Height Measurement for Humans in Motion Using a Camera: A Comparison of Different Methods

Ngoc Hung Nguyen
School of Electrical and Electronic Engineering
The University of Adelaide
Email: a1187073@student.adelaide.edu.au

Richard Hartley
Australian National University
and National ICT Australia
Email: Richard.Hartley@anu.edu.au

***Abstract*—This paper aims to answer the questions whether human height measurement from video is affected by human motion and how accurately human static height can be measured from a video or from a single image frame that captures a walking subject. In this paper, we present a new approach for measuring human height based on the cross ratio. This approach is performed in parallel with the vanishing point based approach for comparison. The accuracy of each approach is examined by comparing the estimated value to the actual static height value which is measured directly. For the height measurement from video, our analyses show that human height varies significantly during human motion and the highest point in a human stride is the most accurate measurement of the static height. We also model the height variation as a sinusoidal pattern to maximize the accuracy of the estimated height. For the height measurement from a single frame, a correction method is developed in which the length of the human torso with human head and the length of the human leg are separately estimated. We conclude that static human height can be measured accurately, even though the variation of human height is significant during human motion.**

I. Introduction

There are many different human identification technique using image such as face recognition, gait analysis or distinctive features on the human body. However, none of these methods can be applied to identify a subject captured by a normal surveillance camera system due to lack of resolution. Human height measurement arises as a suitable approach to human identification where the image resolution is not good enough for the other methods to be used.

The purpose of this paper is to examine the effect of human perambulation on the result of height measurement. Our main interest is the variation of human height during motion and the accuracy of the measured height values, computed as the average of the heights in all image frames and as the average of the maximum heights. In this paper, we propose a new method based on the cross ratio to perform human height measurement. We also measure height using the vanishing points based method proposed by Criminisi *et al.* [1] as a comparison to the cross ratio based method. To determine the limits of accuracy of video-based height measurement, experiments are carried out with 39 subjects captured in controlled videos. The human height variation over a stride then is modelled as a sinusoidal function of time. For height measurement from a single frame, we develop a correction method based on the length of the human torso with human head and the length of the human leg.

a) **Previous work :** Extensive research has been done in human height measurement using images (for example, [2], [3], [1]). There are two different approaches: projective geometry and 3D reconstruction of the scene. One of the projective geometry based approaches makes use of vertical vanishing points, the vanishing line of the horizontal plane and a reference height to compute human height [1]. The vertical vanishing point is estimated as the intersection of two vertical lines while two sets of parallel horizontal lines with different directions are required to determine the vanishing line of the horizontal plane. The known height of an object in the scene can be used as the reference height. The 3D reconstruction based approach uses a multi-camera system [3]; human height can be easily computed if a 3D model of the scene is reconstructed from several views.

In contrast, there is much less research investigating the variation in height of human in motion. Madden and Piccardi [3] took the motion of subjects into account when calculating the subjects' heights from videos recorded by two disjoint camera views. They concluded that it is possible to distinguish between individuals of different heights. However, their research only focuses on the use of height measurement to match the tracks of the same individual. The tracks match each other when the difference between the average heights of the two tracks is smaller than the least standard deviation. Criminisi *et al.* [1] performed an experiment to evaluate the variation in heights of some subjects in motion as a part of their research in height estimation from video. The result shows that the human height variation is significant and the maximum height does not always match the static height measured when the human is standing up straight. In addition, the average of the maximum heights, which occur at the mid-stride phase [1], might be an appropriate measurement of the static height. However, their conclusions are restricted by the limited number of participants in their experiment.

II. Algorithm Description

In this section, we discuss in details of the new method for human height measurement based on the cross ratio. A brief description of the vanishing points method is also given but further details are discussed by Criminisi *et al.* [1]. Both of the

978-1-4673-2181-5/12/$31.00 ©2012 IEEE

methods are based on projective geometry and only require a single perspective view.

It is to be emphasized that the goal of this paper is to determine the limits of height measurement through video. To do this, we develop a new method for height measurement that makes use of favourable reference points in the scene, of a type frequently present in man-made environments. Such reference points may not be available in every scene, but are used here to make height measurements as accurate as possible in a controlled environment.

A. Cross ratio based method

1) Height Measurement: The cross-ratio is a fundamental concept in geometry. Its critical property is invariance under projective transformation [4]: $R(A,B;C,D) = R(A',B';C',D')$ for points (A,B,C,D) and corresponding points (A',B',C',D') under projective transformation. In other words,

$$\frac{AC \times BD}{BC \times AD} = \frac{A'C' \times B'D'}{B'C' \times A'D'} \tag{1}$$

The cross ratio based method relies on the existence of three

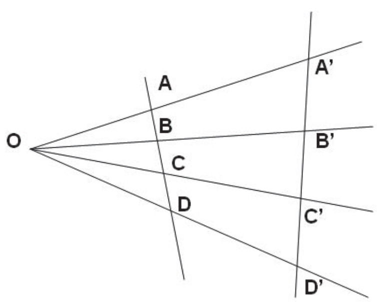

Fig. 1: Cross Ratio

horizontal reference lines, which often occur in the modern architecture, for example, the edge of walls, picture frames, bookshelves or other wall decorations.

