

Recognition and Tracking of the Members of a Moving Human Body*

Costas Panagiotakis and Georgios Tziritas

Department of Computer Science, University of Crete
P.O. Box 2208, Heraklion, Greece

Abstract. We present a method to solve the human silhouette tracking problem using 18 major human points. We used: a simple 2D model for the human silhouette, a linear prediction technique for initializing major points search, geometry anthropometric constraints for determining the search area and color measures for matching human body parts. In addition, we propose a method to solve the problem of human members recognition and 18 major human points detection using silhouette. This result can be used to initialize a human tracking algorithm for real time applications. Our main purpose is to develop a low computation cost algorithm, which can be used independently of camera motion. The output of the tracking algorithm is the position of 18 major human points and a 2D human body extraction. In cases of low quality imaging conditions or low background contrast, the result may be worst. For these cases we defined an appropriate criterion concerning tracking ability.

1 Introduction

The analysis of human motion using computer vision techniques attempts to detect, track and identify people and recognize their activity. There are many applications in many areas such as analysis of athletic events, 3D virtual reality based on real motion, video conferencing, content-based image storage and retrieval, etc.

There has been a significant number of recent papers on human tracking. The following works give an overview of the various task involved in motion analysis of the human body. Aggarwal and Cai [1] focus on major areas relating to interpreting human motion, like motion analysis, body parts tracking and recognition of human activities using image sequences. In a more recent work, Wang, Hu and Tan [6] emphasize on three major issues of human motion analysis systems, namely human detection, tracking and activity understanding. According of them, there are 2D, with or without explicit shape models, and 3D approaches.

Firstly, we consider 2D approaches. The system called W^4 [3] uses a statistical-background model to locate people using stable cameras. W^4 uses a statistical-background model to locate people using stable cameras. W^4 allows multiple person groups and detects and tracks six main body parts of each

* This work was supported by the DXH project.

person using a static-shape model. In our silhouette analysis method we use silhouette statistics like silhouette projections as in the W^4 . Cheng and Moura [2] represent a system that tracks a walking human with monocular video using a 2D model, motion and texture. The human walker can be shaped by a stick model with 12 cone-shaped body parts. The above 2D models are similar to our 2D model mainly in that they both use human joints. Wang, Ning and Tan [7] propose a method to recognize and track a walker using 2D human model and both static and dynamic cues of body biometrics. Kakadiaris and Metaxas [4] present a 3D model-based method for motion estimation of human movement from multiple cameras.

1.1 Objectives

The main objectives of the silhouette analysis method is the the 18 major human points detection using silhouette image. In this problem, there are many solutions because of the projection from 3D in 2D image plane. Our method provides only one of them without using color information. The method is executed automatically, independent of silhouette rotation, scaling, in complex (with overlapping parts) - noisy silhouette, and without any initial information about the silhouette. The 2D human silhouette tracking in the whole sequence, tracking the set of predefined major human points and constructing a 2D segmentation map from them in every image of the sequence. As input, we use color images from a single, uncalibrated and maybe moving camera. The acquisition of realistic human motion by combining anthropometric geometrical constraints, color segmentation using an initial learning method, and time series information using a prediction method. The efficient searching that reduces the complexity and speeds up the algorithm without accuracy loss, is an important feature of our algorithm.

The main hypothesis of our method is that the human head and main body are almost completely visible in the whole sequence. Also, the color of any separate human part should not be changed too much. Colors are represented in the *Lab* color system, because it is closest to human color dissimilarity perception and the distance between colors can be modeled by Euclidean distance. The EMD is based on a distance between histograms. The EMD allows partial matching with very good performance [5].

The paper is organized as follows. The section 2 describes the human silhouette analysis, containing the human members recognition and major human points estimation method. In section 3 the human tracking algorithm is examined. Finally, sections 4, 5 provide experiments results and the discussion.

1.2 Overview

In this section we are going to specify the data collection methods, the main algorithmic steps, the model adopted and the hypothesis used. The silhouette analysis method can be divided into two stages. In the first stage, human members are recognized. In the second stage, the 18 major human points are estimated using

the result of the stage one. Both of the stages have been developed according to the following rule. First, the head is searched, because it has a well defined and static shape. Next, we compute the position of the main body, because it can be easily determined using the head position. Finally, we can compute the position of the legs and the arms. The main steps of the human tracking method are the following.

