



<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper presented at *Proceeding of The Swedish Symposium on Image Analysis (SSBA2013), Gothenburg, Sweden.*

Citation for the original published paper:

Abedan Kondori, F., Yousefi, S., Liu, L. (2013)

Active Human Gesture Capture for Diagnosing and Treating Movement Disorders.

In:

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-83352>

Active Human Gesture Capture for Diagnosing and Treating Movement Disorders

Farid Abedan Kondori, Shahrouz Yousefi, Li Liu

Umeå University, Umeå, Sweden 901 87

Email: {farid.kondori,shahrouz.yousefi,li.liu}@tfe.umu.se

Abstract—Movement disorders prevent many people from enjoying their daily lives. As with other diseases, diagnosis and analysis are key issues in treating such disorders. Computer vision-based motion capture systems are helpful tools for accomplishing this task. However Classical motion tracking systems suffer from several limitations. First they are not cost effective. Second these systems cannot detect minute motions accurately. Finally they are spatially limited to the lab environment where the system is installed. In this project, we propose an innovative solution to solve the above-mentioned issues. Mounting the camera on human body, we build a convenient, low cost motion capture system that can be used by the patient in daily-life activities. We refer to this system as active motion capture, which is not confined to the lab environment. Real-time experiments in our lab revealed the robustness and accuracy of the system.

I. INTRODUCTION

Many people around the world suffer from movement disorders. Movement disorders are neurological conditions that affect the speed, fluency, quality, and ease of movement. As with other diseases, diagnosis and analysis are key issues in treating such disorders. By rapid improvements in computer vision algorithms and software developments, vision-based motion analysis systems have become powerful tools for accomplishing these tasks. In a classical setup for such a system, a patient's movements can be captured in the calibrated laboratory environment, which is equipped with multiple high quality cameras that are installed all over the room for performing full-body three-dimensional motion analysis.

However, despite the fact that these methods can resolve the problem, they suffer from several drawbacks. First of all, the setups are not cost effective due to the use of multiple expensive cameras. Secondly, these systems are not capable of detecting minute motions accurately. Finally, they are confined to the lab environment where the system is installed. This spatial limitation hinders the possibility of observing the subject while moving naturally and freely, the matter that can affect the quality of the result (diagnosis). Investigation of the patient's movement under laboratory conditions results in certain procedural issues that must be considered [1]. There is a basic rule that the measurement tools should not change the function that is being assessed. However it turned out that the patient under observation may be affected or even intimidated by the specific circumstances in the lab environment and may therefore not present his/her natural movement pattern. As a consequence, the patient

will probably try to present the best possible performance under the lab conditions, and this is not comparable to the activities that are being performed in daily life, when not under surveillance of several cameras. Therefore the results of the patient's movement assessment in the lab environment are not completely reliable. In fact, the best way to assess the patient's improvement is to observe the subject's motion in daily life activities.

The innovative solution that we propose here tackles this issue directly. We want to build a convenient, low cost motion capture system that can be used by the patient while performing daily life activities. When it is time for a follow-up check, motion information is already available for the physician, and the decision on the level of improvement is not subjective anymore. This can be achieved by active motion capture. Compared to the traditional motion capture systems (which we will call passive here), active motion capture involves mounting the cameras on the patient's body rather than installing them in a specialized diagnosis environment. It has many advantages over the passive system that are explained in more details in section 3. Real-time experiments have been performed in our lab and the results revealed the robustness and accuracy of the system.

The paper is organized as follows: Section 2 gives an overview of the current human motion analysis systems. A comparison between active and passive motion capture systems is presented in Section 3. Then system overview is described in Section 4. Experimental results are given in Section 5 and finally we present our conclusions.

II. RELATED WORK

Most of the existing human motion tracking and analysis systems can be classified into two categories: position sensing systems and vision-based motion analysis systems.

