This is the published version of a paper presented at 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015), JUL 26-30, 2015, Las Vegas, NV.

Citation for the original published paper:

Halawani, A., ur Réhman, S., Li, H. (2015)
Active Vision for Tremor Disease Monitoring.
In: 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences AHFE 2015 (pp. 2042-2048).
Procedia Manufacturing
https://doi.org/10.1016/j.promfg.2015.07.252

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-109206
Active vision for tremor disease monitoring

Alaa Halawani\textsuperscript{a,c,*}, Shafiq Ur Réhman\textsuperscript{a}, Haibo Li\textsuperscript{b}

\textsuperscript{a}Dept. of Applied Physics & Electronics, Umeå University, Umeå, 90187, Sweden
\textsuperscript{b}School of Computer Science & Communication, Royal Institute of Technology (KTH), Stockholm, 10044, Sweden
\textsuperscript{c}Computer Engineering and Science Dept., Palestine Polytechnic University (PPU), Hebron, Palestine

Abstract

The aim of this work is to introduce a prototype for monitoring tremor diseases using computer vision techniques. While vision has been previously used for this purpose, the system we are introducing differs intrinsically from other traditional systems. The essential difference is characterized by the placement of the camera on the user’s body rather than in front of it, and thus reversing the whole process of motion estimation. This is called active motion tracking. Active vision is simpler in setup and achieves more accurate results compared to traditional arrangements, which we refer to as “passive” here. One main advantage of active tracking is its ability to detect even tiny motions using its simple setup, and that makes it very suitable for monitoring tremor disorders.

Keywords: Active vision; Tremors; SIFT; Motion estimation; Motion tracking

1. Introduction

Movement disorders forbid many people from enjoying their daily lives. As with other diseases, diagnosis and analysis are key issues in treating such disorders. In this work, we present a framework for active vision-based monitoring and analyzing tremor disorders like Parkinson. Human Motion analysis in clinical medicine and therapy has been an active research field since the 1980’s. Most of the existing human motion tracking and analysis systems

* Corresponding author. Tel.: +46-(0)90-786-5862.
\textit{E-mail address:} alaa.halawani@umu.se
can be classified into two categories: Position sensing systems and vision-based tracking systems. In the position sensing paradigm, a set of sensors is mounted to the body of the patient in order to collect motion information. The use of magnetic sensors for human motion estimation was reported in research [1, 2]. The performance of the magnetic sensors is affected by the availability of ferromagnetic materials in the surrounding environment [3]. These materials will disturb the functionality of the sensors, the matter that will result in distorted measurements. Inertial measurement units (IMUs) were also used [4, 5]. An IMU is a combination of accelerometers and gyroscopes. Its major disadvantage is the drift problem that leads to an ever-increasing difference between actual and measured parameters.

Computer vision-based motion capture systems form the other alternative. They rely on cameras as sensors. In a classical setup of such systems, especially designed room is equipped with multiple high-quality cameras that are installed all over the room. The patient stands in the middle of the room with visual markers placed all over his body [6]. As the patient moves, the cameras continuously capture image frames that are used to register the motion by tracking the position of the markers in different frames.

Although such arrangements can accomplish the task, several limitations associated with them can be observed. First of all, the arrangement is not cost-effective due to the use of multiple expensive cameras. Secondly, the tests are all spatially limited to the room where the system is installed. This hinders the possibility of observing the subject while moving naturally and freely, the matter that can affect the quality of the result (diagnosis). In fact, the best possible way to follow the case and notice any developments is to observe the patient’s motion during daily life activities. Again, this is impossible using the classical motion capture systems due to their spatial limitations.

The innovative solution that we use here tackles this issue directly. We want to build a convenient, low-cost vision-based motion capture system that can be used by the tremor patient while practicing 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. This can be achieved by active motion capture. In contrast 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. This way, the mobility of the patient is guaranteed during motion capture and analysis. Moreover, the rapid advancement of the technology allows nowadays for small, low-cost cameras that can be practically used for this purpose. Another advantage of active motion capture over passive one is higher resolution: active motion capture is capable of detecting minute movements much better and more accurately than passive techniques. This was proven both theoretically and practically in our lab. This feature is of particular interest for the tremor-based disorders, as very small motions might need to be detected and analysed.