- Initialization: Dominant colors extraction for every body part based on joint points position. It is executed only in the first image of the sequence.
- Tracking loop for every part sequentially: head, main body, legs and arms. For each subpart it is executed: Color Distances Computation, Visibility Computation, Position Prediction, Search by best matching.

The human tracking method is executed automatically apart from the initialization step. In order to minimize the complexity cost, a sequential algorithm is used for each part or subpart. This is required because the complexity would be $O(N^{36})$ for 18 points in the image plane, where N is the number of image points. Using an appropriate sequential algorithm, we can reduce the complexity to $O(N^2)$. First, we search for the head, because it is the most visible part, and it usually has specific and distinctive color and texture. If we estimate the head position correctly, then we can predict with sufficient accuracy the position of the main body, and finally we can compute the position of the legs and the arms. Following this sequential search, the search area of the body is significantly reduced.

Our method uses color images, because we want to use the whole color information to get better results. We use a 2D human silhouette model, which was created by the position of 18 major human points. Moreover, the human body is divided into 6 parts (Figure 3): head, main body, left arm, right arm, left leg and right leg. Some parts, like legs and arms, have articulated motion. For this reason we divide these parts into subparts. We use two subparts for every arm and three subparts for every leg. The 2D model that we use comprises an ellipse model for the head, and polygon models for the other parts.

2 Human Silhouette Analysis

2.1 Human Body Parts Recognition

In this section we examine the human body members recognition method. The human body is divided into the following six members: head, main body, left leg, right leg, left arm and right arm. The human silhouette pixels will be classified to one of the above members. An example of silhouette image is shown in Figure 1, the white pixels correspond to the human body, while the gray pixels correspond to the background. The member recognition algorithm is sequential. The more visible members are computed first decreasing the search space of the others.

First, the major human body axis is determined using central moments. The silhouette is rotated according to the major axis. The rotation center is the mass



Fig. 1. The second figure shows the silhouette of the human of the first figure. An example of successful execution of the end of head (blue point) localization method (third figure). The left arm, of the right figure, can be distinguished from the main body because the proportion of red boundary pixels is high (fourth figure)

center of the human. Now, the human is located vertically. Next, we estimate the end point of the head (X_d, Y_d) using the following iterative method (Figure 1). The (X_d, Y_d) point is initialized as the human body mass center (m_c, n_c) . The (X_d, Y_d) is changed dynamically having always the meaning of the mean of human points that belongs to the same line. So, the above method can be characterized as dynamic mean method.

Dynamic Mean Method

$$\begin{aligned}
 & X_d = m_c, \quad Y_d = n_c \\
 & \text{repeat } \{ \\
 & \quad X_d = X_d - 1, \quad Y_d = E_{(X_d, y) \in Man}(y) \\
 & \} \quad \text{until } (X_d - 1, Y_d) \in Background
 \end{aligned}$$

The head region will be placed in a rectangle defined by the point (X_d, Y_d) as the end of the head. The maximum height and width of the head are proportional to the human height which can be easily estimated by the silhouette. The two shoulder positions, that will define the down limits of the head region, are determined by the first local minimum of the left and right horizontal silhouette projections.

The main body region is computed using an iterative algorithm similar to the end point of head estimation method. The maximum height and width of the main body are proportional to the human height. Using the human boundary, it can be determined if the arms can be distinguished from the main body (visible arm) and if the legs are distinguished. The rule is the following: if the proportion of the boundary pixels whose the closest background pixel is on the right, computed in an area where is located the left arm, exceeds a threshold, then the left arm can be distinguished from the main body. The above value is the ratio between the red pixels and the red plus green pixels of Figure 1. This knowledge helps in definition of the main body limits.

The legs and arms regions estimation can be done in the same time. An initial approximation is computed using the mass center of the human and the border point D that discriminates the left and right leg. Figure 2 shows an example of

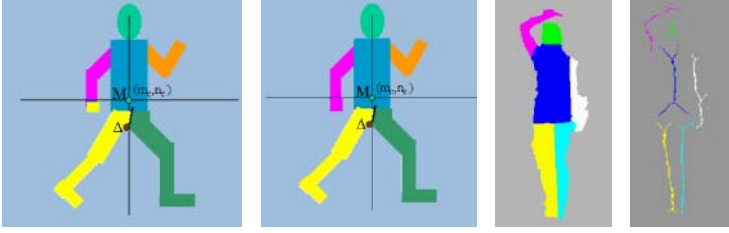


Fig. 2. The initial approximation and the final segmentation result (figures on the left). Human parts and their skeletons (figures on the right)

the initial segmentation. It is possible that more regions than six, that is the number of the human members, will be created. For these cases, we determine first which are the fault regions using the surface and the mass center of the regions. Next, we join the faults regions to the regions which have the most common boundary points with the faults regions. In the Figure 2, the false left leg region is correctly classified to left arm member.

