Kinematics Analysis Multimedia System for Rehabilitation

Minxiang Ye[1], Cheng Yang[1], Vladimir Stankovic[1], Lina Stankovic[1] and Andrew Kerr[2]

[1] Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, UK
[2] Biomedical Engineering Department, University of Strathclyde, Glasgow, UK

Abstract. Driven by recent advances in information and communications technology, tele-rehabilitation services based on multimedia processing are emerging. Gait analysis is common for many rehabilitation programs, being, for example, periodically performed in the post-stroke recovery assessment. Since current optical diagnostic and patient assessment tools tend to be expensive and not portable, this paper proposes a novel marker-based tracking system using a single depth camera which provides a cost-effective solution that enables tele-rehabilitation services from home and local clinics. The proposed system can simultaneously generate motion patterns even within a complex background using the proposed geometric model-based algorithm and autonomously provide gait analysis results using a customised user-friendly application that facilitates seamless navigation through the captured scene and multi-view video data processing, designed using feedback from practitioners to maximise user experience. The locally processed rehabilitation data can be accessed by cross-platform mobile devices using cloud-based services enabling emerging tele-rehabilitation practices.

Keywords: multimedia signal processing, gait analysis, optical marker, multimedia content analysis

1 Introduction

The recent emergence of tele-rehabilitation services that aim to provide clinical rehabilitation diagnostics to patients in the comfort of their own home, calls for a radical shift in rehabilitation technology: from bulky, expensive equipment ideal for large rehabilitation facilities, to portable and affordable technology that can be operated in the home setting. The advances in information and communications technology have been making this shift possible through novel, cheap, and compact communications and media processing tools, with a large development space in terms of user-friendliness and accuracy of diagnostics. Furthermore, post-stroke rehabilitation with pleasant user experience would significantly benefit from tele-rehabilitation.

During typical clinical rehabilitation programs, the walking patterns of stroke patients are periodically assessed [1], which can be performed by using optical

motion capture systems such as VICON [2]. These systems provide accurate walking patterns by tracking markers attached to relevant joints using multiple cameras, at the cost of high expense, operational expertise and large laboratory space.

To provide the service to patients who do not have access to these facilities, simple marker-based or markerless tracking systems using multiple or single RGB camera have been proposed. However, these systems have limitations, such as operational expertise and time-consuming processing [3], and requirement for specific colour of the underlying cloths, such as the single RGB camera systems of [4], [5], [1] and multiple RGB camera systems, such as [6]. An attractive alternative is to use Microsoft (MS) Kinect sensor [7] with its own SDK capable of tracking 25 skeleton joints. However, it is demonstrated in [8], [9], [10] that Kinect's skeleton results are too noisy and not suitable for clinical applications.

In this paper, we develop a novel application using a single depth camera (MS Kinect), combining the benefits of the 3D reconstruction ability of Kinect and high accuracy from VICON-like optical marker-based tracking. The proposed multimedia system provides a convenient solution for tracking multiple retro-reflective markers simultaneously, solving a geometric model-based identification problem even within a complex background. Our proposed geometric model detector automatically locates all labelled markers and constructs the corresponding digital models. We use a blob detector to detect all markers, introduce a novel algorithm to estimate the depth value in the center region of each marker to restore its 3D trajectory, which is in turn used to calculate the required 3D joint angles. In experiments, we adopt the cubic Bezier curve interpolation [11] for gap filling, calculate the joint angles, visualize movement patterns, analyze gait cycle and measure step and stride length, swing and stance phases.

Our designed multimedia application provides visualization for all raw and processed data autonomously and potentially provides an interactive multimedia communication service with mobile terminal devices via cloud servers. Result validation with VICON clearly shows the ability to reconstruct sagittal view gait cycles accurately.

2   Proposed System

The proposed multimedia system enables 3D kinematics reconstruction with high accuracy and robustness, using attached joint marker trajectories estimation to perform automatic computation and kinematics data visualization such as segment angle, movement patterns, gait cycle, step and stride length, swing and stance phases, etc [12]. The system consists of three parts: (1) Pre-processing - cleaning invalid data and building subject reference model before processing. (2) Tracking - detecting and labeling each marker. (3) Kinematics Analysis and visualization - computing and presenting kinematics data. The overall platform visually presents all processed data enabling user interaction using a high level graphic user interface (GUI) framework and is capable of interacting with cloud servers or other remote terminal devices as an inquiry service for tele-rehabilitation.

