A Survey on Soft Biometrics

Ujwala T. Tayade[#1], Seema Biday[*2]
[#]Terna Engineering College, Nerul,
Navi Mumbai, India
[*]Computer Department, Terna Engineering College
Nerul, Navi Mumbai, India
[1]ujwalatayade@gmail.com
[2]scbiday@rediffmail.com

***Abstract*- Soft biometrics have gained more and more interest of the biometry and other communities for various reasons like the need for higher reliability in biometric systems and the great number of advantages coming along with the integration of soft biometrics in the systems. Many existing biometric systems collect ancillary information like gender, age, height, and eye color from the users during enrollment. However, only the primary biometric identifier (fingerprint, face, hand-geometry, etc.) is used for recognition and the ancillary information is rarely utilized. This paper presents the utilization of "soft" biometric traits like gender, height, weight, age, and ethnicity to complement the identity information provided by the primary biometric identifiers. Although soft biometric characteristics lack the distinctiveness and permanence to identify an individual uniquely and reliably, they provide some evidence about the user identity that could be beneficial. Thus, here we emphasized on the use of soft biometrics and provide an overview of soft biometrics integration with the primary biometric system, soft biometric traits, their classification with the related advantages and its limitations.**

Keywords-Primary Biometrics, Soft Biometrics, Fusion, Soft Biometrics Traits.

I. INTRODUCTION

With the widespread use of computers, nowadays biometric identification comes in great demand in many different areas. Biometrics techniques can be largely divided into traditional and the so-called soft biometrics. Whereas traditional biometrics deals with physical characteristics such as face features, eye iris, and fingerprints, soft biometrics is concerned with such information as gender, national origin, and height. Traditional biometrics is versatile and highly accurate. But it is very difficult to get traditional biometric data from a distance and without personal cooperation. Soft biometrics, although featuring less accuracy, can be used much more freely though. Recently, many researchers have been made on human identification using soft biometrics data collected from a distance.

Soft biometric information complements the identity information provided by traditional (primary) biometric identifiers such as fingerprint, face, iris, and voice. Any anatomical or behavioral characteristic that provides some information about the identity of a person, but does not provide sufficient evidence to precisely determine the identity can be referred to as a soft biometric trait. Personal attributes like gender, ethnicity, age, height, weight, eye color, scars, marks, tattoos, and voice accent are examples of soft biometric traits. Hence, utilizing soft biometric traits can improve the recognition accuracy of primary biometric systems. Those traits usually can be 'Captured from a distance' and do not require cooperation from the subject as in traditional biometrics. Thus it helps automatic extraction of soft biometric information. It also helps Filtering of database to a greater extent and hence improves the performance of the traditional biometric system.

II. BIOMETRICS

Biometrics refers to the identification of humans by their characteristics or traits.
It uses different biological features for user authentication:

- Physical traits: fingerprint, face, retina, hand geometry, iris
- Behavioral traits**:** voice, signature, keystrokes.

These traits are called biometric identifiers or simply biometrics. It is used as a form of identification and access control. It is also used to identify individuals in groups that are under surveillance.

Biometric systems automatically recognize individuals based on their physiological and/or behavioral characteristics like fingerprint, face, hand-geometry, iris, retina, palm print, voice, gait, signature, and keystroke dynamics. Biometric systems that use a single trait for recognition, called unimodal biometric systems, are affected by problems like noisy sensor data, non-universality and/or lack of distinctiveness of the chosen biometric trait, unacceptable error rates, and spoof attacks. Some of the problems associated with unimodal biometric systems can be overcome by the use of multimodal biometric systems that combine the evidence obtained from multiple sources. A multimodal biometric system based on different biometric identifiers like fingerprint, iris, face, and hand-geometry can be expected to be more robust to noise, address the problem of non-universality, improve the matching accuracy, and provide reasonable protection against spoof attacks. But there are limitations as well.

- The overall cost involved in building the multimodal system can be high due to the need for multiple high quality sensors and increased storage and computational requirements.
- The system will require a longer verification time thereby causing inconvenience to the users.

A possible solution to the problem of designing a reliable and user-friendly biometric system is to use ancillary information about the user like height, weight, age, gender, ethnicity, and eye color to improve the performance of the primary biometric system.

III. SOFT BIOMETRICS

Soft biometrics have gained more and more interest of the biometry and other communities for various reasons, like the need for higher reliability in biometric systems and the great number of advantages coming along with the integration of soft biometric traits in systems. Here we provide an overview of soft biometric traits, their classification, the related advantages and limitations.

