Volume 2, Issue 5, May 2012 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

Secure Data Storage in Cloud Environment Using Biometrics

K.Govinda[*]
SCSE,VIT University
kgovinda@vit.ac.in

Yannick Ngabirano
SCSE,VIT University
ngabiranoy@gmail.com

***Abstract*— Cloud computing has been envisioned as the next-generation architecture of IT enterprise. In contrast to traditional solutions, where the IT services are under proper physical, logical and personnel controls, cloud computing moves the application software and databases to the large data centers, where the management of the data and services may not be fully trustworthy. This unique attribute, however, poses many new security challenges which have not been well understood. In this article, we focus on cloud secure data storage, which has always been an important aspect of quality of service (QOS). To ensure the secure storage of user's data in the cloud, we propose an effective and flexible bio metric authentication using face.**

***Keywords*— Biometric, Cloud, Storage, Face, KDC.**

I. INTRODUCTION

Biometrics offers new perspectives in high-security applications while supporting natural, user-friendly and fast authentication. Biometric identification considers individual physiological characteristics and/or typical behavioural patterns of a person to validate their authenticity [1]. Compared to established methods of person identification, employing PIN-codes, passwords, magnet- or smart cards, biometric characteristics offer the advantages include they are significant for each individual, They are always available, They cannot be transferred to another person, They cannot be forgotten or stolen, They always vary.**Peter T. Higgins[2]** described this Biometric definition Biometrics are automated methods of recognizing a person based on a physiological or behavioural characteristic. **James L. Wayman[3]** described by A "biometric" technology is an automatic method for the identification or identity verification of an individual based on physiological or behavioural characteristics. Biometric *systems* function to identify individuals by matching a specific personal characteristic, the biometric identifier, with one previously recorded. Another definition of biometrics. A biometric is any *measurable*, robust, distinctive physical characteristic or personal trait that can be used to identify or verify the claimed identity of an individual. Biometric authentication, in the context of this report refers to automated methods of identifying or verifying the identity of a living person. The italicized terms above require explanation. Measurable means that the characteristic or trait can be easily presented to a sensor and converted into a quantifiable digital format. This allows for the automated matching process to occur in a matter of seconds. The robustness of a biometric is a measure of the extent to which the characteristic or trait is subject to significant changes over time.These changes can occur as a result of age, injury, illness, occupational use or chemical exposure. A highly robust biometric does not change significantly over time, A less robust biometric does. For example the iris which changes very little over a person's lifetime is more robust than a voice. Distinctiveness is a measure of the variations or differences in the biometric pattern among the general population. The higher the degree of distinctiveness; the more unique the identifier. The highest degree of distinctiveness implies a unique identifier. A low degree of distinctiveness indicates a biometric pattern found frequently in the general population. The iris and the retina have higher degrees of distinctiveness than hand or finger geometry. The application helps determine the degree of robustness and distinctiveness required. Living person distinguishes biometric authentication from forensics which does not involve real-time identification of a living individual. The term "***Biometrics***" has come to be associated with the automatic identification of a person based on a feature or characteristic [3].These may be based on either:

- A physiological characteristic such as a fingerprint or face
- A behavioural characteristic such as a signature or voice

A variety of methods and techniques are available today, with the most common being are Iris/Retina, Voice Signature ,Fingerprint ,Face. Generally, face, signature and voice are considered to be a lower level of security than fingerprint and iris, with iris scanning being the method of choice for extremely sensitive areas. This does not mean that the others aren't effective, but there is a price difference between the high and low end. It is a good idea to select a system that meets the basic needs. This technology is preferred over traditional methods of identification (passwords/PINs) because it requires that a person actually be present at the point of identification and these features cannot be stolen or forged. It can be used to prevent unauthorized access to ATMs, cell phones, home and office computers, automobiles and secure areas. It can also be used during electronic transactions over the internet or telephone. Biometric systems are not perfect. An authorized user may be rejected by the system while an unauthorized user may gain access to it. The False Rejection Rate (FRR) is usually caused by lighting, climate conditions, low quality equipment or inexperience. The False Acceptance Rate (FAR) is caused by the security standard being too low[4]. The later is far more serious, as it poses a great risk to have unauthorized people gaining access to the systems. The FARs and FRRs vary between biometric techniques, but iris scanning has proven to be the only one that has never had a false acceptance. Biometric systems are becoming commonplace in the society[5]. Though many people have failed to accept these new methods of identification/verification because of the "Big Brother" fear, education and first-hand experience with the technology is slowly winning some people over. One thing to remember is that even though biometrics are a very strong method of security, a single "key" should not suffice. It is generally considered that the use of biometrics or the combination of a biometric with a more traditional method can ensure a higher standard of security

