Real-Time Instrument Tracking in Ultrasound Images for Visual Servoing

Tobias Ortmaier[1], Guillaume Morel[2] and Marie-Aude Vitrani[2]

[1]German Aerospace Center, Institute of Robotics and Mechatronics,
Oberpfaffenhofen, D-82234 Wessling, Germany
[2]Laboratoire de Robotique de Paris (LRP), University of Paris 6,
18, Route du Panorama - BP 61, 92265 Fontenay-aux-Roses Cedex, France
Email: Tobias.Ortmaier@dlr.de

Abstract. Minimally invasive surgery in combination with ultrasound (US) imaging imposes high demands on the surgeon's hand-eye-coordination capabilities. A possible solution to reduce these requirements is minimally invasive robotic surgery in which the instrument is guided by visual servoing towards the goal defined by the surgeon in the US image. This approach requires robust tracking of the instrument in the US image sequences which is known to be difficult due to poor image quality. This paper presents computer vision algorithms and results of experiments for different instrument materials.

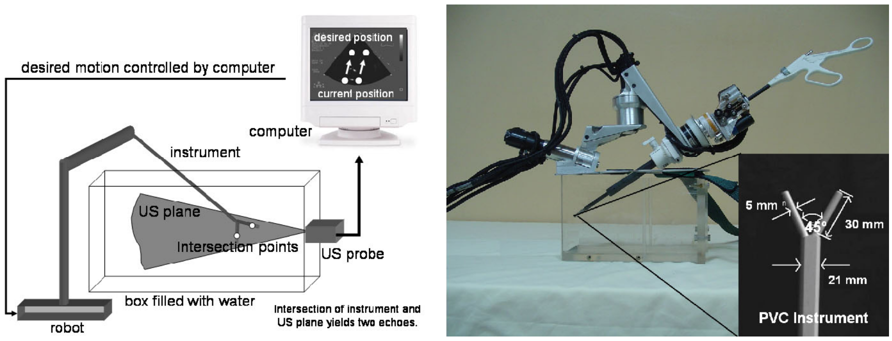
1 Introduction

Ultrasound (US) is an important imaging modality for medical examinations and surgical interventions, as it is cheap, harmless, and allows for real-time image acquisitions. Minimally invasive surgery (MIS) is an operation technique in which the surgeon works with long instruments through small holes. This reduces pain and trauma due to smaller incisions as compared to open surgery. Therefore, a combination of US imaging and MIS leads to a very gentle form of surgery. Unfortunately, this technique imposes high demands on the surgeon, as hand-eye-coordination becomes very difficult. A possible solution is minimally invasive robotic surgery (MIRS) in combination with visual servoing: the surgeon chooses the point of interest in the image plane and the robot moves the instrument towards this goal. The setup is sketched in Fig. 2(a). The proposed approach requires real-time tracking of surgical instruments in US images, which is difficult due to speckles, artefacts, and intensity variations.

A lot of work on tracking and detecting contours in US images can be found in the literature. Most of these approaches work with active contours allowing to deal with deformable structures (e.g. tissue), see [1] as example. Usually, these algorithms require the minimization of an energy function, often involving a gradient based approach, which usually contradicts realtime capabilities.

Only few articles can be found dealing with real-time tracking in US images. Most notable is [2], where different approaches for real-time tracking were compared: correlation and sum of squared differences (SSD) yielded good results

Fig. 1. Experimental Setup.



(a) Visual servoing concept (b) MC2E used for the experiments

but are not applicable in the case presented here, as the echo shape and texture change significantly. The same holds for other tracking techniques based on texture information. The Star and the Star/Kalman [2] approach rely on the detection of edges along rays emanating from a point interior the instrument echo. This approach is sensitive to ambiguities due to edges arising from speckles or echo artefacts. In the work presented here a robust real-time approach is described which is able to cope with large intensity variations and requires neither texture nor geometry information.

