

Colour Based Human Motion Tracking for Home-based Rehabilitation*

Yaqin Tao

Dept. of Computer Science, University of Essex
Wivenhoe Park, Colchester CO4 3SQ, U.K.
Email: ytao@essex.ac.uk

Huosheng Hu

Dept. of Computer Science, University of Essex
Wivenhoe Park, Colchester CO4 3SQ, U.K.
Email: hhu@essex.ac.uk

Abstract - *In this paper, we propose a colour based tracking algorithm for capturing the motion of human body parts for home-based rehabilitation. Different colour belts are attached to the body joints of interest and tracked in the video sequence. The performance of the colour-based tracking algorithm is analysed by comparing the tracking results with the results from a commercial marker-based tracking system, Qualisys. The experiment results show the efficiency of our method in tracking simple target reaching motion of human arms.*

Keywords: Human motion tracking, target reaching, rehabilitation.

1. Introduction

Traditionally patients who sustain a stroke take physiotherapy with the help of physiotherapists or well-trained carers to diagnose if they are doing the rehabilitation properly. However, because of the staffing shortages at the Nation Health Service and the need of a prolonged period of time for the rehabilitation exercise, patients are not receiving enough treatment. We propose to develop a visual tracking system to support the rehabilitation program for the patients at home environments so that the burden of hospitals and the physiotherapists could be relieved.

The standard methods for clinical human motion analysis are marker-based motion tracking method. There are a lot of commercial marker-based human motion capture systems that can be employed for tracking the patients' motion, such as the Qualisys and CODA. However, apart from the difficulties of calibrating both cameras and markers, these systems are too expensive for a daily deployment by stroke patients at home, and too complicated for physiotherapists to interpret the tracking results of patients' motion.

Recently, a lot of attempts have been made to design a marker-free tracking system for the human motion capture [7] [9] [14]. Marker-free tracking systems are very attractive because only conventional cameras are needed instead of special cameras and intrusive markers. However, designing a video system for tracking human motion is a non-trivial task. There are a number of difficulties [4],

including depth ambiguities, occlusion and kinematics singularities, etc. In order to simplify the human motion tracking problem, most human motion tracking algorithms employ a shape model of the subject to support the tracking, others use multi-cameras to improve the robustness. The shape model of a subject varies from the simplest skeleton model to 2D patches models [13], and to the sophisticated 3D volumetric models [6] [8] [4] [10].

Most human motion tracking methods can be formulated into a two-step framework [9], namely feature extraction and feature correspondence. The first step is to extract the image features from each image frame. The most widely used image features are edge, optical flow, silhouette, contour, and template. Others like feature points, colour [5], blobs are also used in motion tracking. And then the correspondence is built up by matching the 2D image features to the model data. The correspondence analysis is often supported by prediction algorithms such as Kalman Filter and Condensation [11] (Particle Filter). Hogg used edge information to track a walking person [6]. The body segments of the subject were represented by cylinders and connected at joints. Bregler et al. used the twist and exponential map to track whole body motions within an optical flow framework [3]. Sidenbladh et al. used the optical flow as the tracking primitive and a generative model of image appearance to track human motion in 3D by using a stochastic method [8]. Isard et al. used contour to track different moving objects such as the hand, the upper body of a human and a leaf [11].

The home-based rehabilitation requires a visual-tracking system that should be cheap, accurate and run in real time. Both the marker-based tracking method and the video-tracking (marker-free) method can only fulfil parts of these requirements. In other words, the marker-tracking system can provide the required tracking accuracy, but the system is too expensive for home-based rehabilitation applications. On the other hand, a video-tracking system is relatively cheaper but its robustness and accuracy need improvement. Therefore, it becomes necessary to build a tracking system that has the advantages from both tracking systems. In this paper, we propose to develop a visual tracking algorithm for home-based rehabilitation in order to achieve real-time performance and sufficient accuracy.

