set limit is reached, where in the user is blocked from the system.

4.3.Advantages

- Ear biometric authentication is 2 factored by design.
- Combines visual recognition and as well as acoustic recognition.
- The ear is resistant to change with age.
- Difficult to counterfeit.

4.4.Disadvantages

- Lacks robust research as compared to Iris, Retinal or Facial scans.
- Not as reliable as Retinal or Iris scans.
- Since this is a form of passive biometrics, the presence of hair over the ear can yield unfavorable results.

5.Conclusion

We conclude this study with the knowledge that there is a strong correlation between Computer Vision concepts and biometrics.

We proposed simple models for Retinal, Iris and Ear scans. Each of which have their own set of advantages as well as disadvantages.

We learnt that Retinal Scans, are the most accurate form of biometrics, but it’s not used as commercially as Iris Scans due to the perception of invasiveness as well as the high cost of the scanning equipment incurred.

We established, that ear scans are a form of passive biometrics.

We also found out that despite ear scans being two factored by design (Visual and Acoustic), is not as commercially viable as Retinal or Iris scan, due to the fact that Hair covering the ears, is a major pain point for this class of Biometrics.

Thus concludes are study into the application of Computer Vision in Authentication.

6.Citations

- [1].Figure 1: https://eyeinstituteaz.com/eye-care-specialties/retina-mesa/
- [2].Figure 2: Research Article A Novel Retinal Identification System By Hadi Farzin and Hamid Abrishami-Moghaddam from: *Department of Electrical Engineering, K.N. Toosi University of Technology, Seyed Khandan, 16315-1355 Tehran, Iran*
  Mohammad-Shahram Moin from: *Iran Telecommunication Research Center, North Kargar, 14399-55471 Tehran, Iran*
- [3].https://en.wikipedia.org/wiki/Mask_(computing)
- [4].https://en.wikipedia.org/wiki/Polar_coordinate_system
- [5].https://en.wikipedia.org/wiki/Iris_(anatomy)
- [6].https://www.mayoclinic.org/healthy-lifestyle/adult-health/multimedia/eyes/sls-20076536?s=4
- [7].Canny Edge Detector
  https://en.wikipedia.org/wiki/Canny_edge_detector
- [8].Daugman’s Rubber Sheet Model
  https://www.researchgate.net/figure/Daugmans-rubber-sheet-model_fig1_235338091
- [9].Garbor Filter
  https://en.wikipedia.org/wiki/Gabor_filter
- [10].Deformable Contours
  http://www.cs.utexas.edu/~grauman/courses/spring2011/slides/lecture10_snakes.pdf
- [11].Gaussian Pyramid
  http://persci.mit.edu/pub_pdfs/RCA84.pdf
- [12].Canny Edge Detection in OpenCV
  https://docs.opencv.org/master/da/d22/tutorial_py_canny.html
- [13]. Voronoi Diagram
  http://jwilson.coe.uga.edu/EMAT6680Fa08/Kuzle/Math%20in%20Context/Voronoi%20diagrams.html
- Papers Referenced:
  - Research Article A Novel Retinal Identification System By Hadi Farzin and Hamid Abrishami-Moghaddam from: *Department of Electrical Engineering, K.N. Toosi University of Technology, Seyed Khandan, 16315-1355 Tehran, Iran*
    Mohammad-Shahram Moin from: *Iran Telecommunication Research Center, North Kargar, 14399-55471 Tehran, Iran*
  - IRIS RECOGNITION BY USING IMAGE PROCESSING TECHNIQUES
    By- Mohamed Alhamrouni from Atilim University

  https://www.researchgate.net/publication/314194215_IRIS_RECOGNITION_BY_USING_IMAGE_PROCESSING_TECHNIQUES

  - Ear biometrics in computer vision: By Wilhelm Burger. https://www.researchgate.net/publication/224068480_Ear_biometrics_in_computer_vision

- Other References:
  - A Survey on Ear Biometrics - by *AYMAN ABAZA, WVHTC Foundation ARUN ROSS, West Virginia University CHRISTINA HEBERT, and MARY ANN F. HARRISON, WVHTC Foundation MARK S. NIXON, University of Southampton.*
  - Iris recognition system - technical overview. By :- *Elijah Omidiora Ladoke Akintola University of Technology, Falohun A.s Ladoke Akintola University of Technology, John Ojo Ladoke Akintola University of Technology.*
  - Daugman, G. J., (2002). How iris works. Proceedings of 2002 International Conference on Image Processing.(ICIP), pp 1.33 -1.36.
  - Xinchao Wang - Computer Vision Lecture Slide 1.