International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 3 Issue 5, August 2019 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Gait Recognition for Person Identification using Statistics of SURF

Khaing Zarchi Htun[1], Sai Maung Maung Zaw[2]
[1]Research and Development Image Processing Lab, [2]Faculty of Computer System and Technologies
[1,2]University of Computer Studies, Mandalay (UCSM), Mandalay, Myanmar

How to cite this paper: Khaing Zarchi Htun | Sai Maung Maung Zaw "Gait Recognition for Person Identification using Statistics of SURF" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-3 | Issue-5, August 2019, pp.1415-1422, **https://doi.org/10.31142/ijtsrd26609**

IJTSRD26609

ABSTRACT

In recent years, the use of gait for human identification is a new biometric technology intended to play an increasingly important role in visual surveillance applications. Gait is a less unobtrusive biometric recognition that it identifies people from a distance without any interaction or cooperation with the subject. However, the effects of “covariates factors” such as changes in viewing angles, shoe styles, walking surfaces, carrying conditions, and elapsed time make gait recognition problems more challenging for research. Therefore, discriminative features extraction process from video frame sequences is challenging. This system proposes statistical gait features on Speeded-Up Robust Features (SURF) to represent the biometric gait feature for human identification. This system chooses the most suitable gait features to diminish the effects of “covariate factors” so human identification accuracy is effectiveness. Support Vector Machine (SVM) classifier evaluated the discriminatory ability of gait pattern classification on CASIA-B (Multi-view Gait Dataset).

KEYWORDS: *Speed Up Robust Feature (SURF); Gait Recognition; Statistical Gait Feature; Support Vector Machine (SVM)*

I. INTRODUCTION

Biometrics is a study that automatically identifies people who use unique physical or behavioral characteristics.

An individual's biometric identification distinguishes individuals focused on their physical or behavioral characteristics such as voice, face, gait, fingerprint and iris. [1] Biometrics is becoming more and more important today and is widely accepted because it is unique and will not be lost over time. An individual's biometric identification distinguishes individuals centered on their behavioral and/or physical characteristics for example fingerprint, voice, face, gait and iris. These two biometric technologies are broadly used in forensics, safety, clinical analysis, monitoring and other applications area.

In essence, gait recognition can be separated into two broad groups: model-based methods and model-free. [2] Model-based methods typically simulate the structure and motion of the human body and highlight features to match the components of the model.

Model-free approach focuses on either shape of silhouettes or the whole motion of human bodies. The method without the model focuses on both the shape of the contour and the motion of the entire human body. In this way, the largest connected area in the foreground of the image is considered to be the contour of the human body. [3] It is insensitive to contour quality and has lower computational costs than model methods.

The proposed method is centered on statistical gait features extracted result of Speeded Up Robust Features (SURF) from the binary image, roughness image and gray scale image. Feature extraction method is selected discriminating gait features from three different images to get the high recognition accuracy results for intra-class variation. Evaluating the performance of the proposed system is constructed on the Correct Classification Rate (CCR) of CASIA-B gait database. In this paper, in order to exceed these limitations, we propose a new gait features to identify the human body by changing the conditions of the clothes, carrying condition or varying the angle of view.

II. RELATED WORK

In the earlier period, several gait recognition approaches have been proposed for human identification it can be separated into two broad groups such as model-based and model-free approach. Model-based approach that applicable for human models and uses gait parameters that are updated over time to represent gait. [4] Model-free approach using motion information extracted directly from the silhouette. Recent studies on gait recognition seem to prefer methods model-free, mainly because of better performance than model-based methods, as well as noise immunity and low computational costs.

Johnson and Bobick (2001) proposed a multi-view gait recognition method for gait recognition. Static body parameters consider as the measurements taken from the static gait frames.[5] They use walking action to extract relative body parameters and do not directly evaluate based dynamic gait patterns. The static parameters as height, the distance between head and pelvis, the maximum distance between pelvis and feet, and the distance between the feet. The view invariant is static body parameters which appropriate for recognition.

Lee and Grimson (2002) described the gait silhouettes divide into seven regions. [6] Each region fitted with ellipses and the centroid, aspect ratio of major and minor axis of the ellipse and the orientation of major axis of the ellipse take as gait parameters these extracted as features from each region. From all the silhouettes of a gait cycle extracted gait features these were well organized and were used for gait recognition. F. Tafazzoli and R. Safabakhsh (2010) proposed the shape model divides the subject's body into three regions: the torso, the head, and the extremities to obtain static parameters such as body size, center of gravity coordinates, and gait cycle. [7] The motion model consists of four parts: the head, the torso, the legs and the arms, used to estimate dynamic parameters. The method uses an active contour model to determine the boundaries of each limb. Each limb is modeled as two canes, representing the thighs and together with the tibia at the knee joint and their rotational models form a dynamic walking function. The dynamic Hough transform is used to study the effects of weaponry on gait detection using NNC.

Jasmine Anitha and S. M. Deepa (2014) Video tracking is the process of using a camera to position a moving object (or multiple objects) over time. The algorithm analyzes consecutive video frames as the video is being tracked. [8] They described combine algorithm to improve the tracking efficiency by using SURF descriptor with Harris corner detector. The SURF function descriptor works by reducing the search space of possible points of interest within the pyramid of large scale spatial images. Use the corner detector to locate interesting points in the image. Using the Harris angle algorithm along the SURF function descriptor can improve tracking efficiency.