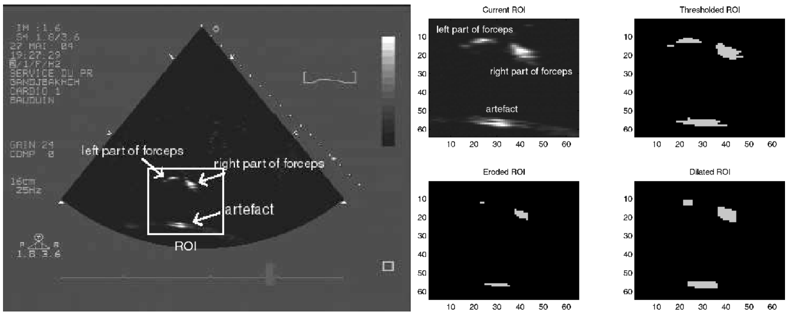
2 Experimental Setup and Image Processing

Robot MC2E (french acronym for compact manipulator for endoscopic surgery, see [3] and Fig. 2(b)) is used to move custom made instruments (materials: polyvinyl chloride (PVC), polyurethane (PUR), nylon, and plexiglas; see Fig. 2(b) for PVC instrument) in a plastic box filled with water. The robot is placed such that the instrument axes intersect with the ultrasound plane. The instrument shape is similar to a pair of forceps (see Fig. 2(b)), therefore two distinct echoes (thereafter also refered as regions or blobs) can be seen in the US image.

As ultrasound device a HP SONOS 5500 is used. The video output is connected to the BT878 frame grabber card of the image processing computer (Linux PC, P4 with 2.8 GHz and 1 GByte of RAM). Images are captured at frame rate (40 ms, 25 Hz) and converted into gray images with pixel intensity between 0 and 255 and a resolution of 384×288 pixels. The vision algorithms described below are applied to calculate the center of gravity $\boldsymbol{p}_1 = [x_1, y_1]^T$ and $\boldsymbol{p}_2 = [x_2, y_2]^T$ of the two regions caused by the instrument echos (see Fig. 3(a)).

First, a median filter is used to suppress small artefacts and noise. Afterwards, an adaptive optimal threshold k_{opt} according to [4] is applied in real-time to ac-

Fig. 2. Instrument echo, artefacts, and image processing steps.



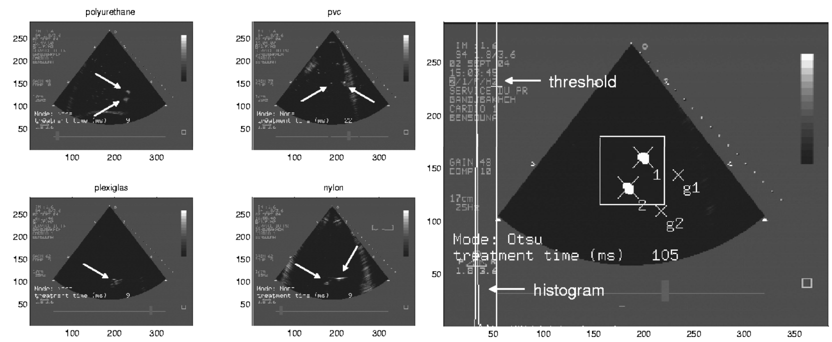
(a) Ultrasound image with ROI (b) Image processing steps

count for intensity variations of the instrument echo. Furthermore, the image treatment becomes independent of the current parameter settings of the usltrasound device. After thresholding, erosion and dilation (i.e. opening) are applied. This separates the instrument echoes from close artefacts and thus helps to increase tracking accuracy. Furthermore, small speckles with high intensity which were unintentionally detected are suppressed. Results for the different image processing steps are shown in Fig. 3(b). Thereafter, a fast MMX (intel's multi media extension) labelling algorithm based on run length coding is applied [5].

Usually, more than the two regions corresponding to the instrument echoes are detected in the selected ROI. To overcome ambiguities and to identify the correct regions a two step heuristic based on region size and velocity of the region center of gravity is applied. First, the region size s_i (in pixel) is considered. Only regions which satisfy $s_{\text{lower}} < s_i < s_{\text{upper}}$ are further examined. This eliminates small regions due to speckles or artefacts as well as large regions due to other surgical devices or prominent organ structures. Second, the distances between the current region positions and the region positions in the precedent frame (i.e. the velocity of the regions) are considered. The two regions having the smallest distance are selected as the instrument echoes. This is based on the fact, that in MIRS the instrument moves slowly. During the experiments this heuristic proved to be robust with respect to the beforementioned disturbances.

Finally, the center $\boldsymbol{c} = [c_x, c_y]^T$ of the ROI is updated, taking the position of the detected instrument echoes $\boldsymbol{p}_i$ into account: $\boldsymbol{c} = 1/2(\boldsymbol{p}_1 + \boldsymbol{p}_2)$. This allows to keep the ROI size small and therefore reduces computation time and possible ambiguities.

