Computers in Biology and Medicine 112 (2019) 103370

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Texture differentiation using audio signal analysis with robotic interventional instruments

C.H. Chen[a,*], T. Sühn[a], M. Kalmar[a], I. Maldonado[a], C. Wex[b], R. Croner[b], A. Boese[a], M. Friebe[a], A. Illanes[a]

[a] INKA Intelligente Katheter, Otto-von-Guericke University, Magdeburg, Germany
[b] Clinic for General, Visceral, Vascular and Transplant Surgery, Otto-von-Guericke University, Magdeburg, Germany

ARTICLE INFO

Keywords:
Robotic minimally invasive surgery
Signal processing
Haptic feedback
Acoustic emission
Non-invasive sensor placement

ABSTRACT

Robotic minimally invasive surgery (RMIS) has played an important role in the last decades. In traditional surgery, surgeons rely on palpation using their hands. However, during RMIS, surgeons use the visual-haptics technique to compensate the missing sense of touch. Various sensors have been widely used to retrieve this natural sense, but there are still issues like integration, costs, sterilization and the small sensing area that prevent such approaches from being applied. A new method based on acoustic emission has been recently proposed for acquiring audio information from tool-tissue interaction during minimally invasive procedures that provide user guidance feedback. In this work the concept was adapted for acquiring audio information from a RMIS grasper and a first proof of concept is presented. Interactions of the grasper with various artificial and biological texture samples were recorded and analyzed using advanced signal processing and a clear correlation between audio spectral components and the tested texture were identified.

1. Introduction

Robotic minimally invasive surgery (RMIS) has become an important medical application in the last decades. For selected applications and interventional procedures, the high precision and stability provided by the robot is beneficial compared to conventional minimally invasive surgery (MIS). However, more assistance given by the robotic technology results in a reduced physical sensation feedback.

Traditionally, surgeons have a direct or indirect tissue response by touching the organs with their hands. This palpation technique provides surgeons with mechanical information of the observed tissues such as boundaries, stiffness, surface roughness and it is even used for pulsation detection underneath the surface. This information helps the surgeons to accurately locate target tissues during surgery and to verify the pre-operative diagnosis derived from medical imaging, biopsy test, or biochemical examination.

Currently, RMIS using robotic surgical systems such as da Vinci system, still cannot provide any haptic and tactile feedback to the surgeons during tissue manipulation. Therefore, the traditional way of organ palpation or tissue mechanical properties determination cannot be performed as in open surgical procedures (Fig. 1). Intraoperative blood vessel detection is another important function that helps prevent accidental vascular injuries. These injuries can lead to internal bleeding, organ failure, or even to the patient's death and they occur mostly in procedures that involve dissection tasks around major vessels, particularly in urologic and gynecologic laparoscopic surgeries. A recent study analyzed data from 14 years of adverse events in robotic surgery [1] and reported that 7% of deaths were attributed to surgeon/staff mistakes involving incorrect instrument change or accidental cuts of arteries. Other works showed that the incidence of vascular injuries during anatomic pulmonary resection was around 2.4% [2] and in laparoscopic urologic procedures between 1.6% and 4.7% [3]. A haptic sense can reduce the risk of accidental injury to a major vessel (aorta, vena cava, or iliac vessels) while cutting tissues. However, this critical function is not available in RMIS.

In a robotic surgery without the sense of touch, surgeons currently use the so-called visual-haptics technique [4]. Previous non-RMIS surgical experiences helps the surgeons to visually assess tissue stiffness or roughness according to the tissue deformity or tissue appearance changes as observed on the endoscopic view [5]. However, this method is highly subjective and varies between different surgeons. Besides, it is hard to share the experience to other surgeons due to the lack of objective quantification and standardization.

To help surgeons retrieve this missing sense of touch during RMIS,

* Corresponding author.
E-mail address: chien-hsi.chen@ovgu.de (C.H. Chen).

https://doi.org/10.1016/j.compbiomed.2019.103370
Received 15 March 2019; Received in revised form 25 July 2019; Accepted 25 July 2019

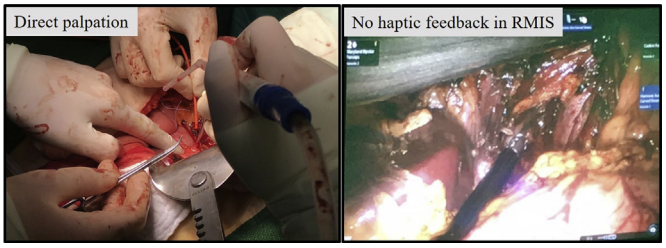

Fig. 1. Comparison between traditional open surgery and RMIS.

various sensors have been investigated. Most of them use a force sensor located at the tip of the robotic surgical instrument [6]. For example, interaction forces during suturing of sponges were measured in Ref. [7]. One of the main drawbacks of these approaches is that the force sensors have to be integrated or embedded into the surgical tools, influencing the tool cost and sterilization procedures negatively.

With tactile sensors a more delicate measurement can be provided. Tactile sensors mainly consist of a pressure sensor array to detect the changing surface or roughness when they are interacting with objects [8]. BioTac, an advanced tactile sensor involving multiple pressure sensors and an additional temperature sensor, was tested on a real da Vinci surgical robot [9,10]. Although these sensors returned a level of touching sensation, their principle technology approach comes with the same problems of integration and sterilization, since the transducer needs to be located close to the device tip in direct contact with organs. In addition, the size of these solutions is relatively large due to the need of peripheral electronic components and connection cables.

Apart from force sensors or tactile sensors that acquire mechanical signals, acoustic emission (AE) sensing techniques have been extensively used for process observation in different research and development fields, such as mechanical machining and non-destructive testing. AE sensing is known to present important advantages such as low-cost and non-invasiveness. It is also a passive sensing technique involving both fast reactions and high sensitivity to subtle changes in processes. Previous studies have shown the feasibility of AE for acquiring feedback information noninvasively from minimally invasive tools [11–13]. This method comes with the advantage to avoid direct contact between sensors and the tissue at the surgical site. The audio signal produced by the tool-tissue interaction can be transmitted over the structure of the instrument and picked up by an audio sensor at its proximal end. This concept reduces the system complexity and allows non-invasive placement of the sensor. This innovative approach has been tested in the meantime for use with biopsy needles and guide wires [11–13], where audio signatures were extracted to detect tissue boundaries during the insertion of a needle and to characterize a guide wire perforation. Recently, this approach has been adapted for obtaining feedback information from a da Vinci grasper instrument for analyzing the feasibility of using audio in RMIS [14]. Preliminary tests have been performed and the obtained results showed a significant potential for using audio as a tool for acquiring haptic information during robotic surgery.

The aim of this work was to verify, if the AE sensing technique can be applied on the robotic surgical tools for differentiating textures, when the instrument palpates different tissues by analyzing the audio signals resulting from the tissue-tip interaction. Based on the results a solid and ready-to-use setup for implementation in the real operating room was is proposed and presented in this paper.

2. Methods

Traditional MIS requires clinical staff to hold interventional tools and for use directly at the surgery table. However, a da Vinci surgical robot with endoscope and robotic surgical tools is remotely controlled by surgeons sitting at a console (Fig. 2). There is no direct contact between the surgeon and the robotic arm containing the tool. These are things to be considered in the experimental setup.

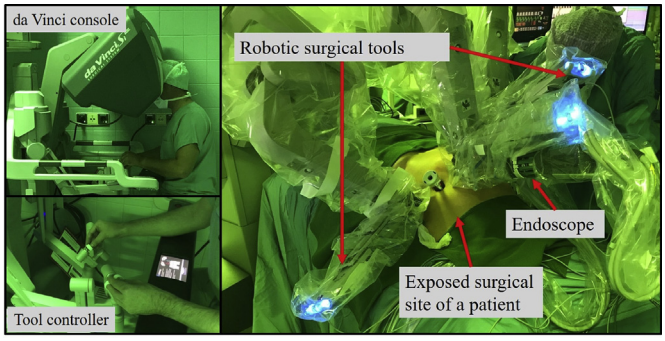

Fig. 2. The setup of da Vinci surgical robot during real surgery.

2.1. Experimental setup

To ensure a consistent signal quality and with that a high acquisition and test reliability, a stable platform holding a da Vinci Prograsp Forceps (Intuitive Surgical, California, USA) was built as shown in Fig. 3. This platform allows a linear movement of the Forceps over surfaces. The control unit of the Forceps (red box) was clamped at the height of 30 cm on a stable stand, which can be manipulated to perform a smooth horizontal pivot. Without additional force pressing downward, the Forceps tip rested with its own weight on the surface of the table where the experiments were performed. Four different textures, two synthetic and two biological textures, were used for the experiments. The two synthetic textures (called textures A and B in the sequel) had dimensions of 36 cm × 8.5 cm each and were attached to the surface of the cork pin board as shown in Fig. 4. Texture A was a 2 mm-thick soft cloth, texture B a 0.5 mm-thick Denim cloth. The microscopic image (SZ61 Stereomicroscope, Olympus, Japan) in Fig. 4 shows that the surface of Texture B was covered with protruding round shape structures. On the other hand, Texture A was more homogeneous on the surface. The two biological tissues used for testing were a pig liver and a pork fillet as shown in Fig. 4.

For audio acquisition the Digital Acoustic Visualization Interactive Device (DAVID) was designed, as shown in Fig. 3. An audio sensor (Adafruit I2S MEMS microphone SPH0645LM4H–B, Knowles, Illinois, USA) was used for this experiment. This sensor has a built-in amplifier and an on-board analog-to-digital converter. The acquired audio signals were sent via cable to a hosting device (Raspberry Pi 3 Model B+, Raspberry Pi Foundation, UK) in standard I2S format.

To attach the sensor properly, an inspection of the internal structure

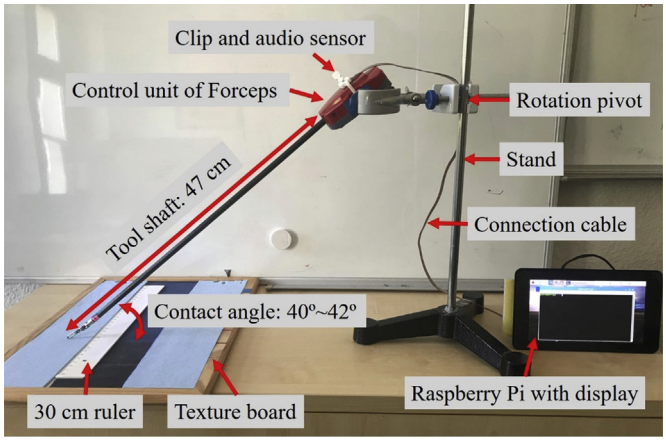

Fig. 3. Overview of the experiment setup.

