International Journal of Enhanced Research in Science, Technology & Engineering
ISSN: 2319-7463, Vol. 4 Issue 10, October -2015

# Front-view Gait Recognition Approach using View Angle Transformation in Real Surveillance Environments

Jiang Yinjun[1,2], Wang JIanxin[1],Guokehua[1], Heng Gong[3], Heng Wu[4]

[1]School of Information Science and Engineering, Central South University, Changsha, China
[2]Department of computer science, Civil Aviation Flight University of China, Sichuan, China
[3]Hunan Public Security Bureau, Changsha, China
[4]Department of computer science, Texas Tech University, Lubbock, USA

**ABSTRACT**

**During surveillance of real environments, front-view gaits of people walking are more common than lateral-view gaits. The former is also more suitable for recognition and tracking with human eyes. The proposed front-view gait recognition method is based on the computation of the centroids of gait energy images and normalized conversion of perspectives. This method extracts and recognizes gait features under the same view angle after the transformation of view angles. Experiment results show that the front-view gaits of people walking can be identified even in the case of monocular surveillance cameras for which the parameters are unknown.**

**Keywords: gait recognition, gait energy image, front-view gait, view angle transformation.**

## 1. INTRODUCTION

People's gaits are formed by a combination of their walking styles and walking postures. Due to the distinctive characteristics of gaits, which possess advantages such as being non-invasive (not requiring contact), being hard to disguise or misrepresent, and long-distance acquisition [1-2], they have tremendous potential for biological recognition for monitoring and access control applications. This is especially true in the case of criminal investigations, as an individual's gait possesses a constancy over time, video evidence of gaits obtained from the security cameras at the scenes of crimes have become an important form, or even the only form of biological recognition evidence. At present, gait recognition systems have already become one of the tools for evidence collection, whereby evidence obtained from gait analysis has already been used in court prosecutions [3].

When a computer system is used to recognize a walker's gait, as lateral-view gaits (as shown in Figure 1, pictures (a) and (b) are lateral-view gaits while pictures (c) and (d) are front-view gaits) are better than front-view gaits at reflecting the cyclic motion characteristics of human walking, such as amplitude and frequency of gait, phase, speed and other movement parameters, the recognition rates of lateral-view gait are significantly higher than the front-view recognition rates. However, compared with lateral-view gaits, front-view gaits are more effective at reflecting spatial and temporal characteristics such as a pedestrian's body contours, the swinging posture of the two arms, the degree of swaying of the head and shoulders, etc., as well as characteristics of the color distribution of the pedestrian's clothing. These characteristics are more conducive for the recognition and tracking of people with the human eye.


**Fig.1. Lateral-view gait images and front-view gait images**

Therefore, in actual outdoor environments, surveillance cameras are generally installed at a height of 3-5 meters above the ground, with a downward angle of 10 to 15 degrees, pointing in the directions of passing pedestrian traffic. The optical axis and the direction of the pedestrian's movement are kept in approximate alignment, thus obtaining the pedestrian's front-view gait. Based on statistical analysis of 1785 pedestrian videos from 35 surveillance cameras in the CSUPD [14] database, 1481 of the videos show front-view gaits, which is 83% of the total. Therefore, compared with lateral-view gait recognition, research on front-view gait recognition also has important practical significance. Therefore, based on the foundation of Gait Energy Image centroids calculation of pedestrian and the application of standardized transformation to view angle, this article proposes a front-view gait recognition method for application in the actual surveillance environments.

## 2. RELATED WORK

In recent years, considering the spatial and temporal characteristics of pedestrians, two types of recognition methods have primarily been put forward for front-view gaits: recognition methods based on the edge characteristics of front-view gait silhouette image[4-5], and recognition methods based on the region characteristics of front-view gait silhouette images[6-7].