There are three required elements for height measurement by the cross ratio based method:

- The camera height.
- Three horizontal reference lines with known heights on the same vertical plane (the *vertical reference plane*).
- Two vertical lines to identify the vertical vanishing point.

a) **Computing human height :** The geometry of the real scene is shown in figure 2, in which the top of human head T and its projection B on the ground correspond to the image points t and b; the points T' and B', which lie on the vertical reference plane, also correspond to the image points t and b. Interestingly, the human height H can not be computed directly by using the cross ratio. However, the human height H can be computed from the heights of points T' and B' ($h_{T'}$ and $h_{B'}$) which are computed using the cross ratio. The process is described as the follows.

$$\frac{T'B'}{TB} = \frac{CB'}{CB} = \frac{CB+BB'}{CB} = 1 + \frac{BB'}{CB} \tag{2}$$

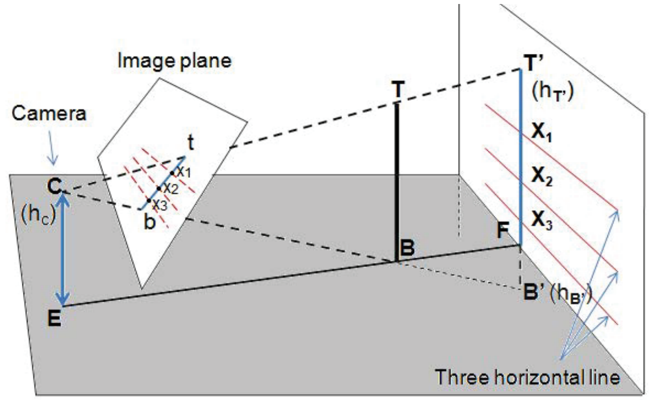

Fig. 2: The geometry of the real scene

$$\frac{FB'}{CE} = \frac{BB'}{CB} \tag{3}$$

where $TB = H$, the human height; $CE = h_C$, the camera height; $FB' = h_{B'}$, $FT' = h_{T'}$ and $T'B' = h_{T'} + h_{B'}$. Note that the point B' lies under the surface of the ground, the $h_{B'}$ mentioned in this case is the absolute value of B' height. By a simple algebraic rearrangement on (2) and (3), the human height H is

$$H = \frac{(h_{T'} + h_{B'})h_C}{h_C + h_{B'}} \tag{4}$$

b) **Computing $h_{T'}$ and $h_{B'}$:** The points T', B' and the three horizontal reference lines all lie on the reference plane, hence the cross ratio can be applied to compute the T' and B' heights. The points x_1, x_2, x_3 in the image correspond to X_1, X_2, X_3 in the space where X_1, X_2, X_3 are the intersections of the vertical line through T' and B' with the three horizontal reference lines respectively. A cross ratio defined by the set of points x_1, x_2, x_3, t can be easily computed from the image. The cross ratio is preserved under projective transformation. Hence, we have

$$\frac{tx_2 \times x_1x_3}{x_1x_2 \times tx_3} = \frac{TX_2 \times X_1X_3}{X_1X_2 \times TX_3} = \frac{(h_{T'} - h_2)(h_1 - h_3)}{(h_1 - h_2)(h_{T'} - h_3)} \tag{5}$$

where h_1, h_2, h_3 are the heights of the three reference lines which can be measured in the real scene.

From equation (5), $h_{T'}$ can be easily computed. Similarly, $h_{B'}$ can be also obtained by the same process.

c) **Computing camera height :** Camera height can be measured accurately by performing a camera calibration assessment. The camera calibration toolbox, which was developed at the California Institute of Technology [5], can be used to estimate the intrinsic and extrinsic parameters of camera. However, only the camera height is needed in this case, thus a simple assessment is conducted to measure the camera height. A reference object with known height is allocated between the camera and the reference plane as shown in figure 3. The camera height is computed as the following equation:

$$h_C = h_Y + (H_{ref} - h_Y)\frac{D}{a} \tag{6}$$

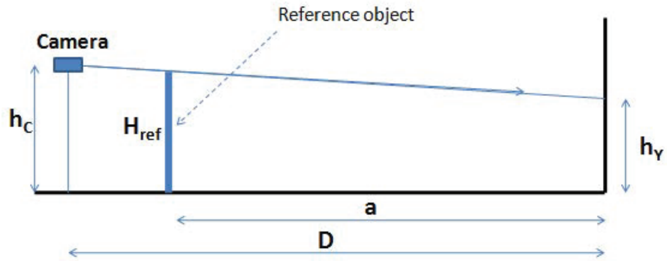

Fig. 3: Camera height assessment

where h_Y on the reference plane can be measured by the same process of measuring $h_{T'}$ and $h_{B'}$.

Note that cameras are commonly allocated stationary at specific locations and their directions are also fixed. In this case, camera height remains unchanged. In addition, if a camera is controllable, its height might have been changed; but the condition of the camera, in the instant when image is captured, could be reconstructed basing on the background information on the image, hence the camera height in that instant can be measured.

2) The Top of the Head and Its Projection on the Ground: The projection of the top point on the ground can be found as the feet position when subject stands up straight. However, it becomes more challenging if the subject is walking because the feet locations are apart from each other. Madden and Piccardi [3] determined the projection point B by the intersection between the vertical line through the top point T and the line formed by the heel locations, but both heels are not often in contact with the floor at the same time. For that reason, this approach is not applicable for all image frames of video. In this paper, we introduce another approach to determined the projection point basing on a normal manner of human motion. Humans normally walk straight to reach their targets as a conditioned reflex except for avoiding an obstruction or for a specific purpose. In this condition, a walking path can be determined accurately by the best-fit line of a set of the points where the heels make contact with the floor. Note that there are two walking paths, one for the left foot and one for the right foot. The projection point B is then determined as the midpoint of two intersections between the two walking paths and the vertical line through the top point T.

