



A gait skeleton model extraction method based on the fusion between vision and tactility

Shijie Dai^{1,2} · Rui Wang^{1,2} · Huibo Zhang^{1,2}

Published online: 29 September 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

In the traditional gait skeleton model, the fixed length proportions among different bones cause the loss of model personality, and the shoulder and hip points that are blocked by the body are difficult to be extracted. For those problems, a method based on the spatial and temporal fusion between vision and tactility is proposed, by which a equal proportion 3D gait skeleton model can be restructured in the camera coordinate system accurately through a single-frame image. In the geometric analysis process, some logical assumptions are proposed according to human anatomy and the laws of human movement, and an effective method is proposed for the extraction of thigh slope. The experimental result shows that the consistent equal proportion models, in which single bone length error is eventually controlled within ± 5 mm, can be extracted in different position under the premise of rapidity with the aid of this method. The extracted shoulder and hip points also meet the real human body skeleton structure, which lays the foundation for the integrity and rationality of the entire skeletal model.

Keywords 3D gait skeleton model · Spatial and temporal fusion · Shoulder and hip points · Equal proportion restructure

1 Introduction

Gait recognition is to identify individuals based on their walking posture, medical studies show that human gait has differences, and each person has their own unique walking style [1]. Compared to other biometrics (fingerprints, faces, irises, etc.), gait recognition has the advantages of being non-contact, noninvasive, easy to perceive, difficult to hide and difficult to camouflage. The existing methods can be divided into three categories: structural characterization, non-structural characterization and fusion characterization [2]. Structural characterization method, also known as model-based method, is one of the important directions of gait recognition. Although the accuracy is slightly lower than the other two methods, it has great advantages in rapid recognition.

Zhang [3] created a five-link bipedal human model, which is used together with the human body's height vector and the base height vector to characterize gait. Lu [4] proposed a body layered deformable model that divides the body into several blocks and calculated 22 parameters including body shape and dynamic features as the gait features. Lin [5] performed gait recognition based on the extracted angle and other related parameters in the 2D contour image. These are the early exploration of the human body model for gait recognition. Feature extraction and recognition are performed by the kinematic data of some bones, but the feature is less, and the importance of bone length is neglected, which makes the recognition error larger.

Hamzaçebi [6] proposed a new MD-SLIP template based on the general spring-loaded SLIP template for leg movements, which enhances the stability of gait tracking. Wang [7] established a skeletal model from the joint point coordinates automatically acquired by Kinect which can obtain infrared images. Wang [8] proposed a new method based on global optimization for searching these correspondences, which performs favorably methods in highly cluttered backgrounds. Zhang [9] introduced the contourlet transform into the gradient vector flow (GVF) Snake model to effectively extract edge features. Zhang [10] created an initial 3D model using the example-oriented radial basis function model that maps

✉ Huibo Zhang
zhanghb@hebut.edu.cn

¹ School of Mechanical Engineering, Hebei University of Technology, Tianjin, China

² Hebei Key Laboratory of Robot Perception and Human–Robot Interaction, Hebei University of Technology, Tianjin, China

the set of 30 measurements to the body shape space and is established based on the given examples. Kusakunniran and Wu [11] utilized procrustes shape analysis and related similarity measure methods for identification problems caused by speed changes. These studies reflect the process of making the human body model more accurate in identifying and tracking, where the researchers overcome the influences of walking speed, complex backgrounds, etc. However, the selection of the shoulder and hip joints is not accurate enough for overusing the anatomical bone proportion and the results of body model pattern recognition, and the application in gait recognition is still insufficient.

Particularly, people have done a lot of research and made great progress in the analysis of human body models based on HumanEva datasets. Sigal [12] described a baseline algorithm for 3D articulated tracking, which help establish the current state of the art in human pose estimation and tracking. Poppe [13] presented an example-based approach to pose recovery, using histograms of oriented gradients as image descriptors to test on the HumanEva-I and HumanEva-II datasets, which provides the 3D error of human joints. Zhang [14] extended the Gaussian process latent variable model (GPLVM) for JGPM learning, where two heuristic topological priors, a torus and a cylinder, are considered and several JGPMs of different degrees of freedom (DoFs) are introduced for comparative analysis to estimate 3D gait kinematics. Jahangiri [15] proposed a method to generate multiple hypotheses for human 3D pose all of them consistent with the 2D detection of joints in a monocular RGB image. These studies provide tracking and prediction of human 3D models and analysis of joints errors, but lack of spatial position calculations and extension to the application for gait recognition.

Establishing a model with high-precision bone length and angle is the purposes and key to the model-based method, which plays a decisive role in the subsequent feature extraction and recognition process [3–19]. From the above research, we can find that: (1) The traditional method of skeletal model applied to gait recognition lacks the application of bone length and the reasonable determination of hip and shoulder joints. (2) In human pattern recognition, the extraction of certain joints (shoulders, wrists, knees and ankles) has reached a certain degree of precision. However, there is a lack of an accurate transform from the planar position to the spatial position and an extension to gait recognition.

In traditional spatial positioning only by vision, at least two cameras are required. Each camera produces a calibration error after performing spatial calibration. When calculating the spatial position, these two errors are not a simple superposition, but a random coupling error which is difficult to analyze and will be larger if the sampling moments of the two cameras are not uniform. Therefore, tactility is

introduced to replace a vision in this paper. the positioning error in the tactility is almost negligible due to the force particularity of the second phalangeal end, so there is only one camera calibration error, and it is easy to analyze and reduce. The shape of tactile sensor is a square with a side length of 8.4 mm, and when used for spatial positioning, it can be ensured that the positioning error of Z-axis coordinates is less than 4.2 mm at any distance.

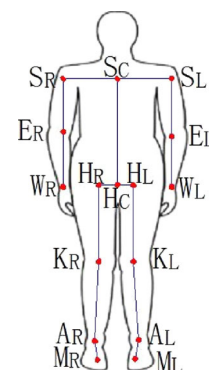
Out of these considerations, a method for extracting a gait model based on the spatial and temporal fusion between vision and tactility is proposed to rapidly establish a dedicated 3D gait skeleton model for individuals, which is based on the accurate plane coordinates obtained by human body model pattern recognition and the equal length of the bones on the left and right sides in human anatomy. Moreover, the use of tactility lays the foundation for the study of gait recognition through dynamic response of the ground support to human motion.

