Alfredo Illanes*, Anna Schaufler, Thomas Sühn, Axel Boese, Roland Croner and Michael Friebe

Surgical audio information as base for haptic feedback in robotic-assisted procedures

<https://doi.org/10.1515/cdbme-2020-0036>

Abstract: This work aims to demonstrate the feasibility that haptic information can be acquired from a da Vinci robotic tool using audio sensing according to sensor placement requirements in a real clinical scenario. For that, two potential audio sensor locations were studied using an experimental setup for performing, in a repeatable way, interactions of a da Vinci forceps with three different tissues. The obtained audio signals were assessed in terms of their resulting signal-to-noise-ratio (SNR) and their capability to distinguish between different tissues. A spectral energy distribution analysis using Discrete Wavelet Transformation was performed to extract signal signatures from the tested tissues. Results show that a high SNR was obtained in most of the audio recordings acquired from both studied positions. Additionally, evident spectral energy-related patterns could be extracted from the audio signals allowing us to distinguish between different palpated tissues.

Keywords: audio analysis; da Vinci robot; haptic feedback; minimally invasive surgery; robotic assisted surgery.

Introduction

Minimally invasive surgical procedures are increasingly performed with the use of robotic systems. Compared to conventional laparoscopy, robotic assistance systems allow increased precision but lacks completely of haptic sensation because of the indirect interaction with tissue through remotely controlled instruments. This limitation

can result in risks of injuries to critical structures such as vessels.

Different approaches have been presented in order to provide surgeons with haptic information. They are mainly based on the direct or indirect measurement of force or pressure. For direct measurements, single sensors [1, 2] or sensor arrays [3, 4] are installed in the instrument components directly interacting with the patient's inner organs. This imposes serious design limitations for fulfilling clinical requirements. Sensor technology can also be integrated into instruments for indirect force measurements, for example, on the shaft of the instrument [5, 6]. However, these measurements can only be used for extracting static one-point information, and for palpation purposes, dynamic information acquisition is required.

A novel method for guiding medical interventional devices using an audio sensor attached to the tool's proximal end has been presented in [7, 8]. Audio has shown promising results for acquiring non-invasively haptic information from medical tools. An audio-based guidance device could be used as a sort of plug-and-play device without the necessity of rebuild specialized instruments and use the already existing ones. Moreover, since the sensing device is located at the proximal end of instruments, no sensor is needed to be directly connected with the patient's organs.

This method has been successfully applied to acquire information from a forceps of a da Vinci surgical robot in three main scenarios: pulsation detection for the presence of vessel identification, palpation of underneath bony structures, and texture differentiation during palpation of different types of tissue [9, 10]. However, in these studies, the sensor unit was attached to the robotic tool in a location inside the sterile zone, making it complex to fulfill the requirements for real use in a clinical environment. In this work, two potential clinically feasible locations of the audio sensor unit are studied. The main objective is to show that haptic information can be obtained using audio sensing without violating the clinical framework and to demonstrate the feasibility of the concept for real scenarios. For that, an experimental setup for producing interactions between a da Vinci instrument and three different tissues was implemented. The acquired audio

*Corresponding author: Alfredo Illanes, Otto-von-Guericke University Magdeburg, Medical Faculty, Magdeburg, Germany, E-mail: alfredo.illanes@med.ovgu.de

Anna Schaufler, Thomas Sühn, Axel Boese and Michael Friebe, Otto-von-Guericke University Magdeburg, Medical Faculty, Magdeburg, Germany

Roland Croner, Clinic for General, Visceral, Vascular and Transplant Surgery, Otto-von-Guericke University, Magdeburg, Germany

recordings were then analysed in terms of their resulting signal-to-noise-ratio (SNR) and their capability to distinguish between different tissues.

Materials and methods

Due to the requirement for attachment of the measuring unit in the non-sterile area, audio measurements in the operating room cannot be performed directly on the instrument, as in the setup presented in [10]. During procedures with a da Vinci robot, the robotic arms are wrapped in a sterile drape during the procedure to avoid contamination of the surgical field by the robot's non-sterile arms. Therefore, to make audio guidance in robotic tools feasible, a suitable position for the placement of the sensor must be identified that conforms to the requirements of non-invasiveness (sensor located at the proximal end of the instrument) and location in the non-sterile zone. As shown in Figure 1, two possible locations are studied in this work: lateral adapter frame (left of Figure 1) and lower adapter edge (right of Figure 1), called in the sequel as locations L_{lat} and L_{low} , respectively. The locations are evaluated in terms of acquired meaningful audio signal information concerning the instrument tip/tissue interaction process during the palpation of different tissues. For that, the acquired audio signals are evaluated according to three parameters: their Signal to Noise Ratio, their capability to show different dynamics when different tissues are palpated, and their dynamical stability.

