Interactive Mirror for Smart Home

Chidambaram Sethukkarasi, Vijayadharan SuseelaKumari HariKrishnan, Karuppiah PalAmutha
National Ubiquitous Computing Research Centre
Centre for Development of Advanced Computing
Chennai, India
{ctsethu, harikrishnans, palamuthak}@cdac.in

Raja Pitchiah
Department of Electronics and Information Technology
Government of India
Delhi, India
pitchiah@mit.gov.in

Abstract—**This paper describes the design and development of a smart artifact called "Interactive Mirror" for smart home users. The idea is to transform a normal mirror into an intelligent artifact by embedding various technologies to support users in their daily activities. This paper explains the state of the art technologies for building the intelligent mirror. It identifies the user using facial recognition technique and provides services such as recognizing emotions, progress representation of measured health parameters, height identification, identify garments, suggest garments with suitable color, and reminds important events. The prototype is developed, and demonstrated in ubiquitous computing laboratory. The algorithms are being tested in the deployed environment and the results are discussed in detail in this paper. Initial user studies indicated a high appeal of the Interactive Mirror features.**

Keywords - Ubiquitous Computing; Interactive mirror; Face Recognition; Emotion Recognition; Human Height Identification; Smart Artifact; RFID (Radio Frequency IDentification); Garment Identification; Garment Suggestion.

I. Introduction

A home environment consists of variety of devices and there exists huge information due to technology advancements. An effort towards ubiquitous computing is to intelligently collect or sense the information from the environment and make it smarter. Nowadays, not only the computers but also any devices like mobile phone, PDAs, tablets have network connectivity and offer various intelligent services to us. The devices used in our daily life can be made smart enough to assist the smart home users. In general, people normally spend considerable time in front of mirror and it has been considered as an ideal artifact for embedding intelligence for demonstrating proposed interactive mirror concept. The general approach is to extract information about the scene using computer vision, and use this information to update a scene model to be rendered using computer graphics.

An effort towards this results in the development of an interactive mirror [1], which augments a normal mirror with intelligence and provides value added services. To identify the user, we propose to utilize facial recognition technology since it is a non contact based recognition method. The mirror assists the individuals in aiding a healthier life style by providing feedback on the measurement of basic health parameters. The user's garments are identified using RFID technology and its details/descriptions are displayed in the mirror. The mirror also guides user in selecting a suitable garment color according to their skin color. The system has been designed, developed, and deployed in ubiquitous research lab.

The contributions of this paper are: 1) conceptualization, design and development of interactive mirror prototype. 2) Emotion Recognition. 3) Human height identification using image processing technique. 4) Garment Identification using RFID technology. 5) Guidance in selecting suitable garment based on skin color. 6) Accuracy improvements of image processing modules.

The reminder of this paper is organized as follows. Section II briefly comments on some related work. The system and functional overview are described in Section III and IV respectively. Section V presents implementation details of the prototype. The experimental results are discussed in Section VI. Conclusion and future work are given in Section VII.

II. Related Work

In several investigations, smart homes have been developed by combining monitor and mirror systems. The AwareMirror [2] is an augmented display that is placed in the bathroom for presenting personalized information to the user. It detects the person's position using proximity sensor and identify using RFID tag embedded in toothbrush. It provides useful information such as closest schedule, transportation information, and the weather forecast. The mirror is constructed by attaching an acrylic board in front of a monitor. Using tooth brush as a tool for identification might not produce accurate results since it needs to be replaced more frequently and tagged properly. This paper uses face recognition technique, a biometric identification system for person identification.

The Memory Mirror [3], developed at the Everyday Computing Lab, aims at helping people remember tasks that have to be repeated. It uses a camera and face-recognition software to identify different users in a home. The drugs are tagged with RFID tags and readers to locate and keep track of drug usage. It alerts consumers if they have taken the wrong bottle or if it is the right bottle at the wrong time. Besides, the cabinet enables patients to monitor blood pressure, heart rate and cholesterol levels, and share this information with their doctor via the Internet. The cabinet also provides a trend chart, and if this one shows a problem tendency, the system will suggest the user to make an appointment with their doctor. The drug usage factor heavily depends on RFID technology and all items need to be tagged properly. The proposed system facilitates users to set remainders. When user comes in front of the mirror, he/she will be reminded. In our system, user needs to set reminders, which will be reminded by the mirror when they use the system.

A mirror that detects and analyzes a human behavior, is demonstrated under persuasive mirror [4]. It makes use of behavioral data in order to provide its user with continuous visual feedback on their behavior. A mirror that provides a natural means of interaction through which the residents can control the household smart appliances and access personalized services is described in [5]. It uses face recognition to authenticate the user and provides widget based interface to access data feeds and other services. Tomoki Hayashi, Hideaki Uchiyama, Julien Pilet and Hideo Saito proposed an Interactive Digital Mirror [6], which captures the ambient light with a camera, extract information about the scene, and display appropriate information to the user, combining real-time computer vision systems with realistic computer graphic. The computer vision routine includes human face detection, tracking and 3D head pose estimation.

Philips research lab demonstrated an Intelligent Bathroom Mirror [7], which supports people need and enable people with new possibilities in the daily use of the bathroom space. It comprises two display mirrors connected to the PC having Internet access, a TV tuner, a wirelessly connected electric toothbrush, a weight and height sensors and two video cameras. It provides personalized services according to the user's preferences such as children can watch their favorite cartoon while brushing their teeth, provide live TV feeds, monitor the latest weather, check the traffic information, provides health information and so on. Personnel recognition has been proposed by using weight, height, etc. Two methods of interaction without physically touching the mirror have been demonstrated. The bathroom lighting comprise of 50 light sources of different kind. Various light sources have been used, which generates light of different color and temperature. Similar to this intelligent mirror, [8] presents i-mirror, which attempts to create an interactive information environment within a mirror interface in a natural way. A special optical system is designed using ac camera, a projector, mirror and a screen to bring a mirror like interface. The three main characteristic of i-mirror are: it shows images in dark, younger/older views of a person and memory to playback the older scenes. The mirror records the scene, which may lead to privacy issues.

A Magic mirror [9] can function like a good friend who listens to the user's questions and automatically responds to their request. It is an interactive multimedia mirror system, which includes speech recognition, speech synthesis, face detection/modified/recognition, 3D virtual genius, hidden LCD (Liquid Crystal Display) mirror, and camera, performs simple syndication to capture information about peripherals and network connections. It can detect an user's feeling based on speech and image recognition features to select the appropriate music and speech to alter the user's mood. Our interactive mirror is also an intelligent mirror, which recognizes the user's mood and attempts to assist the user in leading a healthier lifestyle by providing feedback on measured health parameters.

One of the main usages of mirror is to see how we look like in particular attire? Ching-I Cheng and Damon Shing-Min Liu developed an Intelligent Dressing Advice System [10] to help women choose correct attire for attending a specific occasion. Fuzzy logic rules were used to search good matches in the garment database and showing the matched results. Our application analyzes the skin color of the user and suggests a suitable list of garments based on both the color parameter and occasions to wear.

In comparison with the other works described above, our work is different in that we aim to develop a system for assisting smart home users in their daily activities such as selecting suitable garment, recognizing emotions, and represents the progress in health parameters. The normal mirror is transformed into a smart device by preserving its metaphor in addition to embedding technologies.

III. System Overview

The Interactive Mirror [1] shown in Figure 1 comprises of a LCD display placed behind a dielectric coated mirror, a camera, a weight measurement platform, and a RFID reader.

