A Gender and Ethnicity Identification System in Nigeria using the Fingerprint Technology

Sunny Orike, *Member, IAENG;* Vincent I.E. Anireh, and Ademuyiwa S. Ibironke

Abstract—**Identification in Nigeria and indeed many other developing nations is still lagging behind in terms of individuality, as the conventional ways currently in use (driver's license, international passport, national identity card, passwords, etc) have not fully employed the use of individual-specific technologies. This has caused many untold hardships for the users: some forget their passwords, others misplace their identity cards, some identical relations even swap their photo identity cards, etc. However, the innovation of biometric identification technology brought a lot of improvements to these conventional ways. This makes use of human traits such as face, fingerprint, iris, voice, etc, in identifying and verifying the identity of humans possible without the need of numerical or letter password, or identity card. In this paper, we propose the use of fingerprint technology to capture the fingerprints of a group of people in other to identify and verify their identities, gender and ethnic groups through the use of trained classifiers. The classifier detects shapes (e.g., cores and deltas) and then determines the pattern type of the images. We evaluate the performance of the technology using metrics such as sensitivity, precision, false positive rate, recognition accuracy and recognition time. The result shows that over 98% test cases accurately identified person's ethnicity and gender with an average recognition time of 2 seconds.**

Index Terms—**Biometric, Classifier, Ethnicity, Gender, Neural Network**

I. Introduction

PERSONAL recognition/identification is very important in our daily lives. We always have to prove our identities for getting access to bank account, entering protected sites, drawing cash from Automated Teller Machines, logging in to computers, etc. Generally, we identify ourselves and gain access by physically carrying passports, keys, access cards or by remembering passwords, secret codes and personal identification numbers (PINs) [1].

Conventional personal identification techniques could either be knowledge-based technique (password, PIN, etc) or token-based technique (driver's license, passport, identity card, etc). The problem with the former is that it is prone to fraud, as passwords may be guessed; secret codes and PINs can easily be forgotten, compromised, shared, or observed; while the latter could be lost, cloned or stolen [2].

Biometric solves the problem faced at the level of both the token-based and knowledge-based identification approaches as it attempts to answer the questions: "who are you?" and "are you who you claim to be?" It is a unique, measurable characteristic or trait of a human being for automatically recognizing or verifying identity [3]. Biometric characteristics are generally more difficult to duplicate and they are naturally present with the user without extra memorization or storage efforts. It is the science of measuring physiological and behavioral characteristics that uniquely identify individuals. Instances of physiological traits include face, iris, fingerprint, vascular images, and hand geometry; while that of behavioral traits include voice pattern, signature dynamics, gait pattern, keystrokes, etc. All these traits are captured by specialized devices and converted through sophisticated algorithms into mathematical representations or templates, which are used as references against which an individual's identity is verified.

Based on the varieties of the information available from the fingerprint we are able to process its identity along with gender, age and ethnicity. Fingerprint is an impression of friction ridges, from the surface of the finger-tip. It has been used for personal identification for many decades; more recently becoming automated due to advancements in the computing capabilities [4]. Fingerprints have some important characteristics that make them invaluable evidence in crime scene investigations: (1) It is unique to a particular individual, and no two fingerprints possess exactly the same set of characteristics; (2) It does not change over the course of person's lifetime (even after superficial injury to the fingers); (3) The patterns can be classified, and those classifications then used to narrow the range of suspects [4].

The fingerprint technology, a physiological biometrics technology is an Artificial Neural Network concept. It is the representation of the epidermis of a finger, and consists of pattern of interleaved ridges and valleys. Its evidence is undoubtedly the most reliable and acceptable forensics evidence till date, even in the court of law [5]. Now, this technology is being used in several other applications such as access control for high security installations, credit card usage verification and so on. However, due to the unique nature of fingerprints, it has become increasingly popular for personal identification and verification.