2.2 Major Human Points Estimation

In the second stage, the 18 major human points are estimated using the segmented to human members silhouette. The method uses the skeleton of each human member in order to reduce the computation complexity, as the joint points are located to skeleton points. The skeleton is defined as the set of points whose distance from the nearest boundary is locally maximum. An example of human members skeletons are shown in the Figure 2. The points are computed sequentially. The easier defined points are computed first decreasing the search space of the others.

First, the center of the head is computed as the mass center of the head region. The neck point is defined as the mean of the boundary between head region and main body region. The two shoulders points are computed by minimizing an appropriate function F . The function domain is an isosceles triangle whose vertex is the neck point and its base vertexes are the two shoulders points. The function is minimized when the triangle base is maximized and the triangle height is minimized at the same time. The 18 major points formulations are defined in Figure 3. The points (9), (10) of the main body are computed using the main body height and width. The points (11), (15) of the main body are defined by the mean of boundary between main body region and left or right leg region respectively. The ankle point is computed first. We compute the farthest point K of skeleton points from point (9) using one line segment that should belongs to silhouette. The ankle point A is defined as the farthest point, of the not visited skeleton points, from K using one line segment. The knee point K is estimated by minimizing the function $G(X)$ which is defined by the following equation. Let F be the point (9) of the main body. Let the function $d(X, AF)$ be the minimum distance of point p from the line segment AF . $G(X) =$



Fig. 3. The 18 major human points, the 2D human model, the left leg distance image of the right image

$(|XF| - |XA|)^2 - 0.2 \cdot d^2(X, AF)$. If the point X is located close to the middle of AF and close to the knee angle at the same time, then the function G will be minimized. Finally, the end of leg point E is computed using the knee and ankle points. The E point search area is the left leg skeleton points which are close to the ankle point. The E point is defined as the skeleton point, whose distance from the point K is maximum. In each arm, we have to compute two points, the elbow point and the end of arm point. The computation of the left arm points is similar to the computation of the right arm points.

3 Human Tracking Method

3.1 Color Distance Images

The color distance image is an important measure used in human tracking method. We compute them for every part and subpart. The gray value at a given pixel of the color distance image corresponds to the minimum distance between the color of the considered pixel and the nearest color among the dominant colors of the part or subpart, where the point is assumed to belong. The dominant colors are computed once during the initialization process, where the segmentation of the body parts can be user guided. In Figure 3 an example of color distance image is illustrated.

3.2 Threshold – Visibility Computation

Our method is based on color segmentation using the color distance image. Hence a threshold is needed. In addition we define a visibility measure for evaluating the ability of a given part detection in its changing background. In this section, we examine the automatic threshold and visibility measure computation. These computations are executed for every new frame of the sequence. The algorithm is related to color information, so the color distance images are used. We need to update these values, because the image colors probably could be changed (shadings, color effects). The threshold and the visibility information computation are unsupervised. These methods use the dominant colors and the color distance images of the part or subpart. The threshold corresponds to the maximum value

of the color distance image for a pixel belonging to the considered part. First, the histogram (H) of the color distance image is computed. The threshold T_g and the visibility measure V_g for a part or subpart g depend on the distance distribution.

Let I_g be the Color Distance Image of part g and S be the part surface in pixels. Let us define two thresholds W_1 and W_2 , so that the pixels of the color distance image that have lower value than W_1 , probably correspond to pixels of the foreground, else, if they have higher value than W_2 , probably correspond to pixels of the background. So, the maximum of T_g is W_2 and its minimum value is W_1 . A constant c_g , with values between 2 and 4, depending on the part size, is used in the definition of T_g . Let us define first the cumulative empirical probability: $f(n) = \sum_{k=0}^n H(k)$. Then the two limit values are defined as follows:

$$W_1 = \min_t(t : f(t) > S) \quad W_2 = \min_t(t : f(t) > c_g \cdot S) \quad (1)$$

Finally, we define the threshold value T_g automatically as the mean of W_1 and W_2 . The visibility measure (V_g) is relevant to the ability of a human part detection (foreground/background contrast). It is defined separately for every part. The visibility measure could take any real positive value. If a part has null visibility measure, then it will be almost impossible to extract it.

