Human Identification at a Distance Using Body Shape Information

N K A M Rashid¹, M F Yahya and A A Shafie
Department of Mechatronics Engineering, Faculty of Engineering, International Islamic University Malaysia, 53100, Kuala Lumpur, Malaysia

E-mail: nahrul@iium.edu.my

Abstract. Shape of human body is unique from one person to another. This paper presents an intelligent system approach for human identification at a distance using human body shape information. The body features used are the head, shoulder, and trunk. Image processing techniques for detection of these body features were developed in this work. Then, the features are recognized using fuzzy logic approach and used as inputs to a recognition system based on a multilayer neural network. The developed system is only applicable for recognizing a person from its frontal view and specifically constrained to male gender to simplify the algorithm. In this research, the accuracy for human identification using the proposed method is 77.5%. Thus, it is proved that human can be identified at a distance using body shape information.

1. Introduction
Human identification is critical in today’s society for enhancing public safety and privacy protection. Many methods published so far for human identification at a distance focused on face and gait like [1], [2], and [3]. However, these methods have their own vulnerabilities such as face is prone to fake facial images as demonstrated in [4] and gait is susceptible to occlusion as in [5]. For that reason, developing a new algorithm by utilizing information other than the face or gait is highly anticipated. One of the available information if seen from a far is the shape of human body. Basically, the shape of human body offers distinctive features not only at the upper part of the body which is the shape of the head but also the slant of the shoulder and the physique of the trunk. These distinctive features when combined form a set of body features which are unique from one person to another. In this paper, an intelligent system approach by utilizing fuzzy logic and multilayer neural network is developed to recognize human at a distance based from their body shape information.

2. Methodology
The development of the proposed human identification system consists of three stages. At the initial stage, the process is concerned about extracting body features data from human subjects (i.e. data of the head, shoulder, and trunk). In the second stage, based on the acquired data, fuzzy logic is utilized in a way such that each body feature is classified into its respective classes. The classes are determined based on the shape of the head, the slant of the shoulder, and the physique of the trunk. Then, in the final stage, the information gained from the fuzzy logic implementation is fed into an artificial neural network. This is so that the network can learn and recognize the pattern of the information which belongs to that of a specific human subject.

¹ To whom any correspondence should be addressed.
2.1. Human subject body features data acquisition

In order for a system to perform human recognition, it needs a priori data about the human subject. The priori data is collected based on the body features of human specifically the head, shoulder, and trunk. The head, shoulder, and trunk are detected using a developed set of image processing techniques. Figure 1 shows the original image, the detected shape of body features and their corresponding parameters which will be processed further for recognition purposes. Parameters $H_w$, $H_h$, $H_{bp}$, and $H_{wp}$ in figure 1 (b) symbolize the width of the head, the height of the head, the accumulation of black pixels, and the accumulation of white pixels in the bounding box which encloses the head respectively. In figure 1 (c), parameter $\gamma_s$ represents the topmost angle of the marked triangle and figure 1 (d) shows $T_w$ and $T_h$ which are the parameters which represent the width and the height of the trunk respectively.

![Figure 1](image1.jpg)
(a) original image (b) head (c) shoulder (d) trunk

2.2. Body features classification using fuzzy logic

Since human have ambiguous body shape, the accurate mathematical models for classifying their features are difficult to establish. Fortunately, fuzzy logic provides an alternative approach for modeling problems which are imprecisely defined by mathematical models. In this research, the overall application of fuzzy logic concepts is utilized using a fuzzy inference system (FIS). A FIS defines a nonlinear mapping of the input data vector of body features into a scalar output (the type of body feature). Three FISs are developed for classifying the shape of the head, the slant of the shoulder, and the physique of the trunk. Additionally, triangular membership function is utilized for all of the developed membership functions.

2.2.1. The shape of the head. From cropping and observing several processed images of human head, its shape can be classified either as oval vertical, square, circle, or oval horizontal as shown in figure 2.

