

# Person Recognition Using Soft Biometric

Lavanya Srinivasan (✉ [L.Srinivasan@soton.ac.uk](mailto:L.Srinivasan@soton.ac.uk))

University of Southampton

---

## Research

**Keywords:** Gait, surveillance, security, recognition, semantic.

**Posted Date:** September 28th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-929836/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Person Recognition using soft biometric

Lavanya Srinivasan, University of Southampton, United Kingdom.

L.Srinivasan@soton.ac.uk

**Abstract**-Global security concerns raised the multiplication of video surveillance devices. Intelligent systems identify the person captured at the different cameras, angles, views, background, and wearing different accessories. Gait is measured at a distance without human cooperation. In this work, gait recognition is increased by concatenating semantic features with traditional features. Combining features are dimensionally reduced and classified using classifiers. This method motivates in future to increase the gait recognition with less false positive detection.

**Keywords** Gait; surveillance; security; recognition; semantic.

## I Introduction

Human detection and reidentification are the main research interests in computer vision, especially in video surveillance and security. Human detection, re-identification are challenging tasks that depend upon many different parameters such as backgrounds, views, angles, distance from the cameras, accessories, and the environment. Biometrics provide an automated method to identify people based on their physical or behavioural characteristics. In recent years, the increased threat of terrorist activities and the ever-growing surveillance infrastructure has driven the development of biometrics that operates at a distance. These could recognize people from surveillance footage without their cooperation. This is crucial to quickly identifying known criminals or suspects. Gait biometrics is the most popular long-distance biometrics.

Human is detected using support vector machine classifier trained on histograms of oriented gradients [1], [2], [3]. Census

transform histogram features are extracted to identify humans [4]. Pedestrian extraction in thermal Image used a Histogram of local intensity differences and texture weighted HOG features [5]. Cascade of boosted classifiers is used for real-time human detection [6]. The presented method utilizes standard human detection techniques [7]. Detection of the human method based on hot spots and discrete cosine transform [8]. Human motion detection based on Fisher's ratio [9] background subtraction.

Throughout history the use of human descriptions obtained from eyewitnesses has instigated the identification and apprehension of suspects, humans naturally use labels estimations of physical attributes to describe people. This work aims to use soft biometrics to bridge this gap. Jain et al. [10] defined soft biometric traits as "characteristics that provide some information about the individual but lack the distinctiveness and permanence to sufficiently differentiate any two individuals. Soft biometric recognition is a technique that uses semantic descriptions as features to identify subjects [11]. Soft biometric attributes include height, weight, gender, and skin colour used to identify a person in practical applications [12]. A model using the human semantic description of soft biometrics to identify subjects was proposed in [13], where soft biometric features were used to enrich the recognition method. Nineteen body features were investigated in [14], and the results demonstrated that leg length and arm length can aid recognition. A method proposed in [15] used soft biometric attributes to improve the recognition performance of traditional biometrics.

In this paper, we propose a concatenation of traditional features with semantic features.

The main contributions of our work in the following five aspects:

1. Human detection using YOLO.
2. Background subtraction using GMM.
3. Extracted traditional features from Gait Entropy Image (GEI).
4. Extraction of semantic features based on Elo ranking.
5. Concatenate traditional and soft biometric features to improve recognition accuracy.

The rest of the paper is organized as follows: Section 2 describes the proposed methodology for human detection technique. Background subtraction of the experimental details, results and the comparison of the proposed method with the traditional works are given; and the paper is concluded with future work in Section 3.

## II Methodology

Gait urban video data were recorded using different infrared and visible cameras. The Infrared cameras are Short Wave Infrared (SWIR), Medium Wave Infrared (MWIR), and Long Wave Infrared (LWIR). Thermal imaging plays a role in video surveillance by promising higher robustness to bad weather, illumination changes, and night vision. Compared to the standard RGB visible cameras, thermal infrared cameras are robust in specific conditions such as illumination changes and different climatic conditions. Dispersions of the temperature in the acquired thermal images are utilized for human detection. Thermal images are a low dynamic range and resolution, one channel, low information, and high noise compared to RGB images.