C. BenAbdelkader et al., (2002) [9] used the model-free method in calculates the gait phase of an object by analyzing the width of the bounding box enclosing the motion contour surrounding the silhouette of subject, and uses a Bayesian classification to confirm the identity of the subject. However, the silhouette width is not suitable for calculating the running time of the front view of a moving object.

L. Wang et al., (2003) used contour unwrapping silhouette centroid to convert a binary silhouette into a 1-dimensional (1D) normalized distance signal. [10] Principal component analysis (PCA) is used to reduce the dimensional of the feature space. Centroid obtained in the eigenspace transformation based on Principal Component Analysis (PCA). To increase the identification accuracy based on the subject's physical parameters.]

Ait O Lishani and Larbi Boubchir (2017) proposed a supervised feature extraction method that selects unique features to identify human gait under carrying and clothing situations, thereby effective recognition performance. [11] The characteristics of Haralick take out from the gait energy image (GEI). The proposed method is based on the Haralick features locally selected from the equal regions of the GEI, using the RELIEF selection algorithm for extracting the object to select only the most important objects with the least redundancy. Proposed method evaluated on the CASIA gait database (dataset B) based on changes in clothing and wearing conditions from different perspectives, and experimental results by the KNN classifier with effective results over 80%.

Therefore, this paper presents more effective feature extraction method to extract distinct statistical gait features created on the outcomes of Speed Up Robust Features (SURF) descriptor. Human silhouette image extracted from the background image by means of frame difference background subtraction technique. SURF features described as basic features by using SURF descriptor from silhouette image. Finally, the propose features are extracted the outcomes of SURF from three different types of image: binary image, roughness image and gray scale image for this identification system.

Finally, ten folds cross-validations are verified on the CASIA-B dataset ten times to get new results. Data set is separated into ten subsets and cross-validation is separated into ten subsets. Each validation period subgroup is nine training set, and the remaining one is a testing set. After executing this validation, it orders exact measurement of the classification accuracy. Support Vector machine (SVM) classifier applied for proposed system as a result of its advanced identification accuracy.

The remainder of this paper: Part III is a review of the proposed system. Section IV provides a detailed description of the gait recognition process and describes the proposed features. The final section describes detail analysis of propose gait-based identification system. The propose feature is very simple, so it can significantly recognize the gait feature for identifying person.

III. THE PROPOSED SYSTEM OVERVIEW

Firstly, the proposed system is detected foreground image or moving object from the background section on the input video. In background subtraction step silhouette image is acquired by subtracting the binary person frame from the binary background frame. It reduces less memory space and execution time. The second step consists of two phases that are the interest point detection and description for three different images. In the first phase, for each interest points, SURF detector detect image and then return collection of interest points. For each interest point, the descriptor calculate feature vector to describe the surround region of each point in second phase. These two phases require re-computing the entire image.

For feature extraction step, this system is proposed discriminating twelve statistical gait features computed from the results SURF descriptor of Binary image (BW), Roughness(R) and Gray (G) image under carrying bag and wearing conditions over eleven various view angles to increase the recognition performance.

These twelve statistical gait features are Mean (BW), Root Mean Square(RmsBW), Skewness (SkBW), Kurtosis (KuBW), Mean(R), Root Mean Square (RmsR), Skewness (SkR), Kurtosis (KuR), Mean (G), Root Mean Square (RmsG), Skewness (SkG), Kurtosis (KuG). Finally the extracted gait signals are comparing with gait signals that are stored in a database. Support Vector Machine (SVM) is suitable to examine the capability of the extracted statistical gait eatures. The flow of the system outline is displayed in Figure.1.

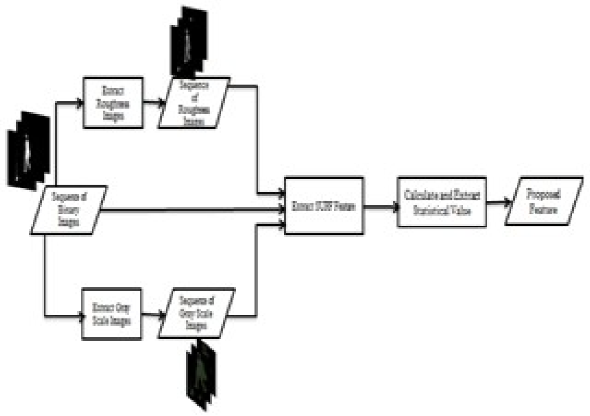


Figure-1: Overview of the Proposed System

IV. PROPOSED SYSTEM

Human identification based on gait recognition has become an attractive research area in computer vision, video surveillance and healthcare system. This system presents model-free approach to extract statistical gait features from the SURF feature descriptor results. The proposed system selects the significant features for human recognition to reduce the influence of intra-class variation in order to increase recognition performance.

A. Moving Object Detection And Silhouette Extraction

The first step of proposed system used frame difference background subtraction method to extract the silhouette images. This method detects and extracts moving object from the background scene for each frame of input video sequence. In this step, original image is converted to binary image using threadsholding level 0.3 for person frame and 0.25 for background frame. Silhouette image is acquired from the binary person frame is subtracted from the binary background frame. The noise and small objects remove from the binary image to get the human silhouette image successively. Background subtraction method easily adapts to the changing background compared with the other methods. This figure shows silhouette images for input video sequence.