The main steps of the visibility measure (V_g) computation are the following. At first two average values, v_1, v_2 , are computed,

$$v_1 = E(n : f(n) < S) \quad v_2 = E(n : ((c_g - 1) \cdot S < f(n) < c_g \cdot S)) \quad (2)$$

These values play the same role as the W_1 and W_2 for the threshold determination. The visibility measure of a part is proportional to the difference between v_1, v_2 (Equation (3)). When v_1 takes values higher than 6, then the part in the current image will probably differ from its dominant colors. So, the part visibility measure should be decreased (third factor in Equation (3)). The second factor in Equation (3) has similar meaning as the third.

$$V_g = (v_2 - v_1) \cdot (1 - e^{-(\frac{v_2}{9})^2}) \cdot e^{-(\frac{v_1}{9})^2} \quad (3)$$

If a part has visibility measure more than 2, from our experience we conclude that it will have high contrast against the background.

3.3 Initialization, Prediction

The initialization step is executed only in the first image of the sequence. The user has to give the position of the predefined 18 major human joint points (Figure 3) in the first image of the sequence.

The algorithm is the following. Using these positions in the first image we construct a 2D shape of the human to get the color information of every body part and subpart. The 30 dominant colors of each part and the 15 dominant colors of every subpart are computed using the Self-Organizing Map (SOM) neural network. An appropriate coordinate system is used in order to utilize the



Fig. 4. The head mask, a head tracking result, the body mask and the 3 masks that are used in arm tracking

human motion, with the center of the head as the coordinate center. Relative coordinates are used for the main body points (7 points). We use angle and amplitude, which determine the positions of joints (10 points). The angle and the amplitude of each joint are computed by using its neighbor joint which is closer to the main body.

The prediction model is described following. This model is used in every image of the sequence. The prediction of the position for the 18 major points is computed using the positions of these points in the two previous images of the sequence. The prediction of the 7 main body points (in the relative coordinate system of the head) is given by the positions of these points in the previous image. For the other 10 joint points, we use a fixed coefficient first-order filter to predict the angle in a relative coordinate system. Concerning the prediction of the part length it is likely to be near that in the previous frame.

3.4 Head Tracking

At first, we search for the head because it is the most visible part and it usually has specific color and texture. The head is defined in the 2D model by two points, the center of the head and the neck point. Because of the 2 degrees of freedom of every point, the complexity of the head tracking would be $O(N^4)$. We use a sequential algorithm to reduce this complexity to $O(N^2)$.

According to the 2D model the position of the center of the head determines the translation and the position of the neck point determines the rotation and the size of the head. In most of the cases the motion due to translation is stronger than that due to rotation and scaling, so the points could be computed sequentially. First, we compute the center of the head and then the neck point. We use the same minimization criterion for the two cases.

Instead of minimization we use a matching function which is maximized. This function is defined in the color distance image of head using the head mask (Figure 4). The mask is defined by the center head point and the neck point, as in the 2D model. The mask consisted of an ellipse, covering a semi-ellipse. The semi-ellipse corresponds to the head background and the ellipse corresponds to the head.

First, we shift the mask in order to find the center of the head. The matching function is maximized when the head pixels belong to the ellipse (points with low color distance) and the background head pixels belong to the semi-ellipse (points with high color distance). This position will give us the center of the

head. The I_h is the head color distance image. Let F_h be the set of the head points and B_h be the set of head background points according to the head mask.

The function h is defined by the percentage of the head points according to the mask, which have color distance lower than T_{head} . The function b is defined by the percentage of the background points according to the mask, which have color distance lower than T_{head} . When h or $1 - b$ becomes low (close to 0) then the matching is extremely wrong. Otherwise, when h and $1 - b$ are close to 1 the matching is satisfactory.

$$f_h = h \cdot (1 - b) \quad (4)$$

By experiments, the best results are achieved when h is high and b is low at the same time. This is done when function f_h is maximized. If we assume that the head model and background model are independent, then the function f_h is proportional to the possibility of matching both of them.

The neck point is computed following the same procedure. The mask is rotated and scaled in order to maximize the matching function. The point that maximizes the matching function will be the neck point.

3.5 Main Body Tracking

In this section we examine the method of 6 main body points tracking (the neck point was computed in the head tracking method). We use relative coordinates towards to the center of the head because the relative coordinates change slowly. We use a sequential-iterative algorithm in order to reduce complexity. The algorithm is described in the following. A point p is chosen among the 6 main body points. The point p is moved in a small region around its position. The new position of p is the position which maximizes the matching function f_b . The algorithm is completed when the points stop oscillating or the number of iterations exceeds a prefixed number of iterations.