2.1 Preparation

Data acquisition is done by capturing infrared (IR) and depth image sequences using MS Kinect v2 [7]. We adopt a depth-map projection method from [13] to map space coordinates of tracked markers to the camera space. To keep the region of interest on the IR map and remove background, we model offline the scene by a virtual trapezoidal cylindrical model shown in Fig.1 due to blob noise from reflective materials that introduce redundant high intensity pixels into IR images and blank holes into depth images.

In Fig. 1, we define the central point x for subject modeling whose distance L to the Kinect sensor can be estimated by extracting the floor square corners (p1, p2, p3, p4 in Fig.1). The start and end points of a straight walking line and a valid scene region (solid line shown in Fig.1) can be estimated by a field of view angles $\beta_w = 70°$ and $\beta_h = 60°$ and a predefined reliable depth range [14]. In this way, our multimedia application will present the ac-

Fig. 1. Virtual Scene.

tual start and end points within the RGB stream using mapping functions from MS Kinect v2 SDK [7], which simplifies walking line calibration.
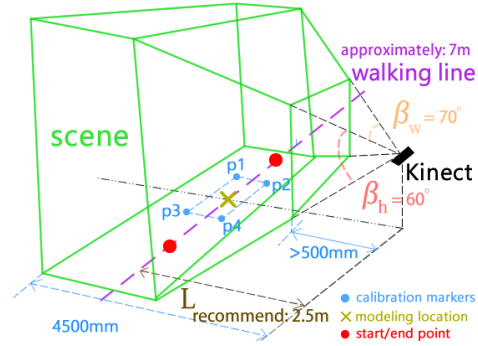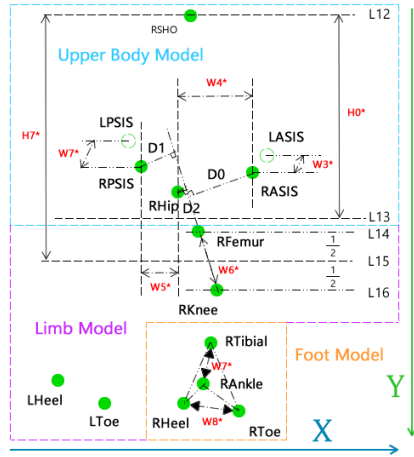

Fig. 2. Sagittal Model (Right Hand Side). 12 visible markers are marked with green circles. 2 partial invisible markers are shown in circle outlines.

Next, we split the defined sagittal subject model into three parts: upper body, limb and foot model. We locate and label all markers by validating their camera space coordinates along with axes X,Y, and measure $H0^*$, $H7^*$, and $W4^*$ to $W8^*$ as sagittal model parameters shown in Fig. 2.

2.2 Tracking

In this section, we introduce the proposed tracker for detecting and identifying the optical markers attached on the subject during the walking exercise, on a frame-by-frame basis using both IR and depth images.

In each frame, each marker is detected using: (1) Blob Detector - detecting centre of each blob. (2) Contour Finder [15] - finding valid contours. (3) Ellipse And Minimum Area Rectangle Fitter [16] [17] - extracting blob information. (4) Our Kernel Cluster Filter - marking valid blobs.

The blob detector converts the acquired IR images into binary images using threshold T following 4 steps: (a) Adopt the contour extraction method of [15]

and ellipse and minimum area rectangle fitter of [16] [17] to detect each blob's centroid. (b) Remove redundant blobs (due to reflection phenomenon) by forming concentric cycles around each blob centroid. Let R_{Marker} be the radius of the marker in pixels, and $Centre$ as blob centroid. Let $C(R_n)$ be the pixel values forming an imaginary circle centred at $Centre$ with radius R_n [in pixels], where $R_n = 1, \ldots, R_{Marker}$ pixels. (c) Apply a histogram filter on all $C(R_n)$ to obtain the significant pixel values κ_n for each concentric cycle. Let κ^m be the mean of κ_n for Marker m. Then, for Marker m find the smallest index i^m such that $\frac{\kappa_i^m}{\kappa_{i+1}^m} > 2$. (d) Let $A^m = \sum_{j=1,i^m} R_j$ and obtain the marker detection threshold T as the mean of all κ^m weighted by $\frac{A^m}{\sum_{l=1}^{M} A^l}$ where M is the total number of visible markers in the floor square.