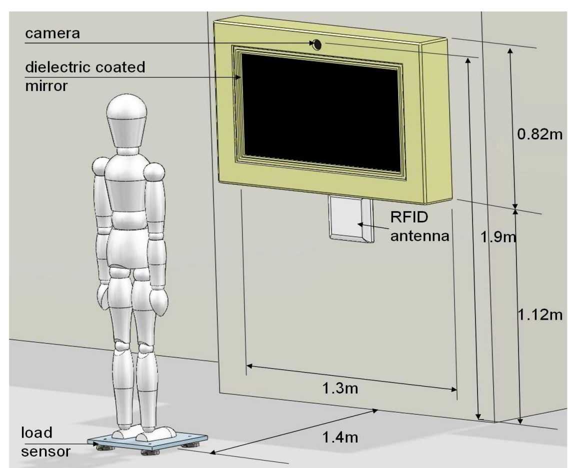

Figure 1. Engineering of Interactive Mirror

The Interactive mirror display can be used in two different modes by placing a dielectric coated mirror over the LCD display. The modes of operation are 1) Normal Mirror: When the display is OFF, it will act as a normal reflective mirror; 2) Interactive Mirror: When the display is ON, it will act as a see through glass and the display can be viewed. A web camera of resolution 640 x 480 is used for capturing the images in front of the mirror. A weighing platform, designed using four load sensors is placed in front of the mirror to measure the person's weight. A RFID reader and an antenna are connected to the system for identifying the tags attached with the garments.

IV. Functional Overview

The system has the following functionalities, as shown in Figure 2.

A. *Person Identification*

Recognizing the user is the first step towards providing personalized services. The system recognizes the user using face recognition technique.

B. *Health Information Services*

The users can measure their health parameters such as weight, height, BMI (Body Mass Index) and BMR (Basal Metabolic Rate) daily with the help of mirror. The system maintains a health database and analyzes the progress of health parameters over the recent days and displayed to the user in the form of 3D graph.

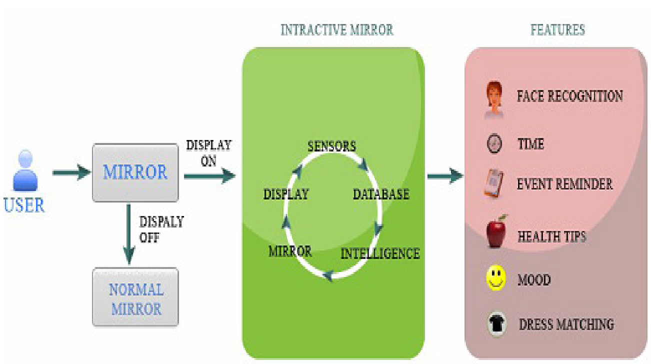

Figure 2. Functional Overview of Interactive Mirror

C. *Clothing advisor*

The mirror assists the user in selecting a suitable garment for a particular occasion out of users garment. It also suggests suitable dress colors based on their skin color.

D. *Recognize emotion*

The emotion of the person is recognized using image processing technique. The emotions like Happy, Sad, Surprise and Normal are identified.

E. *Reminder assistance*

The important messages and reminders are displayed to the recognized users like bill payments, tour plans, meeting schedule, etc. The user needs to set reminders with the help of a GUI (Graphical User Interface).

V. Prototype Development

The prototype is developed using various technologies and tools such as: face recognition, emotion recognition, RFID, PostgreSQL and java programming language. The mirror uses a display with a camera and a weighing platform for user identification and providing personalized services in smart home environment. The system comprises of the following modules: Face recognition, Emotion Recognition, Health Progress Representation, Garment Identification and Suggestion and other features.

A. *Face Recognition*

The mirror identifies the person using face recognition technique. When the user starts using the system, the camera triggers ON and captures the image in front of it. The image contrast is improved by histogram equalization technique before processing by face recognition module.

Naturally, human identifies and recognizes the face based on the characteristics of facial features such as eyes, nose, mouth, and lips. Same process is followed in image processing routines for person identification. The steps involved in face recognition module are shown in Figure 3. The first step is to find faces in an image called as face detection. There exist many techniques such as viola-jones face detection, skin color based detection, LBP (Local Binary Patterns), Adaboost, facial geometry [11], etc. We adopted viola-jones face detection algorithm [12] for our system since it is the most adopted algorithm for face detection in real time. The algorithm has high detection speed with relatively high detection accuracy. It is an especially successful method [13] with very less false positive rate.

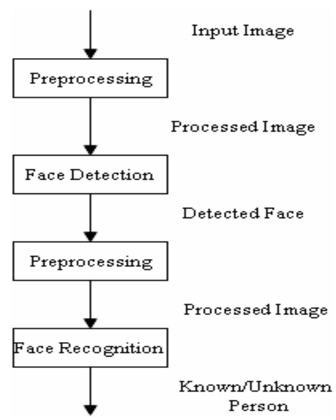

Figure 3. Steps involved in Face Recognition Module

This method makes use of HAAR features, which describes the properties common to human face. The basic principle of the Viola-Jones algorithm [14] is to scan an input image using a sub-window capable of detecting faces. The sub window looks for Haar like features in the image. The standard image processing approach would be to rescale the input image to different sizes and then run the fixed size detector through these images. This approach is a time consuming due to the calculation of the different size images. Here, the detector is rescaled instead of the input image and run the detector many times through the image (each time with a different size). There exists an enormous amount of such features in a sub image. Among all these features few are expected to give almost consistently high values when on top of a face. In order to find these features Viola-Jones uses a modified version of the AdaBoost algorithm [14] developed by Freund and Schapire in 1996. AdaBoost is a machine learning boosting algorithm capable of constructing a strong each containing a strong classifier. The job of each stage is to determine whether a given sub-window is definitely not a face or maybe face. When a sub-window is classified to be a non-face by a given stage it is immediately discarded. Conversely, a sub-window classified as maybe-face is passed on to the next stage in the cascade. It follows that the more stages a given sub-window passes, the higher the chance of sub-window contains a face. A window that passes through all classifiers is classified as face image.

Prior to face recognition, the detected face image has to be pre-processed to remove the background pixels other than the facial features. The detected face image will have few background pixels such as hair pixels other than the facial features. This may affect the recognition accuracy. In order to remove these pixels from consideration, we used ellipse fitting procedure for segmenting facial features region alone from others pixels. We created an elliptical mask (shown in Figure 4a) that has a pixel value of 0 in the region we are interested in and 1 elsewhere. The mask is bitwise ORed with the detected face image (shown in Figure 4b) to segment the interested region of the face image from other pixels. The segmented image (shown in Figure 4c) having facial features is an input for the face recognition module.

a) b) c)

Figure 4. a) Elliptical Mask b) Detected face c) Segmented face image

The next step is face recognition, which is the process of telling whose face is this? There exist a lot of algorithms and methods for recognizing human faces [15]. We were using eigen face recognition [16][17][18] technique, which is one of the global feature based approach. The Eigen face approach has the following advantages and limitations [19] over other methods:

Advantages:

- Recognition is simple, efficient and easy to implement
- No knowledge of geometry or special feature of the face is required
- Little preprocessing work

Limitations:

- Sensitive to scale, pose and illumination variation
- Suitable only for frontal face images
- Good performance under controlled background.