This paper is organized as follows: Section 2 compares active motion tracking to passive one. Actual motion tracking steps are given in Section 3. Section 4 summarizes the results and concluding remarks are given in Section 5.

2. Active vision vs. passive vision for motion tracking

Traditional passive vision-based tracking systems place the camera in front of the moving body. Many difficulties are associated with such configurations. For such systems to work, there is a need to use special markers placed on the joints to be tracked [7, 8, 6]. Sophisticated setups and expensive equipment are required to accomplish the task. If occlusion takes place, some markers cannot be detected which means that detailed motion of some parts cannot be provided. Marker-free systems try to employ computer vision techniques to estimate the motion. A review of such systems can be found in [9]. However, getting rid of markers comes with the price of complicating the estimation process of 3D non-rigid human motion. The system fails if the extracted features are not robust enough. The problem becomes more difficult if the subject appears within a cluttered scene. Issues regarding scale changes (distance of the user to the camera) and light conditions may also affect the performance. Finally, and most importantly, this configuration suffers a resolution problem. The user’s motion causes a change only in a small region within the scene, the matter that makes it difficult to track small movements accurately.

Active motion tracking solves the above-mentioned problems. The camera is to be mounted on the user’s body. Now, instead of capturing the body and its surroundings, the camera captures the view facing the body part it is mounted on. Instead of using special markers or detecting body features, interest points, around which local image
descriptors are computed, these points are examined through consecutive image frames in order to track themotion. We use the SIFT algorithm to extract repeatable interest points and highly discriminative local descriptors [10]. SIFT descriptors are invariant to scale changes and are highly robust against illumination changes, the matter that enhances the flexibility of the system.

Regarding the resolution problem, it was shown in our lab that active motion tracking dramatically enhances the resolution of the tracking system. In fact, the enhancement is in the order of 10 times compared to the traditional passive setup, as the active camera configuration causes a change in the entire image compared to changes limited to small image regions with the camera placed in front of the body (passive). This makes it easier to track even very small movements.

The following theoretical discussion clarifies the concept[11]. It is assumed that the head motion is to be tracked. Figure 1 shows a top view of an abstract head and a camera. When the camera is located in front of the head (passive), it is placed at point A. The other scenario is to place the camera on the head at point B (active). Assume that the head rotates around the y-axis with an angle of $\theta$ degrees. This causes a horizontal motion (change) of $\Delta u$ pixels of the projection of a world point, $P = (X,Y,Z)^T$, in the captured image. Based on the perspective camera model, $\Delta u$ is given by:

$$\Delta u = \frac{f \cdot \Delta X}{Z}$$  \hspace{1cm} (1)

where $f$ represents the focal length and $k$ the pixel size. If the camera is located at point A, then (1) becomes:

$$\Delta u_A = \frac{f \cdot r_1 \cdot \sin \theta}{k \cdot r_2 - r_1 \cdot \cos \theta}$$  \hspace{1cm} (2)

Since $r_1 \cos \theta$ is very small compared to $r_2$, (2) can be written as:

$$\Delta u_A \approx \frac{f \cdot r_1 \cdot \sin \theta}{k \cdot r_2}$$  \hspace{1cm} (3)

![Fig. 1. Top view of a head. The head rotates with angle $\theta$ causing a change in the captured image. The amount of change depends on the camera location (A or B).](image)
Multiplying and dividing by $\cos \theta$ we get:

$$\Delta u_A \approx \frac{f \tan \theta \cdot \frac{r_1 \cdot \cos \theta}{r_2}}{k}$$ (4)

For the case where the camera is mounted on the head (at point B):

$$\Delta u_B = \frac{f \cdot \frac{r_2 \cdot \tan \theta}{r_2}}{k} = \frac{f}{k} \tan \theta$$ (5)

As $r_2 \gg r_1$ it follows from (4) and (5) that:

$$1 \gg \frac{r_1 \cdot \cos \theta}{r_2} \Rightarrow \Delta u_B \gg \Delta u_A$$ (6)

For example, if $f = 3$ mm, $k = 10\mu m$, $r_1 = 10$ cm, $r_2 = 100$ cm, and $\theta = 45^\circ$, then the motion for both cases will be: $\Delta u_A \approx 21$ pixels and $\Delta u_B = 300$ pixels. This shows that it is easier to detect motion when the camera is placed on the user’s body, and demonstrates the superiority of active tracking over passive one.