A. Position sensing systems

In the position sensing paradigm, a set of sensors is mounted to the body of the subject/patient in order to collect motion information and detect changes in body position. Several different types of sensors have been considered. Inertial and magnetic sensors are examples of widely used sensor types. A magnetic sensor, or magnetometer, is a device that is used to measure the strength and direction of a magnetic field. It is sensitive to the earth's magnetic field. The performance of the magnetic

sensors is affected by the availability of ferromagnetic materials in the surrounding environment. Accelerometers and gyroscopes are very well known types of inertial sensors. An accelerometer is a device used to measure physical acceleration experienced by an object. It has been reported that accelerometers are reliable for measuring balance and postural sway, making them suitable for clinical assessment applications [2]. They are, however, sensitive to vibrational artifacts [3]. Another shortcoming of the accelerometers is the lack of information about the rotation around the global Z-axis, and therefore do not give a complete description of human motion [4]. Hence, Gyroscopes, that are capable of measuring angular velocity, can be used in combination with accelerometers in order to give a complete description of orientation [5]. Though it's major disadvantage is the drift problem. New positions are calculated based on previous positions, meaning that any error in the measurements will be accumulated over time.

B. Vision-based motion analysis

Vision-based motion capture systems rely on a camera as an optical sensor. Two different types can be identified: Marker-based and marker-free systems. The idea behind marker-based systems is to place some type of identifiers on the joints to be tracked. Stereo cameras are then used to detect these markers and estimate the motion between consecutive frames. Several commercial systems such Qualisys are available [6]. These systems are rather accurate and they have been used successfully in biomedical applications like gait analysis [7]. However, several drawbacks of such systems can be observed. First of all, they are expensive as the use of special-purpose equipment is required. Secondly, a specially-equipped room is a prerequisite for these systems. This limits the mobility of the user and so they are not suitable for home monitoring applications. Due to occlusion, some markers cannot be detected which means that detailed motion of some parts cannot be provided. Marker-free systems rely only on cameras and try to employ computer vision techniques to estimate the motion. The use of cheap cameras is possible in such systems. However, getting rid of markers comes with the price of complicating the estimation process of 3D non-rigid human motion. This is still an on-going research topic in computer vision, and only partial success in real situations has been achieved [8]. A recent breakthrough in the field of marker-free vision-based motion estimation is Kinect. It provides 3D scene information from a continuously-projected infrared structured light. Using Kinect, human body gestures can be detected and tracked by means of the depth information. Like other passive tracking systems, detecting minute motions could be problematic. Moreover, Kinect-based motion capture systems are spatially limited. Though, it is of interest to know that using vision for motion tracking is at least as good as using a magnetic sensor and sometimes outperforms it [9]. Besides, using the camera provides the patient with the opportunity of recording the whole event if

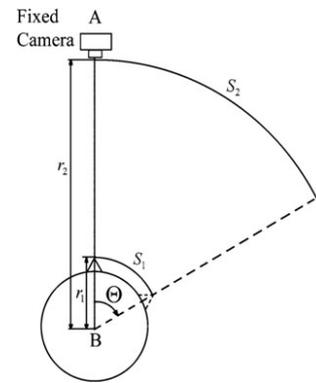


Fig. 1. Top view of a head and a fixed camera. The head turns with angle θ causing a change in the resulted image. The amount of change depends on the camera location (A or B).

needed.

Considering all the mentioned drawbacks in previous implementations, in this paper we present a novel vision-based approach for human motion tracking in biomedical applications. In contrast to passive system, our system involves mounting the cameras on the patient's body rather than installing them in a specialized diagnosis environment. Human motion tracking is achieved by extracting interest points from consecutive frames. Then point correspondences in two consecutive frames will be detected, which will be utilized in human motion tracking.

III. ACTIVE MOTION TRACKING

In this section the concepts of active and passive motion capture systems are clarified, and a technical comparison between these two methods is presented. Conventionally, vision-based human motion tracking systems place the camera in particular point, where the camera can see the user. Thus, the user has to perform desired movements and gestures in the camera's field of view. We address such configuration as the passive motion capture system. However there is another way. In this paper we suggest mounting the camera on the human body and performing motion tracking. Therefore the subject is not limited to be in the camera's field of view. We refer to this system as the active motion capture system.