2 Theory and extraction method of 3D gait skeleton model

The human skeleton is made up of many rigid bones connected by joints. The gait skeleton model presented in this paper (see Fig. 1) simplifies these rigid bones into line segments with joint points as their endpoints. The model mainly considers joints which have relative movement to the center of gravity during walking (*S*: shoulder, *E*: elbow, *W*: wrist, *H*: hip, *K*: knee, *A*: ankle, *M*: phalangeal end). The mid-points of the hips and shoulders (*S_C*: shoulders' midpoint, *H_C*: hips' midpoint) were introduced for simplifying the model, and reasonable assumptions based on human anatomy were proposed:

1. left body (*L*: left body) and right body (*R*: right body) are symmetrical to each other;
2. the same kinds of bones in the left body and right body are equal in length.

Fig. 1 The position schematic of the joint points selected in the gait skeleton model



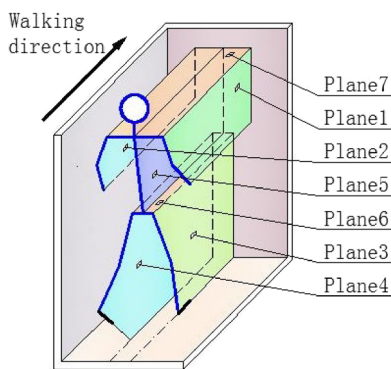


Fig. 2 The plane formed by the swing of each bone of the gait skeleton model in the actual movement, where all the planes are parallel to the walking direction

Each joint has multiple degrees of freedom which causes the complexity of human behavior [4]. In the walking process, these joints generally only have a large variation in one degree of freedom, but have little variation in others. Therefore, the following reasonable accommodations were proposed:

1. the oscillating planes of the bones represented by $K - H$, $K - A$, $E - S$, $E - W$ and $S_C - H_C$ are parallel to the walking direction and perpendicular to the walking plane;
2. the skeletons represented by $H_R - H_L$ and $S_R - S_L$ are parallel to the walking plane.

According to the different spatial positions of different skeletal joints during walking, as shown in Fig. 2, the joints of this gait skeleton model can be assigned to seven planes (except the joint M):

Plane 1: The oscillating plane of the bones represented by $S_R - E_R - W_R$; Plane 2: The oscillating plane of the bones represented by $S_L - E_L - W_L$; Plane 3: The oscillating plane of the bones represented by $H_R - K_R - S_R$; Plane 4: The oscillating plane of the bones represented by $H_L - K_L - S_L$; Plane 5: The symmetry plane of Planes 1–4, including S_C and H_C ; Plane 6: The plane of rotation of the bones represented by $H_R - H_L$; Plane 7: The plane of rotation of the skeleton represented by $S_R - S_L$.

2.1 The spatial and temporal fusion

In the experimental process, visual data and tactility data are measured in different ways. It is difficult to unify the sampling frequency and corresponding start–stop time periods of both parties. In this paper, both data will be unified in time though the coupling relationship between the plantar pressure and the distance of the human feet. After the time is unified, the gait skeleton model will be extracted by spatial fusion.

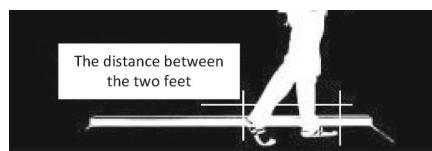


Fig. 3 Schematic of the distance between the two feet during the time fusion

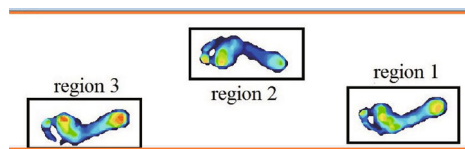


Fig. 4 Definition of pressure region on the plantar pressure integration map

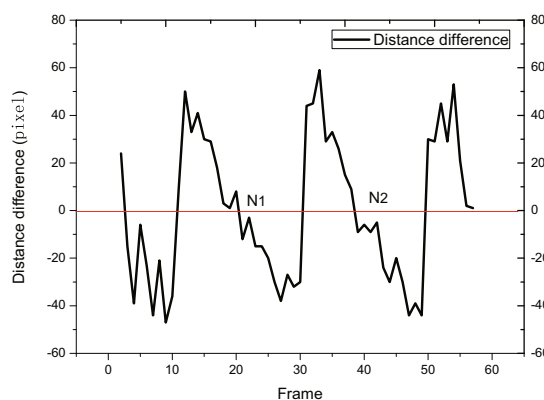


Fig. 5 The distance difference curve of the adjacent frame in the visual date(the distance of present frame—the distance of previous frame)

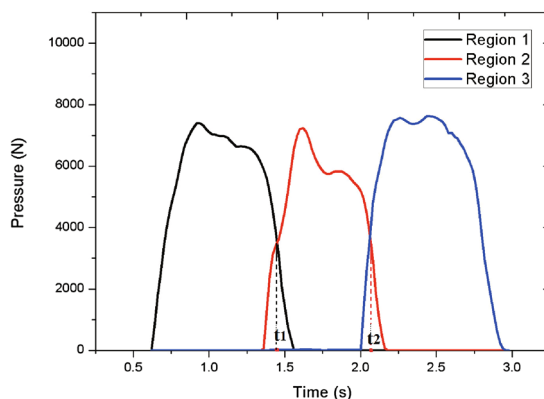


Fig. 6 The plantar pressure curve over time in different regions

2.1.1 The temporal fusion

During walking, the distance between the two feet (the number of pixels from the toes of the front feet to the heels of the other feet, see Fig. 3) and the plantar pressure (see Fig. 4) will change regularly with time, which are repeatable and stable.

