Measurement of surgical motion with a marker free computer vision based system

Cristina Darolti[a], Ralf Bouchard[b] , Stefan Farke[c] , Alexandru Condurache[a] , Ulrich Hofmann[a]

[a]*Institute for Signal Processing, University of Luebeck*
[b]*Clinik für Chirurgie, UKSH, CampusLuebeck*
[c]*Abteilung Chirurgie, Schlosspark-Klinik, Berlin*

Abstract. This paper presents a marker free computer vision based system for tracking the articulated movement of hands in order to measure surgical skill. Tracking is achieved with a fast level set evolution method for active contours based on edge information. The spatial positions of anatomical features like the middle of the arm and of the wrist are computed from stereo images. Our first measurements show that we can track the trajectory of the arm movement with this method.

Keywords: surgical skill measurement, markerless tracking

---

1. Introduction

A surgeon's manual skill is vital for the patient and the doctor; as such, training in surgery must involve surgical skill assessment. Recently, computer systems have been introduced to support the search for objective measures of surgical skill. Two systems are widely used to measure hand motion: simulators (e.g. MIST-VR, daVinci) and electromagnetic based tracker systems (e.g. Imperial College Surgical Assessment Device).

This work explores the use of computer vision based hand tracking systems as an alternative to electromagnetic trackers. Computers eyed with video cameras monitor a surgeon's hand movements. The surgeon is not required to wear any hardware gear so that his movements may remain natural and time may be saved. In stead of marker positions, anatomical feature points like the middle of the wrist, of the palm or finger tips are determined. Their position in space is computed when using two or more cameras for stereoscopic viewing.

The measured spatial coordinates of feature points are processed to determine objective quantities for surgical skill assessment like the number of movements, their speed or the travelled distance [3]. Besides surgical skill assessment, raw motion data may be used to analyse various surgical procedures in order to design new instruments.

In this work we take the example of suturing, a simple procedure in conventional surgery. Motion patterns may include reaching for the needle, positioning/holding the needle, insert/push needle through tissue [4]. A new suturing tool should be an inexpensive instrument which puts the hand of the surgeon inside the patient's body during laparoscopic procedures. We want to know how many degrees of freedom does such an instrument need and how many of the hand's degrees of freedom need to be measured in order to answer this question. This work shows the methods and results for measuring three of the hand's degrees of freedom with a computer vision system.

2. Methods

A workspace has been designed where surgeons can perform a surgical procedure on a dummy object fixed on a transparent plate. The transparent plate is hinged to an aluminium rack to which six multiple synchronized cameras are attached and can be flexibly positioned. The surgical motion can thus be viewed from multiple different angles – including the "from under the table" view. Also, the measurements can be made more robust using data from different sources. Two photo umbrellas are used to illuminate the scene. The monochrome cameras record the working surgeon whilst his arms and hands – further referred to simply as hands – are uncovered.

The goal is to accurately track the boundaries of the uncovered hands and compute the position of hand feature points in each video image. Since placing the surgeon in a uniformly illuminated chamber is not practicable, the challenge is to find the accurate boundaries as shadowed and bright regions ramble on the hands while they move to perform the surgical procedure. In the recorded 8bit images, more than a third of the grey levels occur on the hand surface.

Hand boundaries are found in each frame using active contours [1]. The algorithm starts with an initial contour inside the area of the hand. In an iterative procedure, the position of each contour point is updated by moving the point in the direction of the contour's normal with a speed inverse proportional to the values of edge or/and region boundary features at the currently occupied image position. A speed proportional to the contour's curvature is added to ensure that the final curve will be smooth. The final curve will have minimal length in a metric space where distances are proportional to image properties. Curves are most conveniently represented using non-parametric level sets [5]. In [5], points on the curve constitute the set of points for which the value of an implicit three dimensional function is zero, i.e. the function's zero level set. The implicit function is positive valued for points outside the curve and negative valued for the rest.