In this work, we investigate gender and ethnicity determination methods based on finger related features such as ridge arrangement, fingerprint patterns and the most predominant minutiae features that exist in either a particular gender or ethnic group in Nigeria. A novel work of this nature will help the country to march up with other

Manuscript received February 29, 2016; revised April 7, 2016.

S. Orike is currently the Head, Department of Electrical/Computer Engineering, Rivers State University of Science and Technology, Port Harcourt, Nigeria (+234-8033988451; e-mail: orike.sunny@ust.edu.ng).

V. I. A. Anireh is a Lecturer with the Department of Computer Science, Rivers State University of Science and Technology, Port Harcourt, Nigeria (e-mail: viaemeka@gmail.com).

A. S. Ibironke just completed his MSc at the Department of Information Technology, Boulder, National Open University of Nigeria, (e-mail: ibironkemuyiwa@gmail.com).

developed nations of the world in data collection of her citizens. This will assist and improve her security challenges, and also helps in electoral forensic investigations. The rest of this paper is organized as follows: Section II discusses the concept of gender and ethnicity identification in relation to the present work. Section III investigates the literature on biometric technology as it evolves with time. Section IV describes the design methodology used in this work. Section V details the implementation, results and discussion, including graphical representation; while Section VI concludes the findings of the work.

II. Gender and Ethnicity

Gender identity is a personal experience of one's own gender. This is generally described as one's private sense of being a man or a woman, consisting primarily of the acceptance of membership into a category of people: male or female [6]. All societies have a set of gender categories that can serve as the basis of the formation of a social identity in relation to other members of society. In most societies, there is a basic division between gender attributes assigned to males and females. In all societies, however, some individuals do not identify with some (or all) of the aspects of gender that are assigned to their biological sex**.** Gender and Age information is important to provide investigative leads for finding unknown persons. Existing methods for gender classification have limited use for crime scene investigation because they depend on the availability of teeth, bones, or other identifiable body parts having physical features that allow gender and age estimation by conventional methods [7]. There are a large number of potential application areas where gender recognition is very much involved. These include (but not limited to): biometrics, content-based indexing and searching, human-computer interaction, surveilance systems, etc [8,9].

Ethnicity refers to the idea that one is a member of a particular cultural, national, or racial group that may share some of the following elements: culture, religion, race, language, or place of origin. Two people can share the same race but have different ethnicities. For example, among two black individuals one may be African-American and another may be African-Caribbean [10]. In essence, an ethnic group is a named social category of people based on perceptions of shared social experience or ancestry. Members of the ethnic group see themselves as sharing cultural traditions and history that distinguish them from other groups. Ethnic identity develops in adolescence and is passed from one generation to the next through customs, traditions, language, religious practice, and cultural values.

The ethnicity of the Nigerian state which this work hopes to address is well over three hundred ethnic groups, but they are generally categorized into three main groups i.e. Yoruba, Igbo and Hausa. Identity management is establishing identity of a single person using one or more of the biometric or non-biometric features. Biometric trait is a biological and behavioral characteristic of a person, such as fingerprint, face, gait (the way the subject is walking) and signature [11].

III. Biometric Technologies

The term "biometrics" is both a characteristic and a process. It is a measurable biological and behavioral characteristic that can be used to recognize or identify a person; and also a method of identifying a person based on the measurable characteristics. The biological characteristics involve face recognition, fingerprints, iris, etc; while behavioral characteristics include voice modulations, hand drawing, signature style, keystroke dynamics, etc. These distinctive characteristics usually do not change during the adult life of a person [11]. Biometric systems recognize users based on their physiological and behavioral characteristics, and in essence, it is more reliable and capable in distinguishing between a specific individual and an impostor than any technique based on identification document or a passoword [12].