Pre-processing includes extracting frames from videos, human detection using You Only Look Once (YOLO), and background subtraction. The Gait Energy Entropy features are extracted from the images and classified using K-Nearest Neighbour and Random Forest. The Traditional Gait Energy Entropy features are concatenated with the soft

biometric features to increase the gait recognition accuracy. Figure 1 illustrates the complete process of Gait recognition.

### A. Human detection

YOLO predicts objects inside the bounding boxes. It predicts four coordinates for each bounding box  $I_x, I_y, I_w, I_h$ . The prediction of the boxes corresponds to:

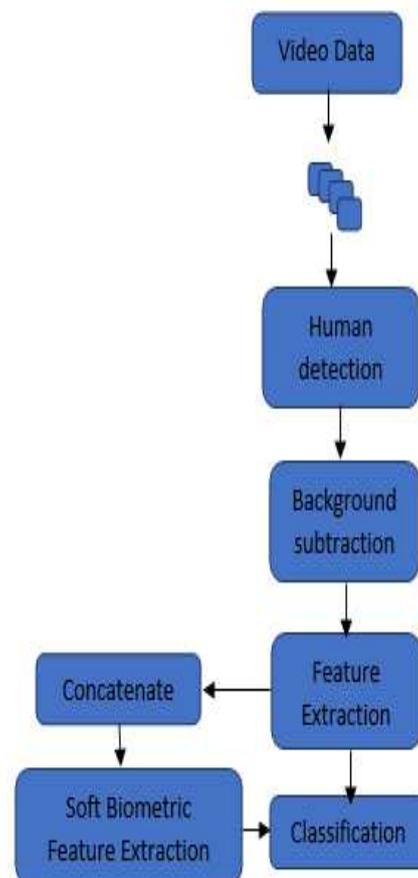
$$b_x = \sigma(I_x) + c_x \quad (1)$$

$$b_y = \sigma(I_y) + c_y \quad (2)$$

$$b_w = p_w e^{I_w} \quad (3)$$

$$b_h = p_h e^{I_h} \quad (4)$$

Bounding box width and height are  $p_w$  and  $p_h$  and the offset at the top left corner of the Image ( $c_x, c_y$ ). Figure 2 illustrates human detection using YOLO.



**Figure 1 Gait Recognition**

### B. Background subtraction

Background subtraction refers to extracting an image foreground from a sequence of images using a statistical model. Object in the background image identified using the absolute difference between consecutive frames, every pixel in the image  $I$  at time  $t$  is compared against the estimated background image  $B$ :

$$T(i, j) = |B(i, j) - I_{(i, j)}^t| < Th \quad (5)$$



**Figure 2 Human detection**

### C. Feature extraction

Traditional features are extracted from the gait entropy images. Soft biometric features are extracted from semantic descriptions of the subjects. The traditional and soft biometric features are concatenated to recognize gait.

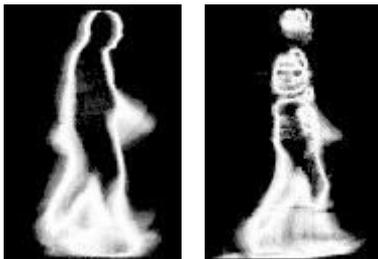
#### *Gait Entropy Image*

Average silhouette image as Gait Energy Image (GEI), Calculated with the given normalized and aligned human walking binary silhouette sequence.