3. Procedure

The camera is installed on the body part or joint to be tracked. Frames are continuously captured and fed to the system to estimate the motion. Motion parameters are estimated between two consecutive frames. Point correspondences between these frames should be formed. This is achieved using the well-known SIFT algorithm [10]. Following is a listing of the main steps of the proposed system:

1. Capture an image frame from the camera, call it $M_1$.
2. Extract SIFT points and descriptors from $M_1$.
3. Capture the next image frame, call it $M_2$.
4. Extract SIFT points and descriptors from $M_2$.
5. Based on the SIFT descriptors, establish point correspondences, $(x', x)$, between points in $M_1$ and $M_2$.
6. Use the five-point algorithm [12] to estimate the essential matrix, $E$, based on $(x', x)$.
7. Use RANSAC to refine the estimation process.
8. Recover the motion parameters, $R$ and $t$, from $E$ [12].
9. Set $M_1 = M_2$.
10. Go to step 3.

The process of motion parameter estimation is briefed as follows: We need to estimate the so-called essential matrix, $E$, between two frames to recover the 3D rotation and translation parameters. Point correspondences, $(x', x)$, are used to estimate $E$ so that the epipolar constraint:

$$x'^T E x = 0$$ (7)

is satisfied. It is assumed that $x'$ and $x$ are pre-multiplied by $K^{-1}$, the inverse of the camera calibration matrix, $K$. See [13] for details.

At least five point correspondences are required to solve for $E$ given $K$, using the five-point algorithm introduced by Nistér in [12]. This algorithm is more robust and stable than other traditional algorithms. As the presence of outliers is inevitable, and as the five point algorithm returns up to 10 different solutions, a RANSAC algorithm [14]...
is used for robust estimation of the essential matrix. Rotation Matrix, \( R \), and translation vector, \( t \), are then recovered from \( E \) as described in details in [12].

4. Results

4.1. Active vs. passive tracking

To show the advantage of the proposed active motion tracking system over traditional passive techniques, we compared the performance of both scenarios. Head motion tracking is taken here as an example, but the discussion applies to any other part or joint. A simple off-the-shelf webcam is used in the tests.

Two cameras were used. One is placed on the head of the user and the other in front of him (Figure 2(a)). As the user moves his head, 3D motion parameters are estimated once from the on-head camera (actively) and once from the frontal camera (passively). The estimated parameters (rotation and translation) of both cases are used to fit a 3D model to the user’s face. The accuracy of the model fitting obviously depends on the estimated parameters. Figure 2 shows some sample results of the model fitting. Four situations are shown where the user rotates his head 0°, -10°, -20° and -40° around the y-axis. Figure 2(b) shows the results for model fitting using active tracking. It can be noticed that the model fits the face accurately in all of the four cases. Results of the passive tracking technique are shown in Figure 2(c). As the estimated parameters in this case are not accurate, the model does not fit the face adequately. This shows that placing the camera on the body enables for more robust motion tracking. If head tremor needs to be monitored, obviously the simple active head motion estimation setup will be chosen over the passive one.

Fig. 2. 3D face model fitting. (a) Experiment setup. (b) Fitting based on motion parameters estimated actively using on-head camera. (c) Fitting based on motion parameters estimated passively using the frontal camera.
4.2. Parkinson monitoring demo

We implemented a demo application in which we use the proposed system for Parkinson tremor monitoring. A camera is mounted on the dorsum of the hand as can be seen in Figure 3. Microsoft Visual C++, together with OpenCV and OpenGL packages were used for programming. To speed up processing, the SiftGPU package was used to extract SIFT keypoints and descriptors. A 3D human model is used to reflect actual human hand movements in real time.