For the edges of gait silhouette images, CSM[4] uses digitalized edge curvature to describe the concavity and convexity characteristics of the outline edges of pedestrians and reflect the lateral swing characteristics of pedestrians' arms, thereby achieving front-view gait recognition. In this method, following the detection of the edges of the pedestrian's outline, one pixel on the edge of the pedestrian's outline is designate as the starting point, and based on the relative positions of the current pixel and its adjacent pixels, the pixel's Freeman vector code value and gradient value are calculated. For the gradient value, a positive gradient value represents that the border of the outline is convex, while a negative gradient value represents concavity of the outline's border. As CSM has a low computation load, it is able to obtain good recognition results under conditions when the border of the moving foreground outline is distinct and the movement direction of the test subjects is fixed. IIFRSD[5] works by extracting characteristics of a series of maximized rectangles covering the person's binary silhouette image to conduct front-view gait recognition. This method divides the maximized rectangle set into three characteristic Bins of width, height and a combination of width and height , then uses Extra-trees to solve higher dimensional problems that occurred during classifications. Classification accuracy is improved through main voting decision rates based in time windows. In real surveillance environments, as the background is complex and varying, it is difficult to to distinctly separate the foreground images, as a result, the edge of the outline image is deformed and filled with noise, which severely impacts the efficiency of recognition.

Unlike methods based on the edge characteristics of silhouette images, recognition methods based on silhouette image region characteristics extract statistical characteristics covering every pixel of the pedestrian's outline. This method extracts the temporal and spatial characteristics of the front-view gait from statistics from multiple gait silhouette images within a gait cycle, and is hence robust against noise and deformation. Jegoon RYU[6] uses a spherical coordinate system $(\rho,\theta,\phi)(0 \le \rho \le R, 0 \le \theta \le \pi, 0 \le \phi \le 2\pi)$ to describe a point cloud of the human body's gait obtained from 3-dimensional surveillance[6], and each MIP formed with one gait cycle is divided into 5 x 12 x 24 bins. The pixel density (point distribution intensity) inside each bin is then calculated, to be used as a recognition characteristic. For application in front-view gait recognition, based on geometric projection principles, the author takes the body's 3-dimensional point cloud and projects it in the direction of the person's forward movement to obtain a 2-dimensional front-view gait point cloud. Kusakunniran et al[7] base their work on a VTM model, using a histogram technique to break Gait Energy Images into 6 regions, and in the lower limb movement regions of the gait images select the region of interest (ROI).

The corresponding relationship of ROI pixel from different view points for the same known object is used to determine the transformation matrix transforming the gait image from one view angle to another view angle. Although these two methods both obtain relatively good recognition results for front-view gait recognition, they both require the gait view angle for the person to be known, whereas in real surveillance environments, a pedestrian's gait view angle always changes. In order to address this issue for the aforementioned methods, for pedestrian's gait in real environments, this article proposes an improved method based on the VTM model[7], whereby the pedestrian's gait image centroid trajectoryis first used to automatically calculate the gait view angle, followed by a view angle transformation to achieve good front-view gait recognition results.

## 3. FRONT-VIEW GAIT RECOGNITION

This article first uses background subtraction and pedestrian detection to detect pedestrians in the surveillance video, and then uses gait energy images to individually calculate the centroids of the pedestrian's head and body. Next it uses segmented line fitting on the movement trajectory formed from the series of silhouette centroids for determining the person's movement view angle and normal view angle transformation. Last, it combines principle component analysis

and the Fisher method to conduct feature extraction and recognition on normalized gait images. The basic process flow is expressed in figure 2:

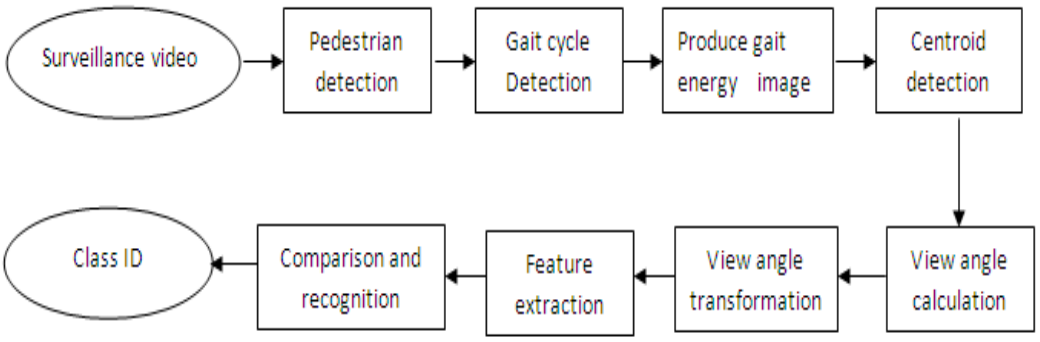



**Fig.2. Front-view gait recognition process flow chart**

For ease of problem description, this article uses the angle $\theta$ between the pedestrian's line of movement and the surveillance camera's optical axis as the definition for view angle (as expressed in figure 3). Gaits with view angles less than or equal to 60 degrees are front-view gaits, while those with view angles greater than 60 degrees are lateral-view gaits. In figure 1, (a) and (b) are lateral-view gait images, whereas (c) and (d) are front-view gait images. The person in (c) is approaching the surveillance camera from afar, which is called a positive front-view gait image. The person in (d) is moving away from the surveillance camera, which is called a back front-view gait image.