The mask that we use is shown in Figure 4. The heptagon points correspond to the main body pixels. The two triangles correspond to the main body background pixels. The heptagon and the two orthogonal triangles are defined by the 7 points of the main body. The matching function (f_b) that we use is similar with the matching function in the head method.

3.6 Leg Tracking

In this section we are going to examine the method for tracking the three leg points. The algorithm is the same for the left and the right leg. We use a sequential algorithm because the normal computation would have $O(N^6)$ complexity. The steps of the algorithm are the following. First, the thigh angles are tracked. Next, shoes and knee angles are tracked. Finally, knee points are tracked.

Thigh Angles Tracking The first step of the method is similar to that of the previous methods. We search the left and the right thigh angle in a search space that is defined by the predictions of the angles. We use the masks of



Fig. 5. The mask of the left thigh angle, the right thigh angle, the left leg, the right leg and the 3 masks that are used in arm tracking

Figure 5 where the thigh point and knee point are represented as circles. The Color matching error is minimized when yellow pixels of the mask correspond to pixels with low left leg Color Distance (left leg) and the red pixels correspond to pixels with high left leg Color Distance (left leg background).

Shoes and Knee Angles Tracking The second step of the method uses the visibility information to combine the shoe color information and the second leg subpart color information. In this section we compute the position of the 2 points of the shoes and the angle of the knee using the masks of Figure 5. The color matching error is minimized when the green pixels of the mask correspond to pixels with low Left leg Color Distance (left leg), the blue pixels of the mask correspond to pixels with low Left shoe Color Distance (left shoe) and the red pixels correspond to pixels with high left leg Color Distance (background). First, we search for the position of the ankle point and the angle in a long area with high searching step. The next step is to search the above features with a lower searching step using the previous method results. Finally, we search for the leg member (the angle and the amplitude of the last leg point), which is changed slowly. If the minimum matching error is lower than a threshold, this means that the shoe is hidden. Let f_s be the matching function (f_s). The I_s is the shoe distance image and the I_l is the second subpart of leg color distance image. Let F_l be the set of the leg points and F_s be the set of shoe points and B_l be the set of leg background points, according to the leg mask. Let l_2 be the second subpart of leg.

The function S is defined by the percentage of the shoe points, according to the mask, which have color distance lower than T_{shoes} . The function L_2 is defined by the percentage of the leg subpart (left or right) points, according to the mask, which have color distance lower than T_{l_2} . The function B is defined by the percentage of the background points, according to the mask, which have shoe color distance lower than T_{shoes} and second subpart of leg color distance lower than T_{l_2} . When B becomes low (close to 0) then the matching is satisfactory. Otherwise, when B is close to 1, the matching is extremely wrong. Function H is defined by the linear combination of S and L_2 .

$$H = \frac{V_{shoes} \cdot S + V_{l_2} \cdot L_2}{V_{shoes} + V_{l_2}} \quad (5)$$

So, H contains the matching information from two models synthesis, while B contains the matching information from one model (background model). A func-



Fig. 6. Final results of human silhouette analysis and of 2D human tracking

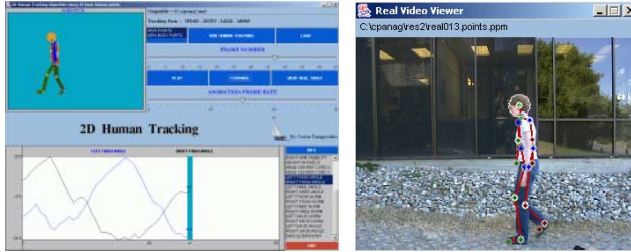


Fig. 7. The Java interface and the Java Real Video Viewer

tion f_s , similar to f_h (Eq. (4)) is maximized, $f_s = H^2 \cdot (1 - B)$. The last step is the knee points tracking. This step is trivial because the knee angle and the thigh angle have been estimated.

3.7 Arm Tracking

In this section, we examine the arm tracking algorithm. The method is the same for left and right arm. We have to compute two points in every arm.

The method will track the arm, if the arm and the main body are not superposed. If the arm is hidden, the method will detect that by high matching error. The algorithm has the following steps. First, the shoulder angle is tracked. Next, the elbow angle and the shoulder point - elbow point distance are estimated. Finally, the elbow point - arm member point distance is estimated. Figure 5 shows the three masks used. The Color matching error is minimized when white pixels of the masks correspond to pixels with low arm Color Distance (arm) and red pixels correspond to pixels with high arm Color Distance (arm background).