![Figure 2](image2.jpg)
(a) oval vertical (b) square (c) circle (d) oval horizontal

From the figure, there are two parameters which can be used to determine the different type of head which are the aspect ratio and the deviation of its shape to the bounding box which encloses the head. The head aspect ratio $R_h$ is defined as
and the head shape deviation to the bounding box is defined as

\[ D_h = \frac{H_{hv}}{H_{ho}} \]  

For both of the parameters \( R_h \) and \( D_h \), their linguistic variables are \{Small, Medium, Big\}. Then, the rules are as tabulated in table 1.

**Table 1. FIS rules for head classification**

| Deviation | Ratio   | Small       | Medium     | Big          |
|-----------|---------|-------------|------------|--------------|
| Small     | Oval_Vertical | Square     | Oval_Horizontal |
| Medium    | Oval_Vertical | Circle    | Oval_Horizontal |
| Big       | Oval_Vertical | Circle    | Oval_Horizontal |

Output of the FIS is a defuzzification value from 1 to 4 where 1, 2, 3, and 4 represent oval vertical, square, circle, and oval horizontal respectively.

2.2.2. The slant of the shoulder. Shoulder slant can be either one of the three classes of very concave down, concave down, or straight as shown in figure 3.

![Figure 3. The slant of the shoulder: (a) very concave down (b) concave down (c) straight](image)

The slant of the shoulder is determined by making use of the constructed triangular geometry using three points between the neck and shoulder of a person. From basic triangular geometry as shown in figure 4, the angle \( \gamma_{pst} \) plays an important role to classify shoulder to its respective classes.

![Figure 4. Triangular geometry](image)

The angle \( \gamma \) at \( P_{st} \) of the triangle by using the law of cosine is found to be

\[ \gamma_{pst} = \cos^{-1}\left(\frac{d_{p_{s1}p_{st}}^2 + d_{p_{s4}p_{st}}^2 - d_{p_{s1}p_{s4}}^2}{2(d_{p_{s1}p_{st}})(d_{p_{s4}p_{st}})}\right) \]
The shoulder angle is defined as

\[ \gamma_s = \gamma_{p\text{st}} \]  

(4)

For the input, let \( \gamma_s \) indicates the shoulder slant angle with linguistic variable be \{Small, Medium, Big\}. For the output membership function (the shoulder type), it has linguistic variable set to \{Very_Concave_Down, Concave_Down, Straight\}. The rules are stated as

\[ R_1: \text{If Angle is Small Then Shoulder Type is Very_Concave_Down} \]
\[ R_2: \text{If Angle is Medium Then Shoulder Type is Concave_Down} \]
\[ R_3: \text{If Angle is Big Then Shoulder Type is Straight} \]

Output of the FIS is a defuzzification value from 1 to 3 where 1, 2, and 3 represent very concave down, concave down, and straight respectively.

2.2.3. The physique of the trunk. In this research, it is discovered that the physique of the trunk can be classified into five classes of very thin, thin, slim, fat, and very fat as depicted in figure 5.

(a)                         (b)              (c)                (d)                       (e)

Figure 5. The physique of the trunk: (a) very thin (b) thin (c) slim (d) fat (e) very fat

The physique of the trunk can be determined by using the aspect ratio of the bounding box which encloses the trunk. The trunk aspect ratio is defined as

\[ T_r = \frac{T_w}{T_h} \]  

(5)

The input to the FIS is defined as \( T_r \) with its subsequent linguistic variable be \{Very_Small, Small, Medium, Big, Very_Big\}. The output membership function (the trunk type) has linguistic variable term set to \{Very_Thin, Thin, Slim, Fat, Very_Fat\}. The rules statement for classifying the type of trunk are stated as

\[ R_1: \text{If Aspect Ratio is Very_Small Then Trunk Type is Very_Thin} \]
\[ R_2: \text{If Aspect Ratio is Small Then Trunk Type is Thin} \]
\[ R_3: \text{If Aspect Ratio is Medium Then Trunk Type is Slim} \]
\[ R_4: \text{If Aspect Ratio is Big Then Trunk Type is Fat} \]
\[ R_5: \text{If Aspect Ratio is Very_Big Then Trunk Type is Very_Fat} \]

Output of the FIS is a defuzzification value from 1 to 5 where 1, 2, 3, 4, and 5 represent very thin, thin, slim, fat, and very fat respectively.