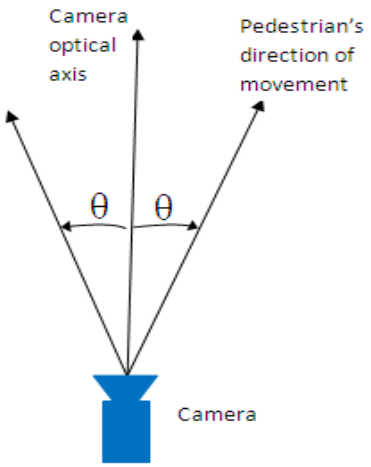



**Fig.3. Gait view angle**

### A. Centroid calculation based on gait energy images

Article[8]determines the pedestrian's direction of movement by calculating the positions of centroids of the head and feet in sequence of frames. This method possesses the following deficiencies: in video frames from real environments, regardless of whether background subtraction, optical flows or other methods are used, it is hard to precisely separate the moving foreground from the background; after separation, the pedestrian's silhouette has a large quantity of blurr, making the outline unclear (as shown in the image in the 5th row on the left side of figure 4), and therefore it is difficult to determine the position of the pedestrian's head, feet and other body parts based on a single image for determining the positions of their corresponding centroids. For the same reasons, the contactless method used by article[9], which use shaar wavelet operators to detect the positions of the knees and ankles, has enormous difficulties in real environments as well.

This article fully utilizes the spatial and temporal characteristics of a pedestrian during walking, and adopts the gait energy image calculation method used in article[10]and calculates the gait energy of a series of binary silhouette images $B_t(x, y)$ within a cycle:

$$G(x,y)=\frac{1}{N}\sum_{t=1}^{N}B_t(x,y) \qquad (1)$$

Here $N$ is the frame count in a complete gait cycle, $t$ is the sequence number of the frame in the cycle, and $x, y$ are the coordinates on the 2-dimensional image plane. Figure 4 shows gait energy images produced according to Eq. 1: The rightmost image shows gait energy images of two different pedestrians, the series of five images to the left are a selection of binary outline images used to produce the gait energy image. In reality, a gait cycle usually includes as many as 20 to 30 silhouette images, not just the five that appear in the figure, the specific number is determined by the camera's shutter speed and the pedestrian's walking speed. Gait energy images can effectively reflect a moving person's outline and the changing condition of the outline during a cycle: the higher the pixel value of a region, the more frequent the movement of the body parts in that region.

**Fig.4. Production process for gait energy images**

Based on gait energy images, a histogram technique is used tosegregate the gait energy images horizontally to separate the head, torso and legs [7]. As shown in Figure 5, the histogram values state the cumulative sum of the pixel values for each row of the gait energy image. After using average values to smooth out the histogram, the horizontal red line located at the first peak from the top (the smallest value) is the dividing line between the head and the torso, and the horizontal green line located at the second maxima (the largest value) is the dividing line between the torso and the legs. From the gait energy image horizontal projection histograms we can see: as under normal conditions, the movements of the legs are forward-backward movements, their left-right movement regions in front-view gait energy images are much smaller than the left-right swings in the upper limb regions . Therefore whether or not the legs are covered up is unlikely to influence the positioning of the green horizontal line in the image relative to the upper limbs.

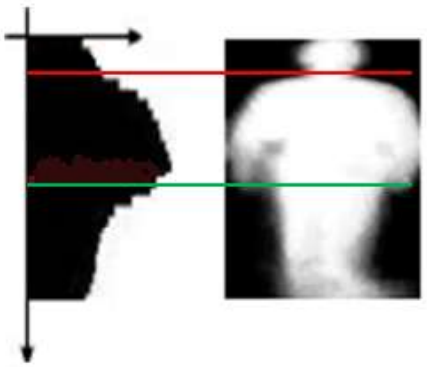


**Fig.5.** Division of the gait energy images

For every separated gait energy image, we apply the centroid calculation method proposed in article [11] for obtaining the position of the centroid for the head, $C^k=(C_x^k,C_y^k)$, and the position for the centroid for the torso, $C^b=(C_x^b,C_y^b)$, and then determine the average of the positions of the two centroids, i.e., the centroid for the person's outline, $S=(S_x,S_y)$, as shown in figure 6. The yellow circle represents the head centroid, the green rhombus represents the torso centroid, and the red triangle represents the centroid of the pedestrian's silhouette. The centroid calculation method is as follows:

$$\begin{cases} M_{pq}^{h} = \sum\sum i^{p} j^{q} G_{t}^{h}(i,j) \\ M_{pq}^{b} = \sum\sum i^{p} j^{q} G_{t}^{b}(i,j) \end{cases} \quad (2)$$

$$\begin{cases} C_{x}^{h} = \dfrac{M_{10}^{h}}{M_{00}^{h}} \\ C_{y}^{h} = \dfrac{M_{01}^{h}}{M_{00}^{h}} \end{cases} \quad (3)$$

$$\begin{cases} C_{x}^{b} = \dfrac{M_{10}^{b}}{M_{00}^{b}} \\ C_{y}^{b} = \dfrac{M_{01}^{b}}{M_{00}^{b}} \end{cases} \quad (4)$$

$$\begin{cases} S_{x} = \dfrac{(C_{x}^{h} + C_{x}^{b})}{2} \\ C_{y}^{b} = \dfrac{(C_{y}^{h} + C_{y}^{b})}{2} \end{cases} \quad (5)$$

Here $G_t^h i, j)$ expresses the pixel value located at (*i*,*j*) on the head of the $t^{th}$ gait image, and $G_t^b i, j)$ expresses the pixel value located at (*i*,*j*) on the torso of the $t^{th}$ gait image.

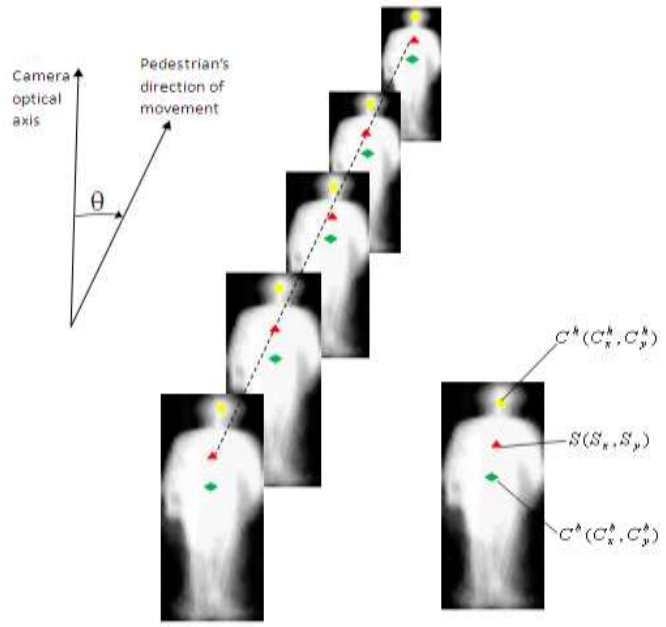



**Fig.6. Centroid position image for a moving person**

### B. Gait image view angle calculation and view angle transformation

People generally follow a straight line when they walk outdoors, however, due to the influence of the direction of the path or obstacles in the path, they often need to change their walking directions. Therefore in surveillance images, the paths of pedestrians are often a crooked line made up of many sections of straight lines . In order to calculate the gait view angle of a pedestrian, we refer to the method described in article[12]whereby we first used the series of silhouette centroids from that person to fit a straight line to every section of the pedestrian's path. Then the angle between each straight line segment and the camera's optical axis is calculated. The key to fitting the straight line segments connected at each end is to find the joint points between each pair of line segments. If the silhouette centroids representing the positions of the

joint points have already been determined, then it is only necessary to link these silhouette centroids together in sequence, thereby obtaining the pedestrian's walking trajectory. The method for determining the joint points is expressed in figure 7: The set of silhouette centroids of person $k$ is defined as $S^k = \{S_i^k\}(i = 1,..,n)$,

Step 1. $L \leftarrow \{1\}$;

Step 2. $b = 1, e = n$;

Step 3.Link $S_b^k$ and $S_e^k$ with a straight line;

Step 4.Calculate the respective distance between points $S_{b+1}^k, S_{b+2}^k, ..., S_{e-1}^k$ and straight line $S_b^k S_e^k$;

Step 5. Determine the point $S_d^k (b < d < e)$ with the greatest distance from the straight line;

Step 6. If distance $D_i$ is greater than the set threshold $T$, then set $L = L \cup \{d\}$,and repeat step 3 with $b = 1.e = d$ and with $b = d, e = e$;

Step 7. $L = L \cup \{n\}$;

$L$ is the set of joint points for the walking trajectory of pedestrian $k$ .The solid blue line in figure 7 represents the walking trajectory after fitting.