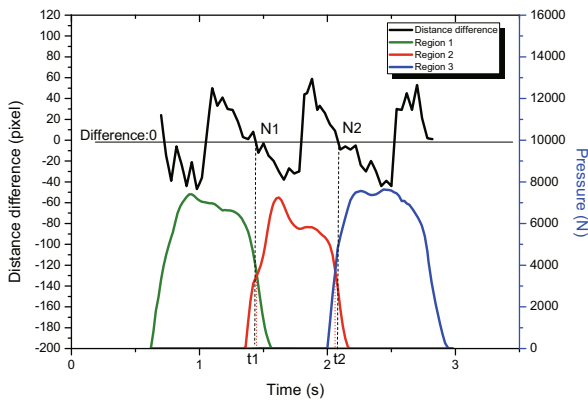


Fig. 7 The fusion curve of width difference and plantar pressure

Figure 5 shows the distance difference curve of the two feet of the adjacent frame. The sampling frequency of the camera (f_c) is 30 frames per second. The positive and negative differences represent the increase or decrease in the distance between the present frame and the previous frame, respectively, so the positive zero points N_1 and N_2 are the critical frames from increase to decrease, and the width of the feet in those frames is the largest.

Figure 6 shows the curve of the plantar pressure over time in the three regions, whose starting time was not a moment with that of Fig. 5. The sampling frequency of the plantar pressure test equipment (f_w) is 50 frames per second. In the walking process, the pressure on the left and right feet is equal at moment t_1 and t_2 . When the pressure on the left and right feet is equal, the heel of one foot and the forefoot of the other foot land at the same time, the inclinations of the two feet are the smallest relative to the ground. Therefore, the moment t_1 and t_2 can be considered as the moment when the distance of two feet is the largest within the allowable range of error.

Compared to the camera, plantar pressure test equipment has higher sampling frequency and higher data accuracy. For this reason, the visual data are integrated into the timeline of the plantar pressure data by using the characteristic moment of the widest feet distance. Define that the plantar pressure of the both feet in Fig. 6 is equal at the moments $t_1, t_2, t_3 \dots t_j, j \in N^+$, and the positive zeros in Fig. 5 appear at the frames $N_1, N_2, N_3 \dots N_i, i \in N^+$. There are multiple characteristic moments in the entire walking process, so the matching function is introduced under the assumption where the moment of first frame N_1 in Fig. 6 corresponds to the moment T_1 in Fig. 5 for improving the precision of the fusion:

$$f(T_1) = \frac{1}{n} \sqrt{\sum_{i=j=1}^n (T_i - t_j)^2} \tag{1}$$

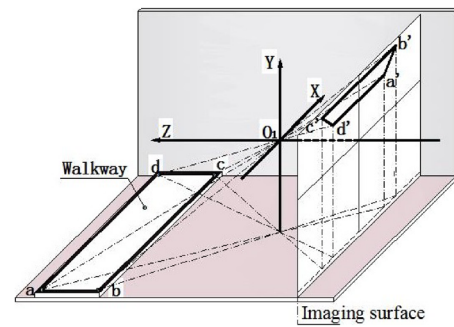


Fig. 8 The imaging process and principle of the Walkway frame in the lens

where n is the number of characteristic moments, $T_i = T_1 + (N_i - N_1)/f_c$ is the moment of each positive zero point in Fig. 5. When the matching function $f(T_1)$ obtains the minimum value, the errors of the corresponding characteristic time points of the visual data and the tactility data are the smallest, and the value of T_1 at this time is set as the moment of the N_1 frame in the visual data, which represents that the process of time fusion is unified. The fusion of visual data (Fig. 5) and tactility data (Fig. 6) is shown in Fig. 7.

2.1.2 The spatial fusion

In the visual imaging process, the actual direction of the object is the same as the imaging direction of the camera screen, but is opposite to the imaging direction on the camera imaging surface (see Figs. 8, 10, 11). The experimental equipment includes a camera lens and the plantar pressure test equipment (hereinafter referred to as Walkway).

The visual and tactility data are spatially fused according to the imaging principle of the black rectangle frame of the Walkway in the lens. As shown in Fig. 8, use the optical center O_1 of the lens as the origin and the optical axis of the lens as Z -axis to establish the coordinate system $O_1 - xyz$. The length of the rectangular frame of Walkway is L , and the width is W . The four vertices of the frame are represented by a, b, c, d , and their coordinates belong to unknown variables. The image points on the imaging plane corresponding to the four vertices are represented by a', b', c', d' , and their coordinates belong to known variables.

When the seven planes of the human skeleton model are parallel or perpendicular to the imaging surface, the calculation time and difficulty of the visual geometry can be reduced. Therefore, the experiment requires that the imaging coordinates of the Walkways frame in the lens satisfy the following relationship:

$$y_{a'}/x_{a'} = y_{b'}/x_{b'} = -y_{c'}/x_{c'} = -y_{d'}/x_{d'} \tag{2}$$

In this condition, the Plane $adcd$ is perpendicular to the Y -axis, the Line bc is parallel to the X -axis. The Line ab and

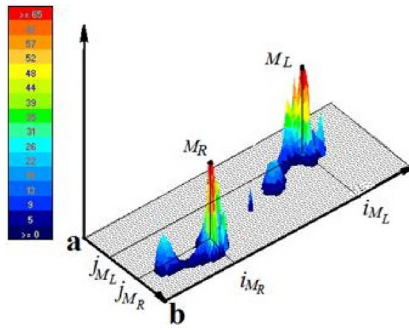


Fig. 9 The selection principle of the M-joint coordinates in the coordinate system $O_2 - xy$ in the tactility data

Line cd are symmetric about Plane yOz . The walking direction at any time during the experiment is processed parallel to the X -axis. According to geometric relations, the seven planes can be divided into two categories:

1. Planes 1–5 are parallel to the imaging plane, referred to as parallel plane.
2. Planes 6–7 are perpendicular to the imaging plane, referred to as vertical plane.