II. Characteristics of Biometrics

Table 1 compares the eight mainstream biometrics in terms of a number of characteristics, ranging from how robust and distinctive [6]. This table is an attempt to assist the reader in categorizing biometrics along important dimensions. Because this industry is still working to establish comprehensive standards and the technology is changing rapidly, however, it is difficult to make assessments with which everyone would agree. The table represents an assessment based on discussions with technologists, vendors, and program managers. The table is not intended to be an aid to those in the market for biometrics, rather it is a guide for the uninitiated. When comparing ways of using biometrics, half can be used for either identification or verification, and the rest can only be used for verification. In particular, hand geometry has only been used for verification applications, such as physical access control and time and attendance verification. In addition, voice recognition, because of the need for enrolment and matching using a pass-phrase, is typically used for verification only. There is considerable variability in terms of robustness and distinctiveness. Fingerprinting is moderately robust, and, although it is distinctive, a small percentage of the population has unusable prints, usually because of age, genetics, injury, occupation, exposure to chemicals, or other occupational hazards. Hand/finger geometry is moderate on the distinctiveness scale, but it is not very robust, while facial recognition is neither highly robust nor distinctive. As for voice recognition, assuming the voice and not the pronunciation is what is being measured; this biometric is moderately robust and distinctive. Iris scans are both highly robust because they are not highly susceptible to day-to-day changes or damages and distinctive because they are randomly formed. Retinal scans are fairly robust and very distinctive. Finally, neither dynamic signature verification nor keystroke dynamics are particularly robust or distinctive as shown in table1, the biometrics vary in terms of how intrusive they are, ranging from those biometrics that require touching to others that can recognize an individual from a distance.

Biometric	**Identify versus Verify**	**Robust**	**Distinctive**	**Intrusive**
Fingerprint	Either	Moderate	High	Touching
Hand/Finger Geometry	Verify	Moderate	Low	Touching
Facial Recognition	Either	Moderate	Moderate	12+ inches
Voice Recognition	Verify	Moderate	Low	Remote
Iris Scan	Either	High	High	12+ inches
Retinal Scan	Either	High	High	1–2 inches
Dynamic Signature Verification	Verify	Low	Moderate	Touching
Keystroke Dynamics	Verify	Low	Low	Touching

Table 1. Biometric characteristics

III. Literature Review

Face recognition is one of the most relevant applications of image analysis. It's a true challenge to build an automated system which equals human ability to recognize faces. Although humans are quite good identifying known faces, we are not very skilled when we must deal with a large amount of unknown faces. The computers, with an almost limitless memory and computational speed, should overcome humans limitations. Face recognition remains as an unsolved problem and a demanded technology - see table2. A simple search with the phrase "face recognition" in the IEEE Digital Library throws 9422 results. 1332 articles in only one year - 2009. There are many different industry areas interested in what it could offer. Some examples include video surveillance, human-machine interaction, photo cameras, virtual reality or law enforcement. This multidisciplinary interest pushes the research and attracts interest from diverse disciplines. Therefore, it's not a problem restricted to computer vision research. Face recognition is a relevant subject in pattern recognition, neural networks, computer graphics, image processing and psychology [8]. In fact, the earliest works on this subject were made in the 1950's in psychology [7]. They came attached to other issues like face expression, interpretation of emotion or perception of gestures. Engineering started to show interest in face recognition in the 1960's. One of the first researches on this subject was Woodrow W. Bledsoe. In 1960, Bledsoe, along other researches, started Panoramic Research, Inc., in Palo Alto, California. The majority of the work done by this company involved AI-related contracts from the U.S. Department of Defense and various intelligence agencies [9]. During 1964 and 1965, Bledsoe, along with Helen Chan and Charles Bisson, worked on using computers to recognize human faces [10]. He continued later his researches at Stanford Research Institute [11]. Bledsoe designed and implemented a semi-automatic system. Some face coordinates were selected by a human operator, and then computers used this information for recognition. He described most of the problems that even 50 years later Face Recognition still suffers - variations in illumination, head rotation, facial expression, aging. Researches on this matter still continue, trying to measure subjective face features as ear size or between-eye distance. For instance, this approach was used in Bell Laboratories by A. Jay Goldstein, Leon D. Harmon and Ann B. Lesk. They described a vector, containing 21 subjective features like ear protrusion, eyebrow weight or nose length, as the basis to recognize faces using pattern classification techniques. In 1973, Fischler and Elschanger tried to measure similar features automatically [12]. Their algorithm used local template matching and a global measure of fit to find and measure facial features. There were other approaches back on the 1970's. Some tried to define a face as a set of geometric parameters and then perform some pattern recognition based on those parameters. But the first one that developed a fully automated face recognition system was Kenade in 1973 [13]. He designed and implemented a face recognition program. It ran in a computer system designed for this purpose. The algorithm extracted sixteen facial parameters automatically. In he's work, Kenade compares this automated extraction to a human or manual extraction, showing only a small difference. He got a correct identification rate of 45-75%. He demonstrated that better results were obtained when irrelevant features were not used. I the 1980's there were a diversity of approaches actively followed, most of them continuing with previous tendencies. Some works tried to improve the methods used measuring subjective features. For instance, Mark Nixon presented a geometric measurement for eye spacing [14]. The template matching approach was improved with strategies such as "deformable templates" . This decade also brought new approaches. Some researchers build face recognition algorithms using artificial neural networks [15]. The first mention to eigenfaces in image processing, a technique that would become the dominant approach in following years, was made by L. Sirovich and M. Kirby in 1986 [16]. Their methods were based on the Principal Component Analysis. Their goal was to represent an image in a lower dimension without losing much information, and then