A typical automated biometric system consists of five components, namely: (i) a sensor or transducer system to collect the biometric data and convert it into digital format; (ii) signal processing modules that build a template with the collected data using certain algorithms; (iii) a database system where these template are stored; (iv) a matching algorithm which compares new templates with those stored in the database; (v) a decision process (either fully automated or human-assisted) which makes a system level decision based on the result from matching algorithm outcome. The authors in [11] compared most popular biometric traits used in commercial recognition systems, such as: fingerprints, hand geometry, hand vein, iris, face, voice, signature, keystroke and gait; including typical applications, advantages and disadvantages. The performance of biometric system depends upon accuracy and speed. Speed depends upon the size of database and is inversely proportional to size of database, where accuracy depends upon the underlying algorithm.

For over a century, fingerprint has been a successful biometric technology used in establishing identities. Initially, fingerprint images were captured by manual impressions using ink and identification process performed manually. This method had limited applications due to poor quality of capturing, short duration template of image, no-proper algorithm for matching the fingerprints, manual inspection lead for inaccuracies in matching process and time inefficient process, etc [11]. A fingerprint is the representation of the epidermis of a finger; it consists of a pattern of interleaved ridges and valleys. Fingertip ridges evolved over the years to allow humans to grasp and grip objects. Like everything in the human body, fingerprint ridges are formed through a combination of genetic and environmental factors. This is the reason why even the fingerprint of identical twins is different [12, 13]. Fingerprints found at crime scenes or developed in the laboratory are categorized by some examiners as patent, latent, or plastic impressions [5].

In this digital age (with high computing power at very low cost), Automatic Fingerprint Recognition System (AFRS) has become part of daily life at in many developed countries, with applications in building access control, border control, computer and network access, e-government, e-commerce, forensic and criminology, etc. However, manual inspection of fingerprints still exists in forensic and criminology as they need specialized ways analyzing the fingerprints. Automatic biometric-based identity system involves two stages: enrolment and recognition stages [11]. In enrolment stage, features of the subject are extracted and stored in a database with unique identifiers. In the recognition stage, the features are matched and recognized against the database entries. There are two typical measures used in assessing fingerprint: False Acceptance Rate (FAR), which indicates the percentage of wrongly accepted fingerprint and False Rejection Rate (FRR), which the percentage of wrongly rejected fingerprint. There have been number of developments attained over a decade not only to bring FAR and FRR to very close to zero, but also to make the recognition process more reliable, robust, highly secured, user friendly and robust. These developments include, but not limited to: high-resolution image capturing sensors, touch-less capturing technology, multi-finger recognitions through fusions and encryption of templates [11]. Before the existence of fingerprint sensors, fingerprints have been in used for data processing using traditional inked-rolled printing process [14, 15]. Efficiency and accuracy of solid-state fingerprint sensor (i.e. live scan) outweighed the traditional method. Classification of fingerprint recognition as a biometric procedure involves identification, verification, authentication and authorization.

The biometric recognition systems process biometric features of a person with the aim of confirming or rejecting that person's identity by using previously gathered reference data. All biometric systems are made up of the following components: data input, pre-processing, feature extraction, classification, and calculation of reference data. In the recent years, the demand for reliable identification system is on a daily increase. Even though identification by means of an object (e.g. an identity card), is still in vogue, nonetheless, it is gradually losing its importance. For this reason, biometry has been getting more important since it combines personal identification with unambiguous and unchangeable characteristics of man. With ever-increasing and ever complex technologies, exact personal identification is imperative. Man has certain unambiguous features which are formed in the earliest phases of human life as part of a random process, and which are different for each individual. One of the first biometric features that was discovered and scientifically investigated was the fingerprint. With advancements of technology, the issue of safety has become more important. For access controls, analyzing fingerprints biometrically has been playing an increasingly important role.

IV. Design Considerations

To achieve the overall goal of this work, we considered the following three factors in the design of the system: (1) it should be trainable - must able to accept several fingerprint templates and also categorize them base on gender and ethnicity; (2) it should be able to classify the biometric templates stored in the database; and (3) it should be able to successfully identify each individual based on their gender and ethnicity. The design architecture is made up of three modules: enrolment, identification and storage modules. The enrolment module consists of fingerprint image acquisition, image pre-processing, feature extraction and template file creation. It is used to get the minutiae sets and stored template information in the database. The identification module detects similar minutiae group from multiple template images generated from the same finger and creates a cluster core set, identifying which finger the test image comes from.