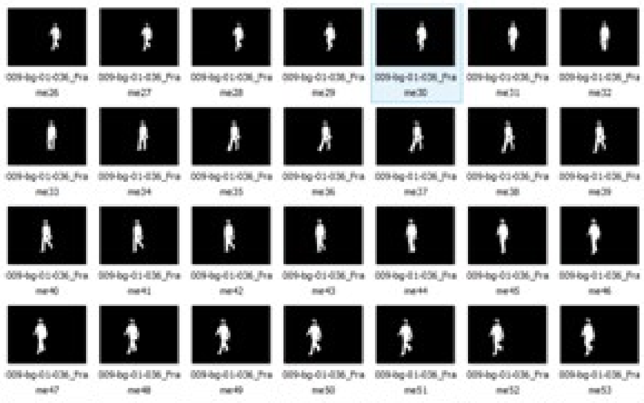


Figure-2: Example of Human Silhouette or Foreground Image

A binary image is a digital image that has only two probable values for every pixel. Generally, the two colors used for binary images are black and white. The color used for objects in the image is the foreground color, and the remaining images are the background colors. Binary image (BW) get from the person binary image subtract from the background binary image.

Waviness is the measurement of the more widely spaced component of roughness image. Roughness measured the intensity difference between the pixels and it used to extract distinct features. Waviness and Roughness image are vertical distribution of pixel values of the image. Waviness image (W) acquires by convolution with the binary image (BW) and Gaussian filter. Roughness image (R) acquires from waviness image subtract from the binary image (BW). RGB image acquires by masking BW image on original image and this RGB image is converted into gray scale image (G)) by forming a weighted sum of the Red (R), Green (G) and Blue(B) components: 0.2989 * R + 0.5870 * G + 0.1140 * B.

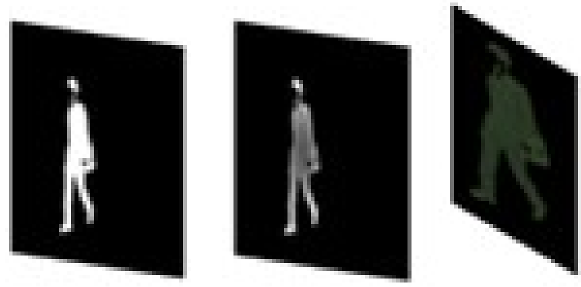


Figure-3: Binary, Roughness and Gray Scale Images

B. Interest Point Detection

SURF algorithm is used to increase the gait recognition system performance. First, the SURF detector is used to find interest feature points in image, and the descriptor retrieves feature vectors for each point of interest. SURF used Integral image technique to overcome the size invariant.

The detection phase uses Hessian-matrix to detect the same points of interest at different scales. [12] Integral image is used to store addition of every pixels intensity value from the input image within rectangular area between point and original image. The formula for integral image is:

$$I_{\Sigma}(X)=\sum_{i=0}^{i\le x}\sum_{j=0}^{j\le y} I(i,j) \qquad (1)$$

After calculating the integral image, only three operations (subtraction or addition) are needed to compute the sum of the pixel intensities on any vertical rectangular region unrestricted of its size. This figure shows the integral image from input image.

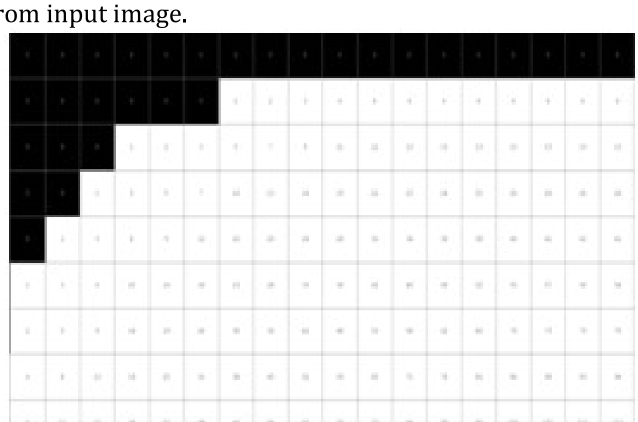


Figure-4: Integral Image for Input Image

As a result of a box filter and an integral image, SURF directly changes the box filter ratio to achieve a relative space. [12]

In image I, a point p = (x, y) is defined by Hessian matrix H (x, σ) in x at scale σ:

$$H(p,\sigma) = \begin{bmatrix} L_{xx}(p,\sigma) & L_{xy}(p,\sigma) \\ L_{xy}(p,\sigma) & L_{yy}(p,\sigma) \end{bmatrix} \quad (2)$$

Where Laplacian of Gaussian L_{xx} (p, σ) is the convolution Gaussian second order derivative $\frac{\partial^2}{\partial x^2} g(\sigma)$ with the image I in point p and also for L_{xy} (p,σ), L_{yy} (p,σ). These interest points are used in human silhouette image.