Due to the round shape of the human head, it is hard to manually identify the top point accurately. Instead of computing the height of an identified top point, the height measurement is performed for every single point on the edge of human head and the human height is found as the maximum value of the computed heights. To do that, an image of the human head first is extracted from the original image and then be filtered by a Canny edge detector. An output of this process is shown in figure 4. For each point on the head edge, its projection on the ground is determined and its height is computed as discussed above. The maximum value is then identified. The whole process is performed automatically except the image of the head is being extracted manually. However, this could be developed to an automatic process.

Fig. 4: The extracted image of human head

a) **The best-fit line :** The best-fit line, discussed by Taylor [6], with the form $y = A + Bx$ of a set of points $(x_1,y_1),\ldots,(x_N,y_N)$ has the coefficients:

$$A = \frac{\sum x^2 \sum y - \sum x \sum xy}{\Delta} \tag{7}$$

$$B = \frac{N \sum xy - \sum x \sum y}{\Delta} \tag{8}$$

where $\Delta = N \sum x^2 - (\sum x)^2$.

B. Vanishing points-based method

The vanishing points based method requires three elements to perform a height measurement:

- The vanishing line of the horizontal plane.
- The vertical vanishing point.
- A reference height.

The detailed description of the method is discussed comprehensively by Criminisi *et al.* in [1].

III. Uncertainty Analysis

A. Propagation of Uncertainties

The propagation of uncertainties through a function f of independent variables $x, y, \ldots, z$, which is discussed by Taylor [6], is computed as

$$\Delta f = \sqrt{\left(\frac{\partial f}{\partial x}\Delta x\right)^2 + \left(\frac{\partial f}{\partial y}\Delta y\right)^2 + \ldots + \left(\frac{\partial f}{\partial z}\Delta z\right)^2} \tag{9}$$

where Δx, Δy,…, Δz are the uncertainties of the variable $x, y, \ldots, z$.

B. Uncertainty in Human Height Measurement

The human height is a function of $h_{T'}$, $h_{B'}$ and h_C as discussed in Section II, hence the uncertainty in the estimated height is computed based on the uncertainties in measuring $h_{T'}$, $h_{B'}$ and h_C.

1) Uncertainty in determining point T' ($\Delta h_{T'}$): Note that T' and T in the space correspond to the same point t in the image, which is automatically determined by searching for a point on the human head edge associated with the maximum height. For that reason, the maximum error in t is 1 pixel. The uncertainty in determining point T' ($\Delta h_{T'}$) is then estimated based on the direction of camera and the distance between the camera and the reference wall.

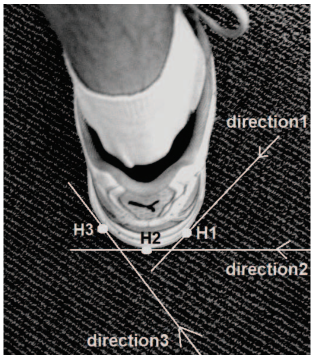

Fig. 5: Error in contact position

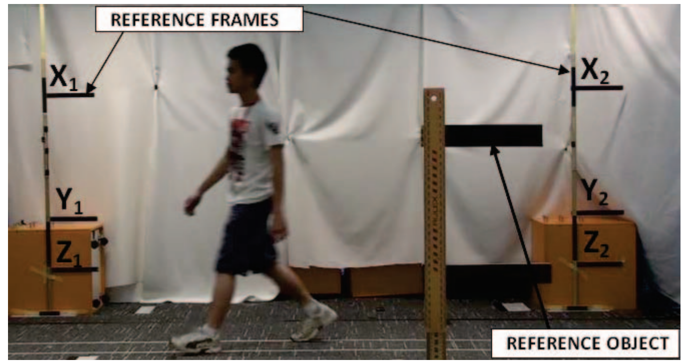

Fig. 6: Experimental Arrangement

2) Uncertainty in determining point B' ($\Delta h_{B'}$): Similarly, B' and B correspond to a same point in the image. The project point B is found as the intersection between the vertical line and the pathway identified by the best-fit line for a set of the points where the heels make contact with the floor. The uncertainty in the contact position leads to the uncertainty in identifying the pathway, hence in B and B'. Let's consider the situation in figure 5, in which the heel actually contacts with the floor at H_2 but the detected contact position might be incorrect depending on the relative position between human and the camera. The uncertainty in contact position is approximated as a half-width of the human foot. In addition, derived from equation (3), we have

$$h_{B'} = h_C \left(\frac{EF}{EB} - 1 \right) \tag{10}$$

The uncertainty in the contact position leads to a uncertainty in EF, hence in $h_{B'}$. The uncertainty $\Delta h_{B'}$ is then computed as the propagation of uncertainties through equation (10).

3) Uncertainty in measuring camera height (Δh_C): The camera height h_C is measured by the camera height assessment, in which h_C is obtained from equation (3). The uncertainty Δh_C is also computed based on the propagation of uncertainties approach.