Experimental setup and data acquisition

An experimental setup was implemented following [10]. The experimental setup was intended to simulate the interaction of a da Vinci Endowrist instrument (Da Vinci Prograsp Forceps, Intuitive Surgical, California, USA) with different texture surfaces. For this purpose, two synthetic materials, felt and denim fabric, mounted on a board base (structure board), were employed as samples (Figure 2a). Additionally, a porcine liver was used as a third biological specimen. As a basic framework for supporting the instrument, a stable stand to which a clamp was attached was used. The clamp was allowed to be displaced laterally so that it could perform a pivoting movement around the vertical axis (Figure 2b). A sterile da Vinci drape was placed over the basic framework, and the integrated instrument adapter of the drape was fixed into the bracket. For the tip/surface interactions, the specimens were placed under the instrument tip so that the normal force applied to them was determined by the weight of the instrument (Figure 2c). The instrument tip was driven into a horizontal movement along the specimen surface by displacing the bracket at an average velocity of 5 cm/s (indicated by the arrow in Figure 2d). Each unidirectional movement across the surface, referred to as *one swipe*, was considered as one tool/tissue interaction event.

For the audio signal acquisition, a MEMS microphone (Adafruit I2S MEMS microphone SPH0645LM4H-B, Knowles, Illinois, USA) was attached to each studied locations (lateral adapter frame and lower adapter edge as shown in Figure 1) employing a double-sided adhesive tape.

Figure 1: Potential locations for sensor placement identified on the sterile adapter. Left: lateral adapter frame and microphone position, middle: adapter frontal view, right: lower adapter edge and microphone position.

Figure 2: (a) texture board with different materials; (b) instrument holder framework; (c) lateral view on the experimental setup; (d) top view on the experimental setup.

For each sensor location and tested specimen, 15 tool/tissue interactions were recorded with a sampling frequency of 44100 Hz (each interaction represents a swipe of the instrument tip over a tissue specimen). Each recording involves a segment with only background noise, followed by the interaction event and finalizing with a new background noise segment. A total of 90 audio recordings were generated and saved into a dataset.

Signal to Noise Ratio computation

The individual SNRs per audio recording are calculated to analyse the SNR resulting from the interaction of the different tested tissues. For that, three segments are extracted from each audio recording: two signal segments preceding and following the interaction event and containing only background noise, and one segment extracted at the middle of the interaction event. Then the background noises energies called E_{n1} and E_{n2} , and the event energy E_{ev} are computed. The SNR ratio for each recording i is finally calculated as $SNR = E_{ev}/(E_{n1} + E_{n2})$.

Spectral energy distribution analysis for distinguishing between tissues

When an interventional instrument interacts with a given tissue, the friction between the tip of the instrument and the tissue results in an audio wave presenting dynamics whose time-varying changes can contain information or patterns of how the tissue *sounds*. For example, let analyse Figure 3, which displays the time-domain and time-scale representations of three different audio signals obtained from one sensor location when sweeps were performed over three different tissues. The time-scale spectrum was obtained with a Continuous

Wavelet Transformation (CWT) using a Morse mother wavelet. The x-axis of the CWT spectrum represents the time, and the y-axis represents pseudo-frequencies ordered in a logarithmic scale between 0 and $f_s/2=22050$ Hz, where f_s correspond to the frequency sampling. It is possible to verify in this figure that even if in the time-domain representation tissues can show similar audio signal behavior (felt and denim), the time-scale spectra are significantly different. This means that the audio signal time-varying characteristics of each tissue are different when the instrument interacts with them.

The main observed difference lies in the spectral energy distribution, which shows different dominant energy frequencies for each tissue. This is the information that we want to exploit to assess the capability of a sensor location to be able to provide audio signals that can be used to distinguish between palpated tissues. For that, the spectrum is first divided into four bands corresponding roughly to pseudo-frequencies of very-low VLF: 5–35 Hz, low LF: 35–250 Hz, middle MF: 250–1500 Hz, and high HF: 1500–9500 Hz frequencies. For each band the energy and the contribution to the total spectral energy of the event was then calculated using the equation E_i/E_t with $i = LF, MLF, MHF, HF$ and $E_t = \sum_i E_i$.

Results

Figure 4 displays the computed SNRs for each of the 90 recordings of the dataset (45 per sensor location and 15 per tissue tested). The audio recordings obtained from the lateral location L_{lat} exhibit high SNRs for nearly all recordings and particularly for the denim tissue. It is important to point out that even if liver tissue is soft

Figure 3: Time-domain and time-scale representations of the audio signal resulting from the interaction of the robotic tool with three different tissues.