The rest of the paper is organised as follows. Section 2 introduces the proposed colour-based human motion tracking method. The implementation of the tracking system is described in Section 3. Our experimental results on the accuracy test of the colour tracking method are presented in Section 4. Finally, conclusions and future work are presented in Section 5.

2. Colour Based Motion Tracking

In this section, we introduce our current work on developing the colour-based motion tracking method.

2.1. Method

Video cameras are employed in our tracking system to capture the subject's motion. In order to track the human motion accurately, different colour belts are attached to different body joints of the subject (see Figure 1). The task of human motion tracking is now simplified to track the colour belts in the input image sequence and the estimated centres of these colour belts are regarded as the joint positions of the corresponding human body parts. The idea of using colour belts attached to the joint positions is inspired by the marker-based tracking system, which uses the markers to highlight the region of interest. We assume that the subject doesn't wear the clothes that have similar colour with the colour belts, and background is less cluttered. Similar work has been done by Zhuang et al.[15]. The major difference is that our system relies solely on colour features, while Zhuang used both colour and edge features.

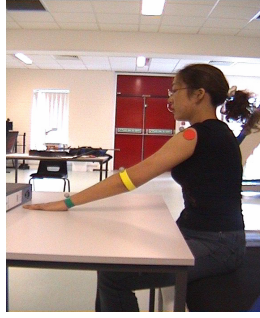


Figure 1 Colour belts

2.2. Colour Model

Colouring of interesting body joints partially simplifies the visual tracking problem and allows fast processing, but colour tracking is sensitive to ambient lighting change. It is important to choose a proper colour space, which can make the colour-based tracking robust to ambient lighting change. There are lots of different colour spaces that have been used in colour-based tracking systems, such as the normalized RGB $[r, g]$, $[H, S, V]$ and $[Y, Cr, Cb]$. According to the comparison results in [2], HSV gives relatively better tracking performance among all the colour spaces. So it is used in our tracking algorithm. For each objective colour k , a region of interest is selected and used to create a colour model M_k before tracking. The following three equations are used to convert RGB to HSV.

$$H = \arccos \frac{\frac{1}{2}((R-G) + (R-B))}{\sqrt{((R-G)^2 + (R-B)(G-B))}} \quad (1)$$

$$S = 1 - 3 \frac{\min(R, G, B)}{R + G + B} \quad (2)$$

$$V = \frac{1}{3}(R + G + B) \quad (3)$$

2.3. Colour Tracking

Colour tracking is essentially a two-step tracking procedure: (1) Segment target objects from each image frame. (2) Build up the correspondence of the same object between consecutive image frames.

In each image frame, a search window W for each colour belt is used to save computational time. The colour models M are actually used as the decision rules to classify image pixels to one of the objects or background. A point P in a search window belongs to the colour belt k if its colour components $f(P_H, P_S, P_V)$ meet the colour model M_k .

$$S_k(P) = \begin{cases} 1 & \text{if } f(P) \in M_k \\ 0 & \end{cases} \quad (4)$$

where S_k is the set of all the pixels belonging to the same object k .

The mean image location of all the pixels is calculated and considered as the position of the corresponding body joint where the colour belt is attached. And it is also used as the centre of the search window in the next frame. If it fails to find the colour belts in the original window, the search window is increased and the search method is repeated. If the algorithm fails to find the colour belts in the whole image (the colour belt may be occluded by other body parts), the joint position in the previous frame is used as the joint position in this frame. This is the simplest strategy. There are more sophisticated algorithms available such as Kalman Filter and Condensation algorithm. Since the subject's motion in the rehabilitation application is slow and simply, this approximation is reasonable.

The correspondence problem is simplified in our tracking system because different colours are used to indicate different body joints. After using this tracking algorithm, we can obtain a 2D position trajectory of each colour belt.

3. Implementation

We have described the colour-based tracking algorithm in the previous section. It is important to know if the proposed tracking algorithm has the acceptable accuracy or has the potential for the rehabilitation project. Here, we chose the target reaching movement to test the accuracy of the colour tracking system. Target reaching motion is mainly the motion of the arm [12], so markers and colour belts are attached to the joints of shoulder, elbow, and wrist (see Figure 1). The original experiment involves only a single video camera, so the target reaching motion is constrained in 2D and performed in the plane parallel to the image plane.