IV. Experiments

A. Experimental Set up

The experiment is arranged as shown in figure 6. The two wood frames $X_1Y_1Z_1$ and $X_2Y_2Z_2$, which are designed and assembled precisely to reduce the physical measurement error, form a reference plane. The lines X_1X_2, Y_1Y_2 and Z_1Z_2 are used as three horizontal reference lines in our experiment.

Thirty nine participants were invited to take part in the experiment. They were required to walk straight in random directions while a camera was used to capture their motions. The recorded videos then were processed to obtain the measured heights of the participants. The actual static height of each participant was carefully measured by a ruler when the participant stood up straight.

B. Experimental Results and Analyses

A sequence of image frames is extracted from each recorded video. The contact positions of the heels with the floor then are manually identified from the sequence of image frames, the walking path is then determined. The subject height in each frame finally is computed by both the cross ratio based method and the vanishing points based method. To observe the height variation, the computed values of height in all image frames is plotted on the same graph. Figure 7 shows the height variation of subject 2 over the first 94 image frames. It can be initially observed that the results obtained by the two methods seem to match each other while the human height variation is quite significant.

1) Statistical Analysis: The above process is performed for all recorded videos of the 39 subjects. Statistical analyses are carried out as the follows.

a) **Height variation :** We define the term *maximum height variation* of subject during motion as the difference between the average value of the maximum heights and the average value of the minimum heights. Figure 8 shows the statistical histograms of the *maximum height variations* over all 39 subjects. The height variation is observed to be considerable, in which 50% of the samples lie between 3.0 cm and 3.9 cm. The histograms have the shapes of normal distributions with a mean of 3.44 cm and a standard deviation of 0.68 cm for the data obtained by the cross ratio method; and a mean of 3.41 cm with a standard deviation of 0.67 cm for the data obtained by the vanishing points method.

b) **Estimated height :** Human height can be estimated as *the average of the heights in all frames* or as *the average of the maximum heights*. Madden and Piccardi [3] used *the average of the heights in all frames* in their research on human height measurement, while Criminisi *et al.* [1] stated that *the average of the maximum heights* is the most representative measurement of the static height. To identify the most accurate approach, we use both of them in our experiments.

Figure 9 shows the statistical histogram of the errors in the estimated heights, which are computed as *the average of the heights in all frames*, with respect to the static heights. The estimated height by this approach is accurate within 1.49 cm RMS error for the data obtained by the cross ratio method,

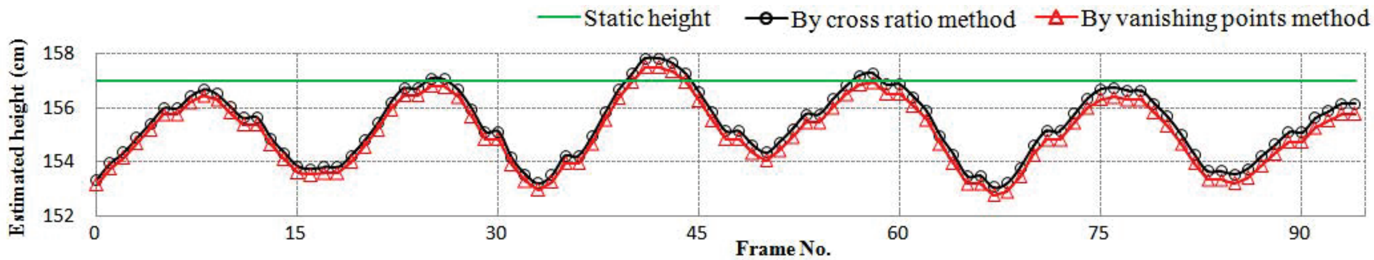

Fig. 7: The estimated heights of subject 2 over the first 94 extracted image frames: the circle shaped points and the triangle shaped points correspond to the heights computed by the cross ratio method and the vanishing points method respectively, while the static height is shown by the horizontal straight line.

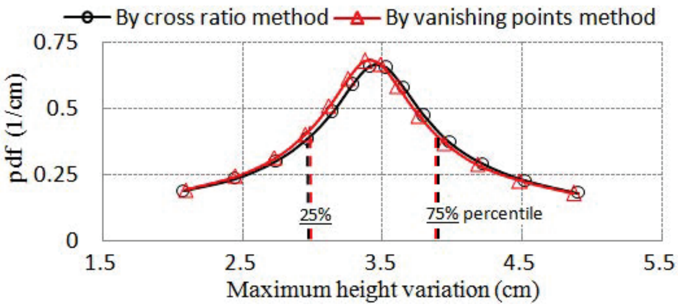

Fig. 8: Histogram of maximum height variations: The graph shows the distribution (over all 39 subjects) of the difference between the maximum and minimum heights measured for the subject during the video. This reflects the extent to which a subject's height varies during the stride, related to the amplitude of the graph in figure 7. The vertical dashed lines correspond to the 25% and 75% percentiles of the distributions. **Note that the vertical scale is normalized so that the total area under the distribution curve is 1 (the notation *pdf* means the probability density function). We also use this vertical scale for other histograms in this paper**.