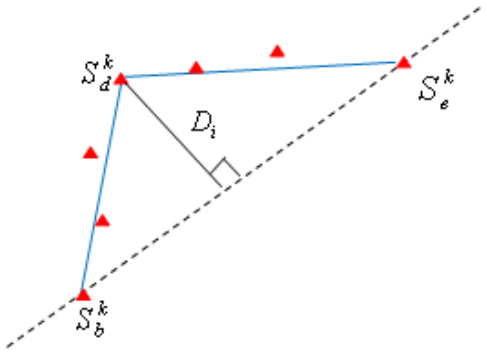



**Fig.7. Linked point determination image**

Using the aforementioned method, the view angle for every straight line segment of the pedestrian's walking path can be calculated. This is the basis for automatically conducting the gait image view angle transformation and rebuilding the gait image sequence. After obtaining the gait energy image view angles, this article uses the VTM model proposed in article[7] for transforming the gait energy image $g_{\theta_i}^k$ of pedestrian k under view angle $\theta_i$ into gait energy image $g_{\theta_j}^k$ under view angle $\theta_j$ :

$$P_{\theta_j}^k \approx f(\theta_i, \theta_j, k, p) = < w, ROI_{p_{\theta i}^k}^{\theta_i} > + b \qquad (6)$$

Here $P_{\theta_j}^k$ expresses the $p$ the pixel of gait energy image $g_{\theta_j}^k$, $ROI_{p_{\theta i}^k}^{\theta_i}$ expresses the region of interest corresponding to pixel $P_{\theta_j}^k$ in $g_{\theta_i}^k$ or multiple related pixels in $g_{\theta_i}^k$. The values of $w$ and $b$ are obtained through training of the sample set $(g_{\theta_i}, g_{\theta_{ji}})$.

### C. Comparison recognition

In order to obtain gait characteristics determined by the differences in the pedestrians' body outline and movement method extracted from the gait energy images and not by the differences in brightness or other factors, we adopt a principle component analysis for reducing the dimensions, followed bythe Fisher method for conducting the feature extraction and target recognition. The Fisher linear recognition method uses the ratios of maximized inter-category distance and intra-category difference to acquire the most optimally differentiated projection matrix. To consider the

similarity of two gait energy images under different view angles, the gait energy images are first transformed into identical view angles, before optimal differentiation matrices are used to carry out projection and calculate the Euclidean distance between the two projected gait energy image vectors. The smaller the distance value, the higher the similarity.

## 4. EXPERIMENTAL RESULTS AND ANALYSIS

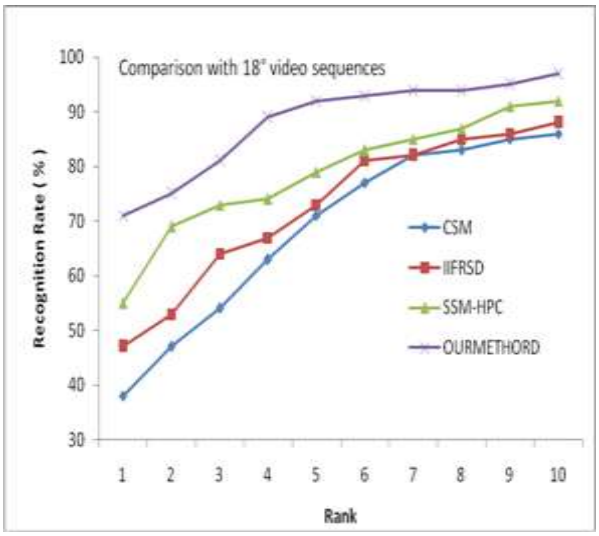
The experiment carried out comparison testing on CSM, IIFRSD, SSM-HPC[6] and the method proposed in this article in two types of databases. The first type of database was the CASIA[13] database, created by the Chinese Academy of Sciences automation. The other database was the CSUPD database created by Central South University's School of Information Science and Engineering. The data in CSUPD were produced based on Hunan province's "Peaceful City" surveillance cameras, primarily containing real surveillance data from areas such as streets, intersections, and public transportation stations. This type of data contains videos of people obtained from cameras with differing parameters, from different view angles, and under different lighting conditions. We carried out experimental sample data selection from the two types of databases based on the following conditions:

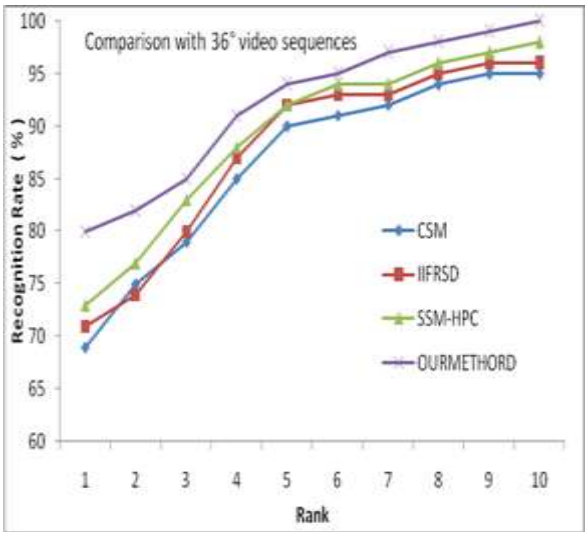
CASIA-B gait database: 124 pedestrian samples were selected, each sample contained 4 front-view or near-front-view view angle video sequences. the 4 view angles were 0°,18°,36°and 54° respectively, each with view angle corresponded to 4 video sequences .

CSUPD gait database: 120 pedestrian samples were selected, with each sample corresponding to 2 video sequences with gait view angle less than 60°. These two sequences werevideo sequences of the same person passing two neighboring surveillance cameras consecutively. Each video sequence lasted 6 to 10 seconds, and included 4 or more gait cycles. The two surveillance cameras are located on the same street at approximately 100 meters apart. The video sequence obtained from the first camera was the target sequence; the video sequence obtained from the second camera was the query sequence. The view angle for each sequence was calculated using the method described previously.

### A. Testing with CASIA

In order to compare the front-view gait recognition ability of each method, for CSM, IIFRSD, SSM-HPC and the method proposed in this article, a pedestrian video sequence with view angle of 36° was used as the query set for searching the target in target sets with view angles of 0°, 18°, 36° and 54° respectively. It can be seen from figure 8that when the view angle of the query set and target set are identical, i.e., when the target set was the 36° video sequence, all of the four methods achieved excellent recognition results. In the top 5 rank the recognition rates of all methods exceeded 90%. However, as the difference between the view angles of the query set and target set increased, i.e. as the target setview angle moved further away from 36°, the recognition rates of the first 3 methods declined with increased speed. When the target set was a video sequence at 18° or 54°, the rates for accurate recognition for the first 3 methods fell below 80%, while the rate for this article's method was above 90%. When the target set was of video sequence at 0°, the rates for accurate recognition for the first 3 methods was below 70%, while the accuracy rate of the method proposed in this article was above 85%. The primary reason for the decline in the accuracy of the recognition rates of the first 3 methods was that when the view angle of the target set and the query setwere not the same, the calculation methods did not carry out view angle transformations on the video sequences before conducting comparison recognition, whereas the proposed method transformed the video sequences from the target set into video sequences with view angles of 36° before carrying out similarity measurement.