4 Experimental Results

The human body members recognition and 18 major human points estimation methods have been tested in more than 100 noisy silhouettes. The algorithm succeeded to recognize the human body members in all the images with high accuracy. The mean error in head, main body major points estimation is about 2% of the total human height, while the arms - legs major points are estimated with less than 200% of the previous error. In Figure 6, some results of human silhouette analysis are shown. The method was implemented in C. The complexity of the total algorithm is $O(N)$ ($N = \#$ pixels of the human). If we did not use skeletons we would have a complexity of $O(N\sqrt{N})$.

The human tracking algorithm has been tested in many video sequences of walking or running people. The quality and analysis of images that we used as input varied from CIF format images with low quality (MPEG1) to high quality images (MPEG2). The complexity of the algorithm is $O(N^2)$ ($N = \#$ pixels of the human). The frame processing rate was about 1 frame per second in Pentium III 800 MHz, using 288×320 images. The results include the position of the 18 major human points in every image of the sequence and 2D segmentation into human parts and subparts using the 2D human model. We examined how the unstable background, the video quality and the image size can affect the final result.

In Figure 6, some results for three sequences are shown. In two of these sequences, the person walked in an unstable background (the camera was moving). In the third sequence, where the person was running and the camera was moving to track him, the arms were not tracked because of their low imaging quality. The algorithm accurately tracked the human in the above sequences.

The method was implemented in C. Also, a Java based interface has been developed for our experiments (Figure 7). The user has many options, like load old tracking results, execute human tracking algorithm, selection of features that will be plotted, viewing of real sequence images, etc.

5 Discussion

In this paper, we have proposed an effective sequential algorithm for human body members recognition and 18 major human points estimation using silhouette. If the estimation of a major point is wrong then the other major points, that use the false point, will be computed wrong. As the results have shown, this is impossible to happen because the points, that are computed first, are the more visible points and its estimation decrease the searching area and the computation cost of the other points.

Also, we have proposed a fast and effective algorithm for human silhouette tracking in video sequence. The method needs the position of the 18 major human points in the first image. The method also detects if any human part is hidden (arm - body, shoes) by using high matching errors. The head tracking is always correct, because of the head's simple shape and motion. The head

is usually the most visible human part. The main body is more difficult to be tracked than the head, because, usually, there are more colors in the main body or the arms may cover it. The most difficult parts to be tracked are the arms and legs, because of their complex motion. In many frames they are hidden and their pixels are smoothed because of motion.

The developed method could be expanded to track more than one person. We can also avoid the initialization step by using background subtraction methods and independent motion detectors. The above methods give us a bitmap image of a moving object (silhouette). Using the extraction method of the 18 major human points from the silhouette we can get the initial position of 18 major human points. One possible extension of the developed method is 3D tracking and the extraction of motion parameters using a stereoscopic system. Another extension can be human-activity recognition (walking, sitting, running, etc). Security system and statistics analysis of human motion systems could be based on our method.

References

- [1] J.K. Aggarwal, Q. Cai. Human Motion Analysis: A Review, *Computer Vision and Image understanding*, vol. 73, no. 3, pp. 428-440, March 1999. 86
- [2] J.C. Cheng, M.F. Moura. Capture and Representation of Human Walking in Live Video Sequences, *IEEE Trans. on Multimedia*, vol. 1, no. 2, pp. 144-156, June 1999. 87
- [3] I. Haritaoglu, D. Harwood, D. Davis. W^4 : Real-Time Surveillance of People and Their Activities, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 809-830, Aug. 2000. 86
- [4] I. Kakadiaris, D. Metaxas. Model-Based Estimation of 3D Human Motion, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp. 1453-1459, Dec. 2000. 87
- [5] J. Puzicha, J. Buhmann, Y. Rubner, C. Tomazi. Empirical Evaluation of Dissimilarity Measures for Color and Texture, *1996 IEEE*, 1999. 87
- [6] L. Wang, W. Hu, T. Tan. Recent developments in human motion analysis, *Pattern Recognition*, vol. 36, no. 3, pp. 585-601, 2003. 86
- [7] L. Wang, H. Ning, T. Tan. Fusion of Static and Dynamic Body Biometrics for Gait Recognition, in *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 14, no. 2, pp. 149-158, Feb. 2004. 87