During marker detection, we apply the contour algorithm from [15] to locate concentric contours $\mathbf{\Omega} = \omega_0, \omega_1, \ldots$ for each blob in the binary image using threshold T. Let $|\mathbf{\Omega}|$ be the number contours in the set $\mathbf{\Omega}$. Then, each marker centroid is located by:

$$Centroid = \begin{cases} Fit_{Ellipse}(\mathbf{\Omega}) & \text{if } |\mathbf{\Omega}| > 5 \\ Fit_{MinRect}(\mathbf{\Omega}) & \text{if } |\mathbf{\Omega}| \in (2, 5) \\ \frac{\omega_0 + \omega_1}{2} & \text{if } |\mathbf{\Omega}| = 2 \text{ and } Fit_{Depth}[Grow(\mathbf{\Omega}, \delta)] = 1 \\ \omega_0 & \text{if } |\mathbf{\Omega}| = 1 \text{ and } Fit_{Depth}[Grow(\mathbf{\Omega}, \delta)] = 0 \end{cases} \qquad (1)$$

where $Fit_{Ellipse}$ and $Fit_{MinRect}$ uses the algorithm of [16], [17] to locate the centroid given each blob's contours $\mathbf{\Omega}$. In order to deal with extreme cases that $|\mathbf{\Omega}| < 3$, we adopt function $Grow$ given by (2) to generate a new rectangular window by applying kernel increment δ on pixels $P_{i,j}$ from contours ω_0 (and ω_1).

$$G = Grow(P_{i,j}, \delta) = \underbrace{Min_{i, P_{i,j} \in \mathbf{\Omega}}(P_{i,j}) - \delta}_{\text{left}}, \underbrace{Min_{j, P_{i,j} \in \mathbf{\Omega}}(P_{i,j}) - \delta}_{\text{top}},$$
$$\underbrace{Max_{i, P_{i,j} \in \mathbf{\Omega}}(P_{i,j}) + \delta}_{\text{right}}, \underbrace{Max_{j, P_{i,j} \in \mathbf{\Omega}}(P_{i,j}) + \delta}_{\text{bottom}} \qquad (2)$$

where δ is set to 3 and 2 for the case $|\mathbf{\Omega}| = 2$ and $|\mathbf{\Omega}| = 1$, respectively. Function Fit_{Depth} in (1) determines the validity of each $Centroid$ by scanning a new window G grown by (2) as:

$$Fit_{Depth} = \begin{cases} 1, & \text{for } N > \tau \\ 0, & \text{for } N = \tau \end{cases} \qquad (3)$$

where we set $\tau = 0$ and N as the number of pixels in G whose corresponding depth value is in the Kinect's reliable range of $(500, 4500]$ mm. As a result, we regard all blobs centred at $Centroid$ as the final detected markers.

To label the detected marker, we introduce a model-based identifier by restoring the depth value using our histogram clustering algorithm, described next. Since optical markers reflect the IR emission, the marker region is full of zero depth values in the depth image and the spatially collocated regions in the IR image will have very high values. Therefore, the task can be simplified into calculating the weighted mean depth from depth-map histogram statistics with our

proposed clustering algorithm, reducing the noise from sensors as well as uncertain overlap conditions. We set the recovery radius $\delta = 3$ for growing the high IR $P_{i,j}$ region by Eq.2 and calculate the weight value as the significant mean relative distance between depth pixels to the centroid in each histogram bin.
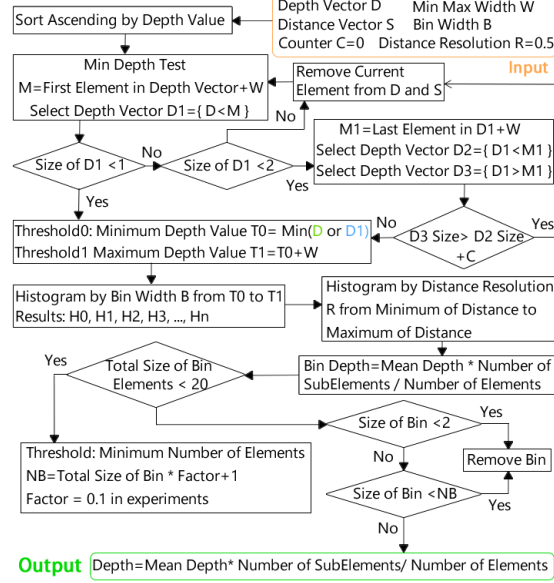

Fig. 3. Histogram Clustering Workflow.

