31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Gait Analysis and Validation Using Voxel Data

Fang Wang, *Student Member IEEE,* Erik Stone, Wenqing Dai, Marjorie Skubic, *Member, IEEE* and James Keller, *Fellow, IEEE*

***Abstract*—In this paper, we present a method for extracting gait parameters including walking speed, step time and step length from a three-dimensional voxel reconstruction, which is built from two calibrated camera views. These parameters are validated with a GAITRite Electronic mat and a Vicon motion capture system. Experiments were conducted in which subjects walked across the GAITRite mat at various speeds while the Vicon cameras recorded the motion of reflective markers attached to subjects' shoes, and our two calibrated cameras captured the images. Excellent agreements were found for walking speed. Step time and step length were also found to have good agreement given the limitation of frame rate and voxel resolution.**

I. Introduction

Gait has been studied in recent years as one of the newest topics in computer vision as it offers potential for applications in various areas, including biometric identification[1], medical diagnostic and therapeutic tools. Clinical research has identified a clear link between human gait characteristics and different medical conditions. But gait is rarely used as a medical diagnosis input due to the difficulties in quantifying a person's gait [2].

We are particularly interested in gait analysis for fall risk assessment in elderly people. Recent studies show that there is significant correlation between walking speed and physical function, and it is reliable to use walking speed as a sole surrogate of the assessment of physical function in the elderly [3]. In addition, elderly people's stride rate tends to increase and their stride length decreases, which results in a higher risk of falling. A change in the gait profile over time may also indicate that a person is more at risk of falling. A gait monitoring system could be used as part of an assessment protocol to screen which elderly people are more at risk of falling. Gait training exercises could then be provided, and the effect on their gait could be measured accurately to determine any improvements.

Most of the current gait assessment systems can be categorized into wearable devices, walk-on devices and vision based devices and techniques. Many such systems are expensive and suitable only for lab or clinic settings, such as the GAITRite Electronic mat, and the Vicon motion capture system. Both systems have been shown to give accurate results [4]. There have been no low-cost passive sensor systems available to provide a reliable and quantifiable method of monitoring the gait parameters in daily activities. This paper describes ongoing work towards building such a monitoring system [5].

We have developed a two-camera system to address gait assessment as described above using inexpensive web cameras. Some studies have already been done on silhouette-based human gait analysis using 2D silhouettes [6] [7]. Walking directions have to be limited because the test images need to be taken from roughly the same viewing angle as the training images. Our approach of using 3D voxel data has eliminated the limitation of a controlled walking path, and is thus suitable for daily assessment in the home environment. Some researchers have taken a different approach in assessing balance impairment by detecting sit to stand strategies using web cameras [8].

Section II describes the gait analysis methodology. Section III presents validation experiments, results and discussion. Section IV summarizes.

II. Gait Analysis Methodology

A. 3-D Voxel Reconstruction from Silhouettes

The ultimate goal for this technology is to monitor an elderly person's gait in his normal daily environment; therefore, privacy needs to be addressed. Instead of using raw images, we perform a silhouette extraction, namely, segmenting the human body from the background with the camera at a fixed location, as the initial stage in the analysis. Before silhouette extraction can occur, an accurate background model must be acquired. The background is defined as any non-human, static object. After the background model is initialized, regions in subsequent images with significantly different characteristics from the background are considered as foreground objects. Areas classified as background are also used to update the background model. Fused texture and color features are used for background subtraction [9].

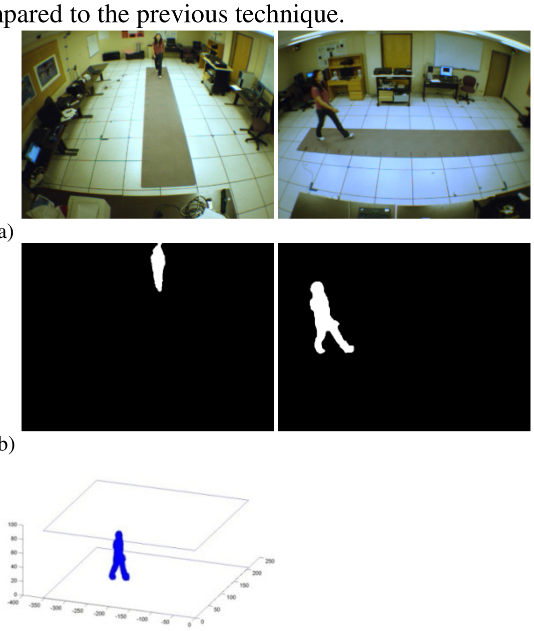
Our three-dimensional human model, called voxel person, described and used in [10], is constructed in voxel (volume element) space by back projecting silhouettes from multiple camera views. A voxel is a three dimensional volume (a non-overlapping cube) resulting from the discretization of the environment. Here, the voxel resolution is 1x1x1 inch.