Position of a plane requires only three space points, so the six points (a, b, c, a', b', c') are used for analysis. The expression for the point imaging principle is described as:

$$(X, Y, Z) = -\gamma(x, y, p) \tag{3}$$

where p represents the camera focal length, γ represents the distance ratio relative to the origin between the space point (X, Y, Z) and its imaging point (x, y, p) . It can be seen from formula 3, when the coordinates of points a', b', c'

are known, the spatial position of points a, b, c is determined by the distance ratio $\gamma_a, \gamma_b, \gamma_c$. According to the condition $l_{ab} = W, l_{bc} = L$, and $\vec{ab} \perp \vec{bc}$. The expressions of $\gamma_a, \gamma_b, \gamma_c$ are described as:

$$\gamma_b = \gamma_c = L / (x_{b'} - x_{c'}) \tag{4}$$

$$\gamma_a = \gamma_b x_{b'} / x_{a'} \tag{5}$$

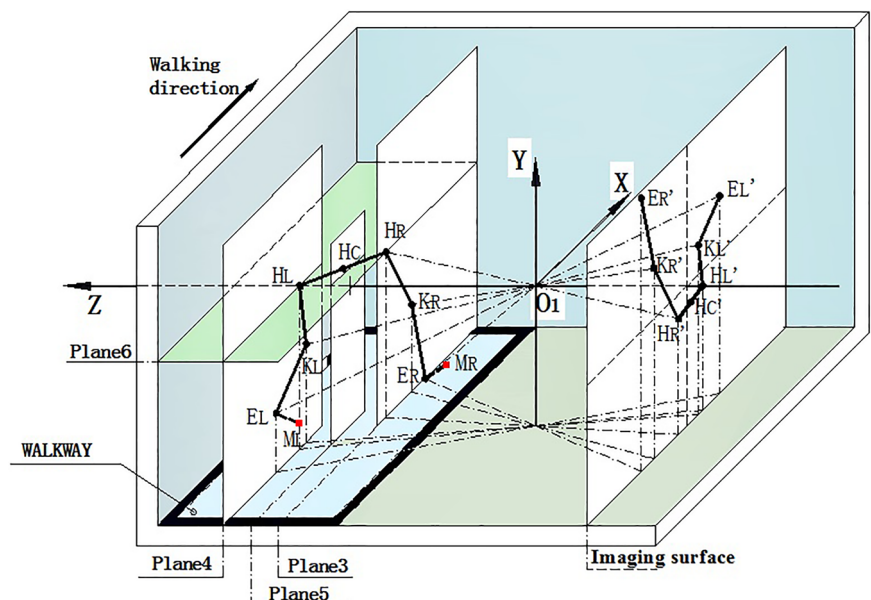
To locate each square pressure sensor (side length is cs) of Walkway, take the Point b as the origin, the Edge bc as X -axis, the Edge ba as Y -axis, and side length cs as the unit length to establish a plane coordinate system $O_2 - xy$. The spatial coordinates of any Point (i, j) in $O_2 - xy$ can be calculated by the space fusion formula:

$$(x, y, z)_{ij} = \gamma_b(x_{b'}, y_{b'}, p) + i[\gamma_a(x_{a'}, y_{a'}, p) - \gamma_b(x_{b'}, y_{b'}, p)]cs/W + j[\gamma_c(x_{c'}, y_{c'}, p) - \gamma_b(x_{b'}, y_{b'}, p)]cs/L. \tag{6}$$

2.2 Spatial geometry analysis

For obtaining the position of the relevant joints as accurately as possible and reducing the error, the extraction order of the skeletal model joints is taken from known to unknown. Unlike hard-to-handle vision data, the position of the human body is more accurately and easily located by tactility. The points of the upper and lower limbs are all in two parallel planes and one vertical plane, which causes that the spatial coordinates of the joints of the upper and lower limbs can be obtained in the same algorithm. In this paper, take the process of obtaining the joint point spatial coordinates of the lower limbs as the example to perform spatial geometry analysis.

Fig. 10 The imaging process and principle of the lower gait skeleton model in vision where Planes 3, 4, 5 and imaging surface are parallel to each other, parallel to the walking direction, perpendicular to the Plane 6



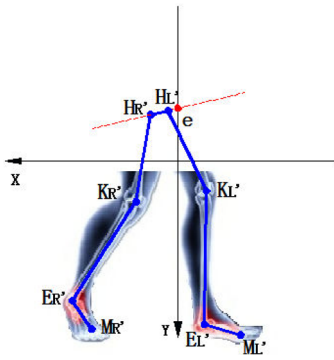


Fig. 11 Imaging of human lower skeleton model in imaging surface

2.2.1 The spatial coordinates of M

In the plantar pressure distribution (see Fig. 9) measured by the Walkway, the highest point means the main stress point, which is considered as the position of the phalangeal end M .

As shown in Fig. 9, the coordinates $(i_{M_R}, j_{M_R}), (i_{M_L}, j_{M_L})$ of the M joint points in $O_2 - xy$ can be obtained by the pressure distribution, which are the position coordinates of the pressure peaks. Through the space fusion formula 6, the two points are transformed from $O_2 - xy$ to $O_1 - xyz$ and the spatial coordinates of M_R and M_L are obtained. The left and right sides of the skeletal model are symmetrical and Plane 5 is a parallel plane, so the Z -axis coordinate expression at any point on Plane 5 is described as:

$$z_{\text{Plane5}} = (z_{M_R} + z_{M_L})/2 \tag{7}$$

where z_{M_R} represents the Z -axis coordinate of M_R , z_{M_L} represents the Z -axis coordinate of M_L .

2.2.2 The spatial coordinates of A and K

The imaging principle of human lower limb is shown in Fig. 10. In $O_1 - xyz$, the length of the bone $K_R - A_R$ is equal to the length of the bone $K_L - A_L$, and the Plane 3, Plane 4 are perpendicular to the Z -axis. When the camera is imaging, the imaging points of joints A_R, A_L, K_R, K_R are represented by A'_R, A'_L, K'_R, K'_R in the image surface and their coordinates can be extracted directly as known variables. According to $\triangle O_1 A_L K_L \sim \triangle O_1 A'_L K'_L$ and $\triangle O_1 A_R K_R \sim \triangle O_1 A'_R K'_R$, the Z -axis coordinate expression of any points of the Plane 3 and Plane 4 is described as:

$$z_{\text{Plane3}} = 2z_{\text{Plane5}} l_{A'_L K'_L} / (l_{A'_R K'_R} + l_{A'_L K'_L}) \tag{8}$$

$$z_{\text{Plane4}} = 2z_{\text{Plane5}} l_{A'_R K'_R} / (l_{A'_R K'_R} + l_{A'_L K'_L}) \tag{9}$$

where $l_{A'_R K'_R}$ represents the length of $A'_R - K'_R$, $l_{A'_L K'_L}$ represents the length of $A'_L - K'_L$. When the values of z_{Plane3} and

z_{Plane4} are calculated, the spatial coordinates of joint points A and K can be obtained by formula 3.