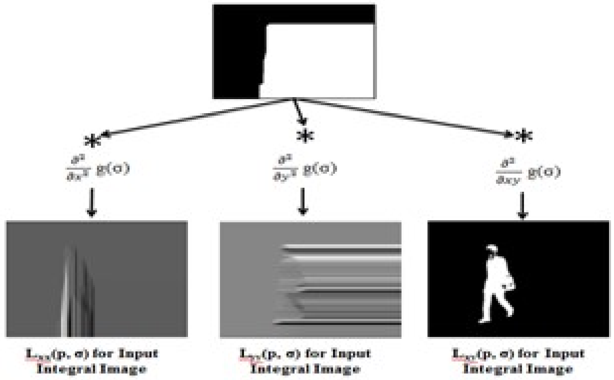


Figure-5: Visual Representation for Hessian matrix of Input Image

Approximate Hessian-matrix is calculated find out distinct interest points of image that is extremely fast because the calculated Hessian matrix is based on the integral image and it reduces computational cost and time. Hessian determinant is used to define the interest point location with this determent maxima value.

$$\det(\mathcal{H}_{approx}) = D_{xx}D_{yy} - (\omega D_{xy})^2 \quad (3)$$

Where D_{xx} is the horizontal response of the second derivative block filter for a given center integral pixel, D_{yy} is vertical filter response, and D_{xy} is the diagonal filter response. Figure-6 shows visual approximated Hessian matrix. [13] An approximation Gaussian box filter size 9×9 with scale σ=1.2 is the bottom level (maximum spatial resolution) for blob-response maps.

$$\mathcal{H}_{\text{approx}}(p,\sigma) = \begin{bmatrix} D_{xx}(p,\sigma) & D_{xy}(p,\sigma) \\ D_{xy}(p,\sigma) & D_{yy}(p,\sigma) \end{bmatrix} \quad (4)$$

Where D_{xx}(p ,σ) , D_{yy} (p ,σ) and D_{xy}(p ,σ) are element of approximated Hessian matrix and these are convolutions of the approximated filters with image I respectively.

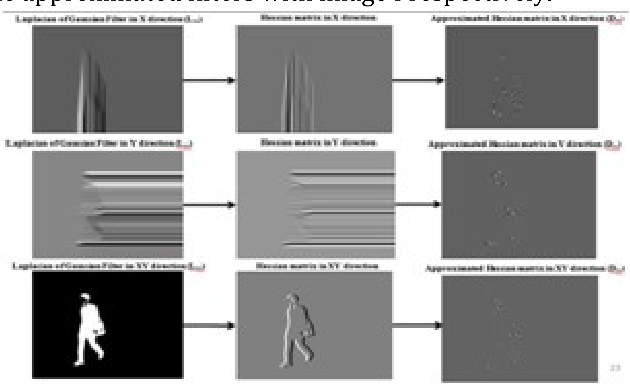


Figure-6: Visual illustration for finding element of the approximated Hessian matrix (D_{xx}(x ,σ) , D_{yy}(x ,σ) and D_{xy}(x ,σ)

The approximate matrix needed to resize a pyramidal scale for matching the interest points across different scales. [14]

Each scale is defined as the image response convolved with a box filters a certain dimension (9x9, 15x15, etc.). Octave denotes a series of response maps or filters of covering a doubling scale.

Octaves	Filter	Scale	Height	Width	Response	Sign of Laplacian
1	9	2	160	120	160x120	1,-1
	15	2	160	120	160x120	1,-1
	21	2	160	120	160x120	1,-1
	27	2	160	120	160x120	1,-1
2	39	4	80	60	80x60	1,-1
	51	4	80	60	80x60	1,-1
3	75	8	40	30	40x30	1,-1
	99	8	40	30	40x30	1,-1
4	147	16	20	15	20x15	1,-1
	195	16	20	15	20x15	1,-1
5	291	32	10	7	10x7	1,-1
	387	32	10	7	10x7	1,-1

Figure-7: Mathematical representation for Response Map of Scale Space Representation

Each response map consists of width, height, box filter scale, filter size, responses and laplacian sign of image. Sign of Laplacian (-1, 1) represented the dissimilarity between dark spots found on a bright background versus bright spots found on a dark background. [15] After creation of response map in different scale space, the next task is to locate the points of interest.

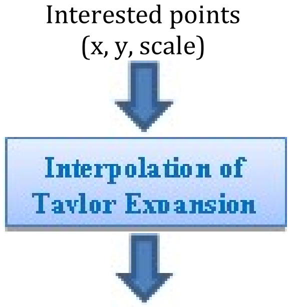
In 3x3x3 region, non-maximum suppression is used to locate and resize the points of interest of image.The process of suppressing a non-maximum consists in finding local maxima within about 8 pixels from itself, its upper and lower response images. If the center pixel has the highest intensity in the search area, it is treated as a local maximum. It then compares the center pixel to a user-defined threshold to exceed the threshold and assumes that the pixel is an interesting point of the local maximum. The method can determine points of interest with x, y (coordinates) and scale only on the second and third layers in each octave.

The interest points are needed to interpolate to get the correct scale and position because the interest points are obtained from the different scale space of images. The interpolation for non-maximum suppression is used to adjust scale and space of these interested points. In essence we have to fit a 3D quadratic expressing the Hessian H(x;y;σ) by using a Taylor expansion for finding extreme by setting the derivative to zero and solving the equation(5) to find $\hat{x}$ = (x;y;σ):

$$H(x) = H + \frac{\partial H^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 H}{\partial x^2} x \quad (5)$$

$$\hat{x} = -\frac{\partial^2 H^{-1}}{\partial x^2} \frac{\partial H}{\partial x} \quad (6)$$

This derivative is calculated from finite differences in the response maps to get correct positions and scale.