In previous work [10], the voxel resolution was 5x5x5 inches. With such a coarse resolution, the camera calibration was not a significant issue. Here, the camera calibration was improved to yield more accurate voxel reconstructions. An intrinsic model of each camera is estimated using the Camera Calibration Toolbox from [11]. Given intrinsic

Manuscript received April 7, 2009. This work was supported in part by the U.S. National Science Foundation under Grant IIS-0703692.

F. Wang, E. Stone, W. Dai ,M. Skubic and James Keller are with the Electrical and Computer Engineering Department at the University of Missouri, Columbia, MO 65211 (e-mails: fwnf5@ mizzou.edu, ees6c6@mizzou.edu, wdb43@mizzou.edu, skubicm@missouri.edu, kellerm@missouri.edu).

models of the cameras, for a given camera configuration, the 6 DOF location (position and orientation) of each camera in the environment is computed independently using correspondences between a set of 5 or more measured 3D locations in the environment and pixels in the camera image. The 6 DOF location of each camera is then optimized such that the pixel projection error of the set of 3D points is minimized. Given the optimized location, along with the intrinsic model, the calibrated view vector of each pixel in each camera can be determined for the purpose of silhouette back projection. This calibration procedure results in significantly more accurate voxel person reconstructions as compared to the previous technique.



(a)

(b)

(c)

Fig. 1. (a) Two raw camera images monitoring the same scene. (b) Human silhouettes (c) The reconstructed three-dimensional voxel person.

B. Extraction of Gait Parameters

Gait Parameters need to be extracted for both the web camera and Vicon systems and compared with the GAITRite results. The extraction method for the Vicon system is presented in Section II .C.

In order to get an accurate comparison, the GAITRite system and its measurement definitions are reviewed as follows. The GAITRite system [4] used in the experiment is an electronic walkway with an effective length of 16 feet. The walkway comprises a series of sensor pads that are inserted in a grid. The sensors are placed 1.27 cm apart (total of 18,432 sensors) and are activated by mechanical pressure. Footfall data from the activated sensor is collected by a series of on-board processors and transferred to the computer through a serial port. The sampling rate of the system is 120Hz. The gait parameter definitions used in the GAITRite system are listed below.

Walking speed: Distance traveled divided by the ambulation time. The ambulation time is defined as the time elapsed between first contact of the first and the last footfalls.

Step time: Time elapsed from first contact of one foot to the first contact of the opposite foot.

Step length: The step length of the right foot is defined as the distance between the center of the left foot to the center of the right foot along the line of progression.

In this paper, unless specified, the step time and step length are the averages of the right and left feet.

The gait parameters are extracted for the web camera system based on the definitions above in order to compare with results obtained from the GAITRite. The extraction methodology for the camera system, however, does not depend on footfalls, but is approximated using the voxel person centroid and distance between the two feet.

i) Walking speed

The voxel person centroid $\vec{C}$ is the average of all the voxels locations $\vec{V}$:

$$\vec{C} = average(\vec{V}) \quad (1)$$

The centroid represents the 3D location of the person at a given time. The distance D a person traveled in 2D space is approximated by adding up the distance the centroid location moved at each frame. Walking speed V, therefore, can be calculated based on the distance traveled divided by time calculated from the frame rate f (5fps) and the number of frames fr_n.

$$V = D * f / fr\_n \quad (2)$$

ii) Step Time

The voxels with a height below 4 inches from the ground plane are used to capture foot motion. They are projected onto the 2D space as shown in Figure 2. The solid line represents the length from the front of one foot to the end of the other foot. It is projected along the walking direction. The walking direction is obtained from the centriod in consecutive frames. As illustrated in Figure 2, this length alternatively expands (shown as peaks) and contracts (shown as valleys) over time as the person's feet spread and close during the gait cycle. The number of steps $step\_n$ is obtained directly from the number of peaks representing the number of gait cycles (Figure 2c). The average step time T is then calculated as the total time divided by the number of steps.