2.2.3 The spatial coordinates of H

Obstructed by the body, the imaging coordinates of the joints H on the imaging surface are difficult to be extracted directly. For this problem, a method based on human anatomy is presented through visual geometry analysis in our experimental environment. As shown in Fig. 10, Plane 3, Plane 4 and Plane 5 are perpendicular to the Z -axis, Plane 6 is perpendicular to the Y -axis, and the length of the skeleton $K_R - H_R$ is the same as the length of the skeleton $K_L - H_L$. By the conditions $\triangle O_1 H_L K_L \sim \triangle O_1 H'_L K'_L$, $\triangle O_1 H_R K_R \sim \triangle O_1 H'_R K'_R$ and $Z_{H_R} = Z_{H_L}$, we can infer that:

$$l_{H'_L K'_L} : l_{H'_R K'_R} = z_{\text{Plane3}} : z_{\text{Plane4}} \tag{10}$$

$$y_{H'_L} : y_{H'_R} = z_{\text{Plane3}} : z_{\text{Plane4}} \tag{11}$$

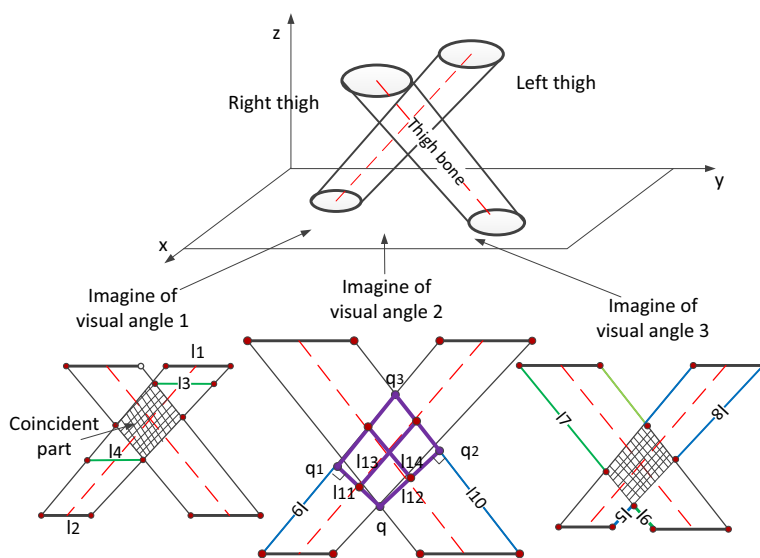
where $l_{H'_R K'_R}$ represents the length of $H'_R - K'_R$, $l_{H'_L K'_L}$ represents the length of $H'_L - K'_L$.

As shown in Fig. 11, the human lower limbs skeleton model forms 5 segments on the imaging surface. For getting the coordinates of joints H'_R and H'_L that blocked by the body, it is necessary to know the coordinates of joints K'_R and K'_L , the slope k_1 of bone $H'_R - K'_R$ and the slope k_2 of bone $H'_L - K'_L$, where the coordinates of joints K'_R and K'_L can be directly extracted as known variables, but the slope k_1 and k_2 which are also blocked by the body are also difficult to be directly extracted. For this problem, a method was proposed by visual geometry analysis (in Fig. 12).

As shown in Fig. 12, there are two oblique truncated cones in the space (representing the left thigh and right thigh, respectively). Assume that the centerline of the top and bottom circle of the truncated cone represents the thigh bone. When viewed from different angles, the position of the two thigh in plane image will be very different, and 3 imagines of different visual angles are taken as examples. In each visual imaging, the plane images form a coincident part, an irregular quadrangle. The position of these quadrilaterals in the visual images will rise or fall regularly with the continuous change of the viewing angle, but the slope of the thigh bone will not change.

In practical applications, the plane line l_5, l_6 regularly exists so short that the error is relatively large, and the plane line l_7, l_8 cannot be directly extracted due to the blocking of the body (as shown in the imagine of visual angle 3). However, the position of the intersection q , the plane line l_9, l_{10} is relatively accurate and easy to be extracted (as shown in the imagine of visual angle 2). Therefore, make the lines l_{11}, l_{12} , respectively, perpendicular to l_9 and l_{10} though the intersection p , and the slope of the line l_{13}, l_{14} connecting

Fig. 12 The principle of selecting the slope of the thigh bone $H - K$, in which the slopes of the l_{13} and l_{14} are selected as the slope of the thigh bone $H - K$



the midpoints of the opposite sides of the new quadrilateral $q_1q_2q_3q_4$ is defined as the slope k_1 and k_2 . Under the assumption that the centerline of the upper and lower bottom circles represents the thigh bone. When the straight line l_3, l_4 is parallel to the line l_1, l_2 (as shown in the imagine of visual angle 1), the line connecting the midpoints of the opposite sides of the quadrilateral is only the thigh bone. Actually, the assumed slope itself does not represent the true value, and the final slope obtained by the above method differs slightly from the assumed slope. Therefore, it is reasonable to represent the slope of the thigh bone with the midpoint line of the new quadrilateral $q_1q_2q_3q_4$. The final result shows that the positions of H joints obtained by the above method meet the actual body structure.

Assume that the slope of the line containing H'_R and H'_L is k , the Y -axis coordinate of the intersection of Line $l_{H'_R H'_L}$ and the Plane yO_1z is e . Through the formula (10, 11), the analytical formulas of k and e can be described as:

$$\frac{P_1k^2 + P_2k + P_3}{P_4k - P_5} = \frac{P_6k^2 + P_7k}{P_8k - P_9} = e \tag{12}$$

where

$$\begin{cases} P_1 = -x_{K'_R}t - x_{K'_L}\lambda \\ P_2 = x_{K'_R}tk_2 + x_{K'_L}\lambda k_1 + y_{K'_R}t + y_{K'_L}\lambda \\ P_3 = -y_{K'_R}tk_2 - y_{K'_L}\lambda k_1 \\ P_4 = t + \lambda \\ P_5 = tk_2 + \lambda k_1 \\ P_6 = y_{K'_R} - y_{K'_L}\lambda - x_{K'_R}k_1 + \lambda x_{K'_L}k_2 \\ P_7 = x_{K'_R}k_1k_2 - \lambda x_{K'_L}k_1k_2 - y_{K'_R}k_2 + \lambda y_{K'_L}k_1 \\ P_8 = k_1 + \lambda k_2 \\ P_9 = k_1k_2 + \lambda k_1k_2 \\ \lambda = z_{Plane3}/z_{Plane4} \\ t = \sqrt{1 + k_1^2}/\sqrt{1 + k_2^2} \end{cases} \tag{13}$$

when k and e are calculated, the spatial coordinate of joint H can be calculated by formula 3.