The system is designed for smart home environment where the background and lighting may not change frequently. Based on this consideration, we adopted eigen face approach for our system. The steps involved in eigen face recognition algorithm are described below. The face objects such as eyes, nose and mouth and its relative distances between these object forms the characteristic features of a face image. These characteristics features are called eigenfaces or principal components. These principal components are identified using a mathematical tool called PCA (Principal Component Analysis). By means of PCA, each original image of the training set is transformed into a corresponding eigenface. An important feature of PCA is that one can reconstruct any original image from the training set by combining the eigenfaces. The original image can be reconstructed with the weighted sum of all eigen faces. This weight specifies to what degree the specific feature (eigenface) is present in the original image. Based on the weights, the face images are grouped into classes. The weight vectors for the training and test images are calculated in training and testing phase respectively. For classification, compute the average distance measure between the training and test image weight vectors using distance measure techniques. The least distance measure is compared against the threshold values t_1 and t_2 to classify the test image as known or unknown person. The details about choosing the threshold values and criteria for classification are described in detail under results section.

B. *Emotion Recognition*

Facial Expression is an inevitable factor in analyzing the emotion of a person. K-nearest neighbor algorithm is used for classifying the input images into four facial expression happy, sad, normal, and surprise. Human Emotions are the most valuable aspects of human life. By the analysis of facial expression and emotional aspect of person, the interactive mirror detects the emotion of a person. The emotional analysis data is recorded for a period of time (say 1 year or 1 month) can be used to identify the mental state of the person such us conditions like depression, fear and enthusiastic factors of user. If required, the details are provided to a doctor. This will be helpful also for depression monitoring of patients.

The camera mounted on the top of the mirror captures the user's image. The captured image is passed through the face detection process to detect and segment the face image. The detected face image will be given as an input to the preprocessor module of facial expression recognition algorithm. The mouth and eyes play an important role in determining or extracting user's emotion. We have extracted mouth features for classifying the expression.

There exist various methods for extracting the facial features [20][21][22][23][24]. The emotion recognition framework based on video sequences and the challenges involved in it are discussed in [25]. The experiments show that the proposed facial expression recognition framework yields relatively little degradation in recognition rate, when faces are partially occluded, or under a variety of levels of noise introduced at the feature tracker level.

Different training and classification methods are being analyzed, which includes Support Vector Machines, Hidden Markov Models, Bayes Classifiers. [26] tells about six classification methods used for facial expression recognition using the above algorithms. The image processing techniques like PCA, LFA (Local Feature Analysis), ICA

(Independent Component Analysis), or FLD (Fisher's Linear Discriminant) are some of the methods used for feature extraction. Lucas-Kanade tracking algorithm for tracking eyebrows and cheeks, canny edge detector for detection of wrinkles are used in preprocessing steps. Approach in [27] describes about using topological mask and similarity measurement for classification of facial expression. Approach in [28] calculates difference image sequence by subtracting the pixel value at the same position (x,y) from video image sequences of adjacent frames and makes use of hidden markov model for classification of facial expression. Approach in [29] creates active appearance model in the form of 2D or 3D mesh of the training sample images and uses the algorithms nearest neighbor and support vector machine for classification of expression. Approach in [30] also extracts 2D images from 3D image sequences and uses the algorithms combined k Nearest Neighbor/rule-based classifier and SNoW (Sparse Network of Winnows) classifier for training and expression classification. Approach in [31] uses SVM (Support Vector Machine) for classification of facial expression for images from video sequences. Approach in [32] uses Tree Augmented Naïve Bayes Classifier.

The popular databases used in the above experiments include Yale database [33], Cohn-Kanade database (video image sequences) [34], MMI –Facial Expression Database [35], The JAFEE (JApanese Female Facial Expression JAFFE) Database [36].

Facial Action Coding System [37] is a widely used method for encoding facial expression based on contraction of facial muscles. It was developed by Paul Ekman and W.V. Friesen in 1970s.

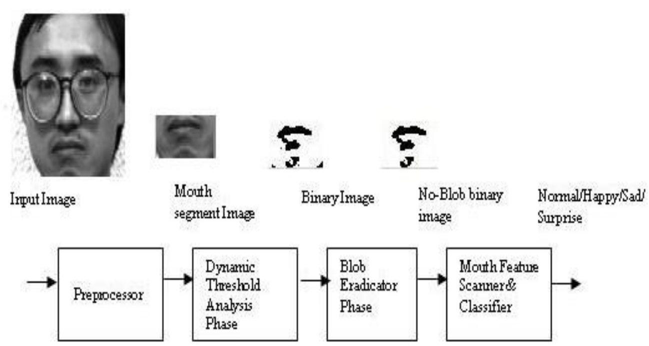

Figure 5. Mouth Feature Extraction.

The steps involved in extracting the mouth features are shown in Figure 5.

1) Preprocessor

The preprocessor takes the input image from the face detection module. From the face image, the mouth region is segmented in order to extract the mouth features. Based on facial geometry property, the mouth region is segmented.

2) Dynamic Threshold Analysis Phase

In the dynamic threshold analysis phase, the segmented mouth image is converted into binary image. Determining the proper threshold value for thresholding an image is too complex. The threshold value for accurate and legible extraction for binary facial images may vary depending on various environmental factors like illumination, the deviation from camera where the person is standing, the facial color of person etc. of the input facial image. To resolve this issue, the dynamic threshold analysis module calculates the mouth segment pixel density factor. Mouth segment pixel density factor is the number of black pixels in the segmented binary mouth image. From our observation, the value of 100 and 230 for mouth segment pixel density factor is the apt for accurate extraction.

Dynamic binary image selection is the method used by the dynamic threshold analysis module. The input image should be threshold with a random threshold value (T) to obtain a binary facial image. Dynamic threshold analysis module calculates the mouth segment density factor of the binary image. If the mouth segment pixel density factor is less than 100, the thresholding process should be repeated with a new threshold value of T_1=T+10, else if the mouth segment pixel density factor is greater than 230, the thresholding process is repeated for the new threshold value of T_1=T-10. The above thresholding process is repeated with new value of T=T_1, until the MSPDF (Mouth Segment Pixel Density Factor) reaches a saturation, i.e., until the condition $100<MSPDF<230$ is achieved.

Let t_c be a threshold increment/decrement factor, the numerical value by which T should be incremented or decremented. In the above occasion, we took the value of t_c as 10. The equation can be written as

$$T_1 = T \pm t_c \tag{1}$$

If we choose t_c as a fixed value like 10, it may lead to infinite repetition of dynamic binary image selection as the Mouth Segment P. This is because MSPDF does not converge to $100<MSPDF<230$ for higher values of t_c. To resolve this issue dynamic threshold analysis phase changes the value of t_c as $t_c = t_c /2$ (t_c is always a natural number). The threshold increment/decrement factor t_c is updated only after every 10 iteration of the dynamic binary image selection described above. The decrease in value of t_c helps the threshold value to converge to a better option of T for which $100<MSPDF<230$. The binary image obtained with MSPDF value between 100 and 230 is used as output of dynamic threshold analysis module, which is further sent for blob detection. We choose MSPDF boundaries lower bound as 100 and upper bound as 230 based on visual observations experienced in clarity of mouth image segment.

3) Blob Eradicator Phase

The obtained binary image may have small unwanted black blobs surrounding the lip. This may be because of non uniform lighting in face image or mustache or beard. To extract the mouth features more accurately, the unwanted blobs needs to be removed. Blob Eradicator use OpenCV blob detection functions to filter all the blobs having blob length less than a count of 12 pixels. After the blobs are removed, the mouth segment image is sent to mouth scanner module.