When using passive configuration, certain issues must be considered. As it was mentioned in section 2, in some cases there is a need to use special markers, or to detect human body gestures. Consequently, the system can fail due to the incorrect marker/gesture detection. Other problems such as cluttered scene, human occlusion, scale variation (user distance to the camera) and illumination can degrade the system performance. Nevertheless, the most essential drawback associated with the passive systems is the resolution problem. Human motion results in changes in a small region of the scene, the fact that increases the burden of detecting small movements accurately [10]. But we believe these challenges easily

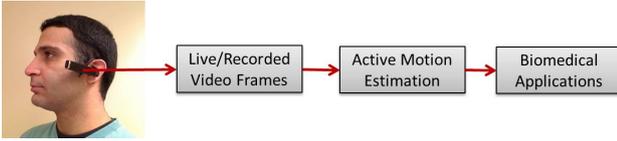


Fig. 2. Active motion tracking system overview.

can be resolved employing active motion tracking. Since the camera is mounted on the user body, there is no need to detect special markers or human gestures to track user motion. Instead, we extract stable key points in the video frames. These points will be tracked in consecutive frames for human motion estimation. In this project SIFT algorithm is used to detect key points [11]. SIFT features are scale invariant, and highly robust against illumination changes. Besides, active motion tracking can dramatically enhance the resolution problem. Based on the experiments in our lab, mounting the camera on human body can enhance the resolution in the order of 10 times compared to the passive setup [10]. In order to simplify the idea, consider a simple rotation around y-axis as it is illustrated in Fig. 1. This figure shows a top view of an abstract human head and a camera. Two possible configurations for human motion tracking are presented, placing the camera at point A, in front of the user (the passive setup) and mounting the camera on the head (the active setup). As the user turns with angle θ , the horizontal change (Δx) in captured images is calculated for both setups based on the perspective camera model. Let's assume $\theta = 45^\circ$, then for the passive motion tracking:

$$\Delta x_1 = f \frac{r_1}{\sqrt{2}r_2 - r_1} \quad (1)$$

and for the active motion tracking:

$$\Delta x_2 = f \frac{r_2}{r_2} \quad (2)$$

$$f \frac{r_1}{\sqrt{2}r_2 - r_1} \ll f \Rightarrow \Delta x_1 \ll \Delta x_2 \quad (3)$$

For example, if $f = 100$, $r_1 = 15cm$, $r_2 = 80cm$, then the change for both cases will be:

$$\Delta x_1 = \left(\frac{0.15}{\sqrt{2} * 0.8 - 0.15} \right) * 100 \approx 15.3 \text{ pixels} \quad (4)$$

$$\Delta x_2 = 100 \text{ pixels} \quad (5)$$

This indicates that motion detection is much easier when mounting the camera on the head, since the active camera configuration causes changes in the entire image while the passive setup often affects a small region of the image.

IV. SYSTEM DESCRIPTION

Fig. 2 depicts active tracking system overview. In this particular scenario, we want to measure patient's head motion. A wearable camera is mounted on the patient's ear. It should be realized that the camera can be either used to record the patient's head movements during daily

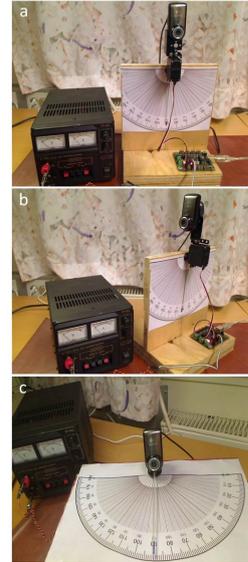


Fig. 3. Electronic measuring device. a) The setup for Z-axis, b) for X-axis, and c) for Y-axis.

life activities for offline analysis, or to provide live video frames for online analysis. As the patient turns his head, the video frames from the camera are fed to the system. Then stable interest points in the scene are extracted. These points are tracked in the next frame to find point correspondences. Afterwards, 3D motion information are recovered. Eventually, this information can be used to facilitate the patient's head motion analysis in biomedical applications.