reconstructing it. Their work would be later the foundation of the proposal of many new face recognition algorithms.

Areas	Applications
Information Security	Access security , User authentication
Access management	Secure access authentication , Permission based systems, Access log or audit trails
Biometrics	Person identification Automated identity verification
Law Enforcement	Video surveillance Suspect identification Simulated aging Forensic Reconstruction of faces from remains
Personal security	Home video surveillance systems Expression interpretation (driver monitoring system)
Entertainment - Leisure	Home video game systems Photo camera applications

Table 2. Applications of face recognition.

The 1990's saw the broad recognition of the mentioned eigenface approach as the basis for the state of the art and the first industrial applications. In 1992 Mathew Turk and Alex Pentland of the MIT presented a work which used eigenfaces for recognition [17]. Their algorithm was able to locate, track and classify a subject's head. Since the 1990's, face recognition area has received a lot of attention, with a noticeable increase in the number of publications. Many approaches have been taken which has lead to different algorithms. Some of the most relevant are PCA, ICA, LDA and their derivatives. Different approaches and algorithms will be discussed later in this work. The technologies using face recognition techniques have also evolved through the years. The first companies to invest in such researches where enforcement agencies - e.g. the Woodrow W. Bledsoe case law. Nowadays diverse enterprises are using face recognition in their products. One good example could be entertainment business. Products like Microsoft's Project Natal or Sony's PlayStation Eye will use face recognition. It will allow a new way to interact with the machine. The idea of detecting people and analysing their gesture is also being used in automotive industry. Companies such as Toyota are developing sleep detectors to increase safety[18]. These and other applications are raising the interest on face recognition. It's narrow initial application area is being widened.

IV. PROPOSED METHOD

In the proposed method we have three modules namely the client, Key Distribution Center and Cloud Server as shown in Fig2.