4) Mouth Feature Scanner and Classifier

The mouth feature scanner module scans the mouth image pixel by pixel along the row pixels. The starting point of scanning is always the middle topmost pixel of mouth segment. Mouth feature scanner divides the mouth segment image into two parts cutting it along the middle pixel column (vertically). Left scanning involves scanning the left portion of the mouth segment to determine left lip tip point A co-ordinates. Right scanning involves scanning right portion of the mouth segment to determine right lip tip point C co-ordinates. The top most first detected black pixel becomes the lip top point B. The points A, B, C in the extracted binary lip image forms a lip triangle ABC shown in Figure 6. The altitude of the triangle ABC gives lip height and the length of side AC gives lip length.

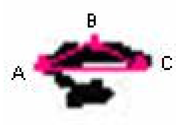

Figure 6. Mouth Triangle.

5) Facial Expression Classification using K-Nearest Neighbour Algorithm

The knn (k-nearest neighbor) algorithm is a method for classifying objects based on closest training examples in the feature space. In k-NN, an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst its k nearest neighbors. The feature space used in this classification is a two dimensional feature space.

The following assumptions have been made: 1) Constant lighting and illumination; 2) Front Face image; 3) No scaling variations for testing of facial expression recognition.

C. Health Progress Representation

The mirror assists the user in leading a healthier life by advising on health parameters such as weight, height, BMI and BMR. The weighing platform measures the weight of the user when he/she starts using it.

There exist several methods to measure the human height, which includes using IR (Infrared) sensor, Ultrasonic sensor and camera. The human height is used as a biometric identity for identifying home residents [38]. Height is typically a weak biometric, but it is well suited for identifying among a few residents in the home, and can potentially be improved by using the history of height measurements. The system has been tested with 20 subjects in 3 homes and that height sensors could potentially achieve at least 95% identification accuracy. A similar work is described by Hsin-Chun Tsai [39], which combines both height measurement and face recognition to identify the person in long distances. The human height identification is used in surveillance application [40] to spot persons coming from the dark area.

We used camera captured image and image processing technique to measure the human height. The steps involved in height detection module (shown in Figure 7) are described as follows.

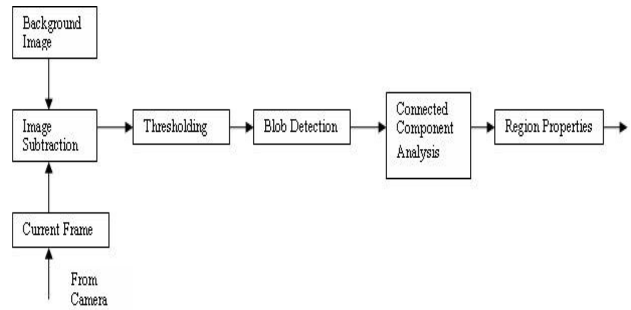

Figure 7. Height Identification

1) Background frame initialization: The image is preprocessed to improve the contrast and remove noise pixels. After preprocessing, a background frame needs to be initialized. There are many ways to obtain the initial background image. For example, take the first frame as the background directly, or the average pixel brightness of the first few frames as the background or using a background image sequences without the prospect of moving objects to estimate the background model parameters.

2) Background Segmentation and thresholding: The current frame denoted as F (shown in Figure 9b) is subtracted from the background image denoted as B (shown in Figure 9a) and if the difference is greater than the threshold value th, then the pixel belongs to foreground otherwise it belongs to background. Mathematically, it can be written in equation as follows:

$$P(x, y) = 255 \quad \text{if } F(x, y) - B(x, y) > th \tag{2}$$

$$P(x, y) = 0 \quad \text{otherwise} \tag{3}$$

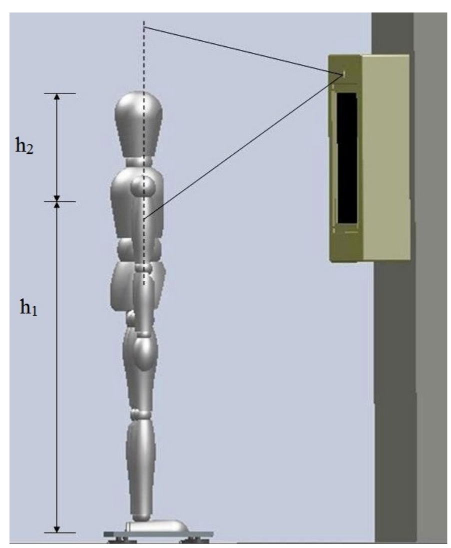

Figure 8. Height Measurement

Figure 9. a) Background Image b) Foreground Image c) Output Image after connected component analysis

3) Blob Detection and Analysis: The Morphological operations such as filtering the holes, filling up of holes, dilation, erosion and removal of smaller blobs are carried out to detect the foreground object. Obtain the height of the blob using connected component labeling and region properties. It gives the height of the human in terms of pixel. The obtained height is converted to actual units. The distance between the camera and human is too short to cover the entire human in an image. The height between the weighing platform and the camera coverage area (h1) is added upon with the obtained height output (h2) to determine the user's height h (h= h1 + h2 as shown in Figure 8).

From the measured weight and height values, parameters such as BMI and BMR were calculated. These parameters were measured and recorded regularly in a health database. BMI is a factor that determines a person's weight according to his height and BMR is the measure of number of calories burned by the body when the user is at rest. This helps in determining how much calories he/she needs to intake in order to maintain an energy balance and balanced diet condition. The measured BMI value is analyzed for the conditions such as normal, overweight, underweight and obesity. Drastic weight change over a short time is the main symptom of identifying certain major diseases in human body. Therefore, the measured values are saved in a health database and its progress over a period of time is reported and displayed to the user in the form of 3D graph.

D. Garment Identification and Suggestion

The textile industries and fashion garments started adopting RFID technology for tagging the garments and other clothes for rapid identification of items throughout its life cycle. Each tag has a unique ID number associated with the garments model name and description, including size, color and fabric, price, material, etc. As the user comes near the mirror wearing a tagged garment/cloth, the system identified it and captures the ID number. The application updates the garment database and the description of that particular garment is retrieved from the database and displayed in the mirror. The user is given feedback on how often he/she is using the particular garment.

The mirror detects the user skin color and attempts to suggest suitable colors for garments. It also tries to suggest according to the occasions such as party, outing, traditional program, competitions, etc.

To suggest a suitable garment color, the skin color of the user needs to be detected and categorized. Skin pixel detection and segmentation is employed in many tasks related to the detection and tracking of humans and human-body parts. The goal of skin pixel detection is to locate the pixels belongs to the skin and discard other pixels in an image.

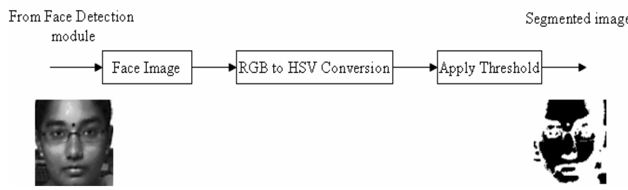

Figure 10. Skin Pixel Segmentation

The simplest way to decide whether a pixel belongs to skin color [41] or not is to explicitly define a boundary based on color channels. Brand and Mason [42] constructed a simple one dimensional skin classifier: a pixel is labeled as a skin if the ratio between its R and G channels is between a lower and an upper bound. The color spaces that are frequently used in studies are RGB, HIS, HSV, TSL and YUV.