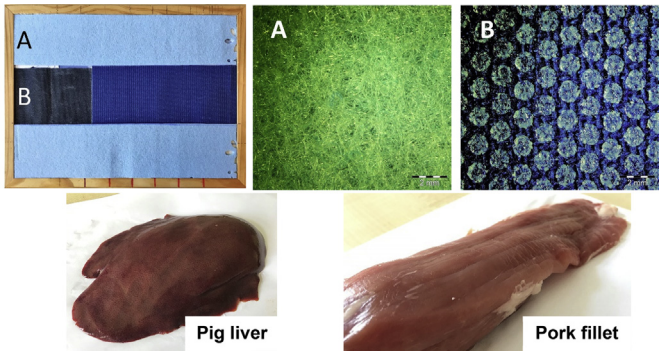

Fig. 4. Cork pin board (upper row left) with texture A (upper row middle, soft cotton) and B (upper row right, Denim). Biological tissues, including Pig liver (bottom row left) and pig fillet (bottom row right).

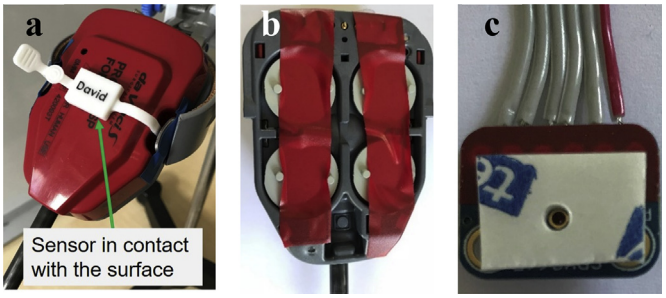

Fig. 5. (a) Digitalized Acoustic Visualization Interactive Device (DAVID) attached on the control unit of the Forceps (b) Fixation of the four controlling wheels to limit the Forceps tip movement (c) Stability enhancement using double-side tape for the audio sensor fixation under the audio sensor.

of the forceps was done and an ideal position for placing the sensor was identified. The surface at the central position of the forceps control unit was right above the hub of all the pivot rotators of the pulling wires inside the instrument. All audio signals produced from the tissue-tool interaction would pass through this position. Therefore, a customized 3D printed clip was designed in compliance with the shape of the surface for a fixed and stable position as shown in Fig. 5a. It was manufactured to avoid any physical modification of the forceps and to ensure no conflict with the original functionality. At the backside of the control unit were four wheels to connect to the robot arm. These wheels were fixed in a neutral position so that the tip would stay in the same orientation during the experiments (Fig. 5b). In order to fill out the small gap between the audio sensor and the surface of the control unit, a 1-mm-thick double-side tape with a 1-mm-diameter hole was attached on top of the audio sensor (Fig. 5c). This setup ensures firm contact to the control unit and prevents noises caused by friction. Another piece of double-side tape was attached on the other side of the sensor to press down for a better fixation.

2.2. Experiment protocol

In order to evaluate the feasibility of this method, an experimental protocol was developed. The tip of the Forceps was moved 30 cm manually on the texture board from left to right within 6 s (i.e. at a velocity around 5 cm/s). This one-direction movement was defined as one "sweep" (Fig. 6). For the biological tissues, due to the limitation given by their natural size and shape, the reachable range was between 8 and 10 cm for the liver and around 20 cm for the pork fillet. In order to keep the results of the experiment consistent and comparable between all the non-biological and biological tissues, one sweep would take around 2–3 s on the liver and around 4 s on the pork fillet to stay close to the previously defined sweep speed of 5 cm/s.

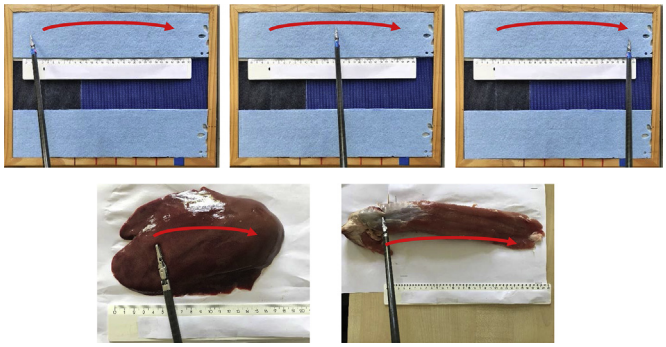

Fig. 6. The "sweep" path of the Forceps tip during the experiment on different textures.

The experiment was designed as follows:

For each trial 10 sweeps were recorded for each texture and the same procedure was then repeated 3 times (i.e. session 1, 2 and 3 as shown in Fig. 7) for a total 12 trials.

2.3. Data acquisition

An example code released by the audio sensor vendor [15] for recording sound on the Raspberry Pi was used for logging the data. The volume gain of the signal was set to *45*% for better signal-to-noise ratio based on the vendor's documentation. The sampling frequency was set to 16 KHz and the recordings were saved as WAV file format.

2.4. Signal processing procedure

The main hypothesis in this work is that different textures should result in different instrument tool-tissue interaction patterns that would be detectable in the dynamics of the proximal acquired audio signal. We further assumed that the acoustic signals produced by the sweeps may have different spectral energy distribution for the different textures due to different levels of texture roughness and softness.

Employing advanced signal processing using a spectral approach, these differences between the audio dynamics of the four tested textures should be detectable. We planned to first implement a sort of spectral averaged model of each texture and then test the resulting model by assessing spectral similarities between the spectrum of test signals and the estimated spectral models of each texture. Higher similarity measurement should be obtained when a test signal and the model come from the same texture.

The diagram of Fig. 8 shows the main steps of the signal processing strategy used for analysis and similarity assessment between the spectra. As an example, we display the main steps only for the synthetic tissues (the procedure is the same for the biological tissues). In order to perform a non-overlapping analysis, the sweeps of sessions 1 and 3 were used for extracting test signals and the sweeps of session 2 were used for extracting signal segments for implementing the spectral model. Each step of the signal processing algorithm is described in the following section.

2.4.1. Sweep identification and random segments extraction

For each recording belonging to a session and texture, the sweep segments are first identified and extracted. For that a CUSUM algorithm [16] is used for detecting abrupt changes in the signal that correspond to onset and offset time instants of a sweep. Then, segments of 1.5 s length are randomly extracted from the different sweep signals. For sessions 1 and 3 corresponding to the test signals, 200 segments were extracted, while for session 2 corresponding to data for implementing the spectral model, 2000 segments were extracted (see in Fig. 8, the outputs of the Random Segment Extraction block).

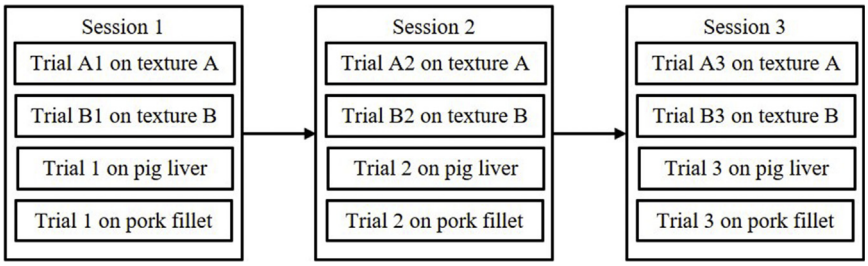

Fig. 7. Protocol overview.

2.4.2. Power spectral density estimation

The Power Spectral Density (PSD) was computed for a total 9600 randomly extracted segments among the four teseted textures (2400 signal segments per texture). In this work two methods were used for estimating the PSD: the standard FFT-based PSD, denoted in the sequel as FFT-PSD, and a parametric spectral estimation using auto-regressive (AR) model, denoted in the sequel as AR-PSD.

AR modelling is a well-known technique for parametrical spectral estimation and many research papers have addressed its advantages over non-parametrical based methods (for detailed information about AR modelling please refer to Ref. [17]). Two characteristics of the audio signals obtained in this work are their significant noise and that they present narrow-band characteristics in the spectrum. One important advantage of AR modelling is that when an appropriate model is selected, it presents a higher spectral resolution even in signals with poor SNR and with less data than classical methods. Additionally, AR modelling is well suited for the analysis of signals presenting peaky narrow-band characteristics.

An AR model assumes that the value of the current sample in a signal can be modeled as a linearly weighted sum of values in the past of the same signal, and a white zero mean noise. For a randomly extracted segment $s[n]$ resulting from the Random Segment Extraction step (see Fig. 8) the AR model can be defined as:

$$s[n] = -\sum_{k=1}^{p} a_k s[n-k] + e[n]$$

where n corresponds to the discrete-time index, $a_k\{k = 1,2,...,p\}$ are the estimated AR parameters, and where $e[n]$ corresponds to a white zero mean noise. The z-transform can be applied and then the AR model transfer function can be expressed as:

$$H(z) = \frac{S(z)}{E(z)} = \frac{1}{1 + \sum_{k=1}^{p} a_k z^{-k}}$$

and finally, the AR-PSD of a randomly extracted segment can be computed by evaluating $H(z)$ around the unit circle in the complex plane, i.e., $z = e^{j2\pi f}$:

$$P_{AR}(f) = \frac{1}{|1 + \sum_{k=1}^{p} a_k e^{-j2\pi fk}|^2}$$

2.4.3. Spectral model implementation and similarity computation with test signals

After the estimation of the PSD (using a periodogram or AR modelling) the spectral model is computed for each texture by averaging the PSDs resulting from the 2000 segments per texture randomly extracted from the recording of Session 2 (see Fig. 8, the outputs of step Spectrums Averaging for the synthetic textures).