3.1. Experimental Set-up

A marker based commercial system, Qualisys, is employed in our experiment for the comparison purpose. The results from our colour-based tracking system are compared to the results from the Qualisys. In order to improve the accuracy of the comparison and exclude the variation of the motion performed by the subject at different time, we capture the subject's motion by using a Qualisys and a colour tracking system simultaneously. Figure 2 shows the experimental set up viewed from the top side. The subject wears both the colour belts and the markers. The video camera and Qualisys cameras 1, 2 capture the subject's motion at the same time. The origin of the Qualisys coordinate system is also shown in Figure 2.

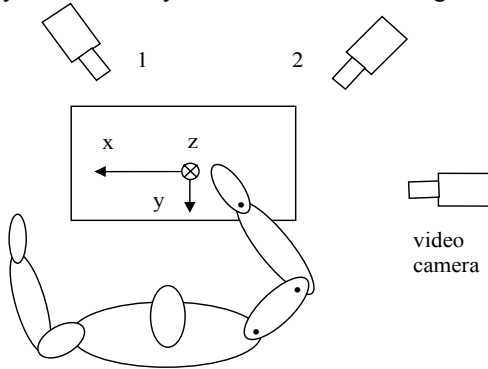


Figure 2 Top view of the experimental set up

3.2. The Qualisys Tracking System

The Qualisys system uses small retro-reflective ball markers, which are attached to the performer's joints (see Figure 3 (a)) and can reflect infrared light. It's easy to detect and locate these markers in the 2D images filmed by different synchronized cameras because they are bright points in the image planes (see Figure 3 (b) (c) (d)). Three digital cameras fixed in the scene capture the arm of the performer here.

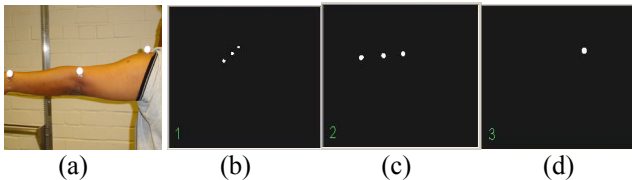


Figure 3 (a) Markers attached to the joints of the performer.
(b)(c)(d) Marker points captured from three cameras.

3D position of each marker (see Figure 4) is calculated from the corresponding 2D marker point in each image by using the stereo correspondence method. At least two cameras should be used in order to obtain 3D data. Each marker is tracked from one frame to the next frame and its 3D position trajectory of each marker can be obtained at this step.

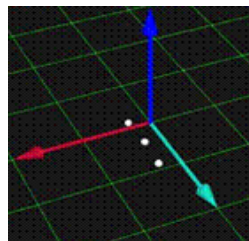


Fig. 4 Markers in 3D

4. Accuracy Analysis

A sequence of the target reaching motion, which lasts about 14.52 second, was captured in our experiments. The tracking result from the Qualisys is the 3D position trajectory of each marker, represented as $(x_m(1), y_m(1), z_m(1))$, $(x_m(2), y_m(2), z_m(2))$, ..., $(x_m(k), y_m(k), z_m(k))$, where m is the data from marker tracking system, and k is the frame index in the image sequence. Since we constrain the motion in 2D, the depth coordinates remain almost constant for each marker during the image sequence and can be ruled out.

Accordingly, the result from the colour tracking system is the 2D position trajectory of each colour belt, denoted as $(x_c(1), y_c(1))$, $(x_c(2), y_c(2))$, ..., $(x_c(j), y_c(j))$, where c represents the data from colour tracking system and j is the frame index.