The Kovac model [43] contains four sub-rules as follows. Pixel is skin color pixel if:

$$R > 95 \text{ and } G > 40 \text{ and } B > 20$$
$$Max(R,G,B) - Min(R,G,B) > 15$$
$$|R - G| > 15$$
$$R > G \text{ and } R > B$$

This rule can be interpreted as the range of R value is from 96 to 255, the range of G value is from 41 to 239, and the range of B value is 21 to 254. Since R value is always greater than G and B, the second rule and third rule are always positive values. Tomaz et al. [44] described that if R-value is too high, and the G and B values are too low, it will result in a pixel more close to red, and should not be considered as skin pixel. In other cases when R < 100 and G < 100 and B < 100, it will result to dark color that may be non-skin pixel, and when

G > 150 and B < 90 or R + G > 400, it will result in yellow like color. Swift's rule [45] is simpler as compared to Kovac's rule and can be described as follows. Pixel is not skin color pixel if:

$$B > R, G < B, G > R, B < 1/4R \text{ or } B > 200$$

The range of R-value is from 4 to 255, the range of B-value is from 1 to 200, and the range of G-value is from 1 to 255. This rule is unable to detect some dark skin color and yellow like color, which is detected as skin color. Finally, a very simple rule was introduced by Saleh [46], which considers only the value of R and G. This rule defines that a pixel is skin pixel when R – G is greater than 20 and less than 80. This rule does not consider a present of B-value that contributed to the whitish color. This rule is also unable to detect dark skin color or skin cover under shadow, and yellow like color and redder color problems, which is detected as skin pixel. We used HSV color space for detecting skin pixels. The steps involved in skin detection module are shown in Figure 10. The face image from the face detection module was converted to HSV color space and the following rule was applied for identifying skin color pixels.

.

$$P(x, y) = 0 \quad \text{if } H < 20, S > 48, V > 80 \qquad (4)$$

$$P(X, y) = 255 \quad \text{otherwise} \qquad (5)$$

The image was then converted to binary image with skin color pixels as black and other pixels as white according to the above rule. Based on segmentation, the average pixel value of the skin color is calculated and classified into three categories dark, medium, and bright. Based on the category, a suitable garment color is suggested for the user. For example, light color garments can be suggested for dark skin complexion, dark color garments for light skin complexion and both light and dark colors for medium complexion. It can also suggest a suitable garment according to the events like marriage function, birthday party, family outing, pilgrim functions, etc. based on the parameters such as price, usage factor, material, and type. The information about the events and occasions are acquired from the event reminder database.

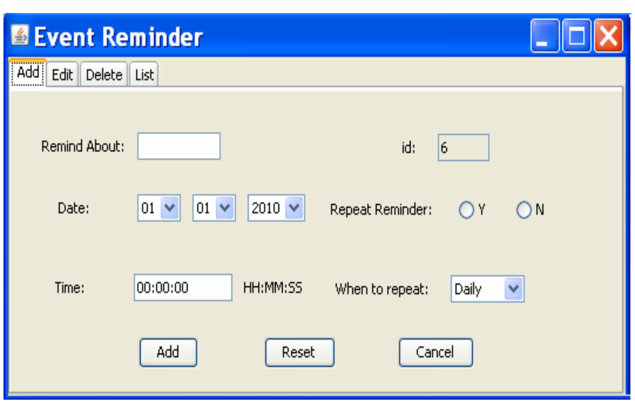

Figure 11. Event Reminder GUI

E. Other Features

The mirror observes how long a user is using it. The weighing platform senses the user presence and absence and calculates the mirror usage time. It can intimate the user if they are getting late for an office, school, or meetings, etc. It also acts as a reminder device, which reminds the user on important things such as Bill payments, prioritized works, birthdays, etc. The user needs to set details such as what to remind, when to remind, repeat reminder, etc. with the help of a GUI shown in Figure 11.

VI. Experimental Results

The prototype has been deployed in our ubiquitous computing laboratory (shown in Figure 12).

A. Face Recognition Algorithm Test Results:

We created our face database with 42 images of 9 different subjects. The images were collected in a semi-controlled environment. To maintain a degree of consistency throughout the database, the same physical setup was used in capturing the images. We maintained constant lighting and the distance between the mirror and weighing platform were kept constant to avoid major scaling variations. The images were collected on different days and at different time. In order to collect front face images, the user were asked to look straight at the mirror and there were no other instructions or restrictions given to the user. Therefore, the images were not exactly frontal face but little variation exists. This created a real practical testing environment.

The face recognition module was trained with 45 images of 9 different subjects. We created a test set of 24 face images with both known and unknown subjects. For distance measure, we explored both euclidean and mahalanobis distance method. The Euclidean Distance is the most widely used distance metric.

Figure 12. Deployment of Interactive Mirror in Ubicomp Laboratory.

The euclidean distance between two points is given as:

$$d(x, y) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (6)$$

The Mahalanobis Distance is a better distance measure when it comes to pattern recognition problems. It takes into account the covariance between the variables and hence removes the problems related to scale and correlation that are inherent with the Euclidean Distance. It is given as:

$$d(x, y) = \sqrt{(x - y)^T S^{-1} (x, y)} \quad (7)$$

Therefore, we adopted Mahalanobis distance method for recognition. The distance measure will give the least distance match or score of the test image out of the training images. When an unknown face image comes up for recognition task, it will still say the test image is recognized as the training image with the lowest score. It is for this purpose that we decided the threshold t_1. Also, to classify non face images with face images, another threshold t_2 for the distance measure was evolved. These threshold values t_1 and t_2 were evolved experimentally and heuristically.

To choose the threshold we chose a set of random images (both face and non-face); we then calculated the distance measure for images of subjects in the database and also for this random set and set the threshold t_2 accordingly. Threshold t_2 decides whether the test image is a face or non face image. If the test image is a face image, then it should fall near some face class in the face space. Again, threshold t_1 decides whether it is a known or unknown face image. When the training set changes, then these threshold values need to be calculated again.

There are four possible combinations on where an input image can lie:

$$(d < t_1) \,\&\&(d < t_2) \quad (8)$$

$$(d > t_1) \,\&\&(d < t_2) \quad (9)$$

$$(d > t_1) \,\&\&(d > t_2) \quad (10)$$

$$(d < t_1) \,\&\&(d > t_2) \quad (11)$$

1. Near a face class and near the face space: The test image is a facial image of a known subject.
2. Near face space but away from face class: The test image is a facial image of an unknown subject.
3. Distant from face space and near face class: The test image is a non face image but still resembles like the one in dataset (False Positive).
4. Distant from both the face space and face class: The test image is not a face image.

The algorithm was tested with 24 images of known and unknown faces and the above said criteria were used for recognition. The results obtained are tabulated in Table I.

TABLE I. FACE RECOGNITION ALGORITHM TEST RESULTS

No of Test Images	Successfully Recognized	Success Percentage
24	**20**	**83%**

The images that produced false results were not frontal face images. Those images were pose varied, tilted and with expressions. Overall, the performance of the face recognition module has found to be reasonable to work with front face images.

Not all the eigenfaces play important role in recognition task. Few eigenfaces may increase the error rate in recognition. The associated eigenvalues allow us to rank the eigenvectors according to their usefulness in characterizing the variation among the images. Therefore, eigenfaces having low eigenvalues can be discarded.