A. Motion estimation

In order to analyze and estimate the head motion, we need to extract stable key points within entire image. Among different feature detectors, SIFT feature detector is used due to its invariance to image transformation [11]. Next, feature point correspondences are found between consecutive frames using pyramidal Lucas-Kanade optical flow algorithm [12]. This method is appropriate for fast motion tracking and has a low computational cost which is of our interest in real time applications. After finding point correspondences, a fundamental matrix for each image pair is computed using RANSAC algorithm [13]. RANSAC is a robust iterative algorithm to detect and remove the wrong matches and improve the performance. Running RANSAC algorithm, a candidate fundamental matrix is computed based on 8-point algorithm, and 3D motion parameters are recovered [14].

V. RESULTS

To report the angular accuracy of the active tracking system, we performed several tests. We developed an electronic measuring device to validate our proposed system (Fig. 3). The electronic device outputs are used as the ground truth to evaluate the active motion tracking system. The device consists of a protractor, a servo motor with an indicator, and a control board connected to a power supply. A normal webcam is also fixed on the



Fig. 4. Active motion tracking demo. As user turns his head, the motion parameters are estimated and used to change the 3d model on the computer screen

servo motor, so its rotation is synchronized with the servo. The servo motor can be operated by C codes through the control board. It can move in two different directions with specified speed, and its true rotation value (the ground truth) is indicated on the protractor. As the servo turns, the captured image frames will be processed and the camera rotation will be estimated by the active tracking system. Then the system outputs are compared to the ground truth to validate the system. Three different setups are used to test the system around X, Y, and Z-axis (Fig. 3 g g a, b, and c). We carried out the tests on an HP machine with an Intel core 2 Duo, 2.93 GHz processor. A Logitech Webcam 905 was used with a resolution of 640X480. Depending on the image content, 280 to 500 SIFT interest points were extracted per image. The system continuously measured the camera motion at the rate of 25 Hz by analyzing interest points. The camera is rotated from 0 to 40 degree around three axes separately, and the mean absolute error is calculated for each turn. The system evaluation was repeated for 100 times for five different motor speeds, and the results are presented in Table I. The error increases as the camera rotates, as it was expected. When the camera turns around X-axis, the number of missed interest points is larger than when rotating around Y and Z-axis. Thus, that error is slightly larger in X-axis. However our system is more accurate and robust compared to most of the current vision-based tracking systems, which aim to provide reasonable motion estimation with a mean absolute error of 5° or less [15]. Taking advantage of the active tracking system, we obtained mean absolute errors of 0.50° , 0.30° , and 0.24° for small rotations (5°), and 2.40° , 1.44° , and 0.72° for large motions (40°) around X, Y and Z-axis respectively.

We also developed another test to show the system usability (Fig. 4). Mounting the camera on user's head, the system estimates the user head motion and records the data. Motion parameters are applied to control a 3D model on the computer screen to visualize the user head motion.

VI. CONCLUSION AND DISCUSSION

We have presented a novel approach for human motion analysis in biomedical applications. To estimate the hu-

TABLE I
SYSTEM EVALUATION DATA SHEET. DATA IN THE LEFT COLUMN ARE ACTUAL ROTATION ANGLES AND THE OTHER COLUMNS ARE MEAN AND STANDARD DEVIATION OF THE ERRORS