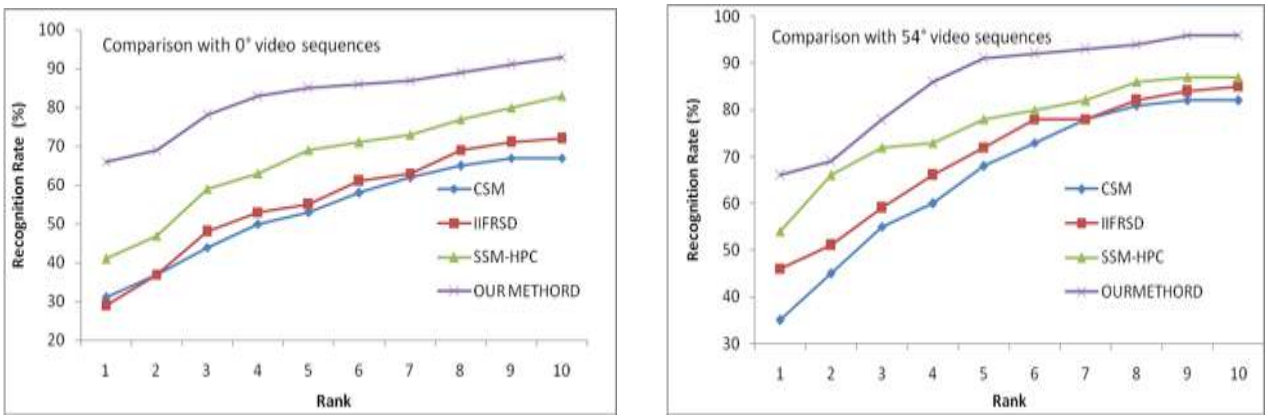
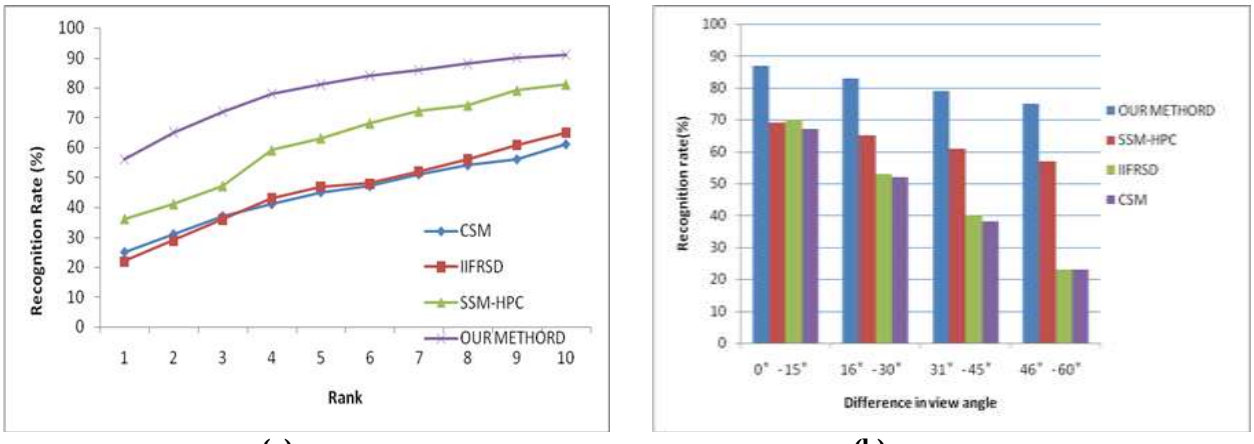



**Fig.8. Comparison of multiple-view angle front-view gait recognition**

## B. Testing with CSUPD

The query set for this test was the video sequences obtained by camera number 2, whereas the video sequences obtained by camera number 1 were used as the target set. The merits and deficiencies of each front-view gait recognition method were measured by comparing the recognition rate of gait images in the target set that corresponded to the gait images in the query set. As it was the same pedestrian who passed through the surveillance range of the two surveillance cameras during a short period of time, there would be no change in clothing and therefore, it was not necessary to consider the influence of clothing on the gait outlines. As in figure 9 (a), the accuracy of recognition rates of targets with the top 5 rank were used as the basis for comparison. The rates for accurate recognition for the first three methods were below 70%, while the rate for the proposed method was 81%. Regardless of which method was used, the recognition rates using CSPD were lower than those with CSUPD.

The reason for this was because the image backgrounds of the video sequences in CSUPD were more complex than those in CASIA; the cameras for CASIA used a level video capture method for taking the video, while CSUPD's cameras used a downward looking capture method for taking the video, such that the pedestrian's outlines were deformed along the vertical axis. In addition to this, among the 4 methods, the recognition rates of CSM[4]and IIFRSD[5] were significantly lower than those of the other two methods. As to the reasons for this, in addition to not having gait view angle transformations, these two methods were based on the recognition on pedestrian's outline border characteristics. As the bodies of pedestrians are 3-dimensional rather than planes, hence the object plane, as well as the content of image planes would differ according to different view angles. The larger the difference in view angle, the greater the differences in the plane outline borders of the object, and the greater the differences in the image plane outline border, the lower the recognition rate, as shown in figure 9 (b).