From the above analysis we can know, the spatial coordinates of all joints can be obtained by the method of this paper, which means that the skeleton model of the human body can also be reconstructed in $O_1 - xyz$. The gait skeleton model obtained by this method is reconstructed according to the actual positions of the joint points, so it is not affected by the visual observation angle and has good versatility.

3 Experiment and result analysis

In this experiment, a white wall is selected as the background, and the experimental equipment include: 1. camera with monocular zoom lens; 2. Walkway Metric 3150TL foot pressure test equipment (Walkway). The focal length of the monocular zoom lens was 27 mm after calibration. The length of Walkways black frame was 2035 mm, and the width was 455 mm. 9152 square sensors whose side length is 8.4 mm are distributed on the Walkway. The sensor of the Walkway uses piezoresistive sensing technology with a maximum measurement value of 862 Kpa. It's detection frequency can reach 50 HZ, and it can even be upgraded to 185 HZ if equipped with a dedicated detection module. The matching software can clearly record and display the 2D and 3D foot pressure contours at each moment and provide a variety of fixed and custom foot pressure analysis functions, such as the real-time position of the pressure center, the integral of the pressure-receiving area, the target area pressure value curve and so on. Although experiments and data processing can be easily carried out through the Walkway, the piezoresistive structure results in the sensor's distribution density and limit pressure values not coexisting. If you want to obtain higher density

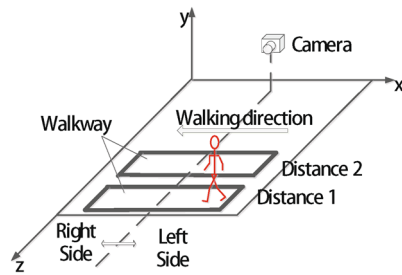


Fig. 13 The relative position of the selected 4 samples and the camera lens

pressure data, then the limit measurement pressure value will inevitably decrease. When used with other devices, if the detection frequencies of the two devices are inconsistent, the data correspondence is difficult due to the fixed frequency of the Walkway, which brings errors to the processing.

During the experiment (see Fig. 13), first place the lens at a distance (Distance 1) away from the background to ensure that the relevant bones and joints of the walking sample are completely imaged in the lens. Then adjust the angle and position of the lens, and make the image of the Walkway frame in the lens satisfy formula 2, which is called position adjustment. After the adjustment was completed, the sample walks from left to right on the Walkway and the data collection was performed at the same time. Finally, change the lens placement distance (distance 2) and repeat the experiment. Figure 14 shows the process of matching the visual data to the timeline of the tactility data, in which the value of T_1 , the moment of the N_1 frame, is 1.433 s when the minimum of $f(T_1)$ is taken.

In the experimental results, we selected four groups of samples from different sides of different distances about one subject to analyze the experimental results (see Figs. 15, 16, 17, 18, 19, 20, 21, 22), where the visual images were background-removed and the process of extracting the slope

Fig. 14 The time fusion schematic of visual data and tactility data, in which there are two moments when the distance between the two feet is the largest, and the fusion process is based on the timeline of the tactility data

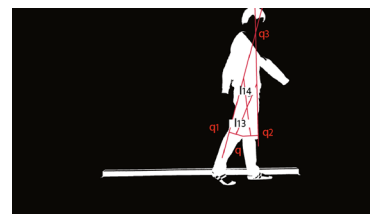
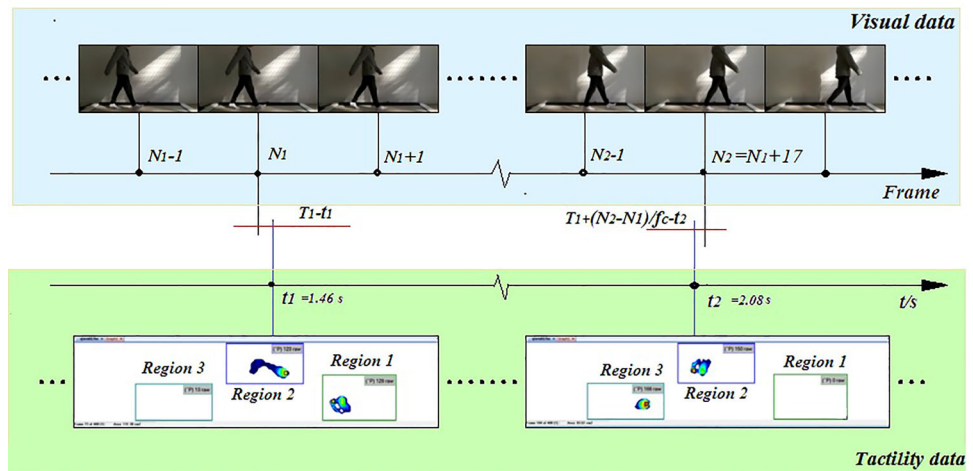


Fig. 15 Visual graphics at right side of distance 1



Fig. 16 The tactile images at right side of distance 1

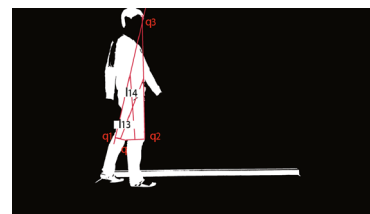


Fig. 17 Visual graphics at left side of distance 1



Fig. 18 The tactile images at left side of distance 1

of the thigh bone was expressed. The relevant joint coordinates and the parameters that can be extracted accurately by vision are shown in Tables 1, 2 and 3, and the coordinates of the second phalangeal end in $O_2 - xy$ that can be extracted by tactility are shown in Table 4.