In order to evaluate the obtained models, similarity metrics using a Pearson coefficient are computed between the PSDs of the single test signals extracted from sessions 1 and 3 and the texture models. For a better visualization of the correlations between the spectra, the Pearson coefficient was arranged in a matrix representing each texture model. Each element (i,j) of the matrix represents a Pearson coefficient obtained between the texture PSD model and the PSD of a test signal

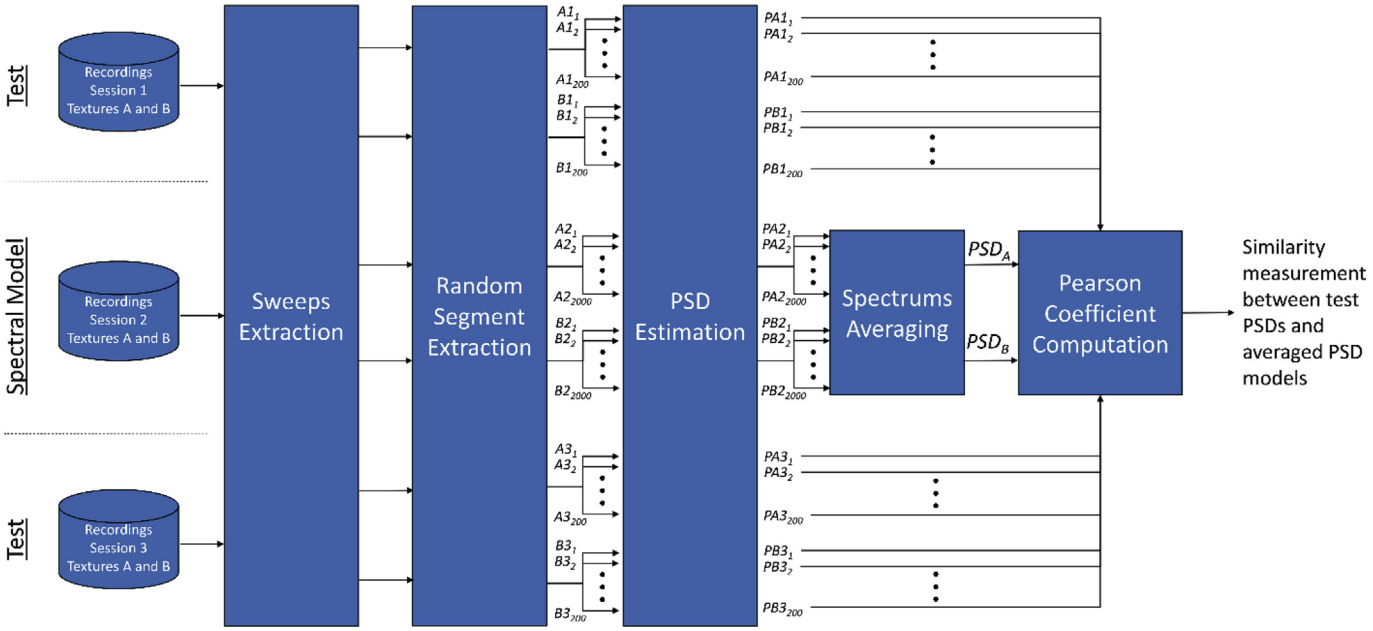

Fig. 8. Main signal processing steps for computation and evaluation of the spectral model of the two synthetic textures. The procedure is the same for the biological textures.

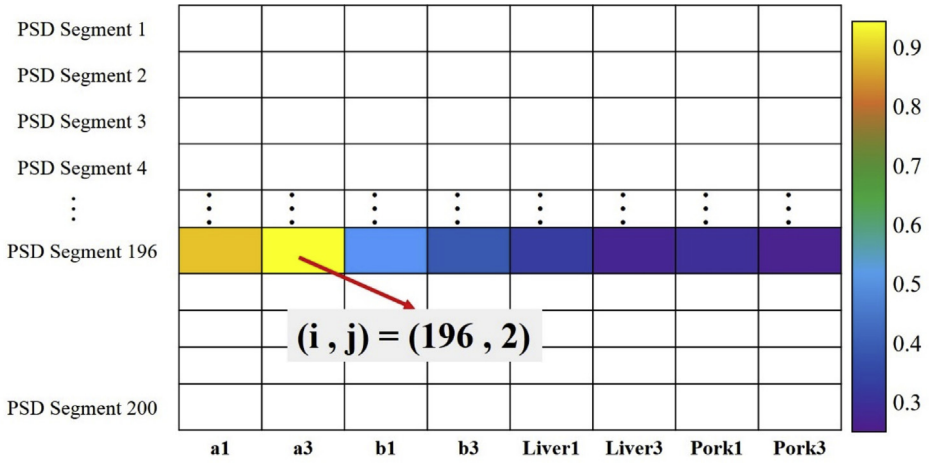

Fig. 9. Example of Pearson coefficient matrix fore PSD texture model evaluation.

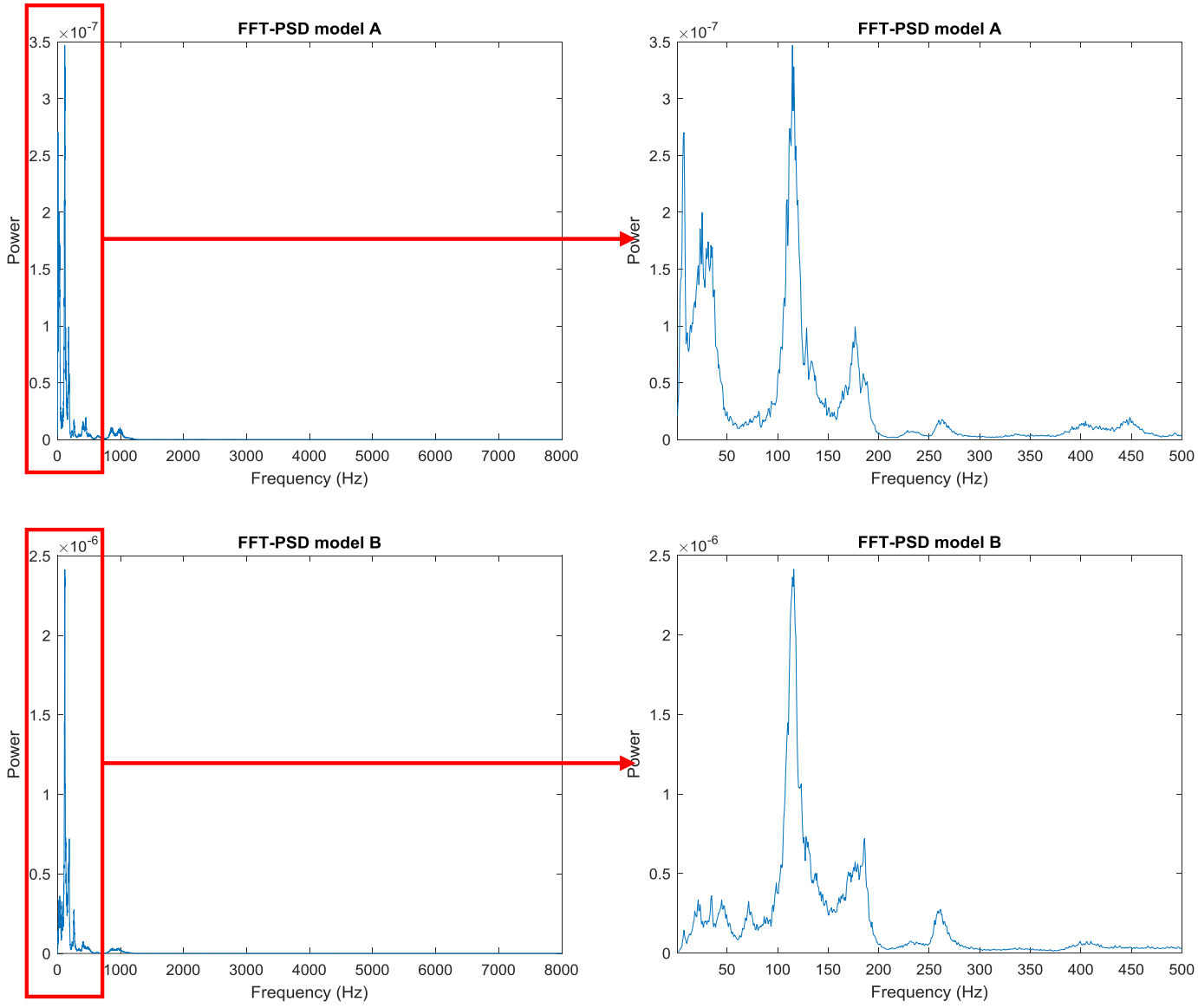

Fig. 10. FFT-PSD models of textures A (top) and B (bottom) in the frequency range 0–8 kHz (left) and 0–500 Hz (right).

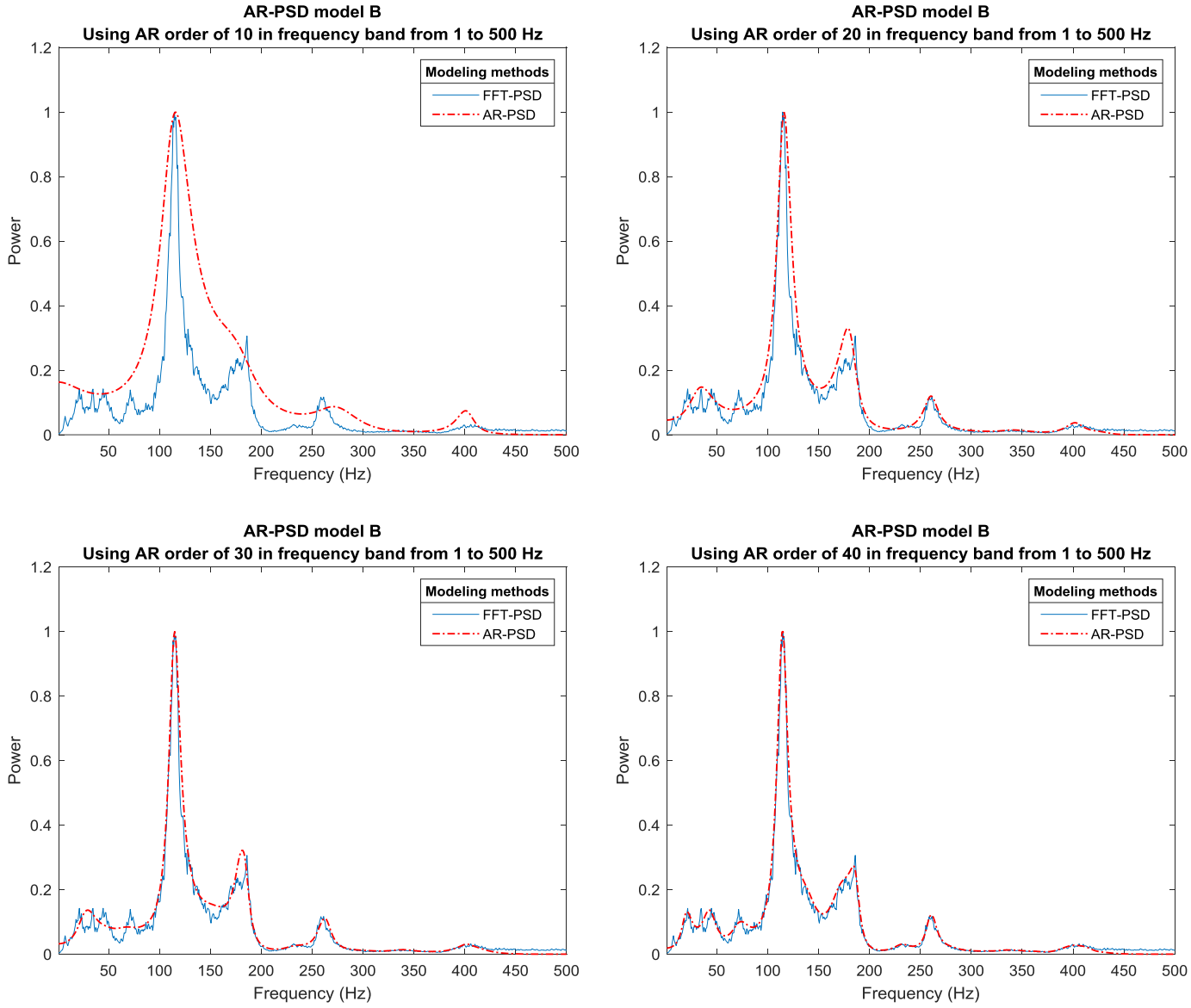

Fig. 11. AR-based PSD spectrums of texture model B for different AR orders ($p = 10, 20, 30, 40$) and comparison with the FFT-based PSD.

belonging to one of the 200 segments of a given texture extracted from sessions 1 or 3. For example, as shown in Fig. 9, the element (i,j) = (196,2) represents the value of the Pearson coefficient between one of the texture PSD models (A, B, liver or pork) and the randomly selected segment No. 196 of session 3 of texture A.