For the Eigenface method (PCA), it has been suggested that by discarding the three most significant principal components, variations due to lighting can be reduced [47]. In [48], experimental results show that the Eigenface method performs better under variable lighting conditions after removing the first three principal components. However, the first several components not only correspond to illumination variations, but also some useful information for discrimination. Besides, since the Eigenface method is highly dependent on the training samples, there is no guarantee that the first three principal components are mainly related to illumination variations and it is evident that discarding first several principal components cannot improve the performance significantly.

The face recognition can be integrated with height measurement to improve the recognition accuracy.

B. Emotion Recognition Algorithm Test Results:

We used 2 face databases for testing our application. The databases used are 1) Yale database and 2) CDAC (Centre for Development of Advanced Computing) Face Database. We used Yale facial expression database images of 15 subjects for training and testing our facial expression recognition algorithm. Yale database images of 10 subjects (Lip length and half of 'mouth opening height' values) are used for training the K-NN classifier. Figure 13 is the knn-classifier feature space plotted after completion of training for Yale Database. We used openCV library functions for training the classifier. To visualize the two dimensional estimated feature space points for KNN (k=6) with Lip length and Lip Height at X-axis and Y-axis, respectively, we used four colors to denote four expressions in the feature space. The red, green, blue and black color represents normal, happy, surprise and sad expression respectively in feature space. The test results of the algorithm with the Yale database images are listed in Table II. The best accuracy level was obtained for the k value of 6.

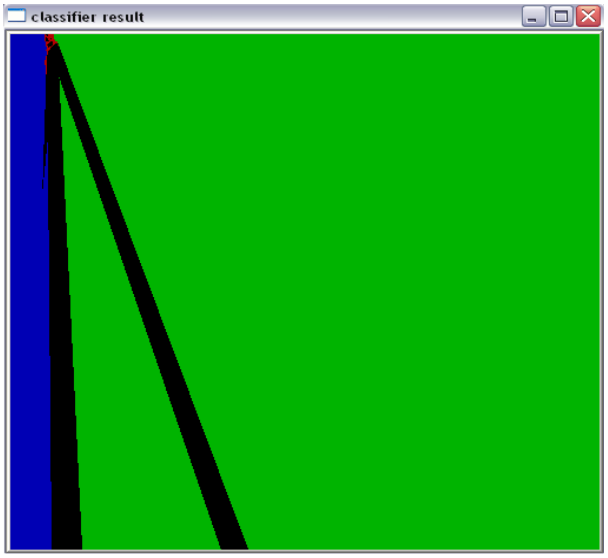

Figure 13. Classification Output (Yale Database Images)

TABLE II. EMOTION RECOGNITION ALGORITHM ACCURACY(YALE DATABASE)

K	Happy	Surprise	Normal	Sad
3	93.33	53.33	66.66	33.33
6	100	60	46.6	26.66

The Emotion recognition algorithm had been tested with CDAC database images in the deployment environment. The algorithm was trained with the mouth features like mouth width and mouth height. The results are presented in Table III. The algorithm was capable enough to recognize the happy expression when compared to other expressions.

Figure 14 describes knn-classifier feature space plotted after completion of training CDAC Database. On K-NN training phase completion, the estimated feature space points are plotted with Lip length and Lip Height at X-axis and Y-axis, respectively. The feature space used is a two dimensional feature space. The k value was selected as 6 for k-NN classifier as it is found to be more accurate when tested with CDAC database images. The training was performed to classify four expressions happy, sad, normal and surprise. Our database consists of 400 images of 5 different subjects. For each expression, there are 20 images per subject. For each expression, 5 images per subject were used for training and 15 images per subject were used for testing. The restricted lab environment with fixed illumination and fixed distance from camera was used for testing.

TABLE III. EMOTION RECOGNITION ALGORITHM ACCURACY (C-DAC DATABASE)

K	Happy	Surprise	Normal	Sad
6	90.67%	52.00%	42.67%	50.66%

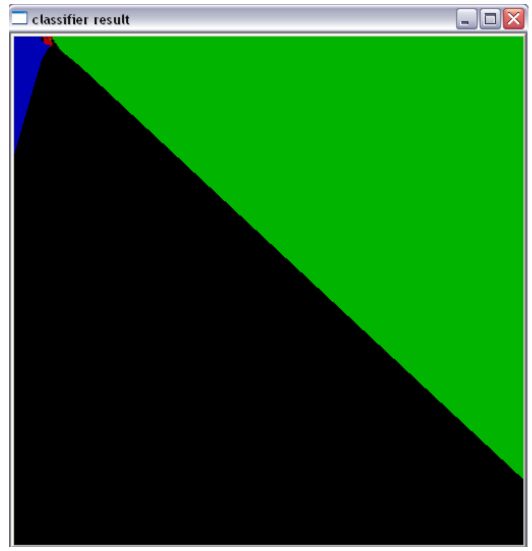

Figure 14. Feature Space Plot (C-DAC database Images)

The following are some of the issues faced in mouth feature extraction:

1) Segmentation Problem:

When segmenting the mouth region, in certain images a small portion of the nose is also included. The nose part is identified as the mouth top point B by the mouth feature scanner algorithm as shown in Figure 15. This reduces the algorithm accuracy.

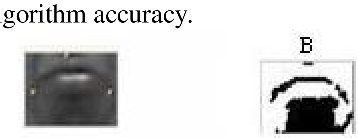

Figure 15. Mouth Feature Extraction – Segmentation Problem

Our facial expression database images for the expressions happy, surprise and sad are shown in Figures 16, 17, and 18, respectively. Our approach uses extraction of lip length and lip height features and using K-Nearest Neighbor algorithm for classification as described in the above sections. We are planning to enhance the classification accuracy using facial action coding units.

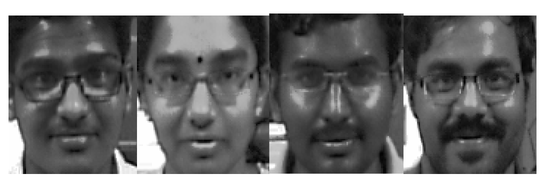

Figure 16. Happy Expression Database Images.

Figure 17. Surprise Expression Database Images.

Figure 18. Sad Expression Database Images.

2) Non-uniform lighting problem:

The mouth portion is not segmented properly during thresholding process, because of non uniform lighting in the face region. Only ¾ length of the original mouth is visible in binary image and the corners are not visible shown in Figure 19. This results in inaccurate feature point extraction.

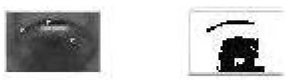

Figure 19. Mouth Feature Extraction – NonUniform Lighting Problem

C. Weight Measurement Accuracy

The four digital load sensors each capable of measuring maximum 50 lb are used to build a weighing platform. The weight measurement accuracy depends on the load sensor accuracy.

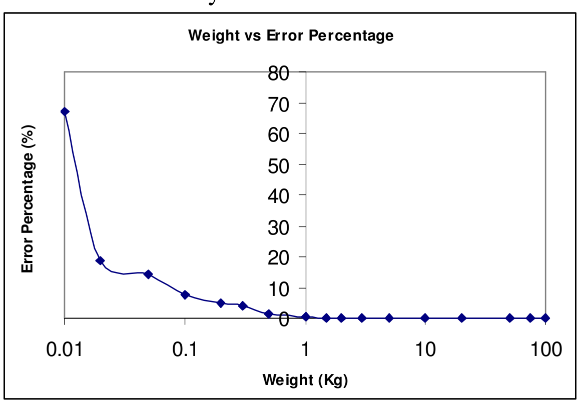

Figure 20. Weight versus error percentage.