No.	X	Y	Scale	Laplacian Sign
1	117.4941	56.5514	4.2667	-1
2	106.1086	201.314	4.2667	-1
3	122.8301	90.8383	8.6667	-1
4	151.8614	136.816	2.2667	-1
5	147.4921	185.259	3.8667	-1
6	147.7223	135.817	4.2667	-1
7	142.4632	152.758	4.1333	1
8	147.5869	191.171	2.9333	-1
9	146.7556	154.577	5.7333	1
10	141.2324	149.213	2.9333	1
----	----	----	----	----
76	79.47125	165.928	10.133	1

Figure-8: Mathematical representation of interest points with Scale and Laplacian for Input Image

Figure-8 shows number of interest points, scale and sign of Laplacian form input human silhouette image.

C. Orientation Assignment

The SURF descriptor describes the pixel intensity distribution in the vicinity of the point of interest. At this stage, the orientation assignment is used to determine the value of the direction for each object (rotation invariance). At the sampling stage, size of Haar wavelets depends on the scale and is set equal to the side length of 4s, and x, y and the scale are displayed on the integral image for fast filtering. These wavelet convolution filters are necessary to calculate six operations based on an integral image in order to get responses in the x and y directions at any scale.

The Haar wavelet responses or gradients are obtained by convolution with a first-order wavelet filter and an integral image in a circular region with radius 6 scales. Apply a Gaussian weighting function to the Haar wavelet response ($\sigma = 2s$) to further emphasize the sample center point. To reduce the effect of distant pixels, multiply the response of the Haar wavelet's result by Gaussian kernel 2s (s= scale).

All of these Gaussian weighted responses are then mapped to the two-dimensional space using the x and y direction responses. [16] The local orientation vector is estimated by computing sum of all responses surrounded by the $\pi / 3$ (or) 60 degree slide orientation window. Create a new object vector by summing the horizontal and vertical responses of all windows. Here, the longest vector (maximum value) is the main direction of the point of interest. Figure-9 and 10 show the process of orientation assignment and value for location of interest point and scale.

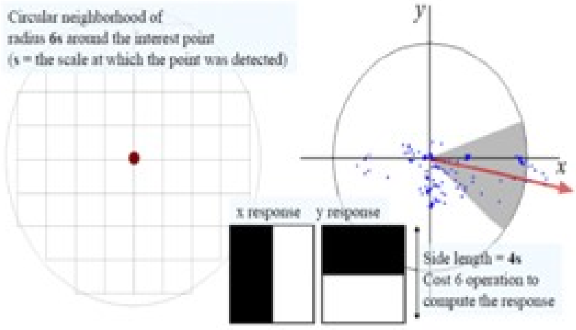


Figure-9: Orientation assignment: The $\pi/3$ sliding window determines the dominant direction of the Gaussian-weighted Haar wavelet response at all points in the circular neighborhood around the point of interest.

No.	X	Y	Scale	Laplacian Sign	Orientation
1	117.4941	56.5514	4.2667	-1	4.3017
2	106.1086	201.314	4.2667	-1	2.9351
3	122.8301	90.8383	8.6667	-1	6.0828
4	151.8614	136.816	2.2667	-1	2.0374
5	147.4921	185.259	3.8667	-1	0.511
6	147.7223	135.817	4.2667	-1	3.1832
7	142.4632	152.758	4.1333	1	3.5773
8	147.5869	191.171	2.9333	-1	2.6117
9	146.7556	154.577	5.7333	1	3.377
10	141.2324	149.213	2.9333	1	3.0045
----	----	----	----	----	----
76	79.47125	165.928	10.133	1	0.0736

Figure-10: Orientation Assignment values for each interest point using x, y, Scale and Laplacian

D. Feature Description

The next step of the proposed system is describing feature vector by calculating the neighborhood of each interest point. The SURF descriptor is focused on the point of interest with a sampling step size of 20s and constitutes a square area aligned with its direction to extract the feature. The area of interest is separated into smaller 4x4 sub-areas. [17] For every sub-region, the Haar wavelet response calculated at 5x5 regular intervals in the rotational direction is a 2s Haar convolution wavelet filter. The Haar wavelet reduces computation time and increases reliability, the size depends on the scale of the function σ. The Haar horizontal and vertical wavelet responses (dx and dy) are multiplied by a Gaussian weight of 3.3 sigma using distance between every pixel in the region and the center point to reduce the effects of geometric deformations and localization errors. The horizontal and vertical directions of the feature path can be alternately rotated.