2.4.4. Parameter settings

Fig. 10 shows, as example, the obtained spectral model for textures A and B in the whole frequency range and also between *0* and 500 Hz. It is possible to verify, that most of the spectral energy is located under 500 Hz. Specifically, for texture A *87%* of the energy is concentrated in the first 500 Hz and for texture B *92%* is concentrated in the first 500 Hz. Therefore, in order to obtain a better assessment in the correlation, the Pearson coefficient was computed only for the first 500 Hz.

Taking this information into account and in order to reduce the complexity of the AR model, the parametrical PSD estimation of the signal was first decimated 16 times obtaining a signal with a frequency sampling of 1 kHz.

The AR model order was selected in such a way that the AR spectrum models the main frequency components in the first 500 Hz. In Fig. 11, the FFT-based PSD model of texture B and the AR-based PSD model for four different model orders are visualized. It is possible to verify, that with an order p = 40 a good delineation of the main frequency component is obtained. Therefore, this order was selected for the subsequent AR analysis. For the estimation of the AR coefficients we used the Yule-Walker method [17]. The algorithm and the analysis were implemented in Matlab R2015b.

3. Results

In this section, we will first show qualitative results concerning how each texture behaves in terms of audio spectrum, how there are some main frequency components that are similar in terms of energy distribution and other components that are characteristic of each texture and that are useful for distinguishing between textures. Then quantitative results are presented using the spectra similarity procedure explained in Section 4.3.

In order to show the spectral behavior of audio recordings belonging to each tested texture and their differences, two sweeps (see Section 4.1) belonging to each of the four textures were concatenated and then a time-variant AR (TV-AR) spectrum was computed to look for changes in the main frequencies of the resulting spectrum. The first two

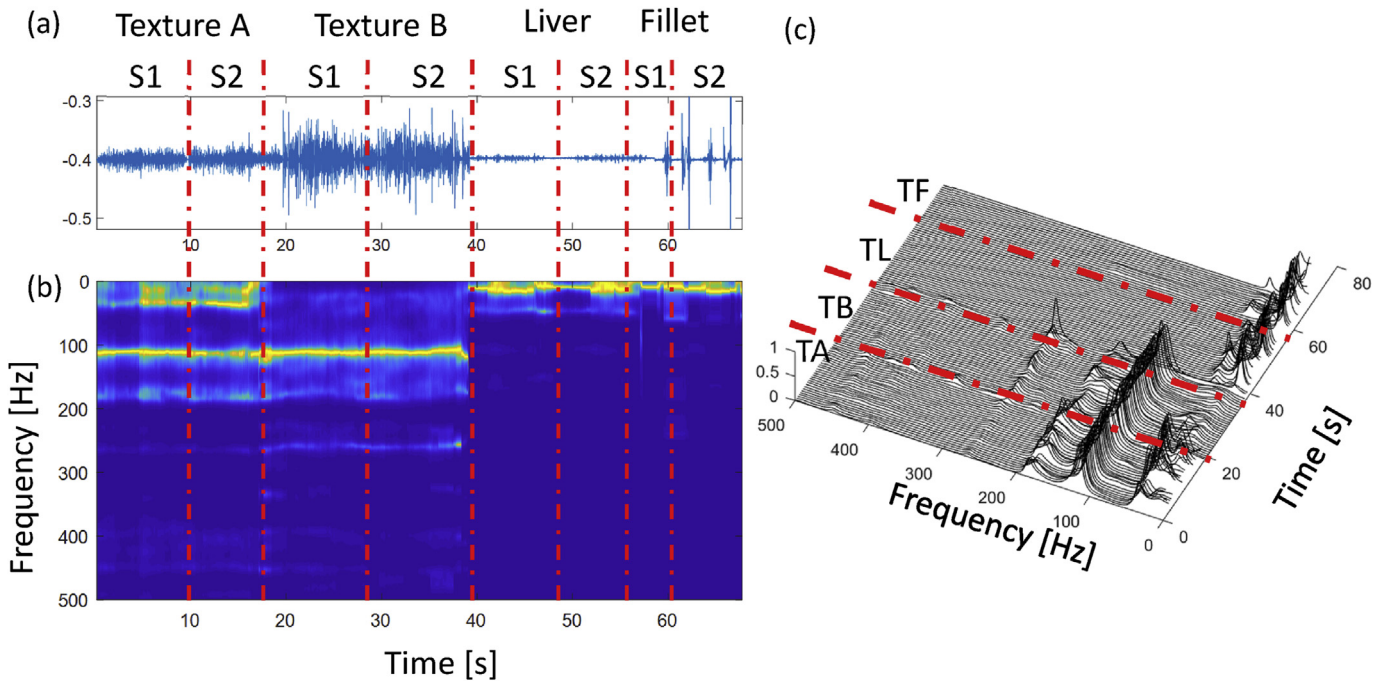

Fig. 12. Time-variant AR analysis of a concatenated audio signal involving two sweeps (S1 and S2) belonging to textures A (TA), B (TB), pig liver (TL) and pork fillet (TF).

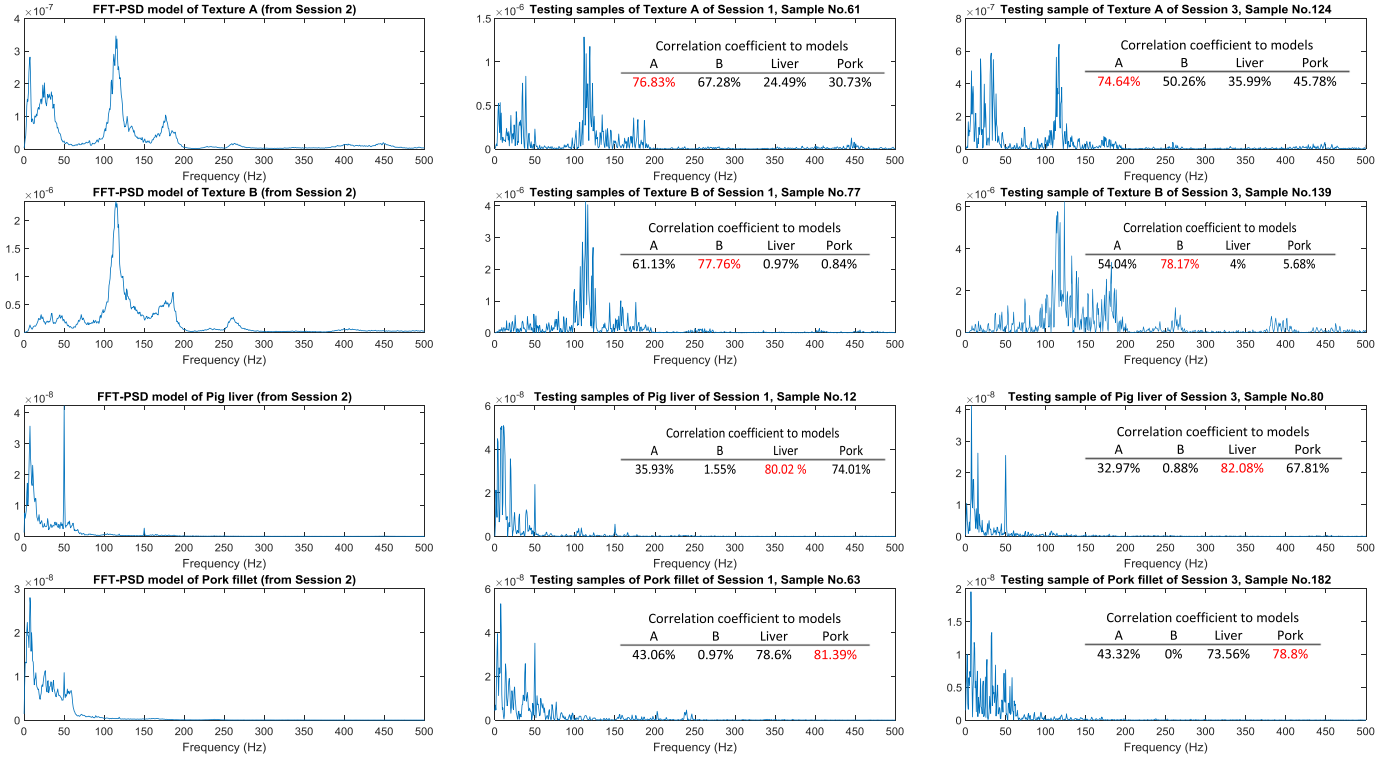

Fig. 13. Correlation examples from FFT-PSD models and random testing samples (frequency band: 0–500 Hz).

concatenated signals belong to sweeps of the audio recordings of Texture A, the third and fourth to sweeps of the audio recordings of Texture B, the fifth and sixth to sweeps of the audio recordings of pork liver, and the seventh and eighth to sweeps of the audio recording of pork fillet. All the sweeps were taken from sessions 1 and 3 for each texted texture.

The TV-AR was computed over a sliding window of 1.5 s duration and an overlap of 90%. In each window a 40th order AR model was applied to estimate the AR parameters using the Yule-Walker method and for each of the windows the AR spectrum was computed. To better visualize the time-variant frequency components, the AR spectrum estimated at each window was normalized between 0 and 1.