The reason active contours have only recently been used in tracking is that they are slow when the implicit function is continuous. However, for discrete implicit functions defined on the image grid, much faster algorithms can then be used for curve evolution[6]. At each exterior pixel in the 8 neighbourhood of the curve a binary decision is made respective to the correctness of the current pixel label. The curve stops when neighbourhood pixels are labelled correctly, provided it is smooth.

In order to decide if a pixel belongs to the imaged hand of the surgeon, a NxN square neighbourhood around the pixel is processed. The SUSAN edge filter [7] is applied to each neighbour. A similarity score between the mask centre and pixels within the filter mask is computed. At edges, it is considered that only a maximum of three quarters of the neighbours will be similar to mask centre – the case of a corner cutting into the object; when the object boundary follows a line the number is halved. Simultaneously, the image region in question is smoothed with a mean shift filter [2]. This helps detect some of the spurious edges inside the hand and confirm real ones at the hand boundary.

A bright patch inside the hand is needed to initialize the active contour. First, a mean background image is computed and subtracted from the current frame. The resulting image contains the surgeon and artefacts around bright objects. This image is in turn segmented with a threshold computed with the Otsu method and the binary result is subjected to connected components labelling. The patches with largest areas prove to be the ones inside of the hand. The fast level set algorithm subsequently finds the correct boundary in the initial frame, which, in turn, will be the area where hand patches will be searched for in the next frame. The procedure is run for two of the video streams recorded by a pair of stereo cameras with similar views. Hand feature points always visible in both cameras can now be defined and their position in pixel coordinates computed. After calibrating the cameras with a multiple camera calibration software [8], the position of hand points in space is triangulated.

The following anatomical features are computed for a suturing task: the middle of the forearm near the elbow, the middle of the wrist and the center of gravity of the palm. The orientation of the hand is determined by computing the orientation of the major axis of an ellipse with the same second moments as the segmented hand. The pixel coordinates of the forearm's middle are at the middle of the perpendicular to this line, where the hand is brightest. The middle of the wrist is at the middle of the perpendicular where the hand is at its narrowest. The palm is computed as the center of gravity of the palm and fingers.

3. Results

Video streams of two experienced surgeons and one trainee have been recorded at approximately 20fps. The resolution of the recorded images is 780x582 where the average length of a hand is 250 pixels. The first 100 frames of each video stream were processed to detect the contours of each hand; only a few frames per video showed partly incorrect arm boundaries; strongly shadowed parts of finger pose a problem in detection. For each hand, the three anatomical feature points mentioned above were determined in each of the two cameras which look down on the scene. The points were subsequently displayed in the original video frames (without detected boundaries) and visually verified by the experimenter. The experimenter seldom disagreed with the results - for the frames with correctly detected arm boundaries - and when, the differences were small (sees figure1, top left, the hand on the left of figure).

The distance travelled by each hand feature was computed for each test subject. Unsurprisingly, this measure was larger for the trainee than for the experienced surgeons. However, plotting each spatial coordinate against time yields curves with many jumps; they are caused by the fact that, for e.g. the middle point of the arm stays the same as the thumb points up or down, whereas a marker on the skin surface would move. The anatomical feature points should be corrected. In our example, the computed spatial point should be displaced by the arm radius towards the axis between ulna and radius.

4. Conclusion

Surgical motion measurement is necessary in skill assessment and motion analysis. A tracker free computer vision based system for measuring surgical motion has advantages over conventional methods. It allows the surgeon to work naturally, it spares time and a large number of points on the hand, up to its entire surface, may be tracked.

In this abstract it has been shown that fast, accurate tracking of hand contours is possible with the described system. The motions of more surgeons performing a conventional suturing task were recorded. Three hand feature points were measured. The object of our future work is to make measurements more robust. To this end the definition of anatomical landmarks should be refined, more points on the surface of the hand may be tracked and measurements from different viewing angles should be integrated.