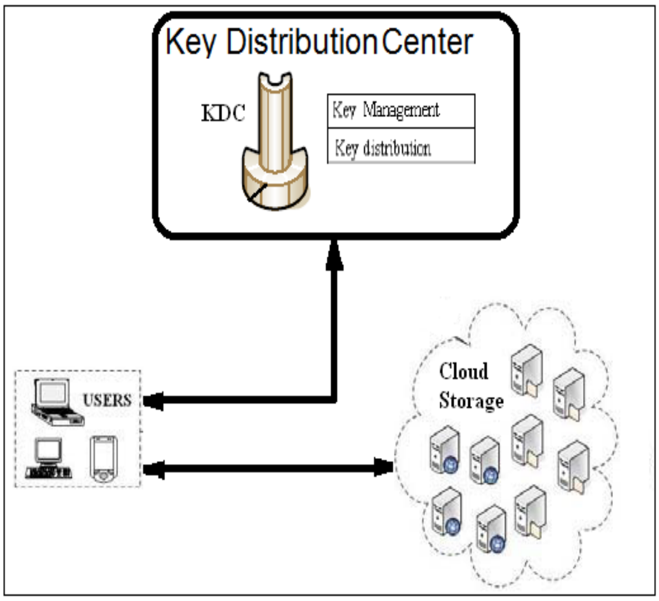


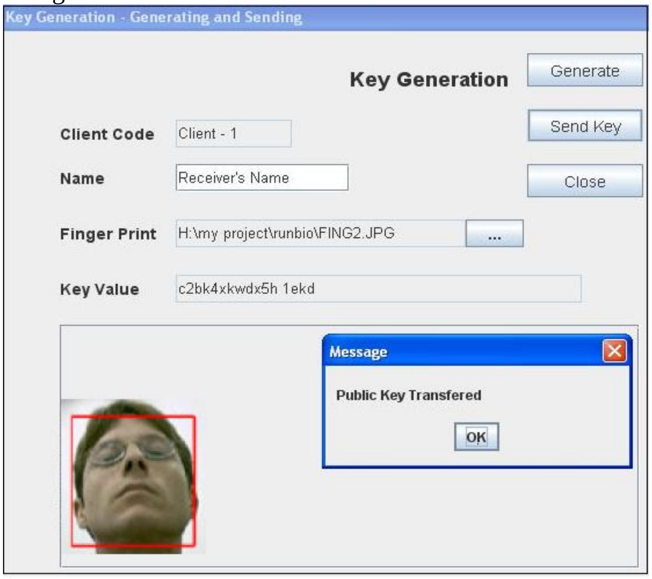
Figure1. System Architecture

A. *Client*

The client application is designed to handle the document transfer process and the key generation process. The client application is divided into four modules. They are the Key generation module, the sender module, the receiver module and the document view module. The key generation module is developed to generate the key from the finger print data. The sender module is designed to encode and send the document to cloud server. The receiver module is designed to receive the decoded documents that are sent by the other client. The document view module maintains all the received documents. The user can view the document content after the decoding process.

B. *Key Generation Module*

This module is designed to generate the public key by using the finger print data as shown in Fig2. The finger print data is given as an image input to the system. The image data value is used to create the key base value. The key base value is used to generate the public key value. The public key value is transferred to the KDC with the client details. The system supports the JPEG and the GIF image formats. The image data is extracted and the pixel matrix is constructed. The key base in generated by using the image data matrix values. The key value is generated by using the key base value using RSA algorithm.



Key Generation

Client Code: Client - 1

Name: Receiver's Name

Finger Print: H:\my project\runbio\FING2.JPG

Key Value: c2bk4xkwdx5h 1ekd

Message

Public Key Transfered

Figure2. Key Generation Module using RSA

C. *RSA algorithm*

The Rivest, Shamir, Adelman (RSA) scheme is a block cipher in which the Plaintext and cipher text are integers between 0 and n-1 for some n. A typical size for n is 1024 bits or 309 decimal digits. Plaintext is encrypted in blocks, with each block having a binary value less than some number n. That is, the block size must be less than or equal to $\log_2(n)$; in practice, the block size is k bits, where $2^k < n \leq 2^{k+1}$. Encryption and decryption are of the following form, for some Plaintext block M and cipher text block C.

$$C = M^e \bmod n$$

$$M = C^d \bmod n = (M^e)^d \bmod n = M^{ed} \bmod n$$

Both sender and receiver must know the value of n. The sender should know the value of e, and the receiver should know the value of d. Thus, this is a public-key encryption algorithm with a public key of KU = { e,u} and a private key of KR ={d,n}. For this algorithm to be satisfactory for public-key encryption, the following requirements must be met.

- It is possible to find values of e, d, n such that $M^{ed} = M \bmod n$ for all M < n.
- It is relatively easy to calculate M^e and C^d for all values of M < n.
- It is infeasible to determine d given e and n.

For now, the first requirement to find a relationship of the form

$$M^{ed} = M \bmod n.$$

Given two prime numbers, p and q, and two integers n and m, such that n=pq and 0<m<n, and arbitrary integer k, the following relationship holds.