Fig. 12 displays the results of the concatenated analysis. In Fig. 12a, the concatenated time-domain audio signals are displayed, and here it is possible to observe that all the textures involve different time-domain patterns in terms of dynamics and energies and that these patterns are

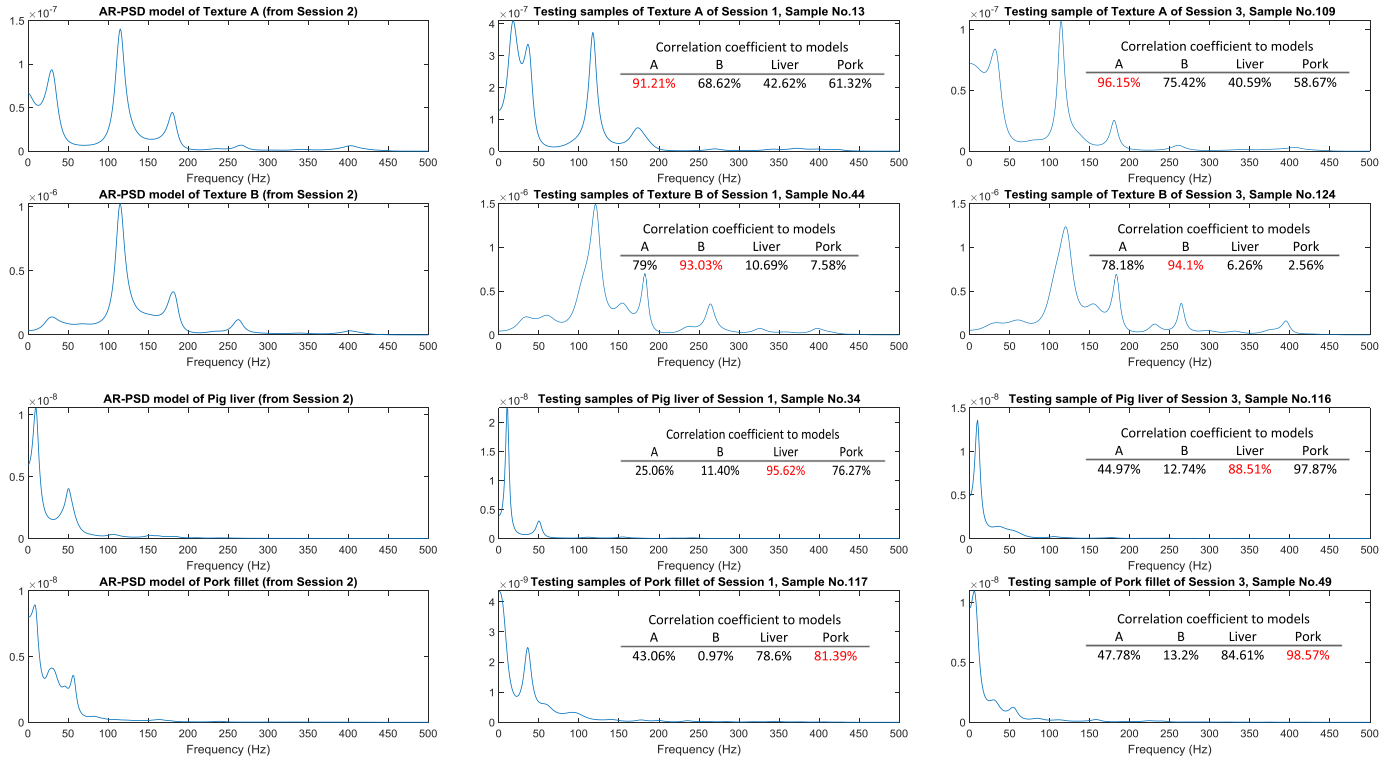

Fig. 14. Correlation examples from AR-PSD models and random testing samples (frequency band: 0–500 Hz).

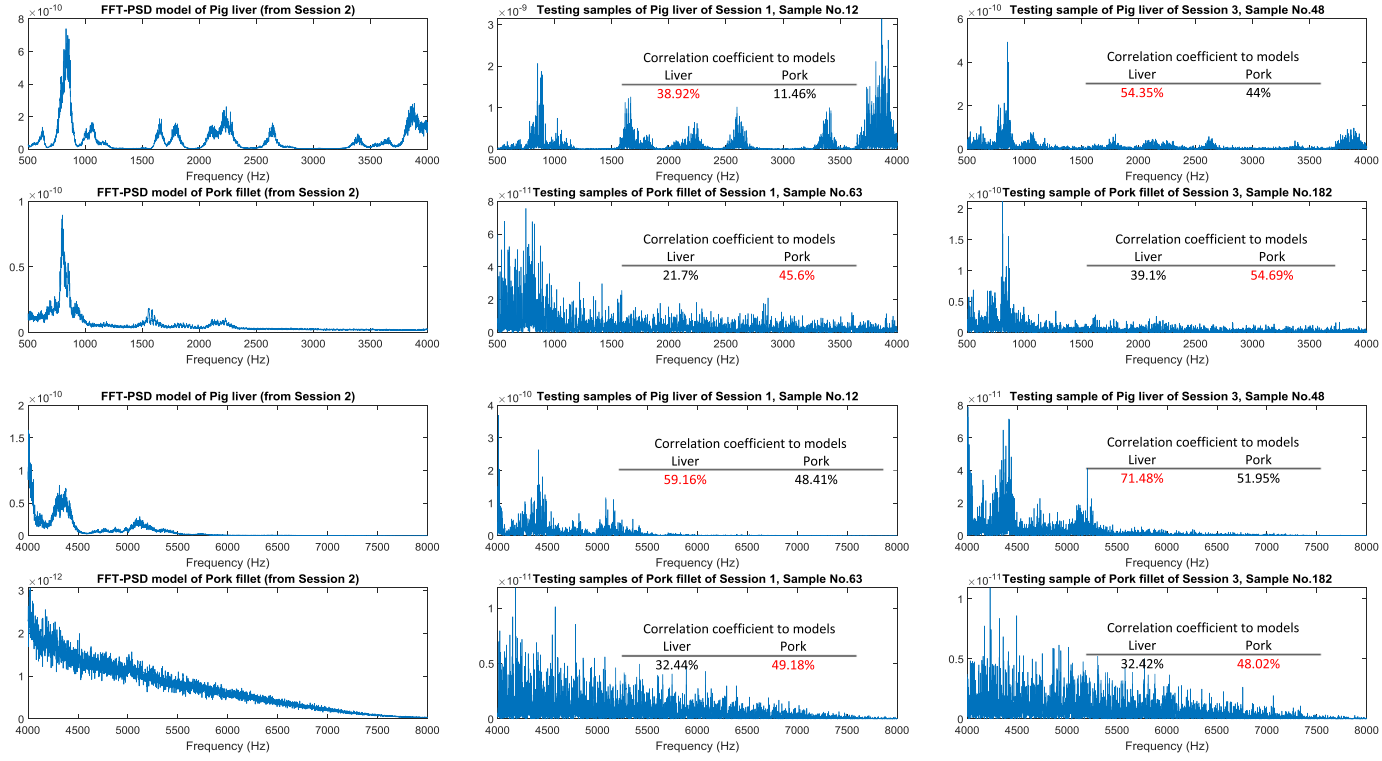

Fig. 15. Correlation examples from FFT-PSD models and random testing samples for pig liver and pork fillet (frequency bands: 500–4000 Hz and 4000–8000 Hz).

stable for sweeps belonging to the same texture. For example, texture B produces audio dynamics of higher energy than Texture A. This can be explained by the fact that texture A is softer than texture B. Also, it was observed that the pork fillet produces more transient dynamics than the other textures since it is the least homogeneous texture of the four tested textures. Fig. 12b shows the TV-AR spectrum synchronous with the time-domain concatenated signal. This spectrum shows different main frequency components when synthetic and biological textures are compared. For the synthetic textures, a main frequency component at around 110 Hz appears to be similarly stable over time for both textures. However other stable frequency components denote the differences between both synthetic textures. It is clear that Texture A

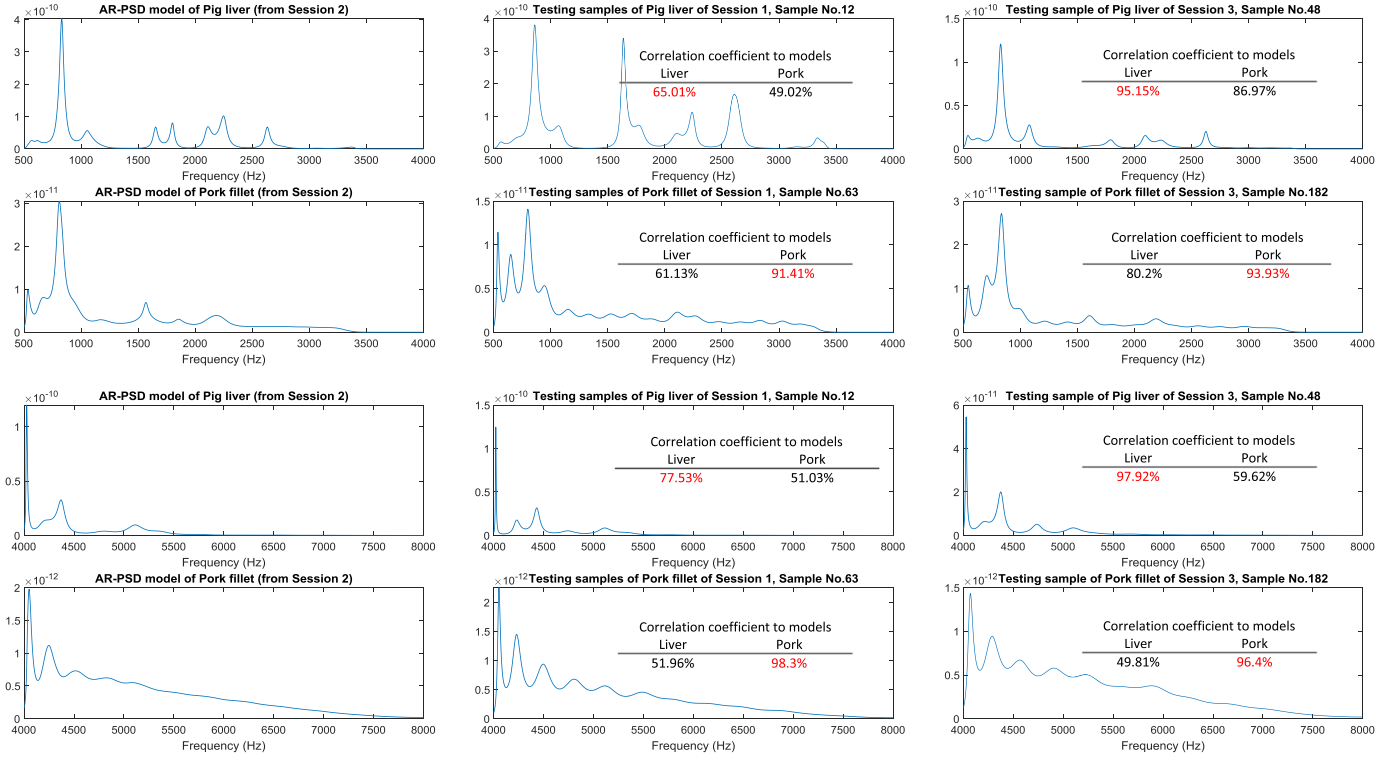

Fig. 16. Correlation examples from AR-PSD models and random testing samples for pig liver and pork fillet (frequency bands: 500–4000 Hz and 4000–8000 Hz).