$$T = (fr\_n/f)/step\_n \quad (3)$$

iii) Step Length

The average step length L is then calculated as:

$$L = D/(step\_n) = V * T \quad (4)$$

C. Capturing Vicon Gait Parameters for Validation

The three-dimensional motion analysis system, Vicon MX, allows for very accurate measurement of movement, using reflective markers and 7 cameras simultaneously. The cameras send out infrared light signals and detect the reflection from the markers attached to each subject's shoes. Based on the angle and time delay between the original and reflected signals, it tracks the movement trajectories of the retro-reflective markers in 3D space.

The trajectory of the markers is projected onto the 2-D floor, as show in Figure 3. The key in analyzing the Vicon data is the accurate detection of the contact of the footfall by calculating the markers location $\vec{M}(t)$ at time t movement:

$$|\vec{M}(t) - \vec{M}(t-1)| < threshold \qquad (5)$$

The threshold is selected to be 1-4 mm in each of the 2D directions depending on the person's gait style.

Once the footfalls are identified, the gait parameters can be extracted based on the GAITRite definitions given in Section II.B. Figure 3 shows an example of the foot marker trajectory, with x marking the footfall locations.

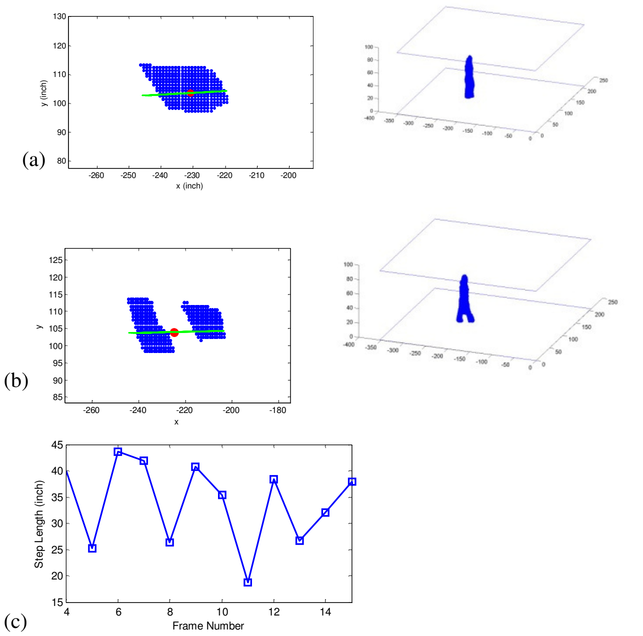


Fig. 2: (a) A projection of a valley frame, i.e., feet are together (b) A projection of a peak frame, i.e., feet are separated (c) Step length variation at different frames.

III. Validation Experiments

A. Experimental Setup

Two stationary cameras were placed in approximately orthogonal locations to record images while the subjects walk across the GAITRite mat. For the experiments, we used Unibrain Fire-i Digital Cameras. The images are recorded at a frame rate of 5 frames per second, with an image resolution of 640x480 pixels.

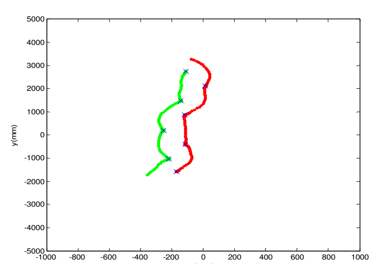


Fig. 3. The projected 2D trajectory of the Vicon markers; x marks the footfall locations.

Thirteen subjects participated in the test. Each subject was tested multiple times for accuracy, as well as various walking patterns, including normal speed, fast, slow, and limping. In total, there are 44 tests. Due to the camera viewing angle limitation, at the camera boundary regions where the subject enters or leaves the scene, the voxel person models are noisy. These frames are eliminated from the data analysis to reduce the errors. Subjects normally take about 8-10 steps to complete the walkway. But when different walking patterns are utilized, such as fast and slow walks, the steps could range from 6 to 16.

B. Results

The test results are shown in Figure 4, and Tables I, II, and III list the comparison differences among the web camera, GAITRite and Vicon systems for walking speed, step time, step length respectively.