2.3. Training of artificial neural network

In this research, since there are multiple inputs of body features information and multiple outputs of Boolean algebra sequence (assigned to each of the human subjects initially), a multilayer network
architecture is best suited to perform mapping of inputs to outputs. Figure 6 shows an architectural design of the network. The network basically consists of three layers; one input layer $L_1$, one hidden layer $L_2$, and one output layer $L_3$.

![Figure 6. A three-layer feed-forward network](image)

$x_1, x_2, \ldots, x_n$ signify the input (from the defuzzification values of the head, shoulder, and trunk) to the multilayer neural network. Each neuron uses sigmoid activation function because it is smooth and bounded function. The output of the network is $y_1, y_2, y_3$ or combination of Boolean algebra set to 000, 001, 010, …, 111. Backpropagation learning algorithm is used to train the network since the inputs (body features information) are non-linearly separable. During training, it is found out that the optimum network have 0.01 mean-squared error with 7078 epochs for which the network is set to have 10 hidden neurons. This network is selected to undergone a test for accuracy of the developed human identification system.

3. Results and Discussion

The human identification accuracy is defined as

$$h_{It} = \frac{n_{si}}{n_{ti}} \times 100\%$$

(1)

where $n_{si}$ is the number of successful identification and $n_{ti}$ is the total number of identification. Before the accuracy test can be performed, the analysis requires experimental setup. A total of 20 video sequences each for 8 people are recorded, where one sequence last for approximately 15 seconds. During the recording, each person was instructed to go towards the camera from multiple directions and stop at a distance of about 2 meters in front of the camera. Then, the images from the video are captured to go through the body features data acquisition stage and afterwards the body features classification using fuzzy logic stage. The initial 10 processed datasets are used as continuous input to train a neural network and another 10 processed datasets are used to test the network. Table 2 shows the tabulated result for successful identification of the human subject.
Table 2. Successful human identification

| Human subject | Successful identification (out of 10) |
|---------------|--------------------------------------|
| HS001         | 7                                    |
| HS002         | 7                                    |
| HS003         | 9                                    |
| HS004         | 8                                    |
| HS005         | 8                                    |
| HS006         | 8                                    |
| HS007         | 7                                    |
| HS008         | 8                                    |

Based on table 2, the final calculated accuracy is found to be 77.5%.

4. Summary
In this paper, an intelligent system approach for human identification at a distance using the body shape information is proposed. The system utilizes three body characteristics of human: the shape of the head, the slant of the shoulder, and the physique of the trunk. The shape of the head can be categorized into four different classes: oval vertical, square, circle, and oval horizontal. The slant of the shoulder can be either one of the class of very concave down, concave down, or straight. The physique of the trunk classes are very thin, thin, slim, fat, and very fat. By using FIS, those classes can be represented as linguistic variables which may yield a unique single value representing each class of body feature. These values are then fed into an artificial neural network to generalize the pattern of given inputs during training process. For testing purposes, the response outputs of neural network are converted into combination of Boolean algebra values to represent or identify specific human subject. During testing of the proposed system, the successful identification rate of 77.5% is achieved. Thus, it is concluded that human can be recognized from a distance using body shape information.

References
[1] Bhanu B and Zhou X 2006 Feature fusion of face and gait for human recognition at a distance in video Int. Conf. on Pattern Recognition vol 4 pp 529-532
[2] Almohammad Manhal Saleh, Salama Gouda Ismail and Mahmoud Tarek Ahmed 2012 Human identification system based on feature level fusion using face and gait biometrics Int. Conf. on Engineering and Technology (ICET) pp 1-5
[3] Kale Amit, Roychowdhury A K and Chellappa R 2004 Fusion of gait and face for human identification Proc. IEEE Conf. on Acoustics, Speech, and Signal Processing vol 5 pp 901-904
[4] Määttä J, Hadid A and Pietikainen M 2012 Face spoofing detection from single images using texture and local shape analysis IET Biometrics vol 1 pp 3-10
[5] Hofmann M, Wolf D and Rigoll G 2001 Identification and reconstruction of complete gait cycles for person identification in crowded scenes Proc. Int. Conf. on Computer Vision Theory and Applications (VISAPP)