As already mentioned, a possible solution to the problem of designing a reliable and user-friendly biometric system is to use soft biometric to improve the performance of the primary biometric system.

Soft biometric traits are those characteristics that provide some information about the individual, but lack the distinctiveness and permanence to sufficiently differentiate any two individuals (Fig. 1 for examples of soft biometric traits). In other words, these characteristics do not have the distinctiveness and permanence to uniquely identify an individual over a period of time and hence we refer them as soft biometric traits.

Fig 1: Examples of Soft Biometrics Traits

A. *Automatic Extraction of Soft Biometric Characteristics*

In order to utilize soft biometrics, there must be a mechanism to automatically extract these features from the user during the recognition phase. As the user interacts with the primary biometric system, the system should be able to automatically measure the soft biometric characteristics like height, weight, age, gender, and ethnicity in a non-obtrusive manner without any interaction with the user. This can be achieved using a special system of sensors. For example, a bundle of infra-red beams could be used to measure the height. The weight sensor could be installed at the place where the users stand while providing the primary biometric. A camera could be used for obtaining the facial image of the user, from which information like age, gender, and ethnicity could be derived. This observed soft biometric information could then be used to supplement the identity information provided by the user's primary biometric identifier. Extensive studies have been made to identify the gender, ethnicity, and pose of the users from their facial images. Another method developed was a demographic classification scheme that extracts faces from unconstrained video sequences and classifies them based on gender and ethnicity. Their demographic classifier was a Perceptron constructed from binary rectangle features. The learning and feature selection modules used a variant of the AdaBoost algorithm. Age determination is a more difficult problem due to the very limited physiological or behavioral changes in the human body as the person grows from one age group to another.

B. *Fusion With Traditional Biometrics System*

A general framework [1] to integrate the information provided by soft biometric a signature with the ones of a primary biometric system is described in Fig. 2. This implementation serves an increase of detection reliability.

The suggested framework can be described as follows. Let $W = \{w1, w2... wn\}$ be the set of the n users enrolled in the database, and let $\mathbf{x}$ be the feature vector corresponding to the primary biometric system. Without loss of generality, we can assume that the output of the primary biometric system is of the form $P(wi|\mathbf{x})$, $i = 1, 2... n$, where $P(wi|\mathbf{x})$ is the probability that the test user is wi given the feature vector $\mathbf{x}$. Let $\mathbf{y} = \{y1, y2 ...ym\}$ be the soft biometric feature vector, the updated probability of user wi, given the primary biometric feature vector $\mathbf{x}$ and the soft biometric feature vector $\mathbf{y}$, $P(wi|\mathbf{x}, \mathbf{y})$, can be calculated using the Bayes rule as:

$$P(w_i|\mathbf{x}, \mathbf{y}) = \frac{p(\mathbf{y}|w_i)\, P(w_i|\mathbf{x})}{\sum_{j=1}^{n} p(\mathbf{y}|w_j)\, P(w_j|\mathbf{x})} \quad (1)$$

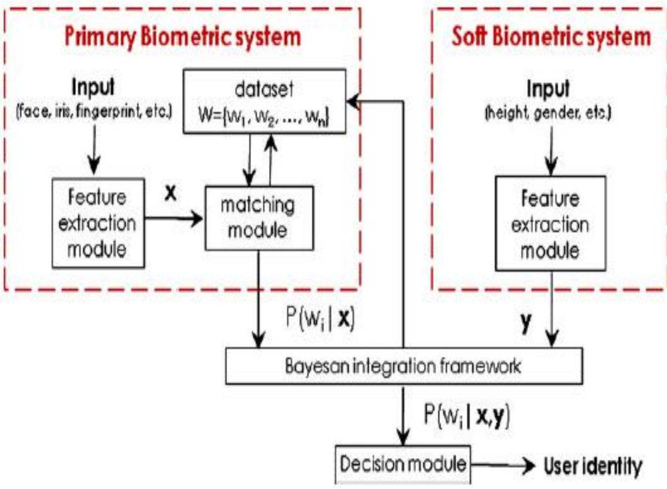

Fig 2: Framework of integration of soft biometrics to improve the accuracy of classical biometric systems.