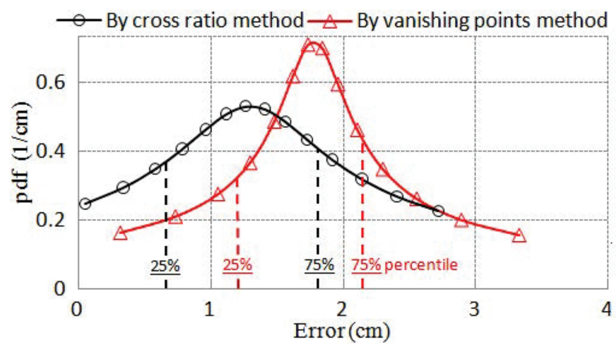

Fig. 9: Histogram of errors in the estimated heights, *as the average of the heights in all frames*, with respect to the static heights: the error is computed by subtracting the estimated height from the static height. The positive value of error indicates that the static height is higher than the estimated height, otherwise the static height is lower. The vertical dashed lines correspond to the 25% and 75% percentiles of the distributions. In the graph, the errors are always positive because *the averages of the heights in all frames* are always lower than the static heights due to the height variation during human motion.

or 1.89 cm RMS error for the data obtained by the vanishing points method.

Figure 10 shows the statistical histogram of the errors in the estimated heights, which are computed as *the average of the maximum heights*, with respect to the static heights. The estimated height by this approach is accurate within 0.63 cm RMS error for the data obtained by the cross ratio method, or 0.62 cm RMS error for the data obtained by the vanishing points method. It is 3 times more accurate than the estimated height obtained as *the average of the heights in all frames*.

The statistical observation shows that *the average of the height in all frames* is systematically lower than the static height as shown in figure 9 due to the effect of human motion. It can be a good approximation of the static height but *the average of the maximum heights* can lead to a more accurate estimation of the static height. We agree with the conclusion of Criminisi *et al.* [1] that *the average of the maximum heights* is the most accurate measurement of the human static height.

c) **Uncertainty analyses :** Figure 11b shows the histogram of the relative uncertainties of the estimated heights over all 39 subjects. The average relative uncertainties are 0.23% and 0.26% for the cross ratio method and the vanishing points method respectively.

d) **Comparisons between the cross ratio method and the vanishing points method :** Comparable results of the estimated heights are obtained by the two methods as seen in the above analyses results. To see the exact difference between them, we compute the relative difference between the estimated heights obtained by the two methods in every single frame, then the average of the relative difference in all frames. Figure 11a shows the histogram of the average relative differences between the the two methods over all 39 subjects, in which the 25% and 75% percentiles of the distribution are

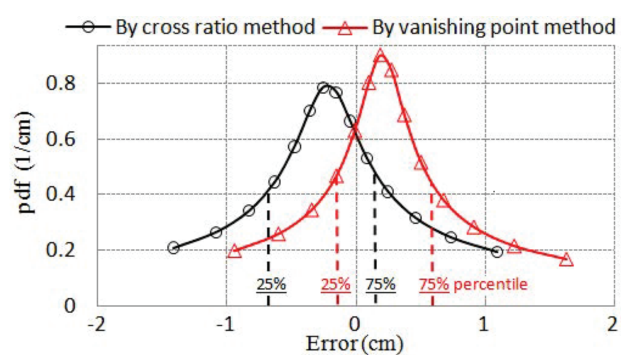

Fig. 10: Histogram of errors in the estimated heights, *as the average of the maximum heights*, with respect to the static heights: the error is computed by subtracting the estimated height from the static height. The positive value of error indicates that the static height is higher than the estimated height, otherwise the static height is lower. The vertical dashed lines correspond to the 25% and 75% percentiles of the distributions. In the graph, the errors spread out over the negative and positive sides as the estimated height can be higher or lower than the static height.

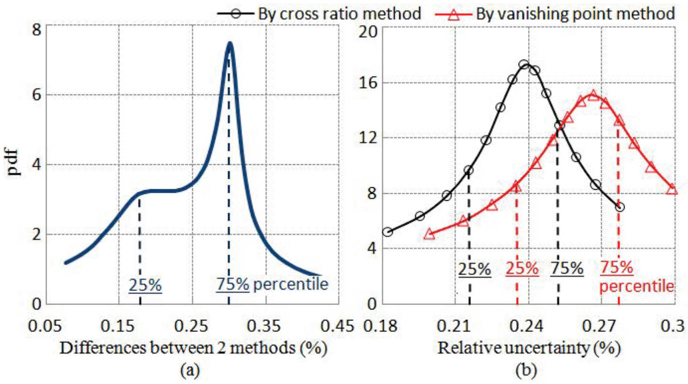

Fig. 11: Histogram of: (a) the difference between the estimated heights obtained the 2 methods; (b) the relative uncertainties of the 2 methods. The vertical dashed lines in the graphs correspond to the 25% and 75% percentiles of the distributions.

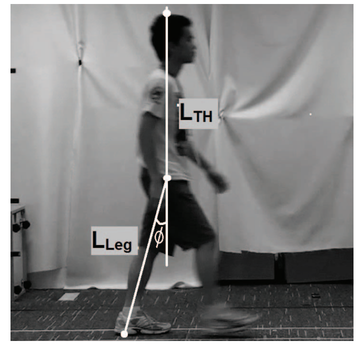

Fig. 12: Modelling as a sinusoidal function

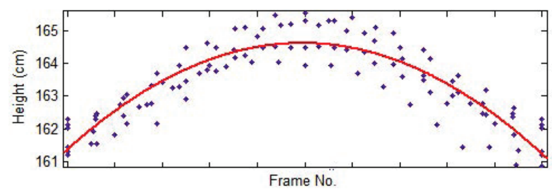

Fig. 13: The best-fit sinusoidal curves for subject 18: the dot points correspond to the estimated heights of the training set, and the red curve corresponds to the determined best-fit sinusoidal function, which is found by using the curve fit toolbox of MATLAB® [7].