Figure 3 shows some samples from one test the demo. The purpose of this test is to check the validity of the tracking system before monitoring the tremors. The real user of the system, with the hand-mounted camera, can be seen to the left of each sample in Figure 3. To the right of it, the corresponding 3D model is shown. As the user moves his hand, the 3D model responds with the same hand movement as can be seen from the figure.

Fig. 4. A screenshot of the tremor monitoring Demo.
Figure 4 shows one screen shot of the tremor monitoring demo. Again, the 3D human model responds to user’s hand movements. To the right of the 3D model, three curves are depicted. These curves show the rotation values around the x-axis, the y-axis, and the z-axis. Spikes in the curves correspond to high rotation values, corresponding to strong tremors. As can be seen in the figure, even minute motions (characterized by low curve values at the beginning and end of the curves) are detected and recorded. Please note that the demo constitutes a proof of concept. Further enhancements and tests are required before usage in real-world scenarios.

5. Conclusion

We introduced a prototype for using active vision-based motion tracking to monitor tremor diseases. Unlike classical setups, cameras are mounted on the patient’s body in active motion tracking. Simpler manipulation and more accurate results are achieved this way. Active motion tracking is high in resolution and can detect small motions. This has been discussed theoretically and confirmed experimentally. As proof of concept, a simple demo for Parkinson tremor monitoring was developed.

References

[1] E. R. Bachmann, R. B. McGhee, X. Yun, M. J. Zyda, Inertial and Magnetic Posture Tracking for Inserting Humans into Networked Virtual Environments, in: Proceedings of the ACM symposium on Virtual reality software and technology, ACM, 2001, pp. 9–16.
[2] O. Suess, S. Suess, S. Mularski, B. Kühn, T. Picht, S. Hammersen, R. Stendel, M. Brock, T. Kombos, Study on the Clinical Application of Pulsed DC Magnetic Technology for Tracking of Intraoperative Head Motion during Frameless Stereotaxy, Head Face Med 2 (1).
[3] H. Zheng, N. D. Black, N. Harris, Position-sensing Technologies for Movement Analysis in Stroke Rehabilitation, Medical and biological engineering and computing 43 (4) (2005) 413–420.
[4] H. Zhou, H. Hu, Y. Tao, Inertial Measurements of Upper Limb Motion, Medical and Biological Engineering and Computing 44 (6) (2006) 479–487.
[5] S. T. Moore, H. G. MacDougall, J.-M. Gracies, H. S. Cohen, W. G. Ondo, Long-term Monitoring of Gait in Parkinson’s Disease, Gait & Posture 26 (2) (2007) 200–207.
[6] E. Delahunt, K. Monaghan, B. Caulfield, Ankle Function during Hopping in Subjects with Functional Instability of the Ankle Joint, Scandinavian journal of medicine & science in sports 17 (6) (2007) 641–648.
[7] R. B. Davis, S. Ounpuu, D. Tyburski, J. R. Gage, A gait Analysis Data Collection and Reduction Technique, Human Movement Science 10 (5) (1991) 575–587.
[8] I. Charlton, P. Tate, P. Smyth, L. Roren, Repeatability of an Optimised Lower Body Model, Gait & Posture 20 (2) (2004) 213–221.
[9] H. Zhou, H. Hu, Human Motion Tracking for Rehabilitation–A Survey, Biomedical Signal Process ing and Control 3 (1) (2008) 1–18.
[10] D. Lowe, Distinctive Image Features from Scale-Invariant Keypoints, International Journal of Computer Vision 60 (2) (2004) 91–110.
[11] A. Halawani, S. Ur Réhman, H. Li, A. Anani, Active Vision for Controlling an Electric Wheelchair, Intelligent Service Robotics 5 (2) (2012) 89–98.
[12] D. Nistér, An Efficient Solution to the Five-Point Relative Pose Problem, IEEE PAMI 26 (6) (2004) 756–777.
[13] R. I. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision, 2nd Edition, Cambridge University Press, 2004.
[14] M. Fischler, R. Bolles, Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, Commun. ACM 24 (6) (1981) 381–395.