In the above formulation it is necessary to take into account that all *m* soft biometric variables are not equally weighted. In practice, some soft biometric variables may contain more information than others. Therefore, it is necessary to introduce a weighting scheme for soft biometric traits based on an index of distinctiveness and permanence, i.e., traits that have smaller variability and larger distinguishing capability will be given more weight in the computation of the final matching probabilities. In the same way, smaller weights should be assigned to the soft biometric traits compared to those assigned to the primary biometric traits. This differential weighting also has another implicit advantage: even if a soft biometric trait of a user is measured incorrectly (e.g., a male user is identified as a female), there is only a small reduction in the posteriori probability and the user is not immediately rejected. In this case, if the primary biometric produces a good match, the user may still be accepted. Only in the case that several soft biometric traits do not match, there is significant reduction in the posteriori probability and the user could be possibly rejected. If the devices that measure the soft biometric traits are reasonably accurate, such a situation has very low probability of occurrence.

In the above given approach we can see the benefits of using gender, ethnicity and height information of the user in addition to fingerprint. The use of these soft biometric signatures leads to an improvement of approximately 5% over the primary biometric system.Similar results can be applied in which body weight and fat measurements can be combined with fingerprint. The results studied show that the total error rate improves from 3.9% to 1.5% when body weight score is fused with fingerprint score.

C. Pruning The Search

Other than improving the performance of a classical biometric system, soft biometric signatures were used in the scientific literature to prune large biometric database in order to improve the search efficiency of the biometric system, as shown in Fig. 3. In this scenario the soft biometric signature is used as a side information to filter the original dataset *W* and to find a subset of the dataset $Z = wj \in W|P(wj|\mathbf{y})$. The selected dataset $Z = \{z1, z2...zp\}$ will contain the *p* (with $p <= n$) elements of *W* that satisfy the soft biometric information **y**. The new filtered dataset will then be used in the classical biometric system to find the user identity based on the probability $P(zi|\mathbf{x})$.

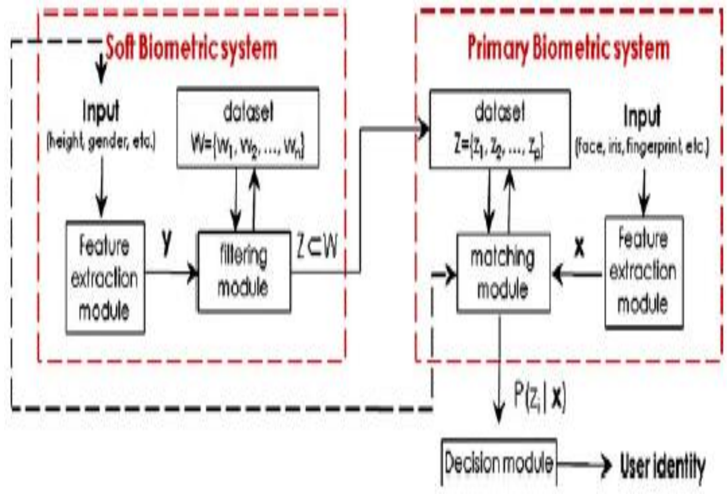

Fig 3: Framework of integration of soft biometrics to improve the search efficiency of classical biometric systems.

As shown in this figure, filtering refers to limiting the number of entries in a database to be searched, based on characteristics of the interacting user. For example, if the user can somehow be identified as a middle-aged male, the search can be restricted only to subjects enrolled in the database with this profile. This greatly improves the speed or the search efficiency of the biometric system. While filtering reduces the time required for identification, errors in filtering can degrade the recognition performance. This is a delicate aspect to take into account when designing the biometric system.

IV. SOFT BIOMETRICS TRAITS

Soft biometric traits are those characteristics that provide some information about the individual, but lack the distinctiveness and permanence to sufficiently differentiate any two individuals.

Traits which accept the above definition include, but are not limited to:

- Physical: skin color, eye color, hair color, presence of beard, presence of moustache, height, weight, etc.
- Behavioral: gait, keystroke.

- Adhered human characteristics: clothes color, tattoos, accessories.

Soft Biometrics inherits a main part of the advantages of Biometrics and furthermore endorses by its own assets. Some of the advantages include non obtrusiveness, the computational and time efficiency and human compliance. Furthermore they do not require enrolment, nor the consent or the cooperation of the observed subject.