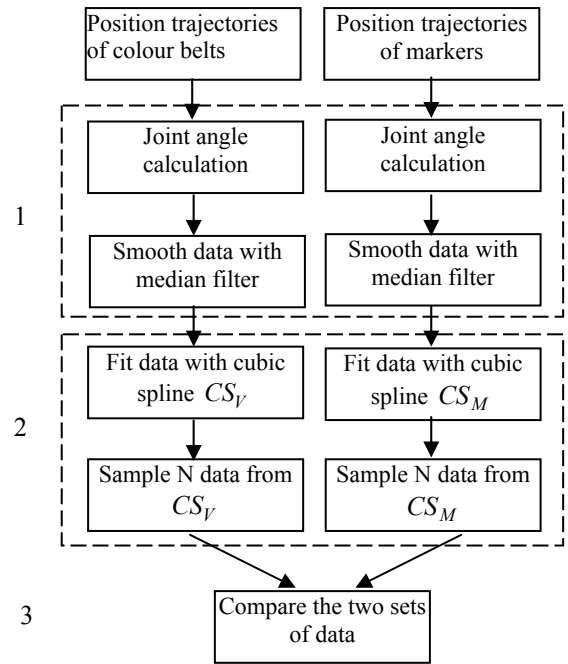


Figure 5 the procedures of data processing

Direct comparison between the marker data and the video data seems unreasonable because the two tracking systems have different coordinate system and units. A series of data processing procedures (see Figure 5) are employed to process the tracking results from each system, and finally use them to compare the tracking performance. These data processing methods are described in detail in the following subsections 1, 2, and 3.

4.1. Angle Calculation and Data Smoothing

In our experiment, we compare the joint angle trajectories instead of the joint position trajectories because of its invariant to the coordinate systems and units. Based on the 2D position trajectory of each joint, we can calculate the shoulder angle and the elbow angle trajectories for the target reaching motion. We define the shoulder angle Θ in the 2D motion situation is the angle between the upper arm and the horizontal axis. And the elbow angle Φ is the angle

between the upper arm and the forearm (see Figure 6). After calculation in each frame, we can get the joint angle trajectories of shoulder and elbow. They are represented as follows for the marker tracking system:

$$(\Theta_m(1), \Theta_m(2), \dots, \Theta_m(k)), (\Phi_m(1), \Phi_m(2), \dots, \Phi_m(k))$$

The procedures are the same for calculating the joint angles in colour based tracking system:

$$(\Theta_c(1), \Theta_c(2), \dots, \Theta_c(j)), (\Phi_c(1), \Phi_c(2), \dots, \Phi_c(j))$$

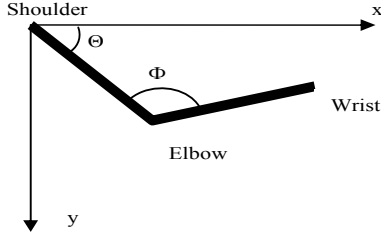
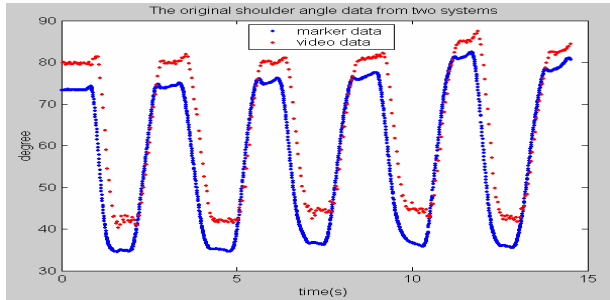
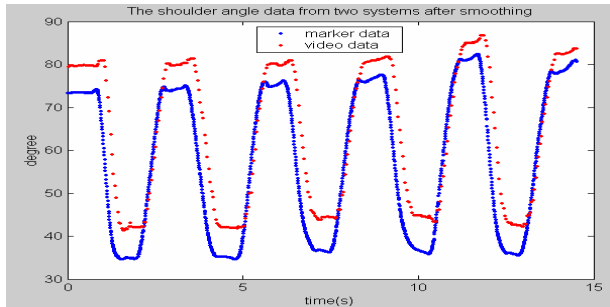