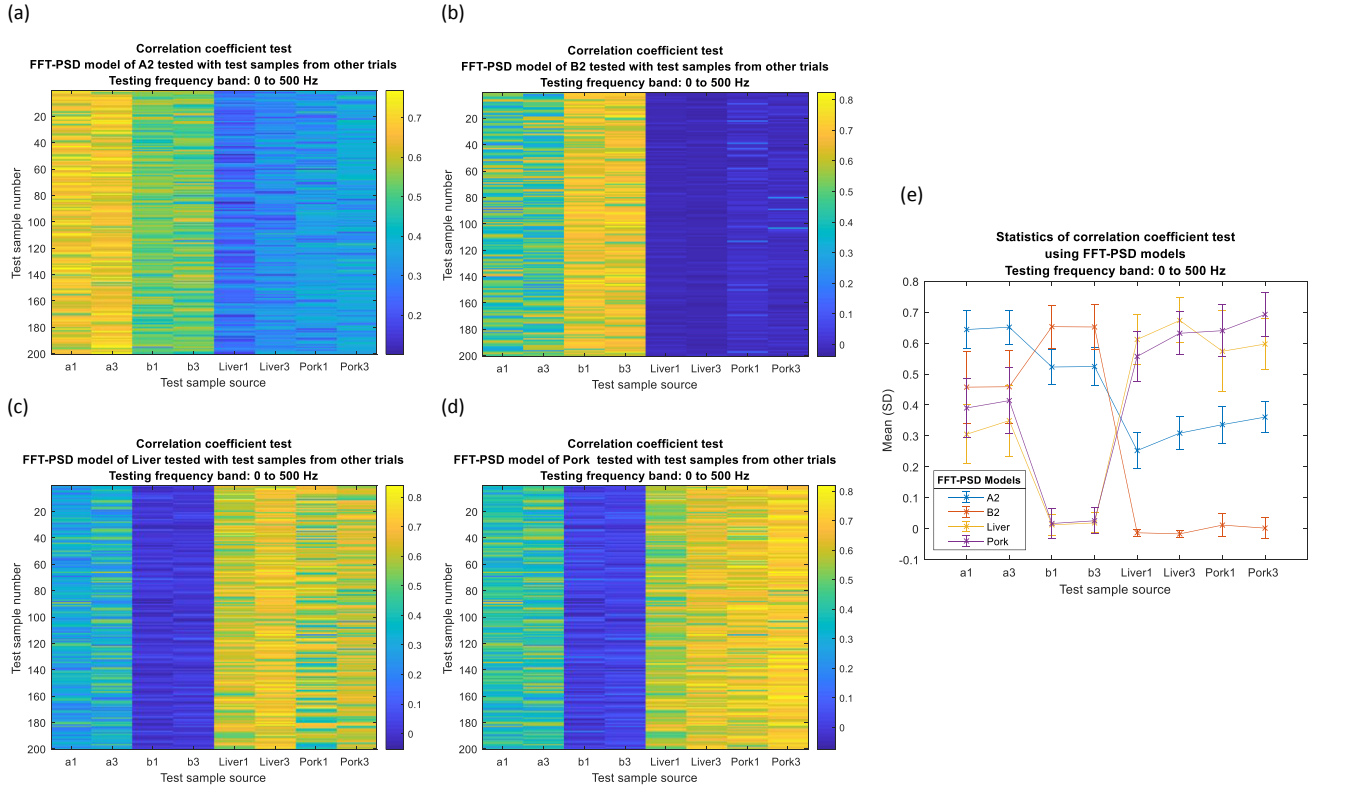

Fig. 17. Overall test results and statistics using FFT-PSD models within 500 Hz frequency band.

involves low frequency components (between 0 and 40 Hz) that are significantly higher than the ones for texture B. This also can be explained by the fact that softer textures should result in lower frequency components. Also, in both texture an important frequency component is present around 250 Hz but with higher energy for texture B. In the case of the biological textures it is possible to observe that both main frequency components are much lower compared to the ones of the synthetic textures. This is explained by the fact that biological textures are generally softer than the synthetic ones. Moreover, the time-frequency behavior of the main components belonging to the biological textures

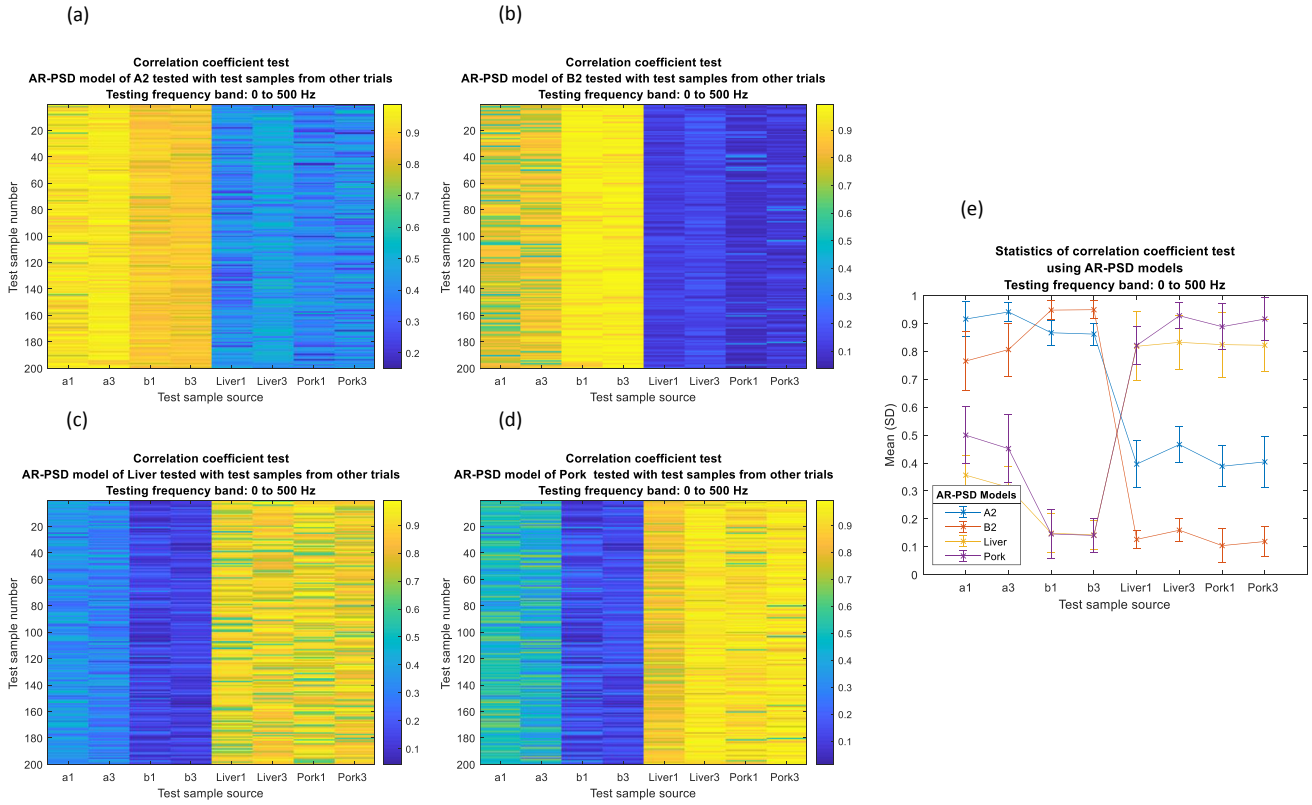

Fig. 18. Test results and statistics using AR-PSD models within 500 Hz frequency band.

Table 1
Correlation test results using FFT-PSD models among all four testing textures with frequency band from 0 to 500 Hz.

Test sample source	FFT-PSD models (mean ± SD)			
	Texture A	Texture B	Pig Liver	Pork fillet
a1	0.64 ± 0.06	0.46 ± 0.12	0.31 ± 0.09	0.39 ± 0.09
a3	0.65 ± 0.05	0.46 ± 0.12	0.35 ± 0.11	0.41 ± 0.11
b1	0.46 ± 0.12	0.65 ± 0.07	0.01 ± 0.03	0.02 ± 0.05
b3	0.46 ± 0.12	0.65 ± 0.07	0.02 ± 0.03	0.03 ± 0.04
liver1	0.25 ± 0.06	−0.01 ± 0.01	0.61 ± 0.08	0.56 ± 0.08
liver3	0.31 ± 0.05	−0.02 ± 0.01	0.67 ± 0.07	0.63 ± 0.07
pork1	0.34 ± 0.06	0.01 ± 0.04	0.57 ± 0.13	0.64 ± 0.08
pork3	0.36 ± 0.05	0.00 ± 0.03	0.6 ± 0.08	0.69 ± 0.07

Table 2
Correlation test results using AR-PSD models among all four testing textures with frequency band from 0 to 500 Hz.

Test sample source	AR-PSD models (mean ± SD)			
	Texture A	Texture B	Pig Liver	Pork fillet
a1	0.91 ± 0.06	0.77 ± 0.11	0.36 ± 0.07	0.50 ± 0.1
a3	0.94 ± 0.03	0.80 ± 0.1	0.31 ± 0.08	0.45 ± 0.12
b1	0.87 ± 0.04	0.95 ± 0.03	0.15 ± 0.07	0.15 ± 0.09
b3	0.86 ± 0.04	0.95 ± 0.03	0.14 ± 0.05	0.14 ± 0.06
liver1	0.4 ± 0.08	0.13 ± 0.03	0.82 ± 0.12	0.82 ± 0.07
liver3	0.47 ± 0.06	0.16 ± 0.04	0.83 ± 0.1	0.93 ± 0.05
pork1	0.39 ± 0.07	0.1 ± 0.06	0.82 ± 0.12	0.92 ± 0.08
pork3	0.4 ± 0.09	0.12 ± 0.06	0.82 ± 0.09	0.92 ± 0.08

are less stable than the synthetic ones since a biological texture is generally less homogeneous than a synthetic texture. In order to better visualize the spectral differences between the textures in Fig. 12c the same TV-AR spectrum is displayed but on a 3D plotting. Here it is possible to observe that the time-frequency differences between both biological textures are not visually evident. However, we will see later that the spectrums for both biological textures in the frequency range 0–500 Hz are different and that this difference increases in higher frequency ranges.