$m^{k\varphi(n)+1} = m^{k(p-1)(q-1)+1} = m \bmod n$

where φ(n) is the Euler totient function, which is the number of positive integers less than n and relatively prime to n. p, q are prime integers

$\varphi(pq) = (p-1)(q-1)$ the relationship if

$ed = k\varphi(n) + 1$

is equivalent to saying

$ed = 1 \bmod \varphi(n)$

$d = e^{--1} \bmod \varphi(n)$

That is e and d are multiplicative inverses mod φ(n).According to the rules of modular arithmetic, this is true only if d (and therefore e) is relatively prime to φ(n). Equivalently, gcd(φ(n),d) = 1.

The ingredient is the following,

p.q, [two prime numbers]	(private, chosen)
n = p*q	(public, calculated)
e, with gcd(φ(n),e) = 1; 1 <e <φ(n)	(public, chosen)
$d = e^{-1} \bmod \varphi(n)$	(private, calculated)

The private key consists of {d, n } and the public key consists of {e, n }.

Suppose that user A has published its public key that user B wishes to send the message M to A. Then B calculates $C = M^e \pmod n$ and transmits C. on receipt of this cipher text , user A decrypts by calculating $M = C^d \pmod n$.

$d = e^{-1} \bmod \varphi(n)$

Therefore,

$ed = 1 \bmod \varphi(n)$

ed is form kφ(n) +1 . but by the corollary to Euler's theorem, given two prime number, p and q, and integers n =pq and M, with 0 < M < n.

$M^{k\varphi(n)+1} = M^{k(p-1)(q-1)+1} = M \bmod n$

So $M^{ed} = M \bmod n$, Now

$C = M^e \bmod n$

$M = C^d \bmod n = (M^e)^d \bmod n = M^{ed} \bmod n = M \bmod n$

Key Generation

Select p,q prime , p≠q	p and q both
Calculate n = p x q	
Calculate ϕ(n)=(p-1)(q-1)	
Select integer e 1 <e < ϕ(n)	gcd(ϕ(n),e) = 1;
Calculate d	$d = e^{-1} \bmod \phi(n)$
Public key	KU = {e, n}
Private key	KR = {d, n}

Encryption

Plaintext	M <n
Ciphertext	$C = M^e \pmod n$

Decryption

Ciphertext	***C***
Plaintext	$M = C^d \pmod n$.

D. Sender Module

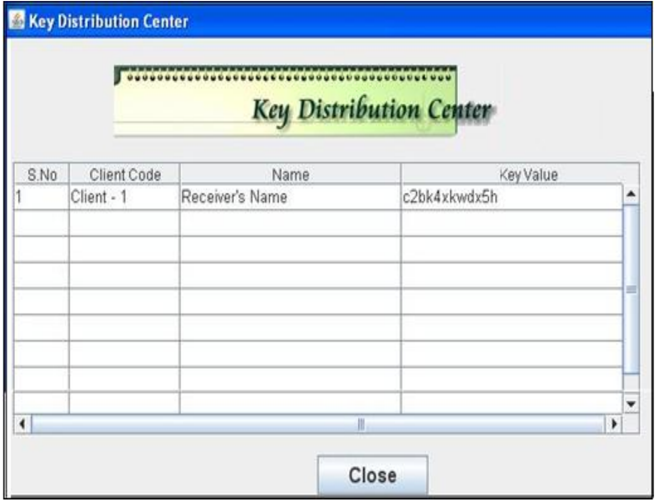
The document encoding and transfer processes are done by using this module. The system is designed with file selection option. The user can select the file name and receiver IP address. The public key value for the destination user is stored in the KDC. Then the key value is applied for the encoding process. After the completion of the encoding process the document is transferred to the destination client

E. Document View Module

This module is designed to view the content of the received documents. The received documents are in the encoded form. The document is decoded before the document view process. The decoding process requires the private key. The private key value is created by using the finger print data. The finger print data is submitted as an image file. The user can select any file from the received file list and decode the content using the private key that is generated by the key generation module.

F. Key Distribution Canter

The role of a KDC is very important in the asymmetric key cryptography. The KDC receives public key values from the clients and stores in its area as shown in Fig3. The KDC is the authorized system that distributes the public key values. The KDC application is designed as a server application. The KDC application has two modules. They are the key management module and the key distribution module. The key management module is designed to receive and maintain the key values. The key distribution module is designed to distribute the public key value based on the client requests.



S.No	Client Code	Name	Key Value
1	Client - 1	Receiver's Name	c2bk4xkwdx5h

G. Key Management Process

The key management module is designed to perform the key maintenance process. The key management module has two main tasks. They are the key receive process and key expiry management process. The key receive process is run as a separate thread. The KDC listen for the key value from the client. This process uses the UDP sockets for the receiving process. The KDC does not make any connection with the client application. All the received public key values are maintained by this module as shown in Fig 4.