Figure 6 Shoulder angle and elbow angle in the image coordinate system

Since the processing are the same for both shoulder and elbow angle data, we will illustrate the procedures by taking the shoulder angle data processing for example in the following sections. Figure 7(a) shows the shoulder angle data obtained from each tracking system. They are calculated according to the method discussed above. In order to smooth the noise, these joint angle data are filtered by a median filter with size of 5 (see Figure 7(b))



(a) Shoulder angle data (marker & video)



(b) Shoulder angle data after smoothing

Figure 7 Shoulder angle data of target reaching motion

4.2. Data Fitting

As we can see from Figure 7, the number of frames of data for the same motion sequence captured by each system

is different. That's because the frame rates of cameras are different, (25f/s in the colour tracking system, and 120f/s in the marker tracking system). In this experiment, the number of image frames from colour tracking system is 363, and 1742 from marker tracking system. In order to compare the tracking results obtained from the two tracking systems, the two sets of data are fitted with a cubic spline [1] respectively.

Cubic spline interpolation is a piecewise interpolation and is very useful to interpolate between known data points (also known as breakpoints) due to its stable and smooth characteristics. It fits a different cubic polynomial between each pair of breakpoints. Considering the shoulder angle data, we get a collection of shoulder angle data $(t_0, \Theta_0), (t_1, \Theta_1), \dots, (t_{i-1}, \Theta_{i-1}), (t_i, \Theta_i), \dots, (t_n, \Theta_n)$, where t is the time variable of the captured motion sequence, n is the index of the number of frames of data and Θ_i represents the shoulder angle at time t_i . A third degree polynomial (see Equation (5)) is constructed between each pair of points (t_{i-1}, Θ_{i-1}) and (t_i, Θ_i) . If we can calculate the 4 unknown coefficients (a_i, b_i, c_i, d_i) for each third order polynomial, we then found a cubic spline for our shoulder angle data set.

According to the properties of a cubic spline that it is continuous, and its first and second derivatives are continuous as well at each breakpoint, equations (6), (7) and (8) can be formulated for each polynomial.

A third order polynomial, where $t_{i-1} \leq t \leq t_i, i \in [1, n]$

$$f_i(t) = a_i + b_i t + c_i t^2 + d_i t^3 \quad (5)$$

The cubic spline is continuous at each breakpoint.

$$f_i(t_i) = f_{i+1}(t_i) = \Theta_i \quad (6)$$

The first and second derivatives are continuous at each breakpoint.

$$f_i'(t_i) = f_{i+1}'(t_i) \quad (7)$$

$$f_i''(t_i) = f_{i+1}''(t_i) \quad (8)$$

According to the analysis above, we can get 4 equations for each cubic polynomial to solve 4 unknown coefficients. However, the first and last cubic polynomials don't have adjoining cubic polynomials beyond the first and last breakpoints, which means Equation (8) is not suitable for the end points. There are a lot of different existing methods to solve the problem. Here we use parabolic run-out splines, which means that the second order derivatives of the spline at the end points are the same as that at the adjacent points.