Fig. 3. Experimental results.



(a) Echoes for materials used (b) Results for PUR

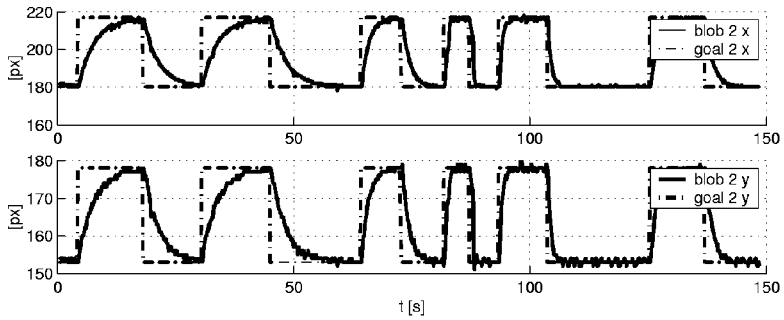
3 Experimental Results

This section presents image processing results for the different materials used, as well as trajectories of the closed visual servoing loop. Examples for the ultrasound images are given in Fig. 4(a). Instrument echoes are indicated by white arrows. Polyurethane and nylon yielded distinctive echoes. Polyvinyl chloride gave rise to frequent speckles and artifacts as a high gain was necessary to obtain significant and stable echoes. In this case, it was hardly possible to track the instrument. The instrument made of plexiglas yielded only one large echo instead of two.

For the algorithms presented in Sect. 2 the following parameters were used: the filter mask of the median filter was 3×3 pixels and so was the mask for the dilation. Erosion was not applied. The region of interest (ROI) was 32×32 pixels. The run length code tolerated gaps up to a length of three pixels in one line. Image processing results for the PUR instrument are given in Fig. 4(b): the two detected blobs, the ROI, and the desired position (g_1 and g_2) are shown. Furthermore, the histogram of the ROI and the calculated threshold according to [4] are presented.

The calculated thresholds for the PUR instrument lie inbetween 50 and 98, for the nylon case results are inbetween 48 and 118. Considering the large variations of the threshold it becomes clear that a fixed threshold leads to unsatisfactory results. Furthermore, the applied method allows for an automated adaptation of the vision algorithms to the chosen instrument material as well as gain and depth correction of the ultrasound device. The desired positions for blob one and two, as well as their measured trajectories for the PUR instrument are shown in Fig. 4. It can be seen, that the measured trajectories converge well towards the desired position with no remaining offset. Trajectories for the nylon instrument

Fig. 4. Desired and measured trajectories for PUR instrument echoes.



are similar and are therefore not shown here. The measurement noise is less than ±1 pixel.

4 Discussion

Reliable tracking of instruments in US image sequences is possible in realtime. During the experiments four different instrument materials were tested: PVC, nylon, plexiglas, and PUR. The instruments made of PUR and nylon yielded good (i.e. significant) echoes. Tracking of the PVC instrument was possible but error prone. The plesxiglas instrument produced only one echo instead of two echoes. The trajectories of the PUR and nylon instrument echoes in the image plane are smooth and only few outliers occurred. Adaptive thresholding is crucial due to large intensity variations in the image sequences and allows to automatically cope with changes of the ultrasound device settings. The overall computation time on a standard PC is below 10 ms per image.

References

1. Akgul YS, Kambhamettu C, Stone M. Task-Specific Contour Tracker for US. In: Workshop on Mathematical Methods in Biomedical Image Analysis; 2000.
2. Abolmaesumi P, Salcudean S, Zhu W, Sirouspour M, DiMaio S. Image-Guided Control of a Robot for Medical Ultrasound. IEEE TRA 2002;18(1).
3. Zemiti N, Ortmaier T, Vitrani MA, Morel G. A Force Controlled Laparoscopic Surgical Robot without Distal Force Sensing. In: Proc. of the ISER 2004; 9th International Symposium on Experimental Robotics. Singapore; 2004.
4. Otsu N. A Threshold Selection Method from Grey-Level Histograms. IEEE Transactions on Systems, Man, and Cybernetics 1979;9(1):62–66.
5. Frese U, Bäuml B, Haidacher S, Schreiber G, Schaefer I, Hähnle M, et al. Off-the-Shelf Vision for a Robotic Ball Catcher. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems; 2001.