A. Model Development

The system model is divided into two phases: fingerprint enrolment and classification. This fingerprint enrolment phases stores biometric features into the database for the first time with the use of a sensor which accepts fingerprints and converts them into grey level fingerprint images. The classification phase is divided into two stages: gender and ethnicity identification. In this phase, the system recognizes an individual by searching the templates of all the users in the database for a match. The classification only succeeds in identifying the individual if the comparison of the biometric sample to a template in the database falls within a previously set threshold because it requires comparing each existing records in the database against the new biometric characteristics. This phase detects similar minutiae group from multiple template images generated from the same finger and thereby create a cluster core set.

Fingerprint recognition systems contain two main modules: *feature extraction* and *feature matching*. Feature extraction detects singular and all other minutiae points which differentiate one fingerprint from another, and impart individuality to each fingerprint. Feature matching involves the actual procedure to identify the unknown person by comparing extracted features from a fingerprint with those from a set of known persons. We first define the region of interest in a fingerprint image. Let *I (x,y)* denote the gray level at pixel *(x,y)* in an *M* x *N* fingerprint image and let *(xc,Yc)* denote the center point. The region of interest is defined as the collection of all the sectors *Si,* where the ith sector *Si* is computed in terms of parameters *(r,0)* as follows:

$$S_i = \{(x,y) | b(T_i+1) \le r < b(T_i+2), \theta_i^l \le \theta_i^h,\} \tag{1}$$

Subject to:

$$1 \le x \le N, 1 \le y \le M$$

The following steps were observed to create the finger code:

1. Preprocessing of the image (to remove noise) by histogram equalization;
2. Core point location using max concavity estimation;
3. Tessellation of circular region around the reference point;
4. Sector wise normalization followed by application of bank of Gabor filters which has following general form in the spatial domain:

$$G(x', y', f, \theta) = \exp\left\{\frac{-1}{2}\left[\frac{x'^2}{\delta_{x'}^2} + \frac{y'^2}{\delta_{y'}^2}\right]\right\}\cos(2\pi f x') \tag{2}$$

Subject to:

$$x' = x\sin\theta + y\cos\theta,$$

$$y' = x\sin\theta - y\cos\theta$$

Where f is the frequency of the sine plane wave along the direction θ (0, 45, 90 and 135 degrees) from the x-axis, δx' and δy' are the space constants of the Gaussian envelope along X' and Y' axes, respectively;

5. Finally, feature code generation by obtaining standard deviation values of all the sectors.

B. *Algorithm*

i. Load trainer set files (both input image and desired output text);
ii. Analyze input image in image coordinates;
iii. Read desired output text from file and convert each fingerprint image to a binary Unicode value to store separately;
iv. For each fingerprint image :
    a) Calculate each fingerprint image by using transfer function G (k) of the Gabor filter;
    b) Filter each of the fingerprint and normalization take place;
    c) Compare results (a) with the desired output corresponding to different bandwidths and modulation frequencies.

Move to the next fingerprint image and repeat step (iv) until all fingerprint images are visited.

V. Implementation, Results and Discussion

We captured ten fingerprints of 1,054 persons of the three main ethnic groups in Nigeria: 673 Yoruba, 179 Igbo and 197 Hausa of both genders (593 male and 461 female) using a fingerprint scanner; and developed biometric capturing software using Java netbeans software development kit. The fingerprints captured were stored in gray scale format. The application captures the name, sex, age range, date of birth, state of origin and the images of the ten fingerprints captured. We trained and tested the fingerprint images using MATLAB. Gabor filter algorithm was used due to its effectiveness in pattern recognition. We also tested the recognition accuracy, average recognition time of individual fingerprint and performance evaluation of the system, and presented results using graphs and charts, showing the percentage of individual ethnicity and their corresponding gender. The performance on trained and tested fingerprints was measured against recognition rate, total training time, false rejection rate and false acceptance rate. The following parameters are used to measure or evaluate the overall performance of the system:

$$\text{False Positive Rate} = \frac{FP}{FP+TN} \quad (3)$$

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (4)$$

$$\text{Specificity} = \frac{TN}{FP+TN} \quad (5)$$

$$\text{Overall Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

If a fingerprint verified is present in a dataset, the result of the fingerprint recognition system is true positive (TP). If a fingerprint verified is absent in a dataset, the result of the fingerprint recognition system is true negative (TN). If the fingerprint recognition system confirms the presence of a non-existing fingerprint, the test result is false positive (FP). If the fingerprint recognition system test suggests that an existing fingerprint is absent in the dataset, the results is false negative (FN). False Positive Rate (FPR): Proportion between FP and all affected fingerprint images. Sensitivity is the ability to identify presence of fingerprint image in the created database, while specificity is the ability to identify absence of fingerprint image in the created database.

We present the results for average recognition time, recognition accuracy, percentage sensitivity, specificity, gender and ethnicity identification in Figs 1-6 respectively. Four different thresholds were used to evaluate the system. Average recognition time and accuracy increase with increase in the threshold value. The percentage sensitivity and specificity are 98% and 66.5% respectively.

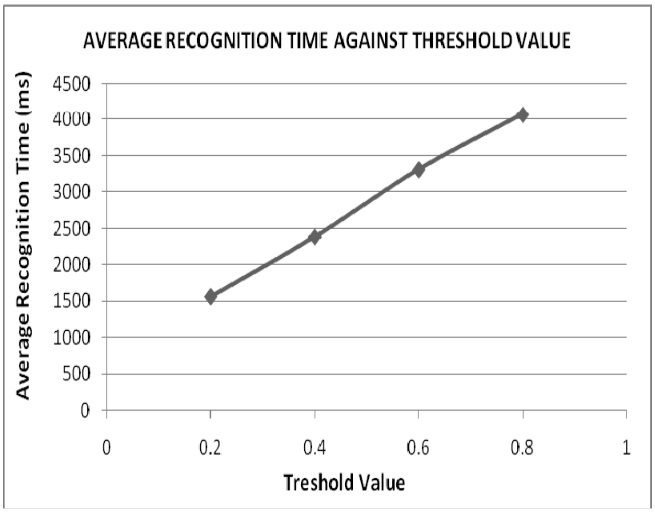

Fig. 1. Average Recognition Time

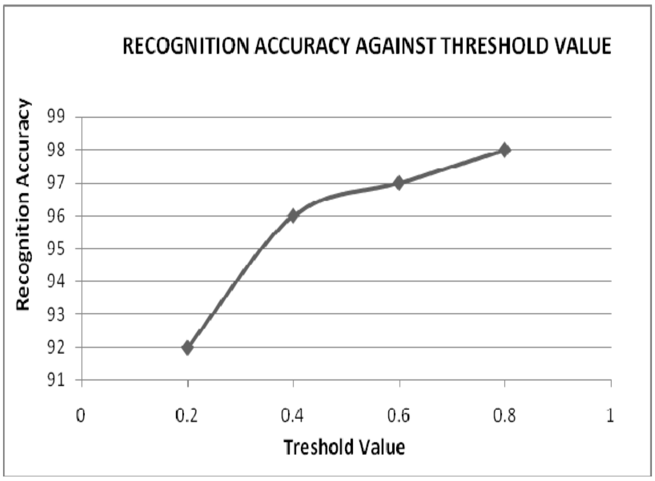

Fig. 2. Recognition Accuracy

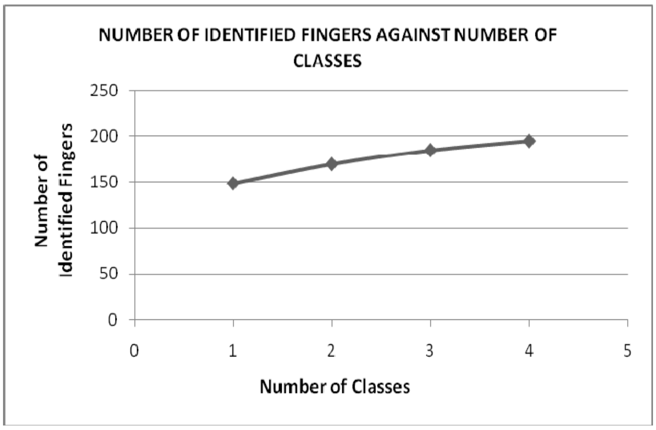

Fig. 3. Percentage Sensitivity for Identified Fingers

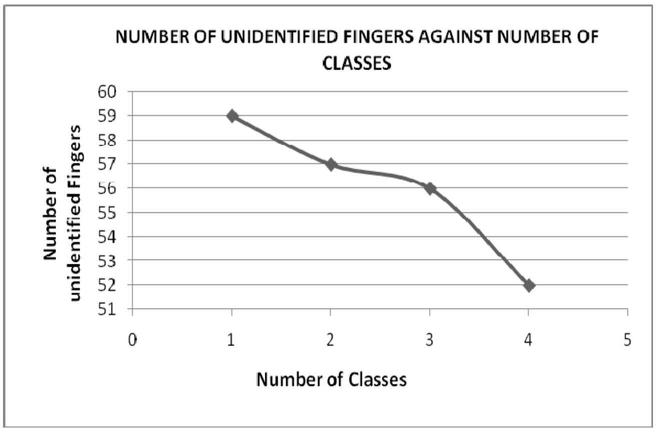

Fig. 4. Percentage Specificity for Unidentified Fingers

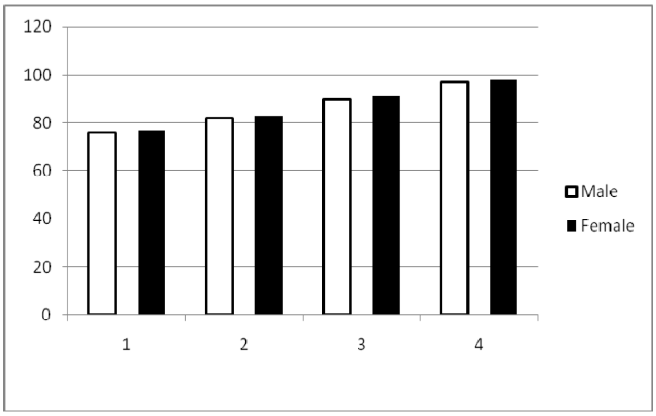

Fig. 5. Gender Identification