$$G(x, y) = \frac{1}{N} \sum_{t=1}^N B(x, y, t), \quad (6)$$

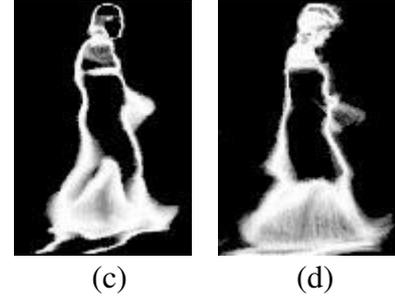
$G(x, y)$  Grey level GEI,

$B(x, y, t)$  Binary Silhouette sequence,



(a)

(b)



(c)

(d)

**Figure 3 Gait Entropy Urban Image (a) LWIR (b) SWIR (c) MWIR (d) Visible.**

$(x, y)$  are values in 2D Image coordinates,  $t$  is the frame number of the sequence,  $N$  is the number of frames in complete cycles of the sequence. Gait entropy image (GEI) is computed from GEI [16]. GEI is computed by calculating Shannon entropy for each pixel in the silhouette images. Figure 3 shows some examples of Gait Entropy Images from our gait dataset. It clearly shows that the dynamic area of the human body, including legs and arms which undergo motions in relation to other body parts, are represented by higher intensity values.

#### *Semantic features*

To allow identification from human descriptions, the physical properties described must be accurate, salient, and reliable. Human descriptions generally consist of two forms of description: categorical labels and continuous estimations. Labels are used to describe inherently categorical traits like ethnicity and gender, can also be used to describe continuous traits. In this work, the soft biometric human description composed of 10 absolute categorical labels and continuous estimations are illustrated in Table 1. The labels have chosen to be universal, distinct, easily discernible at a distance and largely permanent.

The dataset consists of 31 urban subjects with the random five gait images are selected. The number of images per batch is 155 in total. The gait images in one subject are compared to the other two gait subjects.

**Table 1: Traits and corresponding terms**

Trait	Terms
<b>Masculine</b>	[More masculine, Same, Less masculine, Can't decide]
<b>Older</b>	[Older, Same, Younger, Can't decide]
<b>Height</b>	[Taller, Same height, Shorter, Can't decide]
<b>Weight</b>	[Fatter, Same weight, Thinner, Can't decide]
<b>Leg length</b>	[Longer, Same length, Shorter, Can't decide]
<b>Arm length</b>	[Longer, Same length, Shorter, Can't decide]
<b>Neck length</b>	[Longer neck, Same, Shorter neck, Can't decide]
<b>Hunched back</b>	[More, Same, Less, Can't decide]
<b>Head size</b>	[Bigger, Same, Smaller, Can't decide]
<b>Shoe Size</b>	[Larger, Same, Smaller, Can't decide]

For example, image comparison stored in amazon s3 is shown in Figure 4. The gait images are pre-processed by cropping, resized and combined images using the template in Corel Draw Graphic designer software.



**Figure 4 Image comparison (Person A on the Left compared with Person B on the Right)**

The comparison between two subjects was introduced as a more robust method for description. It was then considered to apply to the identification applications. Comparative annotations need to be transformed to convey meaningful subject invariant information. The resulting value is defined as a relative measurement. It can be used as a biometric feature for recognition. In essence, the rating method provides a relative measurement by comparison.

The Elo rating system provides a ranking method based on Thurstone's case for comparative descriptions [15]. The Elo system was initially invented to quantify chess players' skills. Each chess player's capability cannot be measured directly but is usually judged during chess games against other players. Playing chess is much like comparative labels. Relative measurements are made by comparing features, which is the same as comparing the skill of two chess players. The Elo rating system was originally defined as a comparison between two players for chess games. For biometric identification, it is a visual comparison instead. The result of the comparison is the sample that indicates the difference between the two players. The result is used to adjust the player's level.