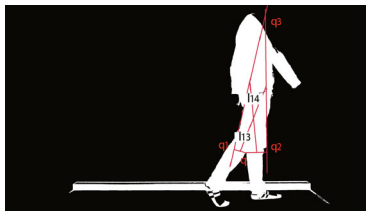


Fig. 19 Visual graphics at right side of distance 2



Fig. 20 The tactile images at right side of distance 2

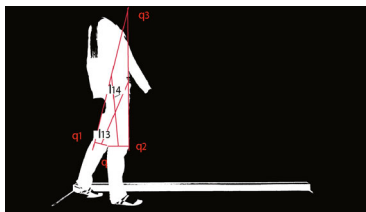


Fig. 21 Visual graphics at left side of distance 2



Fig. 22 The tactile images at left side of distance 2

Table 1 Joint point coordinates that can be directly extracted from two samples of distance 1 (mm)

Joint point	Right side of distance 1		Left side of distance 1	
	x	y	x	y
E_R	-7.452460	-4.890322	8.665001	-4.305875
E_L	-3.477000	-3.979388	4.968015	-3.970222
W_R	-9.523121	-2.909576	8.779359	-1.313805
W_L	-3.189796	-1.928369	3.652465	-1.858532
K_R	-4.066684	2.602730	6.189723	2.878149
K_L	-5.590436	2.630228	7.462935	2.692209
A_R	2.491428	5.981522	6.050922	6.682509
A_L	-6.074492	6.126433	8.922524	5.957952

3.1 Results and analysis

The walking postures of the skeleton models in different locations are complex, and the swing angles of the same bone are also different, which makes the characteristics of the model difficult to directly analyze and compare. Taking into account that the length of each bone does not change with the angle, the accuracy of the experimental results is verified by analyzing the bone length.

Table 2 Joint point coordinates that can be directly extracted from two samples of distance 2 (mm)

Joint point	Right side of distance 2		Left side of distance 2	
	x	y	x	y
E_R	-9.393923	-5.157884	9.780644	-4.500109
E_L	-4.884211	-4.140013	5.889861	-4.672082
W_R	-11.307887	-2.329494	9.250757	-1.227818
W_L	-4.996387	-1.505420	4.033948	-1.868571
K_R	-6.001164	2.927471	7.081452	3.507117
K_L	-7.078833	3.306336	8.652780	3.024806
A_R	-2.813114	7.449841	7.230291	8.588617
A_L	-7.410994	8.184436	10.706854	7.593006

Table 3 The γ values of the Walkway frame at different distances

Location	$\gamma_b(\gamma_c)$	γ_a
Distance 1	86.169681	101.15897
Distance 2	68.420524	84.90186

Table 4 M point coordinates in $O_2 - xy$ that can be directly extracted from Plantar pressure distribution

Location	Right foot		Left foot	
	i	j	i	j
Right side of distance 1	132	9	184	29
Left side of distance 1	47	9	4	33
Right side of distance 2	138	11	183	35
Left side of distance 2	53	11	10	34

Table 5 The random sampling time for solving the spatial skeleton model once

Number	1	2	3	4	5	6	7
Time (μs)	14.6	18.5	11.4	18.8	57.2	20.5	6.7

Table 6 The bone length at distance 1 and the measured values (mm)

Skeleton type	Right side of distance 1	Left side of distance 1	Measured values
$S_R - S_L$	440.289	480.1	450
$S - E$	223.351	236.567	230
$E - W$	242.285	261.709	249
$H_R - H_L$	145.965	163.098	155
$H - K$	324.887	334.081	330
$K - A$	339.705	345.757	345

Table 5 shows the random sampling time for solving the spatial skeleton model by the LabVIEW through the algorithm of this paper according to the data in Tables 1, 2, 3 and

Table 7 The bone length at distance 2 and the measured values (mm)

Skeleton type	Right side of distance 2	Left side of distance 2	Measured values
$S_R - S_L$	443.156	456.967	450
$S - E$	225.725	242.869	230
$E - W$	239.084	260.097	249
$H_R - H_L$	140.092	162.302	155
$H - K$	319.854	335.758	330
$K - A$	335.326	347.36	345

Table 8 The average bone length of the same distance and the measured values (mm)

Skeleton type	Distance 1	Distance 2	Measured values
$S_R - S_L$	460.1945	450.0615	450
$S - E$	229.959	234.297	230
$E - W$	251.997	249.5905	249
$H_R - H_L$	154.5315	151.197	155
$H - K$	329.484	327.806	330
$K - A$	342.731	341.343	345

Table 9 The error rate between the average bone length and the measured values

Skeleton type	Average bone length (mm)	Measured values (mm)	Error (%)
$S_R - S_L$	455.128	450	1.140
$S - E$	232.128	230	0.925
$E - W$	250.7938	249	0.720
$H_R - H_L$	152.8643	155	1.378
$H - K$	328.245	330	0.411
$K - A$	342.037	345	0.859

4. It can be seen that the time is up to $57.2 \mu\text{s}$ and the minimum is $6.7 \mu\text{s}$, which will be slightly different depending on the configuration of the computer or the further function of the program. With the visual acquisition frequency (33 ms for one frame) and the mature human body model pattern recognition technology, this time can fully meet the requirements of fast response.

Tables 6 and 7 show the bone lengths at different locations. It can be seen that the bone lengths of the models extracted from different locations have small fluctuations above and below the actual measured values. The length errors of the shoulder and hip bone are slightly larger than those of other bones, that is because the coordinates of H and S points are obtained by thigh slopes, which means that they have one more calculation and error than coordinates of the other points. Under the same distance, the lengths of the left and

right bones are overall larger or smaller, indicating that the error originates from the initial positioning in the experiment, which can be reduced through another position adjustment.

The average bone lengths of the left and right sides at the same distance are shown in Table 8 where the average values can be well fitted to the measured values of the sample. The error on the left side and the error on the right side complement each other, and it is proved once again that the reason why the overall length of the left or right side is larger or smaller is the error caused by the position adjustment during the experiment. The bone length error rate between the average bone lengths and the measured values is shown in Table 9 where each bone length error rate is eventually controlled within 2%. It is proved that the model extracted by this method can represent the true skeleton of human body to some extent.