The spectral information, which can be extracted from randomly selected sweeps of sessions 1 and 3, is consistent with what we can observe from the obtained models from Session 2 displayed in Fig. 10. Because of that we had the idea to implement a simple similarity computation between the spectral model and the spectrum obtained from a test signal for quantitative differentiation between the textures. Figs. 13 and 14 present some examples of spectral similarity computation using the Pearson coefficient (in percentage of similarity) for the FFT-PSD estimation and for AR-PSD estimation respectively. At the left column of each figure, the obtained models for each texture are displayed, and the middle and right columns of the figure show the PSDs of the test signal, one per texture taken from Session 1 in the middle and from Session 3 on the right. Additionally, the Pearson coefficient values (in percentage) are displayed for similarity assessment of the test signals with the models of all textures.

It can be observed that for both PSD estimations, FFT-based in Fig. 13 and AR-based in Fig. 14, higher Pearson coefficient values are obtained when the test signal and the model come from the same texture. Additionally, as mentioned above, the spectral behavior between synthetic and biological tissue is different and results in Pearson coefficients with completely different values when biological and synthetic textures are compared. However, for distinguishing between biological textures, even if high Pearson values are obtained, the difference between both textures is not as high as between the synthetic textures. Nevertheless, as shown in the spectrums (FFT-based and AR-based) in Figs. 15 and 16, for frequencies higher than 500 Hz the difference between the two biological tissues increases. With that, a further investigation for two higher frequency bands (500–4000 Hz and 4000–8000 Hz) for the biological tissues was conducted using the same

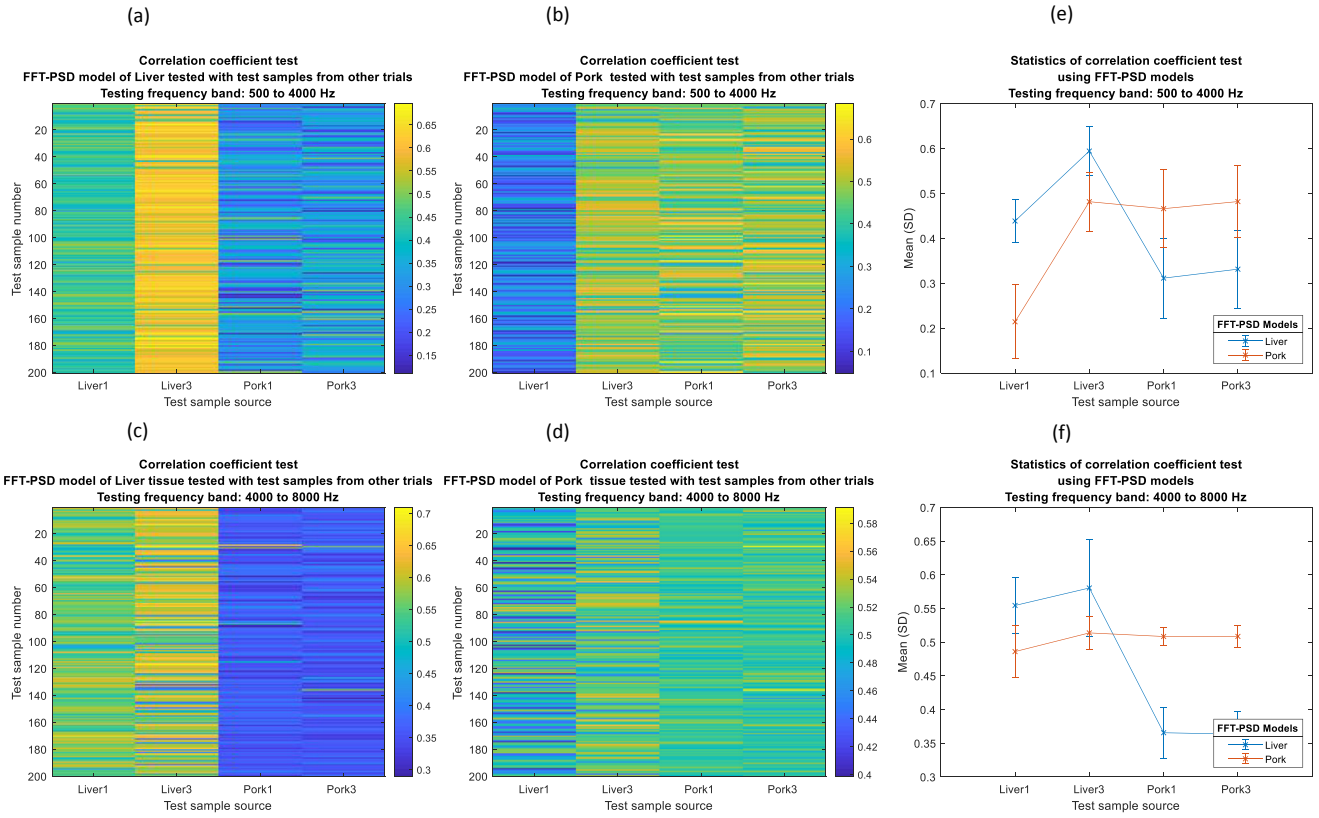

Fig. 19. Test results and statistics from tissues of pig liver and pork fillet using FFT-PSD models in two different frequency bands.

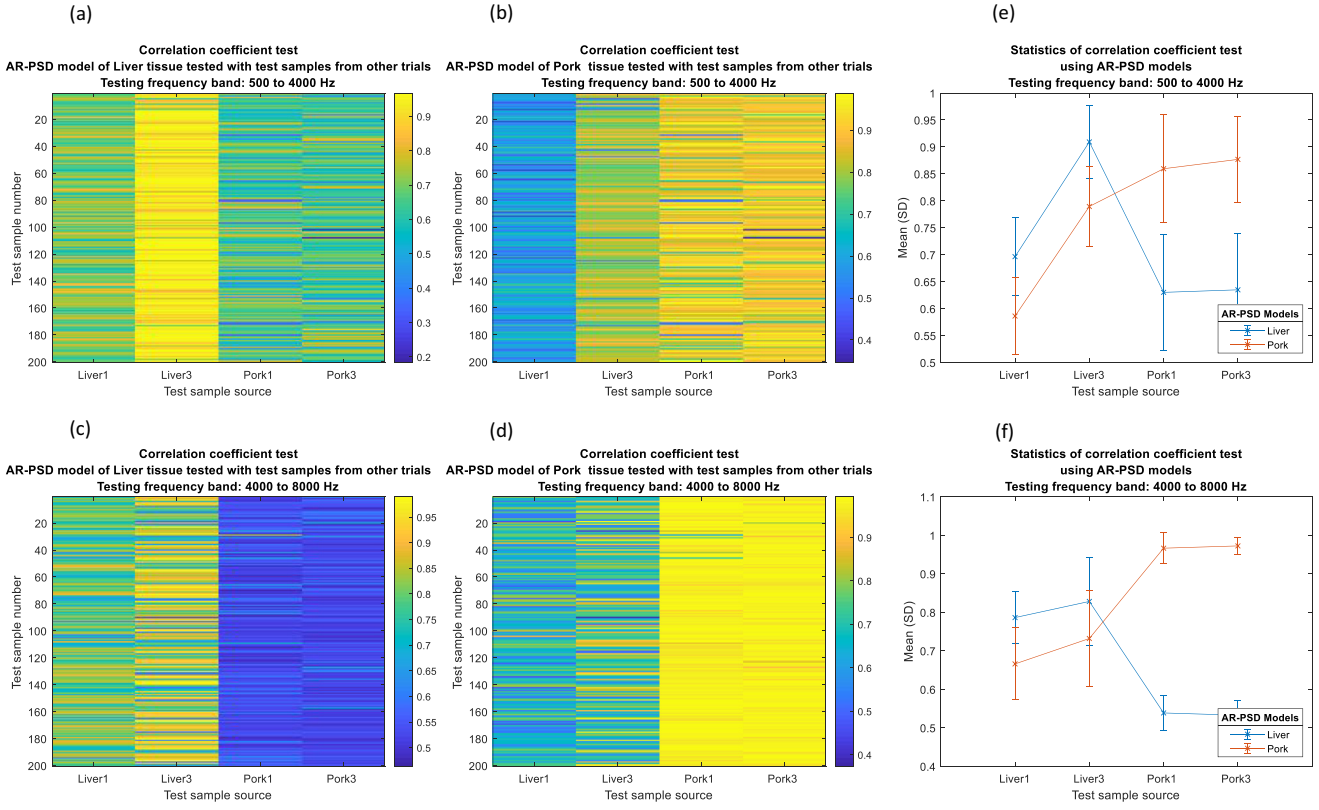

Fig. 20. Test results and statistics from tissues of pig liver and pork fillet using AR-PSD models in two different frequency bands.

Table 3
Correlation test results using FFT-PSD models among biological tissues with two different frequency bands.

Test sample source	FFT -PSD models (mean ± SD)			
	Bands: 500–4000 Hz		Bands: 4000–8000 Hz	
	Pig Liver	Pork fillet	Pig Liver	Pork fillet
liver1	0.44 ± 0.05	0.21 ± 0.08	0.55 ± 0.04	0.49 ± 0.04
liver3	0.59 ± 0.05	0.48 ± 0.07	0.58 ± 0.07	0.51 ± 0.02
pork1	0.31 ± 0.09	0.47 ± 0.09	0.37 ± 0.04	0.51 ± 0.01
pork3	0.33 ± 0.09	0.48 ± 0.08	0.36 ± 0.03	0.51 ± 0.02

Table 4
Correlation test results using AR-PSD models among biological tissues with two different frequency bands.

Test sample source	AR -PSD models (mean ± SD)			
	Bands: 500–4000 Hz		Bands: 4000–8000 Hz	
	Pig Liver	Pork fillet	Pig Liver	Pork fillet
liver1	0.70 ± 0.07	0.59 ± 0.07	0.79 ± 0.07	0.67 ± 0.09
liver3	0.91 ± 0.07	0.79 ± 0.07	0.83 ± 0.11	0.73 ± 0.12
pork1	0.63 ± 0.11	0.86 ± 0.1	0.54 ± 0.05	0.97 ± 0.04
pork3	0.63 ± 0.11	0.88 ± 0.08	0.53 ± 0.04	0.97 ± 0.02

proposed approach.

In order to quantitatively evaluate the suitability of a spectral similarity approach for differentiating between textures, we use the correlation matrix visualization explained in Fig. 9. Figs. 17 and 18 show the Pearson coefficient matrices for FFT-PSD and for AR-PSD respectively in the range 0–500 Hz for the four tested textures. Each matrix in the figure represents the correlation of the test signals with one of the texture models. Figs. 17a and 18a show the correlation matrices with the Pearson coefficient between the test PSD signals and texture PSD model A using FFT-PSD and AR-PSD estimation respectively and Figs. 17b and 18b show the same information but with texture model B. Figs. 17c and 18c shows the correlation matrices of pig liver model and 17d and 18d show the matrices of pork fillet model. It is possible to verify the difference in color intensity when two compared PSDs (model and test) belong to the same texture. Additionally, Figs. 17e and 18e display the average and standard deviation for each of the matrices confirming the difference in Pearson coefficient values when model and test PSDs are taken from the same texture. These average and standard deviation values are shown in Tables 1 and 2 for FFT-PSD and AR-PSD estimation respectively.