The following given Table-I in [1] illustrates a range of characteristics which accept the above stated definition. This given traits list is not exhaustive and will increase with technological progress. We firstly differentiate between soft biometric traits as face, body or accessory categories. Here note that even though classically accessories do not belong to biometry, the above stated definition clearly includes them in the category of soft biometrics. A further argumentation can be the intuitive human use of obvious accessory items as a mean of description and discrimination, for example "the person in the red shirt".

Further significant factors for classifying soft biometric traits are 'distinctiveness and permanence'. Distinctiveness is the strength with which a trait is able to distinguish between individuals. Beard as an example has a low distinctiveness, since it can only be applied to the male part of the population and furthermore has binary categories. The latter, points out a certain correlation between distinctiveness and nature of value. Continuous traits are in general more distinctive than discrete and moreover binary ones.

In this context we want to mention the difference between nature of value and human labeling of traits. While hair color has different nuances and is thus of continuous character, humans tend to label it for convenience purposes as discrete. This approach will as well be followed by soft biometric estimation algorithms, detecting for example hair color in categories (black, blond, brown, etc.) rather than RGB values.

The permanence of a trait plays a major role for the employable application. As an example an application, where identification within a day is required will accept low permanent traits like age, weight or clothes color. The final subdivision subjective perception stands for the ability of humans to unambiguously identify specific soft biometric traits. Again the nature of value plays an important role, since characteristics with binary categories, are generally more straightforward to be sensed than continuous ones. Increased subjective perception of discrete or continuous traits is further due to the not well-defined categories or the different beholder's percipience. In fact the notion of soft biometrics bares subjectivity even in the decision of the nature of value. With other words, colors can be argued to be continues, due to the huge variance in nuances blending into each other or discrete due to the fact that colors can be described by discrete RGB values.

TABLE-I

Soft Biometrics Traits

Soft biometric trait	Face/body/ accessory	Nature of value	Permanence	Distinctiveness	Subjective perception
Skin color	Face	Continuous	Medium	Low	Medium
Hair color	Face	Continuous	Medium	Medium	Medium
Eye color	Face	Continuous	High	Medium	Medium
Beard	Face	Binary	Low/medium	Low	Medium
Moustache	Face	Binary	Low/medium	Low	Medium
Facial measurements	Face	Continuous	High	Medium	Medium/high
Facial shapes	Face	Discrete	High	High	High
Facial feature measurements	Face	Continuous	High	High	Medium/high
Facial feature shapes	Face	Discrete	High	High	High
Make-up	Face	Discrete	Low	Low	Medium
Ethnicity	Face	Discrete	High	Medium	Medium
Marks	Face/body	Discrete	High	Medium/high	Low
Gender	Face/body	Binary	High	Low	Low
Age	Face/body	Continuous	Low/medium	Medium	Medium
Height	Body	Continuous	Medium/high	Medium	Medium
Weight	Body	Continuous	Low/medium	Medium	Medium
Gait	Body	Continuous	Medium	Medium	High
Body measurements	Body	Continuous	Medium/high	Medium/high	Medium
Body shapes	Body	Discrete	Medium	Medium	Medium
Clothes color	Accessory	Discrete	Low	Medium	Medium
Glasses	Accessory	Binary	Low/medium	Low	Low

A. TRAITS

Different types of soft biometric traits on the basis of different studies are given as follows:

1. Facial soft biometrics:

Former work on soft biometrics has been performed predominantly with the aim of preprocessing.

Color based facial soft biometrics:

The color based facial soft biometric traits (eye, skin, and hair color) are the most obvious facial identifiers, generally when we portray unknown individuals. Challenges for skin classification are on the one hand the low spread of different skin colors in color space, and as a consequence, on the other hand the high illumination dependence of classification.

Hair color is detected by similar techniques like skin color and often researched along, but has more broadly scattered color categories. In one of the method used is fuzzy theory which is used to detect faces in color images, where two fuzzy models describe the skin color and hair color, respectively.

Eye color detection, unlike the other color based facial soft biometrics is a relatively new research topic. An advantage of eye color detection is the availability of all necessary information in images used for iris pattern analysis. In the work related to fusion between iris texture and color is proposed where the authors fuse iris and iris color with fingerprint and provide performance improvement in respect with the unimodal systems.