4 Conclusion and future work

In this paper, the gait skeleton model is rapidly extracted through geometric analysis based on the spatial and temporal fusion of vision and tactility. Through the establishment of equal-scale model whose bone length error is eventually controlled within $\pm 5 \text{ mm}$, not only the individualized bone length is ensured, but also the model has good stability at different distances from different angles. Although the joint points were not automatically selected, the blocked shoulder and hip joints were fully analyzed and verified through geometric analysis. The experimental results show that the method proposed in this paper can quickly extract the skeleton model in which the bone lengths are representative.

In the research process of this paper, a high-precision model was obtained through the fusion between vision and tactility, and the tactility played a major role in the spatial positioning accuracy. In the future work, we will use this gait skeleton model to study the gait characteristics of tactile and vision in time. Then the dynamics coupling relationship between vision and tactility will be explored to improve the recognition efficiency and recognition accuracy of gait recognition according to the temporal and spatial fusion method in this paper.

Acknowledgements This work is supported by Master Innovation Funding Project Foundation of Hebei Province, P. R. China (Grant No. CXZZSS2018026), Science and Technology on Space Intelligent Control Laboratory (Grant No: ZDSYS-2017-08), State Key Laboratory of Robotics and System (HIT) (Grant No: SKLRS-2017-KF-15) and Hebei Natural Science Foundation (Grant No: F 2017202243).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures were performed in accordance with the ethical standards of the University of Auckland human participants and ethics approval committee (UAHPEC) and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

References

1. Cho, C.W., Chao, W.H., Lin, S.H.: A vision-based analysis system for gait recognition in patients with Parkinsons disease. *Expert. Syst. Appl.* **36**(3), 7033–7039 (2009)
2. Chai, Y.M., Xia, T., Han, W.Y.: State-of-the-Art on gait recognition. *Comput. Sci.* **39**(6), 10–5 (2012)
3. Zhang, R., Vogler, C., Metaxas, D.: Human gait recognition at sagittal plane. *Image Vis. Comput.* **25**(3), 321–330 (2007)
4. Lu, H., Plataniotis, K.N., Venetsanopoulos, A.N.: A full-body layered deformable model for automatic model-based gait recognition. *Eur. J. Adv. Signal Process.* **2008**(1), 261–317 (2007)
5. Liu, Y.: A gait recognition method based on body skeletal model. *Comput. Eng. Appl.* **09**, 88–92 (2005)
6. Hamzaçebi, H., Ömer, M.: On the periodic gait stability of a multi-actuated spring-mass hopper model via partial feedback linearization. *Nonlinear. Dyn.* **88**(2), 1–20 (2017)
7. Sun, J., Wang, Y., Li, J., et al.: View-invariant gait recognition based on kinect skeleton feature. *Multimed. Tools. Appl.* **2018**(2), 1–27 (2018)
8. Wang, G.F., Wang, B., Zhong, F.: Global optimal searching for textureless 3D object tracking. *Vis. Comput.* **31**(6–8), 979–988 (2015)
9. Zhang, F., Zhang, X., Cao, K., et al.: Contour extraction of gait recognition based on improved GVF Snake model. *Comput. Electr. Eng.* **38**(4), 882–890 (2012)
10. Zhang, Y.Z., Zheng, J.M., Nadia, M.T.: Example-guided anthropometric human body modeling. *Vis. Comput.* **31**(42), 1615–1631 (2015)
11. Kusakunniran, W., Wu, Q., Zhang, J.: Gait recognition across various walking speeds using higher order shape configuration based on a differential composition model. *IEEE Trans. Syst. Man Cybern. Syst.* **42**(6), 1654–68 (2012)
12. Sigal, L., Balan, A.O., Black, M.J.: HumanEva: synchronized video and motion capture dataset and baseline algorithm for evaluation of articulated human motion. *Int. J. Comput. Vis.* **87**(1–2), 4 (2010)
13. Poppe, R.: Evaluating example-based pose estimation: experiments on the humaneva sets. In: *Proceedings of CVPR Workshop on Ehum* (2017)
14. Zhang, X., Ding, M., Fan, G.: Video-based human walking estimation using joint gait and pose manifolds. *IEEE Trans. Circ. Syst. Video Technol.* **27**(7), 1540–1554 (2017)
15. Jahangiri, E., Yuille, A.L.: Generating multiple diverse hypotheses for human 3D pose consistent with 2D joint detections[C]. In: *IEEE International Conference on Computer Vision Workshop*. IEEE Computer Society, pp. 805–814 (2017)
16. Yao, Y., Yu, Z., Sidan, D.: Learning human shape model from multiple databases with correspondence considering kinematic consensus. *Vis. Comput.* **31**(1), 19–33 (2015)
17. Portillo, J., Leyva, R., Sanchez, V.: A view-invariant gait recognition algorithm based on a joint-direct linear discriminant analysis. *Appl. Intell.* **2017**(13), 1–18 (2017)
18. Li, Q., Wang, Y., Sharf, A.: Classification of gait anomalies from kinect. *Vis. Comput.* **2016**, 1–13 (2016)
19. Nordin, M.J., Saadoon, A.: A survey of gait recognition based on skeleton model for human identification. *Res. J. Appl. Sci. Eng. Technol.* **12**(7), 756–763 (2016)



Shijie Dai received the B.S. degree and M.S. degree from School of Mechanical Engineering, Hebei University of Technology, Tianjin. And then he received the Ph.D. degree from Harbin Institute of Technology. He is currently a Professor of Hebei University of Technology. His research interests include signal and information processing and machine vision.



Rui Wang received the B.S. degree from School of Mechanical Engineering, Hebei University of Technology, Tianjin, China. He is currently working toward the M.S. degree in the same university, focusing on the imagine analysis and dynamics about the gait recognition.



Huibo Zhang received the B.S. degree from the Harbin Engineering University. And then received the M.S. degree and Ph.D. degree from Harbin Institute of Technology. He is currently the teacher of School of Mechanical Engineering, Hebei University of Technology. His research interests mainly are the human body dynamics analysis and recognition.