In Fig. 3, we provide the flowchart of the proposed histogram clustering algorithm applied on knee, ankle, heel, and toe markers. Two situations are analyzed based on histogram bin width: (1) Bottom of the nearest pixels with occlusion (such as hip, femur) - using $T_0 = Min(D)$, $T_1 = Max(D_1)$ and $T_0 = T_0 - W$, where D, W, and D_1 are distances shown in Fig. 2. (2) Top of surrounding pixels (such as shoulder), in which case set $T_0 = Min(D_1)$.

We use the subject model defined in Sec. 2.1 to inspect all potential groups for upper body, limb and foot models (see Fig. 2). In particular, we order all markers between L12 and L13 by X-coordinate and validate the distances D0, D1, D2 for obtaining the most likely groups as the first look-up table. Then for the upper limb, we sort markers under L13 by Y-coordinate and X-coordinate and evaluate the nearest six markers to the ground by splitting them up into two groups with the triangular foot model shown in Fig. 2. Finally, we categorize the remaining left markers in the upper limb region by Y-coordinates and solve foot position by examining relative position to the right knee marker and histories over time.

3 Visual Analysis

We design a user-friendly interface for kinematics reconstruction with an interactive multi-view scene manager. For testing, we adopt relative knee angle (based on hip, knee and ankle trajectories) and gait phases (detecting local extremum and inflection ranges) defined in [12] for kinematics analysis. The multi-view tracking snapshot of the software is shown in Fig. 4.

As it can be seen from the snapshot, it is convenient to access the recorded trials or streaming by selecting the tracker tab page. Users can also view the automatic reconstruction process within our multimedia application or manually playback the whole trial. After autonomous analysis and user authentication, the kinematics results (including joint angles, movement patterns, gait phases, measure step and stride length, swing and stance phases) will be generated and
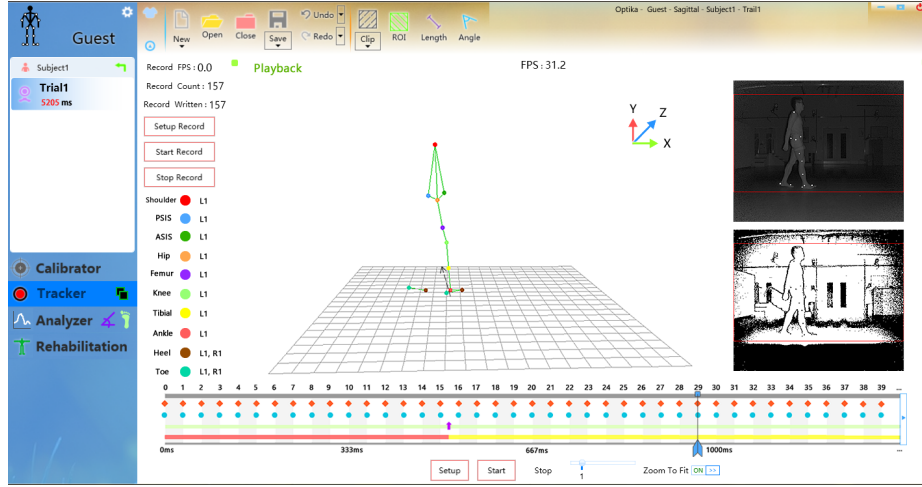
Fig. 4. Multi-view Tracker Snapshot

dispatched to cloud servers. The designed multimedia system provides a platform service for online/offline analysis, interaction and rehabilitation. Therefore, users can easily share all authorized local measurements across platforms with provided high quality multimedia content analysis and visualization as local and remote interaction services for tele-rehabilitation.
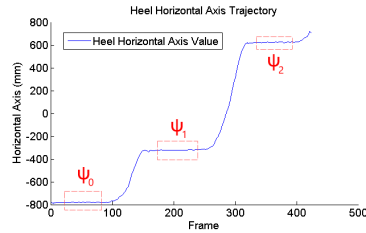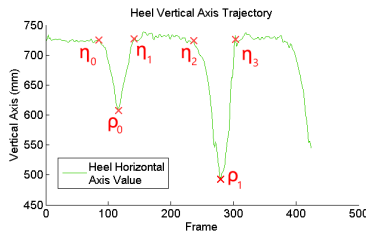