We tested the load sensor accuracy with standard weights. The percentage error for various loads on a 50 lb load sensor is shown in Figure 20. It is concluded that the error rate is less when it is loaded with maximum capacity. Also, it was found that the error rate is found to be high for the weights equal to or less than 1 kg. That is not considered as the major problem, since the weight of the user who will use the system might be of more than 1 kg.

D. Others

The RFID tag performance on different dress materials and human body needs to be analyzed. The proper placement of tags and antenna has to be tested for better read performance. The performance of garment integrated RFID antennas are studied and detailed in [49]. The results show that embroidery technique can be used to fabricate RFID tag and wireless sensor antennas, which are intended to be used in clothing very near the human body.

The rest of the features like mirror usage time and event reminder were found to be useful. The proper placement of RFID antenna and tags in garments need to be evolved to achieve better read performance.

E. User Evaluation

The use case scenario is as follows: When a person enters the dressing room, he/she stands in front of the mirror. The weighing platform detects the user's presence and triggers camera ON to capture the image. The initial step is to recognize the user using face recognition technique with the help of the captured image. The health parameters are measured and saved in the database. The progress of health parameters are analyzed and displayed in the form of 3D graph. The weight is displayed using java speedometer component with a needle moves over a weight scale. The garment details are then identified and its descriptions are shown on the screen and followed by suggestion on suitable garments. This information is also provided in audio output using MBROLA TTS (Text To Speech) engine.

The system is deployed in our UBICOMP (UBIquitous COMPuting) laboratory for testing. The users of age group 24 to 32 used the mirror for a period of say 10 or 12 days and gave feedback. The feedback rating is plotted in Figure 21. The rating is done considering the comfort/convenience level achieved in each service. The rating 5 represents that the user is more comfortable and 1 represents that the user is not at all comfortable.

Few suggestions from the user are as follows:

1) In addition to recognizing the emotions, they want the mirror to entertain them in case they are not in good mood.

2) Want more interactions like voice output other than displaying information.

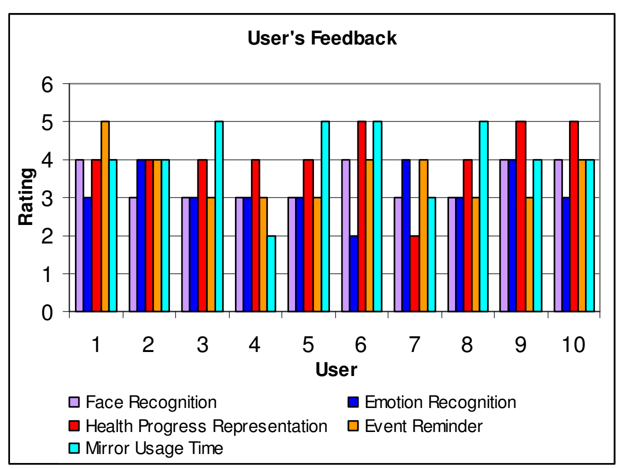

Figure 21. User's Feedback

3) Feeling little incomfort in asking them to particularly stand on the weighing platform.

4) Can this system be adoptable for more than one users using the mirror at same time?

5) Color lighting environment can be created as per the user's preference.

The user's feedback and suggestions were collected and improvements are being done.

VII. Conclusion and Future Work

This paper demonstrated a smart artifact for assisting smart home users in their daily activities. It incorporates intelligence into a normal mirror by embedding image processing and RFID technologies. The mirror recognizes the user using face recognition technique and offers personalized services. These services include recognizing emotions, identifying garments, suggesting suitable garment colors, remind important events, and monitor the progress of health parameters. It functions in two different modes of operation: In normal mode, it acts like a traditional mirror and in interactive mode it acts as an intelligent device that recognizes information and provides personalized services. The prototype has been developed and deployed in ubiquitous computing laboratory and kept for user evaluation. The image processing algorithms were tested in the deployed environment and the results were discussed in detail. The primary user feedbacks were mostly positive and the system is found to be highly satisfied and useful. Our future work includes improvement of image processing algorithms in terms of accuracy, automate the training process of face recognition algorithm, enrich the features using machine learning techniques, introduce additional technologies such as speech recognition, and enable touch screen facility for user interaction. Empirical evaluation of users feedback in the laboratory environment is given in this paper. More exhaustive evaluation of users experience in real-life environments could be carried out as further scope of work. The features of the mirror can be enhanced for deploying in other environments such as Beauty Parlors, Textile shops, and Hotels. However, security and privacy requirements need to be adequately addressed.

Acknowledgment

The work is developed under National Ubiquitous Computing Research Centre, funded by Department of Electronics and Information Technology, Government of India. We would like to thank Department of Electronics and Information Technology for providing us an opportunity to develop this smart artifact.

References

[1] C. Sethukkarasi, V. S. Harikrishnan, and R. Pitchiah, "Design and Development of Interactive Mirror for Aware Home," The First International Conference on Smart Systems, Devices and Technologies, pp. 1-8, 2012.

[2] K. Fujinami, F. Kawsar, and T. Nakajima, "Aware Mirror: A Personalized Display Using a Mirror," Third International Conference, User Interaction, Pervasive Computing, vol. 3468, pp. 315 332, May 2005.

[3] Everyday Computing Lab, Memory Mirror, Available from: http://we-make-money-not-art.com/memory_mirror_a/ 2016.05.18

[4] A. C. Andres del Valle and A. Opalach, "The Persuasive Mirror: computerized persuasion for healthy living," Accenture Technology Labs, France. Available from: http://www. consultoras.org/frontend/movil/descargar.php?idf=6700 2016.05.18

[5] M. A. Hossain, P. K. Atrey, and A. E. Saddik, "Smart Mirror for Ambient Home Environment," 3rd IET International Conference on Intelligent Environments, pp. 589-596, 24-25 Sep 2007.

[6] C. H. Morimoto, "Interactive Digital Mirror," IEEE Proceedings on Computer Graphics and Image Processing, pp. 232-236, 2001.

[7] T. Lashina, "Intelligent Bathroom," Philips Research, Netherlands. Available from: https://www.researchgate.net/publication/ 228881021_Intelligent_bathroom 2016.05.10

[8] K. Ushida, Y. Tanaka, T. Naemura, and H. Harashima, "i-mirror: An interaction/information environment based on a mirror metaphor aiming to install into our life space," Proceedings of the 12th International Conference on Artificial Reality and Telexistence (ICAT2002), pp. 113-118, 2002.

[9] J. Ding, C. Huang, J. Lin, J. Yang, and C. Wu, "Magic Mirror," IEEE International Symposium on Multimedia, pp. 176-185, December 2007.

[10] C. Cheng and D. S. Liu, "Discovering Dressing Knowledge for an Intelligent Dressing Advising System," Fourth IEEE International Conference on Fuzzy Systems and Knowledge Discovery, pp. 339-343, 2007.

[11] H. H. J. Kim, Survey Paper : Face Detection and Face Recognition. Available from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.93.4859 &rep=rep1&type=pdf 2016.05.10

[12] P. Viola and M. J. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," IEEE Conference on Computer Vision and Pattern Recognition, pp. 511-518, 2001.