$$Q_A = 10^{RA/U} \quad (7)$$

$$Q_B = 10^{RB/U} \quad (8)$$

$$E_A = \frac{Q_A}{Q_A + Q_B} \quad (9)$$

$$E_B = \frac{Q_B}{Q_A + Q_B} \quad (10)$$

$$R'_A = R_A + K(S_A - E_A) \quad (11)$$

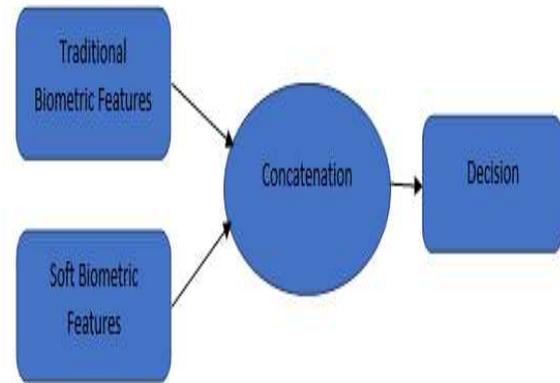
$$R'_B = R_B + K(S_B - E_B) \quad (12)$$

The system uses the result  $S$  of the game to adjust the player's level.  $S_A$  is a compared result between player  $A$  and  $B$ ,  $S_B$  is the inverse of  $S_A$ . Different game results will update a player's level. When  $S$  equals 1, it means winning the game, 0.5 for a draw and 0 for losing.  $E$  is the mathematical expectation

of the game's result, which can be calculated based on the player's level using Eq. (9) and (10). The adjusted difference is controlled by  $K$ , and  $K$  defines the maximum adjusted level value. A constant  $U$  reflects how the player's level impacts the expectation, of which the value is chosen by experience; and set to 400.

### Feature level concatenation

Concatenation is a combination of traditional biometric features and soft biometric features from Elo ranking, which then derives a new score. The soft biometric features are concatenated at the end of the traditional biometric features. The concatenation is to calculate the combined score into a single score. Figure 5 illustrates diagram of a feature level concatenation. The concatenated features are dimensionality reduced using Principal Component Analysis (PCA). The PCA reduce higher dimension features to lower dimensions [17]. The experimental results of traditional and soft biometric are shown in Table 2. The table shows the different combinations of Traditional biometric concatenated with Visible Soft biometrics.



**Figure 5 Feature Level Concatenation**

In Table 2, The experiments are conducted, using K-nearest neighbour (K-NN) [18] with one neighbour, Random Forest (RF) [19] with hundred estimators. In this work for semantic features considered the visible images to recognize, compare easier with other subjects. The recognition accuracy of traditional and soft biometric is higher than traditional biometric. The recognition accuracy of K-NN is higher than the RF classifier.

**Table 2 Results of Traditional and Combined (Traditional+ Visible-Soft) Biometric**

Urban	Accuracy (Traditional)	Accuracy (Traditional+ Visible-Soft biometric)
<b>LWIR</b> K-NN (n=1)	93.1%	95.67%
	Random Forest (n=100)	74.1%
<b>Visible</b> K-NN (n=1)	50%	70%
	Random Forest (n=100)	48.3%
<b>SWIR</b> K-NN (n=1)	78.6%	84.2%
	Random Forest (n=100)	64.3%
<b>MWIR</b> K-NN (n=1)	77.6%	79.41%
	Random Forest (n=100)	77.6%

--	--	--

### III Conclusion

Biometrics plays a major important in video surveillance to identify a person without cooperation, mainly useful for the reidentification of a person required to avoid national security. In traditional biometric, the features are extracted from the Gait entropy image, dimensionally reduced, and classified using the different classifiers. In this work, semantic features considered as soft biometric are concatenated with the traditional biometric to increase the recognition rate. In the future, the semantic features with traditional features increase the person recognition accuracy, reduce false identification.

#### Author's Contribution

Lavanya wrote the manuscript, revised, and edited.

#### Author's information

Lavanya is a research fellow at the University of Southampton, United Kingdom. Her research areas of interests include biometrics, image processing and medical image processing.

#### Funding

Not applicable.

#### Availability of data and materials

Materials used in the manuscript may be requested from the corresponding author.