Fig. 5. Heel Horizontal Axis

Fig. 6. Heel Vertical Axis

For analyzing a gait cycle, we follow the gait phases definitions from [12] and adopt trajectories (100 fps) of heel, ankle, knee and hip markers to calculate step and stride length, stance and swing phases. The main task is to calculate vertical thigh segment angle, tibia segment angle, and find inflection points and local peaks of heel markers' horizontal and vertical axis values shown in Figs. 5 and 6.

The task of measuring step and stride length can be simplified into examining the stable values ψ_0, ψ_1, ψ_2 using window matching for the region between inflection points in Fig. 5. Once the left and right heels horizontal stable values are found, the step and stride length can be calculated using the adjacent stable values over time.

Stance and swing phases are separated by the following two gait events: heel strike and toe off. This can be solved by searching inflection points η_0, η_1, η_2, η_3 and local extremum ρ_0, ρ_1 in Fig. 6 by searching the region between the inflection points from a global minimum to maximum by regrouping iteratively.

4 Experimental Results & Discussion

We tested the system using 40 independent trials for 5 subjects by measuring knee angle α, step length ζ, stride length ξ, stance and swing phases defined in [12], as shown in Fig. 7.
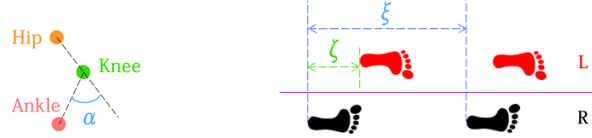

Fig. 7. Knee Angle, Step and Stride Length

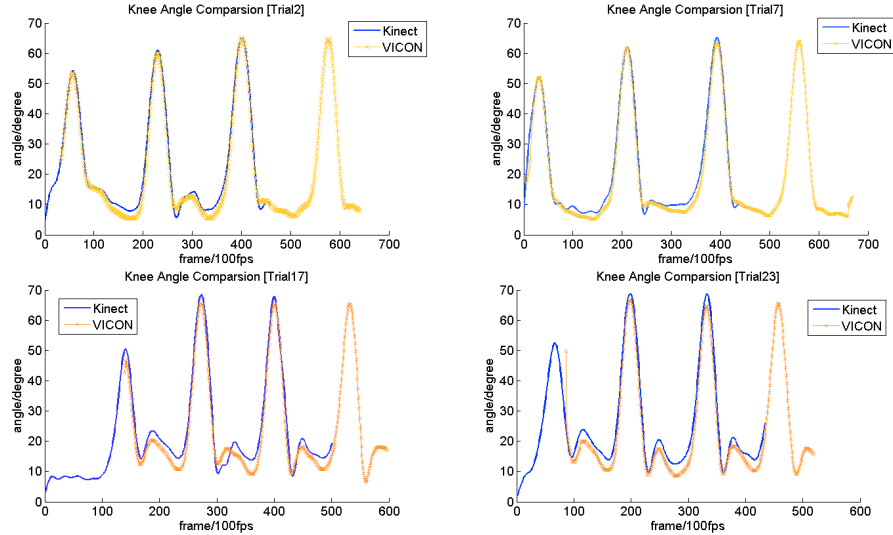Four typical trials from 4 different subjects are shown in Fig. 8

Fig. 8. Relative Knee Angle Comparison with VICON

It is obvious that our sagittal plane knee angle results have high accuracy benchmarked to the state-of-the-art industrial standard VICON. Furthermore, we closely followed the definitions of each gait event [12], and manually selected the key frames from the recorded IR and depth sequences as gait cycle references. Then we validated the step and stride lengths, and stance and swing phases with manually measured values that are based on [12]. The mean percentage error and standard derivation of the error [in percentage] on 40 sets of results are listed in Table 1.