The walking speeds obtained from the tests cover a large variation, ranging from around 40cm/s to 180 cm/s. An excellent agreement among the three systems is achieved. The standard deviation as seen in Table I is around 2, which implies that most of the difference would fall within +/- 6% (3sigma) when comparing the system to each other. As for the test with the largest difference, the difference of the web camera vs. the GAITRite is 7.9%, and the Vicon vs. the GAITRite is 6.4%. However, the web camera has a good match with the Vicon with a difference of 1.4%. One reason that might cause the larger difference is the very low walking speed for this particular test (GAITRite 32.9 cm/s, Vicon 35.0 cm/s, and web camera 35.5 cm/s). Although the absolute difference in walking speed is small, the percentage difference is quite large.

Table II lists the results for step time. The camera frame rate (5fps) needs to be considered when discussing step time. Based on Nyquist-Shannon's theorem, the Nyquist sampling rate needs to be twice the highest signal frequency, which would be translated to 0.4s step time for our camera system. Our results have shown all the subjects' step time is longer than 0.4s (age ranging from early 20's to early 60's). Considering the technology application for elderly people, the current frame rate is adequate to capture the step time accurately. It has been observed that the tests with large differences are the ones with a short step time, fast walking speed, and large step length. In the tests with the largest difference shown in Table II, the subject takes only 6 steps compared to the 8-10 steps normally needed to complete the walkway. In this case, for the same walkway length, there are fewer steps available to analyze. The step time obtained for this test from the GAITRite is 0.43s, 0.44s for Vicon, and 0.40s for the web cameras. The average step time for all subjects tested is around 0.6s. So we believe the large difference is due to a fewer number of steps as well as a short step time.

The average step length for GAITRite is obtained through the average of the step length reading of the right and left feet. And the average step length for both the Vicon and the web camera system is calculated using walking speed multiplied times the step time. Therefore, the average step length depends on the two gait parameters discussed earlier. It is not surprising to find that the step length difference is very similar to the step time difference, which implies that the difference in step length mainly is caused by step time difference among the three systems compared.

Overall, all the gait parameters extracted from the web camera system showed good agreement compared to GAITRite and Vicon. The largest difference occurs only at extreme conditions, such as very low walking speed, or an extremely fast step cycle. Still for the very low walking speed, the web camera and Vicon match well. As for the errors caused by short step time, it can be resolved by having a longer walk distance. Here, it is not a big concern as we are focusing on an elderly population.