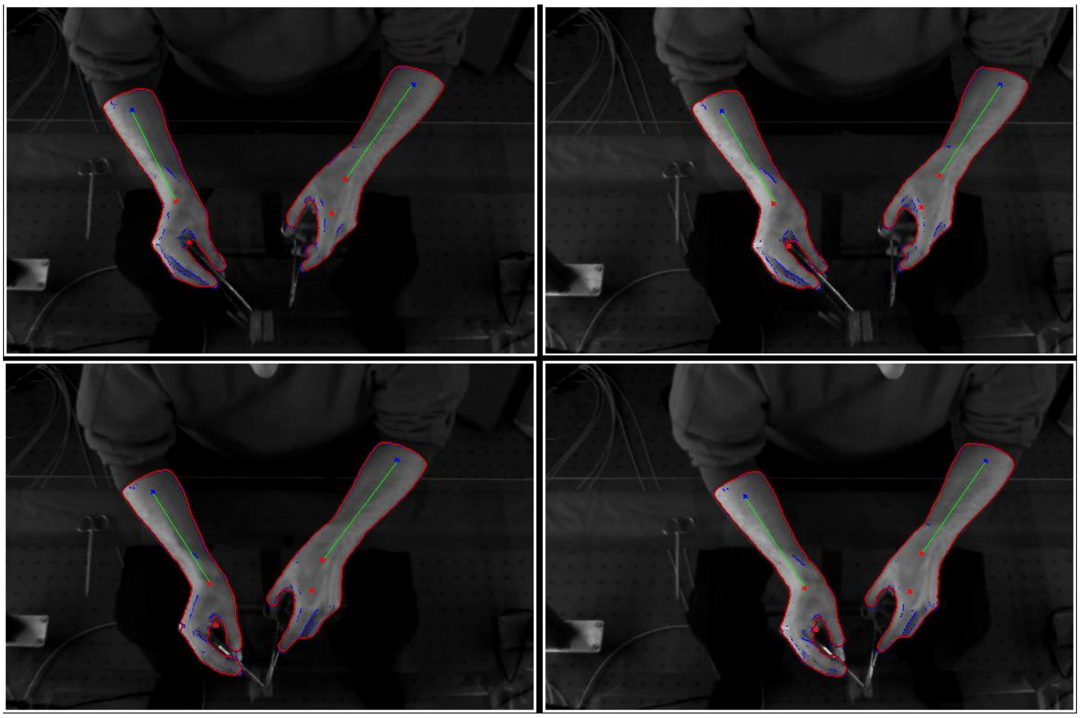

Fig. 1.Top: 3rd frame of two video sequences from similar point of views. Bottom: 31st frame of the same sequence. Red – detected contours. Blue – spurious edges overcome by the smoothing term. Red and blue crosses – computed anatomical feature points.

References

[1] Caselles,V.; Kimmel,R. & Sapiro,G. (1997), 'Geodesic Active Contours', Int. J.Comput. Vision 22(1), 61--79.

[2] Comaniciu,D. & Meer,P. (2002), 'Mean Shift: A Robust Approach toward Feature Space Analysis', IEEE Trans. Pattern Analysis Machine Intell., 24, 603-619.

[3] Datta,V.; Mackay,S.; Darzi,A. & Gillies,D. (2001), 'Motion Analysis in the Assessment of Surgical Skill', Computer Methods in Biomechanics and Biomedical engineering4, 515–523.

[4] Lin,H.; Shafran,I.; Yuh,D. & Hagery,G.D. (2006), 'Towards Automatic Surgical Skill Evaluation: Detection and Segmentation of Robotic-Sugrical Motions', Journal of Computer Aided Surgery 11(5), 20-30.

[5] Osher,S. & Sethian,J.A. (1988), 'Fronts Propagating with Curvature-Dependent Speed: Algorithms Based on Hamilton-Jacobi Formulations', Journal of Computational Physics 79, 12--49.

[6] Shi,Y. & Karl,W.C. (2005),Real-Time Tracking Using Level Sets, Conference on Computer Vision and Pattern Recognition, 34--41.

[7] Smith,S.M. & Brady,J. M. (1997), 'SUSAN—A New Approach to Low Level Image Processing', Int. J. Comput. Vision 23(1), 45--78.

[8] Svoboda,T.; Martinec,D.; Pajdla,T. (2005) A convenient multi-camera self-calibration for virtual environments. PRESENCE:Teleoperators and Virtual Environments.