#### Competing interests

The author declare that she has no competing interests and that there is no conflict of interest regarding the publication of this manuscript.

#### Author details

Lavanya Srinivasan, Research Fellow, School of Electronics and Computer science, University of Southampton, Southampton, United Kingdom, SO17 1BJ.

### References

- [1] T. Surasak, I. Takahiro, C.H. Cheng, C.E. Wang, P.Y. Sheng, "Histogram of oriented gradients for human detection in video," International Conference on Business and Industrial Research, pp. 172-176, 2018.
- [2] Y. Khandhediya, K. Sav, V. Gajjar, "Human detection for night surveillance using adaptive background subtracted image," International Journal of Scientific Research and Development, pp. 185-189, 2017.
- [3] V. Teju, D. Bhavana, "An efficient object detection using OFSA for thermal imaging," The International Journal of Electrical Engineering & Education, 2020.
- [4] I. Riaz, J. Piao, H. Shin, "Human detection by using centrist features for thermal images," International Conference Computer Graphics, Visualization, Computer Vision and Image Processing, Citeseer, 2013.
- [5] D. Zhou, S. Qiu, Y. Song, K. Xia, "A pedestrian extraction algorithm based on single infrared image," Infrared Physics & Technology, Vol. 105, pp. 103-236, 2020.
- [6] Y. Benezeth, B. Emile, H. Laurent, C. Rose-nberger, "A real time human detection system based on far infrared vision," International Conference on Image and Signal Processing, Springer, pp. 76-84, 2008.
- [7] P. Rudol, P. Doherty, "Human body detection and geolocalization for UAV search and rescue missions using color and thermal imagery," IEEE aerospace conference, pp. 1-8, 2008.
- [8] M. Teutsch, T. Muller, M. Huber, J. Beyerer,

“Low resolution person detection with a moving thermal infrared camera by hot spot classification,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 209-216, 2014.

[9] L. Sharma, D.K. Yadav, A. Singh, “Fisher’s linear discriminant ratio-based threshold for moving human detection in thermal video Infrared,” Physics & Technology, vol.78, pp. 118-128, 2016.

[10] A. K. Jain, S. C. Dass K. Nandakumar, “Soft biometric traits for personal recognition systems,” in Biometric Authentication, Springer, pp. 731-738, 2004.

[11] D. Reid, S. Samangoei, C. Chen, M. Nixon, A. Ross, “Soft biometrics for surveillance: an overview,” in Handbook of statistics, vol. 31, pp. 327-352, 2013.

[12] M. S. Nixon, P. L. Correia, K. Nasrollahi, T. B. Moeslund, A. Hadid, M. Tistarelli, “On soft biometrics,” Pattern Recognition Letters, vol. 68, pp. 218-230, 2015.

[13] S. Samangoei, B. Guo, M. S. Nixon, “The use of semantic human description as a soft biometric,” in Biometrics: Theory, Applications and Systems, 2nd IEEE International Conference, pp. 1-7, 2008.

[14] D. A. Reid, M. S. Nixon, S. V. Stevenage, “Soft biometrics; human identification using comparative descriptions,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 36, no. 6, pp. 1216-1228, 2014.

[15] K. Tsukida, M. R. Gupta, “How to analyze paired comparison data,” Technical report UWEETR, University of Washington, College of Engineering, 2011.

[16] S. Lavanya, “Gait Biometric for Person Re-Identification,” Conference on Biomedical Signal, Image and Vision Processing, London, 2021.

[17] Svante Wold, Kim Esbensen, Paul Geladi, “Principal component analysis,” Chemometrics and Intelligent Laboratory Systems, vol.2, Issues 1–3, pp.37-52, 1987.

[18] R.C.Neath, M.S.Johnson, “Discrimination and Classification,” International Encyclopedia of Education (Third Edition), pp. 135-141, 2010.

[19] S. Suthaharan, “Cognitive Computing: Theory and Applications,” Handbook of Statistics, 2016.