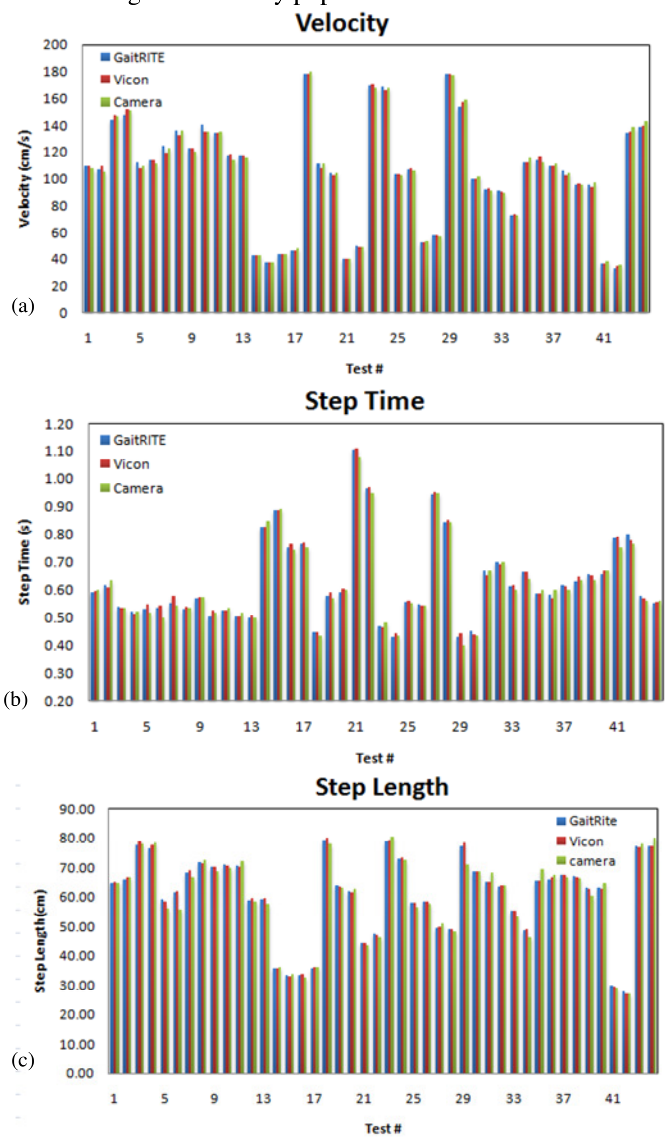


Fig. 4. (a)Velocity (b) Step Time (c) Step Length results for camera, GAITRite and Vicon systems

Table I. Walking speed

	avg diff (%)	max diff (%)	stdev
Camera vs. GAITRite	0.18	7.92	2.32
Camera vs. Vicon	0.29	-3.91	2.07
Vicon vs. GAITRite	-0.09	6.45	1.99

Table II. Step Time

	avg diff (%)	max diff (%)	stdev
Camera vs. GAITRite	-0.63	-6.98	2.46
Camera vs. Vicon	-1.08	-9.09	2.95
Vicon vs. GAITRite	0.48	4.85	1.65

Table III. Step Length

	avg diff (%)	max diff (%)	stdev
Camera vs. GAITRite	-0.71	-9.47	3.04
Camera vs. Vicon	-0.81	-9.91	3.10
Vicon vs. GAITRite	0.10	-2.27	0.88

IV. Summary

We have developed a simple and reliable algorithm to extract gait parameters using low cost web cameras. The results have been validated with GAITRite and Vicon systems. Very good agreement was achieved for walking speed, step time and step length. Using 3D voxel data, the limitation of walk path direction is eliminated, which makes the technology more suitable for gait assessment in a normal daily living environment. Our future work will address asymmetric and shuffle gait patterns.

Acknowledgment

The author would like to acknowledge the members in the Eldertech team at the University of Missouri for their participation in the tests. Special thanks to Dr. Marilyn Rantz, Carmen Abbott and Tanvi Banerjee.

References

[1] M. S. Nixon, T. N. Tan, and R. Chellappa, "Human Identification based on Gait", International Series on Biometrics Springer, 2005.

[2] D. Hodgins, "The Importance of Measuring Human Gait," Medical Device Technology. 2008 Sep;19(5):42, 44-7.

[3] J. Yu, "The Significance of Walking speed in Physical Function among a Group of Community Dwelling Older Adults", Ph.D thesis, University of Missouri-Columbia. 2008.

[4] K. E. Webster, J. E. Wittwer, Julian F. Feller, "Validity of GAITRite Walkway System for the Measurement of Averaged and Individual Step Parameters of Gait", Gait and Posture 22, pp 317-321, 2005.

[5] M. Skubic, G. Alexander, M. Popescu, M. Rantz, and J. Keller, "A Smart Home Application to Eldercare: Current Status and Lessons Learned," *Technology and Health Care*, in press.

[6] R. T. Collins, R. Gross and J. Shi, "Silhouette-based Human Identification from Body Shape and Gait", Proc. IEEE Intl Conf on Automatic Face and Gesture Recognition, 2002.

[7] L. Wang, T. Tan, H. Ning and W. Hu, "Silhouette Analysis-Based Gait Recognition for Human Identification", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 25, no. 12, pp1505-1518

[8] S. Allin, and A. Mihailidis, "Sit to Stand Detection and Analysis", AAAI in Eldercare: New Solutions to Old Problems, Washington DC, November 7-9, 2008.

[9] R. H. Luke, "Moving Object Segmentation from Video Using Fused Color and Texture Features in Indoor Enviroments", technical report, CIRL, University of Missouri, 2008.

[10] D. Anderson, R. H. Luke, J. M. Keller, M. Skubic, M. Rantz and M. Aud , "Linguistic Summarization of Video for Fall Detection Using Voxel Person and Fuzzy Logic," Computer Vision and Image Understanding, 113, pp 80-89, 2008,

[11] Jean-Yves Bouguet, "Camera Calibration Toolbox for Matlab," Intel Corp, 2004, http://www.vision.caltech.edu/bouguetj/calib_doc/.