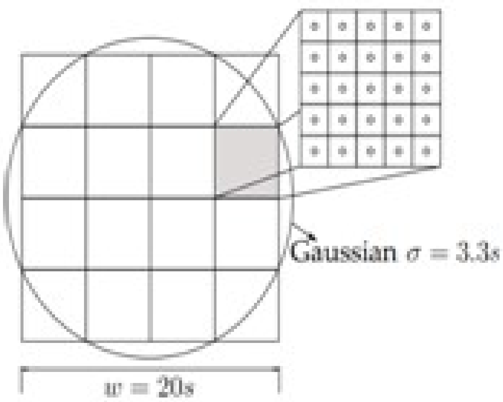


Figure-11: A 20s areas is divided into 4x4 subareas that are sampled 5x5 times to get the wavelet response

In the feature descriptor abstraction, the first step consists of constructing a rectangular area towards the direction determined by the method of centering on the interest point and selecting the direction. [18] This area separated into 4 x 4 small squares. This saves important spatial information. The Haar wavelet response was calculated using 5 × 5 aliquots in each sub-region. Determine the direction of horizontal d_x and vertical d_y. These are the first set of records for each unique vector. Thus, each subfield has a four dimensional vector descriptor for its basic strength structure.

$$v = (\sum dx, \sum dy, \sum |dx|, \sum |dy|) \quad (7)$$

If all 4x4 sub-areas are related, final result is a 64-dimensional vector descriptor. This is usually used in the next similar feature phase. These features distinct because of the number and location of the points selected by the SURF detector and SURF descriptor calculate the objects around these points. Figure 12 shows the wavelet responses computed for each square and figure 13 shows the results of SURF features points from Binary image.

The green square limits any of the 16 sub-areas, and the blue circle indicates the sample point which the wavelet response should be calculated.

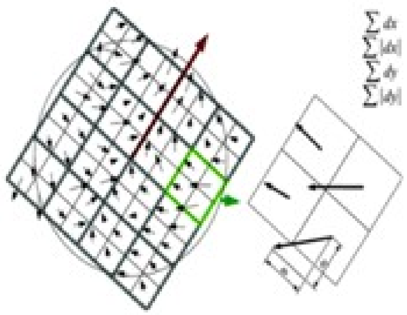


Figure-12: The wavelet responses computed for each square. The 2x2 sub-partitions of each square correspond to the actual descriptor fields. These are the sums dx, |dx|, dy, and |dy| calculated relatively to the orientation grid.

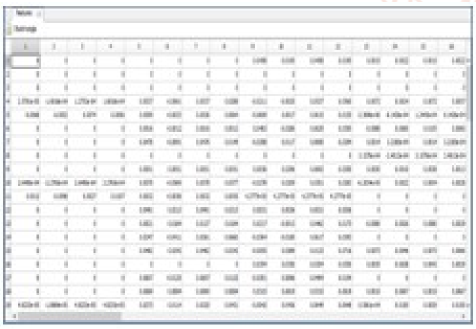


Figure-13: 76 - SURF Features Points from Binary Image

E. Gait Features Extraction

The feature extraction procedure is defined as a collection of features that provide important information efficiently or prominently for analysis and classification. Features points of interest are used when extracting movement parameters from a sequence of gait patterns to show patterns of a person's gait. The feature of point of interest is used when the motion parameters are extracted from a sequence of gait patterns to display the gait pattern of the person.

In this paper, the SURF features are used as the basis features for calculating the individual characteristics of walking. These features extracted from a series of three different type images are binary image, roughness image and gray scale image. For this reason, points of location and the number selected in the SURF detector are different in every individual's image.

This system uses statistical measurement approach to extract gait-specific features to identify people. Statistics is the systematic collection and analysis of numerical data to study the relationships between phenomena and to predict and control their occurrence. There are various statistics such as mean, mode, median, variance, standard deviation, covariance, asymmetry and kurtosis. For the gait feature extraction, twelve statistical gait features are Mean (BW), Root Mean Square(RmsBW), Skewness (SkBW), Kurtosis (KuBW), Mean(R), Root Mean Square (RmsR), Skewness (SkR), Kurtosis (KuR), Mean (G), Root Mean Square (RmsG), Skewness (SkG), Kurtosis (KuG) are used as gait features these features are calculated from the result of SURF descriptor. [19] Statistical measurements identify a value as a representation of the entire distribution. This allows all data to be accurately described using a small number of parameters.

Mean is a basic texture feature that represents the average pixel value of the image. This type of calculation removes random faults and supports to obtain precise result than the result of a single experiment.

$$Mean = \frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij} \quad (8)$$

The standard deviation is the root mean square value of the deviation from the average of the underlying texture.

$$Rms = \sqrt{\frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} |x_{ij} - x_m|^2} \quad (9)$$

Kurtosis and Skewness are called "shape" statistics. That is, they represent the shape of the pixel value distribution.

$$Sk = \left(\sqrt{\frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} |x_{ij} - x_m|^2}\right)^3 \quad (10)$$

$$Ku = \left(\sqrt{\frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} |x_{ij} - x_m|^2}\right)^4 \quad (11)$$

where m, n are number of rows and columns of SURF feature matrix, x_m mean of average feature vector value from SURF feature matrix and x_{ij} mean of feature vector value at i,j coordinate from SURF feature matrix. Figure 14 shows proposed statistical gait features for one person.