**(a)** **(b)**

**Fig.9. A comparison of front-view gait recognition using CSUPD**

Based on the experimental results described above, when compared with existing front-view gait recognition methods, the method provided in this article increases the recognition rate significantly. Particularly in real surveillance environments, the method proposed in this article which is based on gait view angle calculation and automatic view angle transformations of Gait Energy Images has maintained a stable front-view gait recognition rate, even in situations when the gait view angle difference exceeded 45°.

## 5. CONCLUSION

As a biological feature that can be obtained covertly at a distance, gait characteristics have aroused sustained research interest. However, most of the existing methods for biological recognition are based on gait characteristics and are only applicable to lateral-view gaits with view angles of between 45° and 90°, or require calibration of camera parameters during recognition, or multiple cameras obtaining silhouette images from multiple angles in order to carry out effective recognition. Recognition systems that are intended for gait recognition in real surveillance environments should not rely on camera parameter calibration or the movement direction of targets; in addition, for surveillance videos, front-view gaits obtained from downward pointing perspectives also need to be considered carefully. This article focuses on the adoption of gait energy image centroids for calculating pedestrian's movement trajectories to determine moving target gait view angles, following which, gaits that are to undergo recognition are normalized into front-view gaits or lateral-view gaits based on these view angles. Testing showed that when the parameters of the cameras are unknown, compared with other known research methods, the method proposed in this article obtained superior recognition results when used on front-view gait images that are particularly difficult for recognition.

Further research should focus on identifying a more optimal model that is not constrained by environmental conditions for gait spatial and temporal characteristics, based on the 2-dimensional properties of monocular surveillance video, and to make gait recognition more attractive as a universally applicable method through rapid and accurate direction calculation and view angle transformations.

## REFERENCES

[1]. K. Bashir, T. Xiang, and S. Gong, “Gait recognition without subject cooperation, ”Pattern Recognition Lett., vol. 31, no. 13, pp. 2052–2060,2010.
[2]. F. Jean, A. B. Albu, and R. Bergevin, “Towards view-invariant gait modeling: Computing view-normalized body part trajectories,” Pattern Recognition, vol. 42,no. 11,pp. 2936–2949,2009.
[3]. Bouchrika, M. Goffredo, J. N. Carter, and M. S. Nixon, “On using gait in forensic biometrics,” J. Forensic Sci., vol. 56,no. 4,pp. 882–889, 2011.
[4]. M. Soriano, A. Araullo, and C. Saloma, “Curve spreads--a biometric from front-view gait video,” Pattern Recognition Lett., vol. 25,no 14,pp.1595-1602, 2004.
[5]. Barnich, M. V. Droogenbroeck, “Frontal-view gait recognition by intra- and inter-frame rectangle size distribution,” Pattern Recognition Lett., Vol. 30, no. 10, pp. 893-901, 2009.
[6]. JegoonRyu, S. Kamata, "Front view gait recognition using Spherical Space Model with Human Point Clouds," Image Processing (ICIP), 2011 18th IEEE International Conference on , pp.3209 - 3212, 2011.
[7]. W. Kusakunniran, Q. Wu , Z. Jian ,and H.D. Li, “Support vector regression for multi-view gait recognition based on local motion feature selection ,” IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2010.
[8]. F. Jean, R. Bergevin, and A. B. Albu, “Body tracking in human walk from monocular video sequences,” in: Proceedings of the Second Canadian Conference on Computer and Robot Vision, Victoria, BC, Canada, pp. 144–151, 2005.
[9]. M. Goffredo, J. N. Carter, and M. S. Nixon “Front-view gait recognition ,”BTAS 2008. 2nd IEEE International Conference ,2008.
[10]. J. Han and B. Bhanu, “Individual recognition using gait energy image,”IEEE Trans. Pattern Anal. Machine Intell., vol.28,no. 2,pp. 316–322, 2006.
[11]. M. K. Hu, “Visual pattern recognition by moment invariants,” IRE Trans. Information Theory, IT(8), pp.179–187, 1962 .
[12]. F. Jean, R. Bergevin, and A. B. Albu, "Computing View-normalized Body Parts Trajectories", In Proceedings of the Fourth IEEE Canadian Conference on Computer and Robot Vision, (Montreal, QC, Canada), pp. 89-96, May 27-30 2007.
[13]. CASIA Gait Database, http://www.cbsr.ia.ac.cn/china/Gait%20Databases%20CH.asp.
[14]. CSUPD pedestrian database（2013）. http://netlab.csu.edu.cn/bioinformatics/web/csupd.htm.