Figure 4: Obtained SNR per audio recording for the two studied sensor locations.

compared to the other two tested tissues, the obtained SNRs in this tissue are also high. The location L_{low} produces good SNRs for felt and denim tissues, but it shows less sensitivity to some of the tests made with liver tissue.

Figures 5 and 6 presents the four-band energy distribution analysis for both studied locations. Figure 5 displays two examples per tested tissue of energy distributions of an interaction event. We can observe how the energy distribution follows an evident pattern for the two interactions with each tissue. Additionally, both locations show energy patterns that do not vary from one recording to the other one.

Figure 6 shows the energy distribution for the whole dataset for both locations. This figure serves to analyse the stability of the obtained energy distribution patterns. Each energy distribution set (as the ones shown in Figure 5) was arranged in a matrix where each row corresponds to the energy distribution of a single recording. Using this visualization, it is possible to make two important verifications. First, the energy distributions are highly stable in a set of interactions belonging to a same tissue. Second, the energy distributions are different according to the tissue. For example, at the location L_{lat} , the felt tissue presents VLF and LF of similar intensities and is very high compared to

Figure 5: Contribution of the separated bands to the total spectral energy for recordings of the different tissues obtained from the two testes locations.

Figure 6: Energy contributions of the whole audio dataset for both studied locations.

MF and HF. More than 90% of the energy is concentrated in the lower frequency bands. For denim, 80% of the energy is concentrated in the VLF band, while for liver, more than 80% of the energy is concentrated in the LF range. A similar analysis can be done in the L_{low} location, where also clear patterns to distinguish between tissues can be observed.

Conclusion

This work shows that audio with a high SNR and containing important dynamic information of tissues can be obtained from the proximal end of a robotic tool with a clinically realistic sensor location. Two possible sensor placement in the sterile zone of the robotic instrument has been successfully evaluated, showing evident energy distribution patterns for distinguishing interactions of the instrument tip with different tissues. The next step will be to test this setup with a functional da Vinci robot in order to analyse the robustness of the studied signal patterns with the forceps interacting with hard and soft tissues.

Research funding: The author state no funding involved.

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Competing interests: Authors state no conflict of interest.

References

1. Kim U, Lee D-H, Yoon WJ, Hannaford B, Choi HR. Force sensor integrated surgical forceps for minimally invasive robotic surgery. *IEEE Trans Robot* 2015;31:1214–24.
2. Hong MB, Jo Y-H. Design and evaluation of 2-dof compliant forceps with force-sensing capability for minimally invasive robot surgery. *IEEE Trans Robot* 2012;28:932–41.
3. Qasaimeh MA, Sokhanvar S, Dargahi J, Kahrizi M. PvdF-based microfabricated tactile sensor for minimally invasive surgery. *J Microelectromech Syst* 2008;18:195–207.
4. King C-H, Culjat MO, Franco ML, Bisley JW, Carman GP, Dutson EP, et al. A multielement tactile feedback system for robot-assisted minimally invasive surgery. *IEEE Trans Haptics* 2008;2: 52–6.
5. Dalvand MM, Nahavandi S, Fielding M, Mullins J, Najdovski Z, Howe RD. Modular instrument for a haptically-enabled robotic surgical system (herosurg). *IEEE Access* 2018;6: 31974–82.
6. Khadem SM, Behzadipour S, Mirbagheri A, Farahmand F. A modular force-controlled robotic instrument for minimally invasive surgery—efficacy for being used in autonomous grasping against a variable pull force. *Int J Med Robot Comput Assist Surg* 2016;12:620–33.
7. Illanes A, Boese A, Maldonado I, Pashazadeh A, Schaufler A, Navab N, et al. Novel clinical device tracking and tissue event characterization using proximally placed audio signal acquisition and processing. *Sci Rep* 2018;8:1–11.
8. Illanes A, Sühn T, Esmaeili N, Maldonado I, Schaufler A, Chen C-H, et al. Surgical audio guidance surag: extracting non-invasively meaningful guidance information during minimally invasive procedures. In: 2019 IEEE 19th international conference on bioinformatics and bioengineering (BIBE). IEEE, Greece; 2019: 567–70pp.
9. Chen C-H, Sühn T, Illanes A, Maldonado I, Ahmad H, Wex C, et al. Proximally placed signal acquisition sensoric for robotic tissue tool interactions. *Curr Dir Biomed Eng* 2018;4: 67–70.
10. Chen C, Sühn T, Kalmar M, Maldonado I, Wex C, Croner R, et al. Texture differentiation using audio signal analysis with robotic interventional instruments. *Comput Bio Med* 2019;112: 103370.