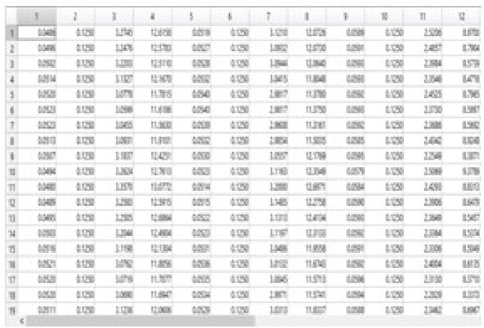


Figure 14: Proposed Statistical Gait Features for One Person

F. Support Vector Machine (SVM)

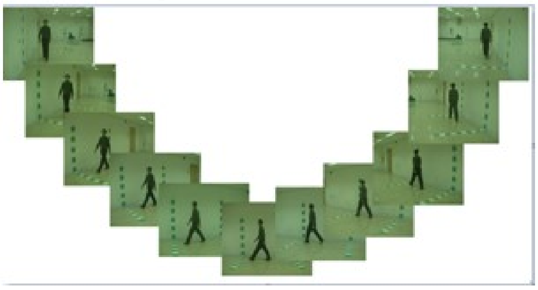
Recently, Support Vector Machines (SVM) has become a powerful classification method in several research fields. This statistical machine learning technique was first introduced by Vapnik in 1995 [19]. This algorithm prevents redefinition by choosing a specific hyperplane from a set of data that can be shared in the feature space. The SVM uses a linear segmentation hyperplane to create a classifier to maximize margins. The width of the field between periods is considered an optimization criterion. The margin is defined as the distance between the nearest points of the classroom data the best hyperplane. Guillon, Boser, and Vapnik show how to produce a nonlinear classifier by kernel functions in the original input space during the 1992 nonlinear separation. First, the SVM converts the original object into a feature space. Various nonlinear mappings can be used to obtain the transform. [20] The kernel function K(x;y) can be selected according to the task. After this conversion, you can easily find the best hyperplane. The achieved hyperplane is the best case for maximum margin.

V. EXPERIMENTAL RESULT AND ANALYSIS

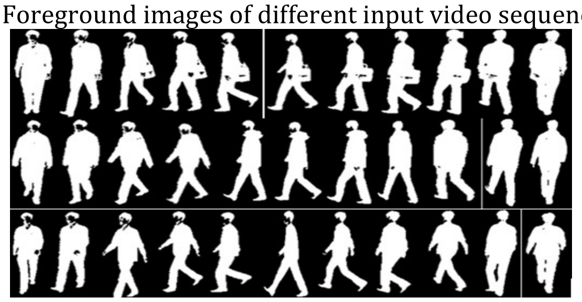
In this paper, CASIA provides multiple views of CASIA-B data. CASIA-B dataset (several walking databases) consists of 124 subjects taken from eleven different view angles from 0 to 180 degrees. Every subject has two dressing sequences, two carrying bag condition and six regular walking sequences. Each frame is shot with a camera by a video resolution of 320 x 240 pixels and a frame rate of 25 frames per second. These videos operate under different lighting conditions with different covariate conditions (carrying luggage, wearing coats, and changing vision), fast walking, normal walking and slow walking speeds from different sides. People have 110 video clips, each containing more than 90 frames. The data set contains 124 people and has 13,640 video0 sequences with a disk size of approximately 17 GB. The proposed method is tested on 11550 video sequences of 105 persons with 11 different view angles. Experimental results display that the proposed approach has the characteristics of high recognition rate and strong robustness.

Ten-fold cross validation is used to classify in separating all features into ten disconnect subgroups. Each disconnects subgroup used for training and testing by performing cross-validation. Cross-validation is used to verify human perception at threshold 10. The proposed system estimation centered on a support vector machine (SVM). This classification provides best classification accuracy for all types of walking in the class.

(a) Sevral videos sequence with eleven view variation



(b) Foreground images of different input video sequences



(c) SURF features points of different foreground images

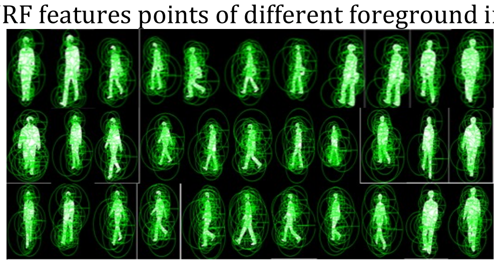


TABLE I. PROPOSED FEATURES AND DESCRIPTIONS

Gait Features	Descriptions
Proposed Statistical Features based SURF	Mean (BW), Root Mean Square(RmsBW), Skewness (SkBW), Kurtosis (KuBW), Mean(R), Root Mean Square (RmsR), Skewness (SkR), Kurtosis (KuR), Mean (G), Root Mean Square (RmsG), Skewness (SkG), Kurtosis (KuG).

Propose features are made using statistical values based on the results of SURF, they are Mean (BW), Root Mean Square(RmsBW), Skewness (SkBW), Kurtosis (KuBW), Mean(R), Root Mean Square (RmsR), Skewness (SkR), Kurtosis (KuR), Mean (G), Root Mean Square (RmsG), Skewness (SkG), Kurtosis (KuG). Proposed statistical gait features and their description shown in Table I.

TABLE II. PERSON IDENTIFICATION RESULTS

Propose Features and its length	Total Number of Video Sequences	Average Identification Accuracy over 10-Fold Cross Validation (%) SVM
Statistic Features from SURF (12)	1100 (10-Persons)	85.6
	5500 (50-Persons)	68.4
	11550 (105-Persons)	62.6

In Table II, The propose features provided to identify people with various intra-class variation (wearing coats, carrying baggage and walking normally). This table shows correct human identification rate of three covariate conditions with different view angles.