Client_id	Public key

Figure4. Shows Clientid and public key pair

The key expiry management module maintains the authenticity of the key values. If the client application is not contacting with in a specific duration of time-stamp then the KDC automatically removes the key values from the key list. The KDC maintains only one key entry for each client id communication within a time slot. The clients can change their key value and update them at any time. In this case the existing key value is removed from the list and the new key value is added into the list.

H. Key Distribution Process

The key distribution module is designed to handle client key requests. All the client applications collect the required public key values from the KDC. The KDC continuously waits for the client's request. The client request is received and parsed to retrieve the client id. After the parsing process the KDC sends the public key value to the requested IP address. All the key request and key distribution process are done by using the UDP.

I. Receiver Module

This module is developed to receive the documents stored on cloud that are transferred by other clients. The receiver checks the list of files to receive and collects the public key from KDC. The documents are received and saved into the download directory. The metadata is stored in a separate file with the sender information. The system supports all file types. All the received files decoded and maintained separately.

V. CONCLUSIONS

In this paper we proposed and implemented secure data storage in cloud environment using face. The security of data is maintained even though the same application is shared by multitenant. To ensure the correctness of users' data in cloud data storage, we proposed an effective and flexible biometric. We believe that data storage security in Cloud Computing, an area full of challenges and of paramount importance, is still in its infancy now, and many research problems are yet to be identified. This can be implemented with more security by using strong biometric measures like eye.

REFERENCES

[1] www.rand.org/publications/MR/MR1237/MR1237.appa.pdf.

[2] www.cost275.gts.tsc.uvigo.es/presentations/COST275_Jain.pdf.

[3] James L. Wayman, *"Biometrics Identification"*, Communications of the ACM, February 2000.

[4] V. Vijaya Kumari and N. Suriyanarayanan, "*Performance Measure of Local Operators in Fingerprint Detection*", Academic Open Internet Journal, vol. 23, pp. 1-7, (2008).

[5] Xifeng Tong, Songbo Liu, Jianhua Huang, and Xianglong Tang, "*Local Relative Location Error Descriptior-Based Fingerprint Minutiae Matching*", the Journal of the Pattern Recognition Letters, vol. 29, pp. 286-294, (2008).

[6] Schneier.B, *"The uses and abuses of biometrics"*. Communications of the ACM, August 1999.

[7] W. Zhao, R. Chellappa, A. Rosenfeld, and P. Phillips. Face recognition: A literature survey. ACM Computing Surveys, pages 399–458, 2003.

[8] J. S. Bruner and R. Tagiuri. The percepton of people. Handbook of Social Psychology, 2(17), 1954.

[9] M. Ballantyne, R. S. Boyer, and L. Hines. Woody bledsoe: His life and legacy. AI Magazine, 17(1):7–20, 1996.

[10] W. W. Bledsoe. The model method in facial recognition. Technical report pri 15, Panoramic Research, Inc .. Palo Alto, California, 1964.

[11] W. W. Bledsoe. Semiautomatic facial recognition. Technical report sriproject 6693, Stanford Research Institute. Menlo Park, California, 1968.

[12] M. Fischler and R. Elschlager. The representation and matching of pictorial structures. IEEE Transactions on Computers, C-22(1):67 –92, 1973.

[13] T. Kenade. Picture Processing System by Computer Complex and Recognition of Human Faces. PhD thesis . Kyoto University, November 1973.

[14] M. Nixon. Eye spacing measurement for facial recognition. Proceedings of the Society of Photo-Optical Instrument Engineers, SPIE, 575(37):279–285, August 1985.

[15] T. J. Stonham. Practical face recognition and verification with wisard. In H. D. Ellis, editor, Aspects of face processing. Kluwer Academic Publishers, 1986.

[16] . Sirovich and M. Kirby. Low-dimensional procedure for the characterization of human faces. Journal of the Optical Society of America A Optics, Image Science and Vision, 4(3):519–524, March 1987.

[17] M. Turk and A. Pentland. Eigenfaces for recognition. Journal of Cognitive Neurosicence, 3(1):71–86, 1991.

[18] B. Dudley. "e3: New info on microsoft's natal – how it works, multiplayer and pc versions". The Seattle Times , June 3 2009.