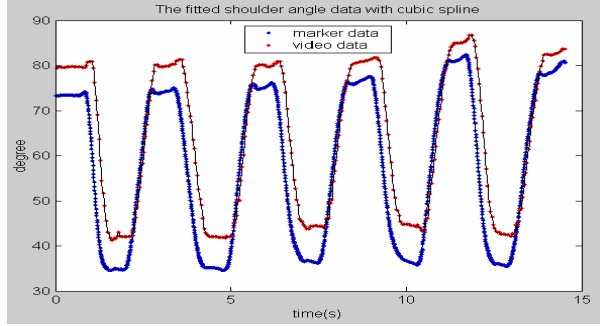
$$f_1''(t_0) = f_1''(t_1) \quad (9)$$

$$f_n''(t_n) = f_n''(t_{n-1}) \quad (10)$$

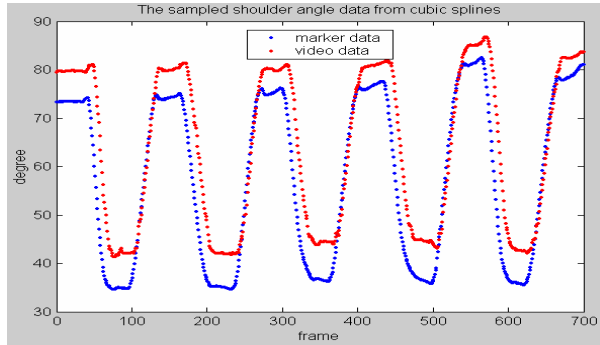
So the data fitting with a cubic spline becomes solving $4*n$ linear equations to get the $4*n$ unknown coefficients. After solving a large set of linear equations, the cubic spline curve of the shoulder angle data can be found. Figure 8 (a) shows the shoulder angle data from each system fitted with cubic spline curves.

We equally sample N data points from the corresponding spline curve. And the extracted data set in each system, which has the same number (700 data points in this

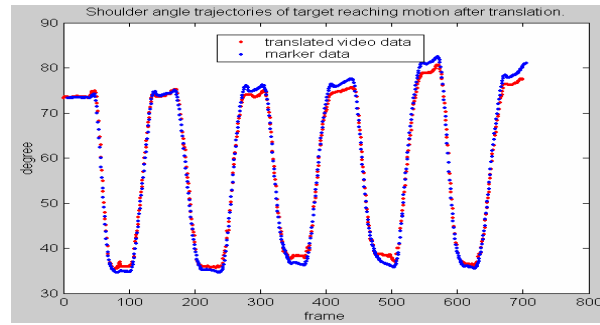
experiment), is used to compare the tracking results. Figure 8(b) illustrates the sampled shoulder angle data after data fitting.



(a) Shoulder angles fitted with cubic splines



(b) Re-sampled shoulder angle data after fitting



(c) Shoulder angle after translation

Figure 8 Shoulder angles of target-reaching motion

After the data processing procedures, these sampled data should be ready for comparison in theory. But as we can see from Figure 8(b), there is a translation transformation between the marker data and the video data. This is because of the layout difference of the markers and the colour belts on the arm (see Figure 1) and the synchronization problem between the two tracking systems. The markers and colour belts are attached slightly away from each other for the same body joint in order to be visible to the Qualisys and video cameras respectively. We can compensate the translation caused by the experiment set up by translating the video data towards the marker data (see Figure 8 (c)).

Figure 9 illustrates the elbow angle trajectories of this motion sequence after being processed by the same procedure as the shoulder angle trajectories. Note that there are total 5 cycles of the target-reaching motion in this

sequence of about 14.5s, so the frequency of the motion is about 2.90s/cycle.

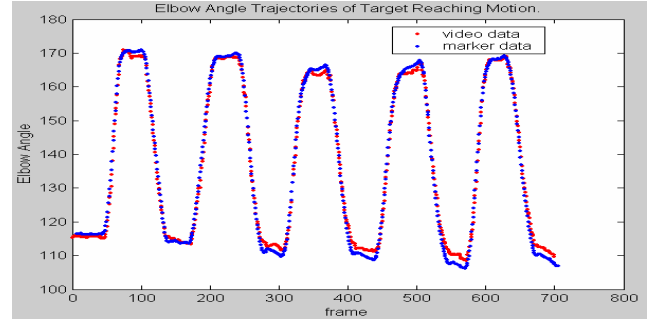


Figure 9 Elbow angle after translation

4.3. Comparison of the two sets of data

Two indices are used to evaluate the performance of two systems:

- The absolute difference

$$d_i = |x_i - y_i| \quad i \in (1, 2, \dots, N) \quad (11)$$

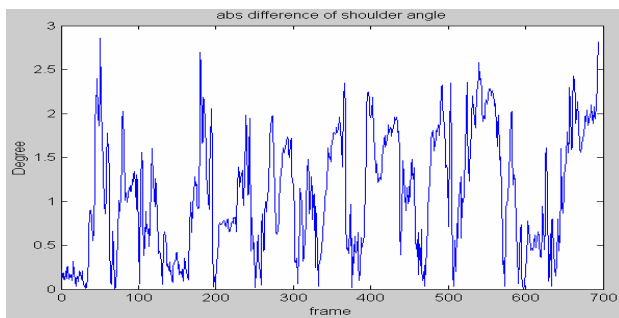
- The correlation coefficient r

$$r = \frac{\sum_{n=0}^{N-1} x_n y_n}{\sqrt{(\sum_{n=0}^{N-1} x_n^2 \sum_{n=0}^{N-1} y_n^2)}} \quad (12)$$

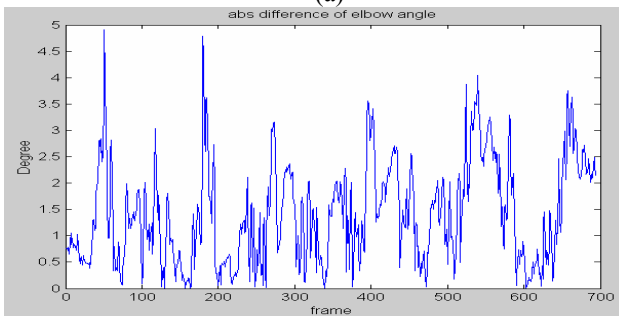
where x , y represent the sampled video data and sampled marker data respectively. N is the number of points in each set.

The absolute differences of the shoulder angle and elbow angle for the two systems are illustrated in Figure 10. Most of the differences for the shoulder angle are less than 2.5 degrees and less than 4 degrees for the elbow angle. The correlation coefficient is also employed to evaluate the similarity of the tracking results. According to Equation (12), the correlation coefficient for the shoulder angle is 0.9998, and 0.9999 for elbow angle. The correlation ratio (>0.99) by using our colour based tracking algorithm is much higher than the similar ratio (>0.95) in [1], which is the result based on the marker data and the medial axis transformation.

The target reaching motions with various speeds are also tested and the experimental results are shown in the Table 1. From table 1, we can see the colour based tracking algorithm does give promising tracking results. All the correlation coefficients are higher than 0.99. The absolute difference increases as the speed of the target reaching motion increases. However, even for the highest speed in our experiments 1.55s/cycle, which is faster than the daily life motion, the absolute difference for the shoulder angle is about 5 degrees and 8 degrees for the elbow angle. For the normal speed motion, 2.25s/cycle and 2.90s/cycle, the absolute difference is lower than 5 degrees. Since the rehabilitation motion of the patients is slower, or at most equivalent to the speed of the daily life, which means that the absolute difference is less than 5 degrees, the colour-based tracking algorithm is suitable for the rehabilitation project under the constraints assumed in this paper.



(a)



(b)

Figure 10 Absolute differences of shoulder & elbow angle

Table 1 Experimental results of the target-reaching motion

Speed	Absolute Difference		Correlation Coefficient	
	Shoulder Angle	Elbow Angle	Shoulder Angle	Elbow Angle
1.55s/c	$\leq 5^\circ$	$\leq 8^\circ$	0.9994	0.9998
2.25s/c	$\leq 4^\circ$	$\leq 5^\circ$	0.9997	0.9999
2.90s/c	$\leq 2.5^\circ$	$\leq 4^\circ$	0.9998	0.9999

5. Conclusions and Future Work

Traditionally, patients' rehabilitation needs the help of physiotherapists or expert carers. We proposed to develop a visual-tracking system for the patients who sustain a stroke to do their rehabilitation exercise at a home environment without the need of physiotherapists' presence. In this paper, the motion tracking module is developed by employing a colour-based human motion tracking algorithm which is simple and fast. Although only the constrained tracking motion of a human arm was currently investigated, it is a good starting point because tracking arm motion is very useful for the rehabilitation process.