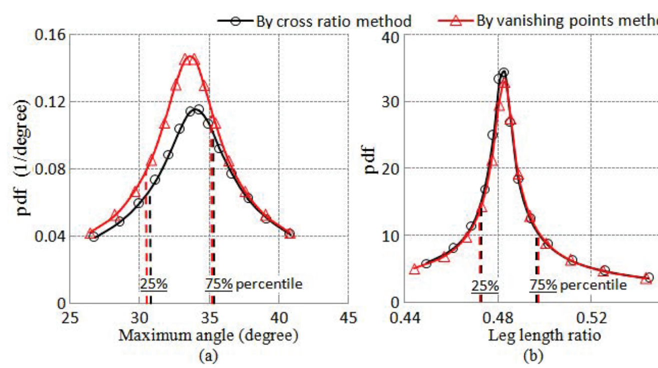

Fig. 14: (a) Histogram of maximum angle between two legs over a stride: The distribution spreads out on a wide range with a mean of 33 degrees and a standard deviation of 3.5 degrees. This is reasonable as the maximum angle depends on the walking style and the speed of each subject; (b) Histogram of leg length ratio: the leg length ratios, computed as $L_{Leg}/(L_{Leg} + L_{TH})$, are quite similar among the experimental subjects, with a mean of 0.49 and a standard deviation of 0.02. Note that the vertical dashed lines in the graphs correspond to the 25% and 75% percentiles of the distributions.

at the relative differences of 0.18% and 0.30%. The difference can be explained by the uncertainties of the two methods.

2) Modelling as A Sinusoidal Function: The average of the maximum heights is proved to be the most accurate measurement of the static height, in which only the maximum height values are taken into consideration regardless of height values in other image frames. Therefore, only $1/20$ (or 5%) of information is used if there are 20 image frames captured during a stride period. To make use of the information in all image frames, we model the height variation over a stride as a sinusoidal pattern. *The highest point of the sinusoidal model* then is defined as a new estimation of the static height.

a) **Sinusoidal model :** The variation of human height can be simply modelled as:

$$h = L_{TH} + L_{\text{Leg}} \cos(\phi) \tag{11}$$

where L_{Leg} is the length of the human leg, L_{TH} is the length of the human torso with human head and ϕ is the angle between a straight leg and the vertical line through human hip as shown in figure 12. Based on the statistical analyses, the number of image frames captured during a stride period is observed to be approximately unchanged, so the subject seems to walk in a constant speed. The angle ϕ can be approximated as a linear function of time, hence as a linear function of frame number because camera has a constant frame rate. Therefore, the variation of human height during a stride as:

$$h = L_{TH} + L_{\text{Leg}} \cos(\omega n - \Theta/2) \tag{12}$$

where Θ is the maximum angle between two legs over a stride and n is frame number.

b) **Validation test of sinusoidal model :** In order to examine the precision of the sinusoidal model, the obtained data is divided into training and test sets. For each subject, the first half of the data is used as the training set and the second half of the data is used as the test set. Note that the subjects walk from the left to the right in the first set but from the right to the left in the second set, and the walking paths in the two data sets are also different. Therefore, the test set is completely independent from the training set.

A best-fit curve of equation (12) is determined basing on the training set using the curve fit toolbox of MATLAB® [7] after all strides captured in video are aligned into a single stride. Figure 13 shows the best-fit sinusoidal curves for the height variations of subject 18. From the determined best-fit curve, we can work out the values of highest point, leg length ratio and maximum angle between two legs for each subject. The statistical histogram of the maximum angle and the leg length ratio are shown in figure 14.

The determined sinusoidal curve is now examined on the test set. The overall testing error is computed as a root mean square value of the relative differences between the estimated values obtained from images and the predicted values obtained

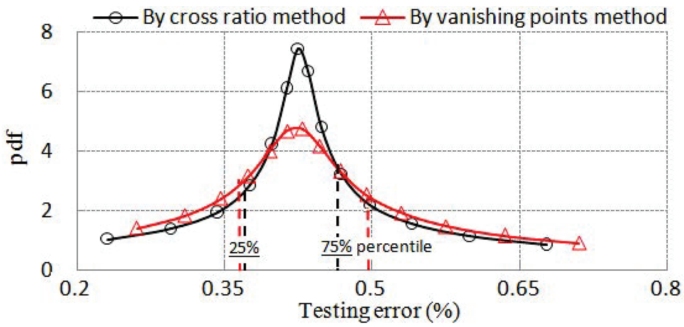

Fig. 15: Histogram of testing error of sinusoidal model: The graph shows the distribution (over all 39 subjects) of the testing error of the best-fit sinusoidal curve (which is obtained from the training set) on the test set. The testing error is computed as a RMS value of the relative differences between the estimated values obtained from the test set and the predicted values obtained from the determined best-fit curve. The vertical dashed lines in the graph correspond to the 25% and 75% percentiles of the distributions.