[13] V. Gupta and D. Sharma, "A Study of Various Face Detection Methods," International Journal of Advanced Research in Computer and Communication Engineering, vol. 3, Issue 5, pp. 6694-6697, May 2014.

[14] O. H. Jensen, "Implementing the Viola-Jones Face Detection Algorithm," IMM-M.Sc.: ISBN 87-643-0008-0, ISSN 1601-233X, 2008.

[15] W. Zhao, R. Chellappa, A. Rosenfeld, and P. Phillips, "Face Recognition : A Literature Survey," ACM Computing Survey, pp. 399-459, 2003.

[16] M. Turk and A. Pentland, "Eigenfaces for Recognition," J. Cognitive Neuroscience, Vol. 3, pp. 71-86, 1991.

[17] M. A. Turk and A. P. Pentland, "Face Recognition using Eigen Faces," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 586-591, 1991.

[18] Face Recognition using Eigenfaces and Distance Classifiers: A Tutorial. Available from: http://onionesquereality.wordpress.com/2009/02/11/face-recognition-using-eigenfaces-and-distance-classifiers-a-tutorial/ 2016.05.31

[19] R. K. Gupta and U. K. Sahu, "Real Time Face Recognition under Different Conditions," International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 3, Issue 1, pp. 86-93, January 2013.

[20] W.K. Teo, L. C. D. Silva, and P. Vadakkepat, "Facial Expression Detection and Recognition System," Journal of the Institution Engineers, pp. 14-26, 2004.

[21] H. Gu, G. Su, and C. Du, "Feature points extraction from Faces," Image Vision and Computing, pp. 154-158, 2003.

[22] R. S. Feris, T. E. D. Campos, and R. M. C. Junior, “Detection and tracking of Face features in video Sequences,” Lecture Notes in Artificial Intelligence, vol. 1793, pp. 129-137, Apr. 2000.

[23] M. Deriche, “A Simple Face Recognition Algorithm using Eigeneyes and a Class-Dependent PCA Implementation,” International Journal of Soft Computing 3(6), pp. 438-442, 2008.

[24] M. Gargesha and S. Panchanathan, “A Hybrid Technique for Facial Feature Point Detection,” Fifth IEEE Southwest Symposium on Image Analysis and Interpretation, pp. 134-138, 2002.

[25] F. Bourel, C. C. Chibelushi, and A. A. Low, “Robust Facial Expression Recognition Using a State-Based Model of Spatially – Localised Facial Dynamics,” Proceedings of the Fifth IEEE international Conference on Automatic Face and Gesture Recognition, pp. 106-111, 2002.

[26] C. C. Chibelushi and F. Bourel, “Facial Expression Recognition: A Brief Tutorial Overview”. Available from: http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/CHIBELUSHI1/CCC_FB_FacExprRecCVonline.pdf 2016.05.31

[27] X. Wei, J. Loi, and L. Yin, “Classifying Facial Expressions Based on Topo-Feature Representation,” Affective Computing, pp. 69-82, 2008. Available from: http://www.intechopen.com/books/affective_computing/classifying_facial_expressions_based_on_topo-feature_representation 2016.05.31

[28] F. Hulsken, F. Wallhoff, and G. Rigoll, “Facial Expression with Pseudo-3D Hidden Markov Models,” 23rd DAGM-Symposium on Pattern Recognition, pp. 291-297, 2001.

[29] S. Lucey, I. Matthews, C. Hu, Z. Ambadar, F. D. L. Torre, and J. Cohn, “AAM Derived Face Representations for Robust Facial Action Recognition,” 7th International Conference on Automatic Face and Gesture Recognition, pp. 155-162, 2006.

[30] M. Valstar, M. Pantic, and I. Patras, “Motion History for Facial Action Detection in Video,” IEEE International Conference on Systems, Man and Cybernetics, pp. 635-640, 2004.

[31] P. Michel, and R. E. Kaliouby, “Real Time Facial Expression Recognition in Video using Support Vector Machines,” The 5th International Conference on Multimodal interfaces, pp. 258-264, 2003.

[32] I. Cohen, N. Sebe, A. Garg, L. Chen, and T. S. Huang, “Facial Expression Recognition From Video Sequences,” IEEE International Conference on Multimedia and Expo, pp. 121-124, 2002.

[33] Yale Face Database. Available from: http://cvc.cs.yale.edu/cvc/projects/yalefaces/yalefaces.html 2016.05.24

[34] Cohn-Kanade AU-Coded Expression Database. Available from: http://www.pitt.edu/~jeffcohn/CKandCK+.htm 2016.05.24

[35] MMI Face Database. Available from: http://www.mmifacedb.com/page/About/ 2016.05.24

[36] The Japanese Female Facial Expression (JAFFE) Database. Available from: http://www.kasrl.org/jaffe.html 2016.05.24

[37] Y. Tian, T. Kanade, and J. F. Cohn, “Recognizing Upper Face Action Units for Facial Expression Analysis,” IEEE Conference on Computer Vision and Pattern Recognition, pp. 294-301, 2002.

[38] V. Srinivasan, J. Stankovic, and K. Whitehouse, "Using height sensors for bio-metric identification in Multi-resident Homes,” Pervasive Computing, pp. 337-354, 2010.

[39] H. Tsai, W. Wang, J. Wang, and J. Wang, “Long Distance Person Identification Using Height Measurement and Face Recognition,” IEEE TENCON 2009, pp. 1-4, 23-26 Jan 2009.

[40] E. Jeges, I. Kispal, and Z. Hornak, “Measuring Human Height using Calibrated Cameras,” IEEE Conference on Human System Interactions, pp. 755-760, 2008.

[41] R. Sohal and T. Kaur, “Skin Pixel Segmentation Using Learning Based Classification: Analysis and Performance Comparison,” International Journal of Engineering Research and Applications, Vol. 4, Issue 5 (Version 3), pp. 49-56, May 2014.

[42] J. Brand and J. S. Mason, “A Comparative Assessment of Three Approaches to Pixel-Level Human Skin Detection,” Proc. IEEE International Conference on Pattern Recognition, vol. 1, pp. 1056-1059, Sep. 2000.

[43] J. Kovac, P. Peer, and F. Solina, "Human skin colour clustering for face detection," IEEE EUROCON 2003, pp. 144-148, 2003.

[44] F. Tomaz, T. Candeias, and H. Shahbazkia, "Improved automatic skin detection in color images," VIIth Digital Image Computing: Techniques and Applications, Sydney, pp. 419-427, 2003.

[45] D. B. Swift, "Evaluating graphic image files for objectionable content," US Patent US 6895111 B1, 2006.

[46] S. A. Al-Shehri, "A simple and novel method for skin detection and face locating and tracking," Asia-Pacific Conference on Computer-Human Interaction (APCHI 2004), LNCS 3101, 2004.

[47] W. Chen, M. J. Er, and S. Wu, “Illumination Compensation and Normalization for Robust Face Recognition Using Discrete Cosine Transform in Logarithm Domain,” IEEE Transactions on systems, man, and cybernetics, Vol. 36, no. 2, pp. 458-466, Apr. 2006.

[48] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, “Eigenfaces versus Fisherfaces: recognition using class specific linear projection,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no. 7, pp. 711–720, Jul. 1997.

[49] K. N. Goh, Y. Y. Chen, and E. S. Lin, “Developing a Smart Wardrobe System,” IEEE Consumer Communications and Networking Conference, pp. 303-307, 2011.