Age:

Age plays an important role for long time employable systems based on face or body and is a challenging and relatively new field. We know that face changes over time, which spans a biometric, forensic, and anthropologic review, and further discusses work on synthesizing images of aged faces. Another study distinguishes children from adults based on the face/iris size ratio. Viola–Jones face detection technique is used, followed by an iterative Canny edge detection and a modified circular Hough transform for iris measuring, with good results.

Gender:

Gender perception and recognition has been immensely researched already in social and cognitive psychology work in the context of face recognition. The latest efforts employ a selection of fused biometric traits to deduce gender information. For example, gait energy images and facial features are fused and classified by support vector machines. Another approach proposes a combined gender and expression recognition system by modeling the face using an Active Appearance Model, feature extraction and finally linear, polynomial and radial based function based support vector machines for classification.

Ethnicity:

In the example for ethnicity recognition of Asian and non-Asian faces, machine learning framework applies a linear discriminant analysis (LDA) and multi scale analysis. A further framework, integrating the LDA analysis for input face images at different scales, further improves the classification performance. Another method for ethnicity recognition approach is based on Gabor Wavelets Transformation, combined with retina sampling for key facial features extraction. Finally support vector machines are used for ethnicity classification that provides very good results, even in the presence of various lighting conditions.

Recent work on facial soft biometrics is performed on scars, facial marks and periocular region which are given as follows:

Soft Biometrics using Facial Marks:

Soft biometric traits embedded in a face (e.g., gender and facial marks) are ancillary information and are not fully distinctive by themselves in face-recognition tasks. However, this information can be explicitly combined with face matching score to improve the overall face-recognition accuracy. Moreover, in certain application domains, e.g., visual surveillance, where a face image is occluded or is captured in off-frontal pose, soft biometric traits can provide even more valuable information for face matching or retrieval. Facial marks can also be useful to differentiate identical twins whose global facial appearances are very similar. The similarities found from soft biometrics can also be useful as a source of evidence in courts of law because they are more descriptive than the numerical matching scores generated by a traditional face matcher. Here we try to utilize demographic information (e.g., gender and ethnicity) and facial marks (e.g., scars, moles, and freckles) for improving face image matching and retrieval performance. An automatic facial mark detection method has been developed that uses 1) the active appearance model for locating primary facial features (e.g., eyes, nose, and mouth), 2) the Laplacian-of-Gaussian blob detection, and 3) morphological operators.

Soft Biometrics using Periocular region:

With periocular biometrics gaining attention recently, the goal of this study is to investigate the effectiveness of local appearance features extracted from the periocular region images for soft biometric classification. It contains extraction of gender and ethnicity information from the periocular region images using grayscale pixel intensities and periocular texture computed by Local Binary Patterns as our features and a SVM classifier. Furthermore, the fusion of the soft biometric information obtained from the stated classification approach with the texture based periocular recognition approach results in an overall performance improvement.

2. Body soft biometrics:

Height, gait, body weight and color of clothes concern the body and are the main traits that can be extracted from a distance. The best distinctiveness is provided by gait detection, which is why gait occasionally is referred to as a classical biometric.

Gait:

Gait is a complex pattern that involves not only some anthropometric parameters but also behavioral information. It is one of the few traits that can be gathered at a distance. In the preliminary experiment on gait analysis, the author used lights attached to the joints of the human body to record subject's gait models. The author demonstrates how observers can recognize walking people familiar to them just by the light traces. Since 1970's many other authors were interested in the topic of automatic gait recognition. For recognition with this analysis both the structural and behavioral characteristics of gait is captured. Another interesting work proposed is where gait is chosen as primary biometric trait to be coupled with "semantic biometrics", that seems to be a very similar concept to soft biometrics. The system merges the results of the signature generated by gait with the one generated by the semantic information in order to identify users of the biometric system.

Height:

For automatic height estimation, foreground and background recognition is necessary, which can be adopted by diverse silhouette extraction techniques used for gait recognition. Height is a trait employed for human tracking or as an aid for other algorithms, like gait. Here single and multiple calibrated camera systems are used for height estimation, respectively. The estimation is performed via the computation of height related to the real world coordinates estimated in camera images.

Weight:

The soft biometrics which involves weight uses a scale to weigh users of a fingerprint recognition system. By exploiting weight and body fat measurements the authors reduce the total error rate of the system by 2.4%. It is clear that weight represents a novel soft biometric trait that still has to be explored especially for what concerns its measurement.

3. Accessory soft biometrics:

The new soft biometrics definition allows the inclusion of accessories among these traits. Accessories can indeed be related to personal characteristics as sight problems in case of glasses, or personal choices as adornment in case of jewelry.