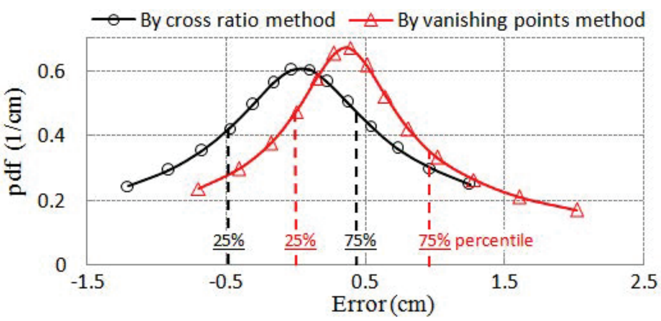

Fig. 16: Histogram of errors in the estimated heights, as *the highest point of the sinusoidal model*, with respect to the static heights: the error is computed by subtracting the estimated height from the static height. The positive value of error indicates that the static height is higher than the estimated height, otherwise it is lower. The vertical lines correspond to the 25% and 75% percentiles of the distributions. This error is comparable to the error in the estimated height, as *the average of the maximum heights*, shown in figure 10.

from the determined best-fit sinusoidal function over all image frames of the test set. Figure 15 shows the statistical histogram of the testing errors over all 39 subjects. We can observe that the testing error is about 2 times larger than the relative uncertainties as shown in figure 11. It is reasonable as the uncertainty occurs in both the training and test sets. Consequently, the sinusoidal function (12) is an appropriate model of the height variation over a stride.

c) **Highest points of the sinusoidal model :** As mentioned above, a new estimation of the human height is defined as *the highest point of the sinusoidal model*. Figure 16 shows the statistical histogram of the errors in this estimation with respect to the static height. The estimated height is accurate within 0.62 cm RMS error for the data obtained by the cross ratio method, or 0.81 cm RMS error for the data obtained by the vanishing points method. This approach gives a comparable accuracy to the approach based on *the average of the maximum heights*.

d) **Height measurement using the first stride:** In our experiments, the number of image frames seems to be sufficient to get the estimated heights accurately obtained as *the average of the maximum heights* or as *the highest point of the sinusoidal model*. A question here is how accurately the estimated heights obtained by these two approaches are when the number of image frames is limited.

We assume that only the first stride in each video is available for measuring the subject heights. In this case, only a frame can be used to compute *the average of the maximum heights* as there is only a maximum height over a stride, while the estimated heights in all frames can be used to compute *the highest point of the sinusoidal model*.

The RMS errors of the estimated heights from the first stride are reported in Table I, in which the estimated heights obtained from the first stride are less accurate than the estimated heights obtained from the whole video. We also observe that the method based on *the highest point of the sinusoidal model* is a bit more accurate than the method based on *the average of the maximum heights* in the case of using data from the first stride only.

3) Height Measurement of Subjects from a Single Frame: The height measurement may need to be performed based on a single image frame only because video is not always available. In this case, the height of head top point is not a good representation of the actual static height as the subject might be at any walking gait phase.

a) **Height correction algorithm :** A corrected height is computed as the sum of the length of the human torso with human head and the length of the human leg which are estimated separately. The length of the human torso with human head is found by subtracting the height of the hip point from the height of the head top point. Note that these heights are estimated as discussed in Section II. The leg length is found by Pythagoras' theorem based on the height of the hip point and the distance FB between the foot location and the projection point of the hip point on the ground as shown in figure 17a. Note that FB can be estimated based on background information in the image but the method may vary in different cases. In our experiments, we measure FB based on a known-length segment R_1R_2 of the reference frame as the follows.

$$FB = \frac{FF_o - BB_o}{\sin\left[\arctan\left(\frac{R_1R_1' - R_2R_2'}{R_1R_2}\right)\right]} \qquad (13)$$

where FF_o, BB_o, R_1R_1', R_2R_2' are the distances between the reference plane and the points F, B, R_1 and R_2 respectively. Figure 17b illustrates how we estimate these distances. Note that R_1' and R_2' are the intersection between the line FB and the perpendicular lines to the reference plane. These lines can be sketched from the vanishing point of parallel lines in that direction while the vanishing point can be found based on sticky-tape lines on the floor.

b) **Accuracy of the corrected heights :** To measure the accuracy of the corrected heights, we estimate the height of each subject from a single frame which is randomly selected from the video. Figure 18 shows the statistical histogram of the errors in the estimated heights, as the height of the top point or as the height obtained by the correction height algorithm, with respect to the static heights. The estimated height of the head