Rotation angle	X-axis	Y-axis	Z-axis
5°	$0.50 \pm 0.41^\circ$	$0.30 \pm 0.35^\circ$	$0.24 \pm 0.31^\circ$
10°	$0.48 \pm 0.63^\circ$	$0.39 \pm 0.46^\circ$	$0.27 \pm 0.36^\circ$
15°	$0.74 \pm 0.50^\circ$	$0.48 \pm 0.52^\circ$	$0.36 \pm 0.45^\circ$
20°	$0.78 \pm 0.74^\circ$	$0.59 \pm 0.60^\circ$	$0.45 \pm 0.50^\circ$
25°	$1.49 \pm 0.68^\circ$	$0.62 \pm 0.61^\circ$	$0.45 \pm 0.44^\circ$
30°	$1.50 \pm 1.41^\circ$	$0.69 \pm 0.74^\circ$	$0.47 \pm 0.60^\circ$
35°	$1.92 \pm 1.94^\circ$	$0.98 \pm 1.01^\circ$	$0.64 \pm 0.77^\circ$
40°	$2.40 \pm 2.72^\circ$	$1.44 \pm 1.13^\circ$	$0.72 \pm 0.78^\circ$

man motion, the camera is mounted on the user's body rather than in front of it. Using active motion capture system, the main issues of human motion estimation are tackled. In this way, higher resolution and more accurate motion estimation are achieved, which have been demonstrated through theoretical analysis and practical experiments. The experimental results illustrate the robustness and efficiency of the proposed system. Though the system was used to estimate head motion, it can be utilized to recover human body motion.

REFERENCES

- [1] B. Rosenhahn, R. Klette, and D. Metaxas, *Human motion: understanding, modeling, capture and animation*, ser. Computational imaging and vision. Springer, 2008.
- [2] F. Foerster, "Detection of posture and motion by accelerometry: a validation study in ambulatory monitoring," *Computers in Human Behavior*, vol. 15, no. 5, pp. 571–583, 1999.
- [3] C. Bouten, K. Koekoek, M. Verduin, R. Kodde, and J. Janssen, "A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity," *Biomedical Engineering, IEEE Transactions on*, vol. 44, no. 3, pp. 136–147, march 1997.
- [4] H. J. Luinge and P. H. Veltink, "Measuring orientation of human body segments using miniature gyroscopes and accelerometers," *Medical & biological engineering & computing*, vol. 43, no. 2, pp. 273–282, Mar. 2005.
- [5] B. Kemp, A. J. Janssen, and B. Van Der Kamp, "Body position can be monitored in 3d using miniature accelerometers and earth-magnetic field sensors," *Electroencephalography and Clinical Neurophysiology*, vol. 109, no. 6, pp. 484–488, 1998.
- [6] [Online]. Available: <http://www.qualisys.se/>
- [7] I. Davis, S. Ounpuu, D. Tyburski, and J. R. Gage, "A gait analysis data collection and reduction technique," *Human Movement Science*, vol. 10, no. 5, pp. 575–587, Oct. 1991.
- [8] H. Zhou and H. Hu, "Human motion tracking for rehabilitation-a survey," *Biomedical Signal Processing And Control*, vol. 3, no. 1, pp. 1–18, 2008.
- [9] Z. Yao and H. Li, "Is a magnetic sensor capable of evaluating a vision-based face tracking system?" in *Computer Vision and Pattern Recognition Workshop, 2004. CVPRW '04. Conference on*, june 2004, p. 74.
- [10] Z. Yao, "Model-based coding - initialization, parameter extraction and evaluation," Ph.D. dissertation, 2005.
- [11] D. Lowe, "Distinctive image features from scale-invariant keypoints," 2004, *int. J. of Computer Vision* 60, 2, 91.110.
- [12] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," 1981, pp. 674–679.
- [13] R. B. M. Fischler, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," 1987, readings in computer vision: issues, problems, principles, and paradigms, 726.740.
- [14] A. Z. R. I. Hartley, "Multiple view geometry," 2004, cambridge University Press, Cambridge, UK.
- [15] E. Murphy-Chutorian and M. Trivedi, "Head pose estimation in computer vision: A survey," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 31, no. 4, pp. 607–626, april 2009.