TABLE III. CORRECT GAIT CLASSIFICATION RESULTS

Intra-class Variation and Gait Features Length	Total Number of Video Sequences	Average Classification Accuracy over 10-Fold Cross Validation (%) SVM
Carrying Bag (12)	2508 (114-Persons)	50.2
Wearing Coat (12)	2332 (107-Persons)	52.6
ormal Walking (12)	7524 (114-Persons)	82.5

Table III shows propose features classification results for each covariate type (Carrying condition, dress and Normal Walking with different speed) with eleven viewpoints. For two carrying bags, the classification accuracy rate of 114 people is 50.2%. When wearing two coats, 52.6% of the recognition accuracy was confirmed in 2332 video sequences (107 people). Under normal six walking conditions, the video sequence tested at 7524 (114 people) gave 82.5% SVM classification accuracy. In this table, the accuracy of the covariate normal walking SVM classifier is superior to other covariates and other classification methods.

VI. CONCLUSION

This paper describes gait based on human recognition. The propose system is appropriate for monitoring and security areas. This system presents twelve statistical gait features these are tested under three covariance factors with eleven various view angles to obtain advanced discrimination identification accuracy. The proposed feature is modest, the person identification accuracy is appropriate, but it needs to obtain the maximum gait classification accuracy, while considering other important features. Further research will focus on more effective method to extract gait features to measure similarity and effective classifier.

ACKNOWLEDEDGMENT

The authors would like to thanks Institute of Automation, Chinese Academy of Sciences (CASIA) for supporting the datasets.

REFERENCES

[1] Abraham A., Lloret Mauri J., Buford J., Suzuki J., Thampi S M, "Advances in Computing and Communications", Springer Science and Business Media LLC, 2011.

[2] Imad Rida, Somaya Almaadeed, Ahmed Bouridane. "Improved gait recognition based on gait energy images", 26th International Conference on Microelectronics (ICM), 2014.

[3] Qian, Z., Cao, L., Su, W., Wang, T., Yang, "Recent Advances in Computer Science and Information Engineering", Springer Science and Business Media LLC, 2012.

[4] Wang, M. She, S. Nahavandi and A. Kouzani , "A Review of Vision-based Gait Recognition Methods for Human Identification", Digital Image Computing: Techniques and Applications, 2010.

[5] Johnson AY, Bobick AF. A multi -view method for gait recognition using static body parameters. Audio- & Video -Based Biometric Person Authentication (AVBPA 2001); 6 -8 June 2001; Halmstad, Sweden. Berlin: Springer; 2001. pp. 301 -311.

[6] Lee and Grimson. Gait Analysis for Recognition and Classification. Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition, June 2002.

[7] F. Tafazzoli and R. Safabakhsh, "Model-based human gait recognition using leg and arm movements," Eng. Appl. Artif. Intell., vol. 23, pp. 1237-1246, 2010.

[8] J.Jasmine Anitha and S.M.Deepa, "Tracking and Recognition of Objects using SURF Descriptor and Harris Corner Detection", International Journal of Current Engineering and Technology (IJCET),Vol.4, No.2 (April 2014).

[9] C. BenAbdelkader, R. Culter, and L. Davis, "Stride and cadence as a biometric in automatic person identification and verification," in: Proc. IEEE Int. Conf. Autom.Face Gesture Recog., Washington, DC, USA, 2002, pp. 372 - 377.

[10] L. Wang, T. Tan, H. Ning, and W. Hu, "Silhouette analysis-based gait recognition for human identification," IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 12, pp.1505-1518, 2003.

[11] Ait O. Lishani, Larbi Boubchir, Emad Khalifa, Ahmed Bouridane, "Gabor filter bank-based GEI features for human Gait recognition", 2016 39th International Conference on Telecommunications and Signal Processing.

[12] H. Bay, T. Tuytelaars, and L. Van Gool., "Surf: Speeded up robust feature", European Conference on Computer Vision, 1:404 417, 2006.

[13] Md. Ashiqur Rahman, Shamim Hasan, S.M.Rafizul Haque, "Creation of Video Summary with the Extracted Salient Frames using Color Moment, Color Histogram and Speeded up Robust Features", International Journal of Information Technology and Computer Science, 2018.

[14] Surya Prakash. "An Efficient Ear Recognition Technique Invariant To Illumination And Pose", Telecommunication Systems, 09/03/2011.

[15] John H.R. Maunsell, Geoffrey M.Ghose, John A. Assad, Carrie J.Mcadams, Christen Elizabeth Boudreau, Rett D. Noerager, "Visual response latencies of magnocellular and parvocellular LGN neurons in macaque monkeys", Visual Neuroscience, 1999.

[16] Leqing Zhu, "Finger Knuckle Print Recognition Based On SURF Algorithm", 2011 Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), 2011.

[17] M. Ebrahimi and W. W. Mayol-Cueva, "SUSurE: Speeded Up Surround Extrema Feature Detector and Descriptor for Realtime Applications", Computer Vision and Pattern Recognition Workshops, CVPR Workshops, 2009.

[18] D. C.C.Tam, "SURF: Speeded Up Robust Features", CRV Tutorial Day 2010.

[19] Blackledgey, J. M. and Dubovitskiyz, D. A .," Texture Classification Using Fractal Geometry for the Diagnosis of Skin Cancers ", EG UK Theory and Practice of Computer Graphics, 1-8, 2009.

[20] Brieman, L., "Bagging Predictors. Machine Learning", 24,123 -140, 1996. https://doi.org/10.1007/BF00058655