Eye glasses detection:

For eye glasses detection perform edge detection on a preprocessed gray level image. Certain face areas are observed and an indicator for glasses is searched for. The most successful identifier region for glasses is found to be the nose part of the glasses, between the eyes. A different approach for glasses extraction is employed, where a face model is established based on the Delaunay triangulation. A 3D method to detect glasses frames is presented, where 3D features are obtained by a trinocular stereo vision system. The best results on glasses detection up to now are achieved on thermal images.

4. Combined soft biometrics:

Since soft biometrics is individually not distinctive and permanent, a combination of traits could overcome those limits. Many recent papers deal with fusion of classical biometry and soft biometry or exclusively with fusion of soft biometric traits.

In some methods height and appearance are extracted from videos and exploited in a multiple camera video surveillance scenario in order to track the subjects that cross the surveillance network. This framework obtains information on primary biometric traits like face and finger print and secondary biometric traits like height and clothing needed for identification from video surveillance camera and fingerprint sensor in short distance to determine the access of the subject at the entrance of the building .Although height and clothing color are not as permanent and reliable but can be obtained from entrance camera and stored in the database and is used for additional biometric information. If the user is determined as unauthorized, the entry of the user will be controlled. So information on face, height and clothing color is obtained only by video surveillance camera.

V. CONCLUSION

Thus we studied the use of soft biometrics and how to utilize the ancillary user information like gender, age, facial marks, and color of eye, skin, and hair to complement the identity information provided by the traditional biometric system. Although these soft biometric characteristics have some limitations that mean they are not as permanent and reliable as the traditional biometric identifiers like fingerprint, they provide some important information about the identity of the user that can result into higher accuracy in establishing the user identity.

In this paper we also studied the general framework for integrating the soft biometric information with the output of the primary biometric system. Then identified and classified soft biometric traits and further elaborated the related advantages, limitations, scientific work and applications. In future we can say that the utility and range of soft biometrics systems are expected to increase, as the number of traits increases, and as the underlying image capture processing technology improves.

REFERENCES

[1] Antitza Dantcheva, Carmelo Velardo, Angela D'Angelo, Jean-Luc Dugelay, "Bag of Soft biometrics for person identification new trends and challenges", *Springer Science Business Media, LLC 2010.*

[2] A. K. Jain, S. C. Dass, and K. Nandakumar, "Soft biometric traits for personal recognition systems," in *Proc. ICBA*, vol. 3072, LNCS, pp. 731–738, 2004.

[3] A. K. Jain, S. C. Dass, and K. Nandakumar, "Can soft biometric traits assist user recognition?" *Proc. SPIE*, vol. 5404, pp. 561–572, 2004.

[4] J. Lyle, P. Miller, S. Pundlik, D. Woodard, "Soft biometric classification using periocular region features", in: *IEEE 4th International Conference on Bio- metrics:Theory, Applications, and Systems*, 2010.

[5] J. Lyle, P. Miller, S. Pundlik, D. Woodard, "Soft biometric classification using local appearance periocular region features", *Pattern recognition, Elsevier, 2012*.

[6] Unsang Park and Anil K. Jain, "Face Matching and Retrieval Using Soft Biometrics", *IEEE Transactions on information forensics and security*, vol.5, no.3, September 2010.

[7] A. K. Jain and U. Park, "Facial marks: Soft biometric for face recognition," in *Proc. IEEE ICIP*, pp. 1–4, 2009.

[8] Karl Ricanek Jr. and Benjamin Barbour," What are soft biometrics and how they can be used?" *Published by the IEEE Computer Society*, September 2011.

[9] B. Klare and A. K. Jain, "Sketch to photo matching: A feature-based approach," in *Proc. SPIE, Biometric Technology for Human Identification VII*, pp. 1–10, 2010.

[10] Min-Gu Kim, Hae-MinMoon Yongwha Chung, and Sung BumPan, "A Survey and Proposed Framework on the Soft Biometrics Technique for Human Identification in Intelligent Video Surveillance System", *Research Article, Journal of Biomedicine and Biotechnology Volume*, 2012.

[11] Koichiro Niinuma, Unsang Park and Anil K. Jain, "Soft Biometric Traits for Continuous User Authentication, *IEEE Transactions on Information forensics and security, vol.5, no.3, Dec 2010.*