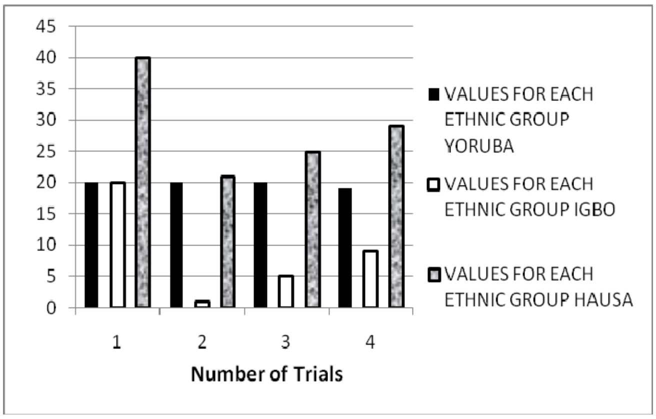

Fig. 6. Ethnicity Identification

VI. Conclusion

In this work, we designed and implemented a gender and ethnicity identification system using fingerprint technology for the three main ethnic tribes in Nigeria: Yoruba, Igbo and Hausa. We captured ten fingerprints of 1,054 persons: 673 Yoruba, 179 Igbo and 197 Hausa of both genders (593 male and 461 female) using a fingerprint scanner; and developed biometric capturing software using Java netbeans software development kit. The system was trained and tested with fingerprint patterns. Four different thresholds were used to evaluate the system; and results were presented for average recognition time, recognition accuracy, percentage sensitivity, specificity, gender and ethnicity identification. The percentage sensitivity and specificity are 98% and 66.5% respectively, with an average recognition time of less than 2 seconds. The system is highly recommended for electronic voting and forensic investigation in Nigeria

References

[1] F. Anwar, Md. A. Rahman and Md. S. Azad, “Multi-biometric Systems Based Verification Technique,” European Journal of Scientific Research, vol. 34, no. 2, pp. 260 – 270, 2009.

[2] C. Sinha, “Gender Classification from Facial Images using PCA and SVM,” Department of Biotechnology and Medical Engineering National Institute of Technology Rourkela, Rourkela, India, 2013.

[3] L. L. Ramenzoni and S. R. Line, “Automated Biometrics-Based Personal Identification of the Hunter-Schreger Bands of Dental Enamel,” *Proceedings of Biological Sciences,* vol. 273, no. 1590, pp. 1155-1158, 2006.

[4] K. Ritu and G. M. Susmita G. M., “International Journal of Advances in Engineering and Technology,” *Fingerprint Based Gender Identification using Frequency Domain Analysis,* vol. 3, no. 1, pp. 295 – 299, 2012.