To improve the performance of the colour tracking algorithm and make it more robust and reliable, the colour model problem and the occlusion problem will be investigated in our future work. An adaptive colour model is to be employed to deal with the lighting condition change in the environment. The occlusion problem is one of the most difficult problems in human motion tracking, which is to be dealt with by using multi-cameras and the Kalman Filter. Some kinematic constraints and motion constraints of the human body can also be used to solve the ambiguities occurred in the motion tracking.

Acknowledgments

The authors would like to thank Ms. Harun, Hafizah H, Dr Martin Sellens and Prof. Ralph Beneke in the Biological Sciences Department to allow us to use their Qualisys system. Our thanks also go to Dr Nigel Harris at Bath University and the other members of EPSRC EQUAL Smart Rehabilitation Consortium for useful discussion.

References

- [1] A. Bharatkumar, K. Daigle, Q. Cai, J. K. Aggarwal, "Lower Limb Kinematics of Human Walking with the Medial Axis Transformation". Proc. of Workshop on Motion of Non-Rigid and Articulated Objects, Austin, Texas, USA, 1994
- [2] B. D. Zarit, B. J. Super, and F. K. H. Quek, "Comparison of Five Colour Models in Skin Pixel Classification". Proceedings of ICCV'99 Int'l Workshop on recognition, analysis and tracking of faces and gestures in Real-Time systems, 58-63, 1999.
- [3] C. Bregler, and J. Malik, "Tracking People with Twists and Exponential Maps". Proc. of IEEE Int. Conf. on Computer Vision and Pattern Recognition, 1998.
- [4] C. Sminchisescu, "Estimation Algorithms for Ambiguous Visual Models 3D Human Modeling & Motion Reconstruction in Monocular Video Sequences". PhD Thesis, Inst. National Politechnique de Grenoble (INRIA), July 2002.
- [5] C. Wren, A. Azarbayejani, T. Darrel, and A. Pentland, "Pfinder: Real-time tracking of the human body". Proceedings of SPIE, Bellingham, WA, 1995.
- [6] D. Hogg, "Model-based vision: a program to see a walking person". J. of Image and Vision Computing, 1(1): 5-20, 1983.
- [7] D. M. Gavrilu, "The Visual Analysis of Human Movement: A Survey". Journal of Computer Vision and Image Understanding, Vol.73, No1, pages 82-98, 1999.
- [8] H. Sidenbladh, M. Black, and D. Fleet, "Stochastic Tracking of 3D Human Figures Using 2D Image Motion". European Conf. on Computer Vision, 2000.
- [9] J. K. Aggarwal, and Q. Cai, "Human Motion Analysis: A Review". Journal of Computer Vision and Image Understanding. 1999.
- [10] J. M. Rehg, and T. Kanade, "Model-Based Tracking of Self-Occluding Articulated Objects". Proceedings of International Conference on Computer Vision, Cambridge, MA, pages 612-617, June 20-23, 1995.
- [11] M. Isard and A. Blake, "CONDENSATION -- conditional density propagation for visual tracking". Int. Journal of Computer Vision, 29, 1, pages 5--28, 1998.
- [12] N. Yang, M. Zhang, C. Huang, and D. Jin, "Motion Quality Evaluation of Upper Limb Target-Reaching Movements". Medical Engineering & Physics 24, pages 115-120, 2002.
- [13] S. X. Ju, M. Black, and Y. Yacoub, "Cardboard people: A parameterised model of articulated motion". 2nd Int. Conf. on Automatic Face- and Gesture-Recognition, Killington, Vermont, pages 38-44, Oct 1996.
- [14] T. Moeslund, and E. Granum, "A Survey of Computer Vision-Based Human Motion Capture", Computer Vision and Image Understanding (81), No 3, pages 231-268, 2001.
- [15] Y. Zhuang, Q. Zhu, and Y. Pan, "Hierarchical Model Based Human Motion Tracking". ICIP 2000