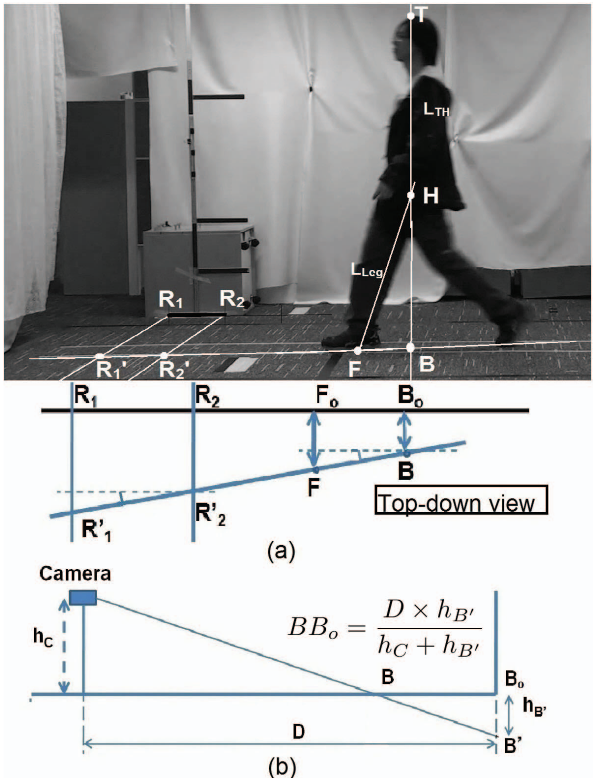

Fig. 17: (a) Height correction algorithm: a corrected height is computed as the sum of L_{Leg} and L_{TH} which are estimated separately, see text for more details; (b) Illustration of measuring the distance between point B on ground and the reference plane: D is the distance between the camera and the reference plane, h_C is the camera height and $h_{B'}$ is the height of the point B' on the reference plane which is measured based on the cross ratio method (detail discussed in Section II). The same process is applied to measure the distance from F, R_1 and R_2 to the reference plane.

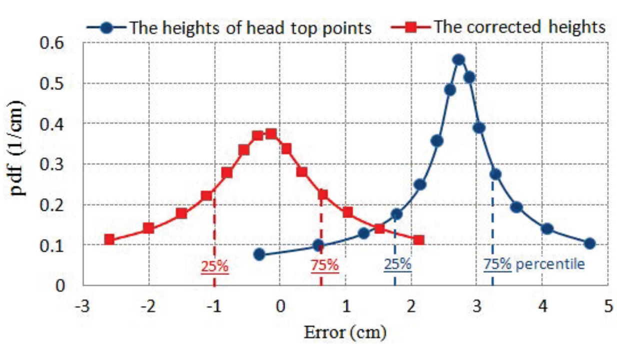

Fig. 18: Histogram of errors in estimated heights from a single frame: The graph shows the distributions of the errors in the estimated heights with respect to the static heights, in which the blue circle-marked curve corresponds to the estimated heights obtained as the height of the head top point and the red square-marked curve corresponds to the estimated heights obtained by using the above correction algorithm. Note that the error is computed by subtracting the estimated height from the static height. The positive value of error indicates that the static height is higher than the estimated height, otherwise the static height is lower. The vertical dashed lines correspond to the 25% and 75% percentiles of the distribution.

top point is observed to be systematically 2.75 cm less than the static height. The RMS error in the head top point height with respect to the static height is 2.71 cm. This systematic error is almost completely removed if we use the height correction algorithm discussed above, where the errors in the corrected height with respect to the static height are normally distributed with a mean of -0.29 cm. The RMS error in the corrected height is 1.19 cm, which is 2.2 times more accurate than the estimated height obtained as the height of the head top point.

V. Conclusion

Our experiments with 39 subjects show that static human height can be accurately measured, even though the variation of human height is significant during human motion. For the height measurement from video, the highest point in a human stride is proved to be the most accurate measurement of the static height. It can be measured as the average of the maximum heights or as the highest point in the sinusoidal model of the height variation. The two approaches give comparable results but the highest point of the sinusoidal model is observed to be more accurate if the number of image frames is limited. The limits of accuracy in the different approaches are summarized in Table I. For the height measurement from a single frame, a height correction algorithm based on the length of the human torso with human head and the length of the human leg leads to an estimation of height accurate to within 1.19 cm. It is also observed in our experiments that the different methods of height estimation (ours and Criminisi's [1]) give comparable results.

TABLE I: SUMMARY OF **RMS ERRORS** IN DIFFERENT APPROACHES

Measurement from video	Method	
	Cross ratio	Vanishing points
Average heights in all frames	1.49 cm	1.89 cm
Average of maximum heights	0.63 cm	0.62 cm
Highest point on sinusoidal model	0.62 cm	0.81 cm
Data from the first stride		
Average of maximum heights	0.91 cm	1.02 cm
Highest point on sinusoidal model	0.79 cm	0.92 cm
Measurement from a single frame		
Height of the head top point	2.71 cm	
Corrected height	1.19 cm	

Acknowledgment

The research was supported by the Australia National University. We also would like to thank NICTA staffs and students for their supports by participating in our experiment.

References

[1] A. Criminisi, A. Zisserman, L. Van Gool, S. Bramble, and D. Compton, "A new approach to obtain height measurements from video," in *Proc. of SPIE*, vol. 3576, Boston, Massachussets, USA, November 1998.
[2] A. Criminisi, I. Reid, and A. Zisserman, "Single view metrology," *International Journal of Computer Vision*, vol. 40, pp. 123–148, 2000.
[3] C. Madden and M. Piccardi, "Height measurement as a session-based biometric for people matching across disjoint camera views," in *IEEE Conference on Image and Vision Computing*, 2005, pp. 282–286.
[4] C. Springer, *Geometry and Analysis of Projective Spaces*. San Francisco, USA: W. H. Freeman and Company, 1964.
[5] C. I. of Technology. (2012, jan) Camera calibration toolbox for matlab. [Online]. Available: http://www.vision.caltech.edu/bouguetj/calib_doc/
[6] J. R. Taylor, *An introduction to error analysis*, 2nd ed. Sausalito, California, USA: Univeristy Sceince Book, 1992.
[7] MathWorks. (2012, jun) Interactive curve and surface fitting::interactive fitting (curve fitting toolbox^{TM}). [Online]. Available: http://www.mathworks.com.au/help/toolbox/curvefit/brviv3f-1.html/