[5] H. C. Lee, and R. E. Gaensslen, “Methods of Latent Fingerprint Development,” *in Advances in Fingerprint Technology*, 2nd ed.; Lee, H. C., Gaensslen, R. E., Eds.; CRC Press, pp 105 – 175, 2001.

[6] N. R. Carlson and C. D. Heth, "Sensation," in *Carlson, N. R. Carlson and C. D. Heth, Psychology: The Science of Behaviour,* (4th ed.), Pearson, pp. 140 – 141, 2009.

[7] P. Gnanasivam and S. Muttan, “Estimation of Age Through Fingerprints Using Wavelet Transform and Singular Value Decomposition,” *International Journal of Biometrics and Bioinformatics*, vol. 6, no. 2, pp 58 – 67, 2012.

[8] S. A. Khan, M. Ahmad, M. Nazir and N. Riaz, “A Comparative Analysis of Gender Classification Techniques,” Middle-East Journal of Scientific Research, vol. 20, no. 1, pp. 1-13, 2014.

[9] O.F.W. Onifade and K.T. Bamigbade, “A Multifactored Model of Soft and Hard Biometric Trait for ease of Retrieval," in *World Congress on Computer Applications and Information Systems (WCCAIS),* 2014.

[10] A. Butler, “The ACT for Youth Online Presentation: Adolescent Ethnic and Racial Identity Development,” Cornell University, 2010.

[11] M. Nadarajah and T. Celalettin, “Fingerprint Biometric for Identity Management,” *International Journal of Industrial Engineering and Management,* vol. 2, no 2, pp. 39-44, 2011.

[12] K. J. Anil and N. Karthik, “Integrating Faces, Fingerprints, and Soft Biometric Traits for User Recognition,” *Proceedings of Biometric Authentication Workshop,* LNCS 3087, pp. 259-269, 2004.

[13] D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar, “Handbook of Fingerprint Recognition,” Second Edition, Springer, New York, 2009.

[14] E. R. Henry, “Classification and Uses of Fingerprints,” George Routledge and Sons Limited, 1990.

[15] Y. Chang, J. Yang, D. Chen, and R. Yan, “People Identification with Limited Labels,” in *Privacy-Protected Video, International Conference on Multimedia & Expo*, 2006.