Table 1. Performance of the proposed method for measuring step and stride length, and stance and swing phases

Error	Step	Stride	Stance	Swing
Mean(%)	1.05	1.17	1.82	1.10
Std(%)	5.33	4.76	5.83	4.37

5 Conclusion

The proposed system provides a mature solution for building up a mobile multimedia interactive service for visualization, presentation and tele-rehabilitation.

It is more attractive than typical optical motion capture systems since it is portable and cheaper for clinic and home use. As a convenient indoors kinematics diagnosis system, it only uses single MS Kinect v2 and needs low operational expertise within a user-friendly application. Validation results indicate good agreement for the sagittal plane gait analysis benchmarked with the commercial industry standard VICON.

References

1. Yang, C., Ugbolue, U., Carse, B., Stankovic, V., Stankovic, L., Rowe, P.: Multiple marker tracking in a single-camera system for gait analysis. In: ICIP IEEE Int. Conf. Image Processing. pp. 3128–3131. IEEE (Oct 2013)
2. VICON: Gait Analysis (2015), http://www.vicon.com
3. Ugbolue, U.C., et al.: The evaluation of an inexpensive, 2d, video based gait assessment system for clinical use. Gait & Posture 38, 483–489 (July 2013)
4. Leu, A., Ristic-Durrant, D., Graser, A.: A robust markerless vision-based human gait analysis system. In: Applied Computational Intelligence and Informatics (SACI), 2011 6th IEEE International Symposium on. pp. 415–420 (May 2011)
5. Liao, T.Y., Miaou, S.G., Li, Y.R.: A vision-based walking posture analysis system without markers. In: Signal Processing Systems (ICSPS), 2010 2nd International Conference on. vol. 3, pp. V3–254–V3–258 (July 2010)
6. Li, Y.R., Miaou, S.G., Hung, C., Sese, J.: A gait analysis system using two cameras with orthogonal view. In: Multimedia Technology (ICMT), 2011 International Conference on. pp. 2841–2844 (July 2011)
7. Kinect for Window software development kit (May 2015), http://www.microsoft.com/en-us/kinectforwindowsdev
8. kook Jun, S., Zhou, X., Ramsey, D.K., Krovi, V.N.: A comparative study of human motion capture and analysis tools, http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.391.73
9. Nguyen, H., Meunier, J.: Gait Analysis from Video: Camcorders vs. Kinect, Lecture Notes in Computer Science, vol. 8815, book section 8, pp. 66–73. Springer International Publishing (2014)
10. Clark, R., Vernon, S., Mentiplay, B., Miller, K., McGinley, J., Pua, Y., Paterson, K., Bower, K.: Instrumenting gait assessment using the kinect in people living with stroke: reliability and association with balance tests. Journal of NeuroEngineering and Rehabilitation 12(1), 15 (2015)
11. Agosto, M.K.: Bezier curves. In: Computer Graphics and Geometric Modelling: Implementation & Algorithms. pp. 396–404 (2005)
12. D.Casey Kerrigan, M.S., N.Wen, M.: Gait analysis. In: Rehabilitation Medicine: Principles and Practice. pp. 167–174 (1998)
13. Lachat, E., Macher, H., Mittet, M.A., Landes, T., Grussenmeyer, P.: First experiences with kinect v2 sensor for close range 3d modelling. ISPRS - Int. Archives of the Photogr., Remote Sens. Spatial Inform.Sciences XL-5/W4, 93–100 (2015)
14. Kinect for Windows features: Kinect sensor key features and benefits (May 2015), https://www.microsoft.com/en-us/kinectforwindows/meetkinect/features.aspx
15. Suzuki, S., be, K.: Topological structural analysis of digitized binary images by border following. Comp. Vision, Graphics, and Image Proc. 30(1), 32 – 46 (1985)
16. Fitzgibbon, A.W., Fisher, R.B.: A buyer's guide to conic fitting. In: Proc. of the 6th British Conf. Machine Vision (Vol. 2). pp. 513–522. BMVC, BMVA Press (1995)
17. Toussaint, G.: Solving geometric problems with the rotating calipers. In: IEEE MELECON83 (May 1983)