Further evaluation was performed in higher frequencies in order to evaluate if it is possible to better differentiate between the two biological tissues in other frequency bands. For that we implemented additional correlation matrices in the frequency bands 500–4000 Hz and 4000–8000 Hz. This is shown in Figs. 19 and 20, where the correlation matrices with the Pearson coefficient between the test PSD signals and PSD model from liver and pork fillet using FFT-PSD and AR-PSD estimation in the two frequency bands are displayed. In these two figures, the first row shows the correlation matrices for testing in frequency band of 500–4000 Hz with liver and pork. The second row shows the correlation matrices for testing in the frequency band 4000–8000 Hz. The right column of each row displays the average and standard deviation for each of the matrices confirming the difference in Pearson coefficient values when model and test PSDs are taken from the same texture. More detailed average Pearson coefficient values can be found in Tables 1–4.

4. Discussion

The preliminary results verified that it is possible to apply AE sensing technique to differentiate textures based on audio signal analysis on a robotic surgical tool. The proposed setup in this work also showed that it is feasible to place a sensor proximally on a surgical device. This could be further developed into a plug-and-work module that can be easily applied on a surgical robot in the real operating room or to other surgical devices that would benefit from audio sensing.

The results show that different textures produce different time-domain and frequency-domain characteristics. In this study, a simple frequency components similarity approach was used for comparing audio signals arising from a given texture and a spectral model of a texture. In an ideal setup, a high correlation would be expected when the model and the test samples belong to the same texture. On the contrary, if the model and the test samples are from different textures, then the correlation should be low since the roughness level and the materials are not equal. In the comparison of both synthetic textures, a relatively larger signal amplitude was observed in the sweeps on texture B. The explanation for this finding may result from the different stiffness and thickness among the textures. Texture A is 2 mm-thick while texture B is 0.5 mm and directly below the texture is the hard cork material. In addition, texture A has a soft and spongy-like structure, which may act as a damper and lead to a smaller amplitude during the experiment. In addition, audio signals produced by a soft texture could be expected to be more prominent in the lower frequency band, which was also observed in the experiments. As expected, the two tested biological tissues involve mainly lower frequency components compared to the synthetic textures. Although both biological tissues behave similar in the low frequency range going from 0 to 500 Hz, statistically the similarity between model and tested signal shows a degree of distinguishability between both textures. Moreover, if the analysis is made in higher frequency ranges, more pronounced differences are evident that could help to distinguish both soft tissues. Thus, as the results revealed, the audio signal processing strategy used in this work indicated and confirmed the potential for differentiating textures, even for soft tissues.

In this work, FFT-based and parametrical AR-based spectral estimations were used to differentiate textures. The results show that when the model and testing samples are from the same texture, the correlation is relatively higher. According to the results, it seems that by using FFT-based PSD model, a higher probability of successful texture differentiation can be achieved because the difference of the mean correlation is generally larger. However, the presented results were done in a controlled and quiet room with minimal ambient noise, which is quite different from a real surgery room. We can expect that under real conditions the audio artifacts coming from other sources will be higher causing a lower SNR of the audio signal. Additionally, the way the instrument can be moved over a surface is limited by the trocar and anatomy in a real robotic surgery. Thus, the quantity of data and dynamics obtained per sweep (or palpation in real case) may be significantly reduced. Under this condition, standard FFT-based techniques are not anymore suitable for obtaining good estimation of the spectrums and in this case AR parametrical modelling could be a viable alternative.

Finally, it is important to mention that the proposed approach can overcome the main drawbacks of the solutions that have already been proposed for tactile feedback in RMIS. In fact, sensors for extracting kinesthetic information [18–20] and tactile sensors for extracting information related to the skin sensation [18,21,22] generally require the sensing device to be placed at the instrument tip or next to it. This increases both costs and complexity of the instruments because the sensor embedded at the tip of the instrument can change its mechanical properties. Additional wires and sensors result in a reduction of the major functionality of being a “good interventional instrument”. These are the main reason why there are only very few devices approved for

use during clinical in-vivo studies. None of them is actually used in current robotic assisted procedures. Since our approach acquires information from the proximal end and is non-invasive makes the feedback information could be acquired without the necessity of altering the surgical devices. As the additional component would not need to be in direct contact with the organs a simple clip-on device could be feasible.

5. Conclusion

In this study, an innovative method to differentiate textures using audio signal transmitted from the tip of a Forceps was presented. The stability and the feasibility were proven in a preliminarily step. The results showed that this method could serve as an alternative solution for the present lack of sense and touch in robotic surgery. A further study that assesses the palpation functionality using the same principle on a real da Vinci surgical robot is in preparation. In addition, the acoustic dynamics of biological tissue phantoms should also be researched to find out the suitable parameters for implementing this acoustic palpation system.

With the DAVID setup, surgeons could receive more information about a surface and rigidity of an organ they are touching during robotic surgeries (e.g. the Forceps in this work). This can be an additional and complementary information to the existing visual feedback.

The textures tested in this work correspond to two biological tissues and two synthetic ones. The four tested textures present a clear signature in the audio signal with patterns that can allow a classification of these textures. As a first step in this work, the feasibility of extracting spectral characteristics for distinguishing between textures was shown, and in a further step this data will be used for extracting significant features for classifying tissue using machine learning approaches. In the future, a wider choice of bio-phantoms mimicking real surgical tissue-tool interactions should be tested with the same setup to prove the value of audio acquisition for clinical applications. This can include the detection of pulsating structures like embedded vessels, the interaction with other instruments or the correlation of audio signals with forces and impacts on the instrument tip. Other applications that should be explored in the future with this approach concern tissue characterization for distinguishing between healthy and tumor tissues and also, since force sensing is important in every laparoscopic surgery it would be of high interest to translate audio information into a measure of applied force. This could be done through assimilating or correlating the changes in the acoustic transfer function by using an impulse excitation input (at the proximal end of the instrument) and by estimating the transfer function parameters using the response audio signal to the stimulus.

The current experiments were done in a silent room without any actively running electronic devices close by and without any verbal noise. This is obviously an ideal environment for the signal acquisition and not very realistic compared to real surgery setups. All ambient noises may cause changes to our signal and a more advanced signal conditioning and processing strategy would then be necessary to assure the quality of the signal.

Since the main goal of this paper was to show feasibility of the acoustic emission approach, a simple spectral method for distinguishing between textures was proposed, implemented, and successfully tested. Differentiating between a larger number of textures would produce more complex audio patterns and require additional or other advance signal processing techniques, such as pole-based or non-linear analysis using higher order statistics, for extracting robust features that can feed a classification approach.

As a future step, it is necessary to study how to translate or map the extracted/decoded information from the audio into a friendly and easy feedback information to the surgeons.

Conflicts of interest

None Declared.

Acknowledgment

This work was financially supported by the Federal Ministry of Education and Research (BMBF) in context of the “INKA” project (Grant number 03IPT7100X and by EFRE funding in context of the ego.-INKUBATOR program (ZS/2016/09//81061/IK 01/2015)).

References

[1] H. Alemzadeh, J. Raman, N. Leveson, Z. Kalbarczyk, R.K. Iyer, Adverse events in robotic surgery: a retrospective study of 14 Years of FDA data, PLoS One 11 (2016) e0151470https://doi.org/10.1371/journal.pone.0151470.
[2] R.J. Cerfolio, K.M. Bess, B. Wei, D.J. Minnich, Incidence, results, and our current intraoperative technique to control major vascular injuries during minimally invasive robotic thoracic surgery, Ann. Thorac. Surg. 102 (2016) 394–399 https://doi.org/10.1016/j.athoracsur.2016.02.004.
[3] D. Tare, P. Maria, R. Ghavamian, Vascular complications in laparoscopic and robotic urologic surgery, in: R. Ghavamian (Ed.), Complications of Laparoscopic and Robotic Urologic Surgery, Springer New York, New York, NY, 2010, pp. 45–58.
[4] A.M. Okamura, Methods for haptic feedback in teleoperated robot-assisted surgery, Ind. Robot 31 (2004) 499–508 https://doi.org/10.1108/01439910410566362.
[5] M.E. Hagen, J.J. Meehan, I. Inan, P. Morel, Visual clues act as a substitute for haptic feedback in robotic surgery, Surg. Endosc. 22 (2008) 1505–1508 https://doi.org/10.1007/s00464-007-9683-0.
[6] J.B. Gafford, S.B. Kesner, R.J. Wood, C.J. Walsh, Force-sensing surgical grasper enabled by pop-up book MEMS, IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, Tokyo, 2013, pp. 2552–2558 2013/11/3 - 2013/11/7.
[7] Z. Hu, C.-H. Yoon, S.B. Park, Y.-H. Jo, Design of a haptic device with grasp and push-pull force feedback for a master-slave surgical robot, Int. J. Comput. Assist. Radiol. Surg. 11 (2016) 1361–1369 https://doi.org/10.1007/s11548-015-1324-9.
[8] M.O. Culjat, C.-H. King, M.L. Franco, C.E. Lewis, J.W. Bisley, E.P. Dutson, W.S. Grundfest, A tactile feedback system for robotic surgery, Conf. Proc. IEEE Eng. Med. Biol. Soc. (2008) 1930–1934 2008 https://doi.org/10.1109/IEMBS.2008.4649565.
[9] J.C.T. Hui, K.J. Kuchenbecker, Evaluating the BioTac's ability to detect and characterize lumps in simulated tissue, in: M. Auvray, C. Duriez (Eds.), Haptics: Neuroscience, Devices, Modeling, and Applications, Springer Berlin Heidelberg, Berlin, Heidelberg, 2014, pp. 295–302.
[10] C. Pacchierotti, D. Prattichizzo, K.J. Kuchenbecker, Cutaneous feedback of fingertip deformation and vibration for palpation in robotic surgery, IEEE Trans. Biomed. Eng. 63 (2016) 278–287 https://doi.org/10.1109/TBME.2015.2455932.
[11] A. Illanes, A. Boese, I. Maldonado, A. Pashazadeh, A. Schaufler, N. Navab, M. Friebe, Novel clinical device tracking and tissue event characterization using proximally placed audio signal acquisition and processing, Sci. Rep. 8 (2018) 12070 https://doi.org/10.1038/s41598-018-30641-0.
[12] A. Illanes, A. Boese, I. Maldonado Zambrano, A. Pashazadeh, A. Schaufler, N. Navab, M. Friebe, Interventional Device Guidance Support with Proximal Audio Emission Signal Acquisition, (2018).
[13] N. Mahmoodian, A. Schaufler, A. Pashazadeh, A. Boese, M. Friebe, A. Illanes, Proximal detection of guide wire perforation using feature extraction from bispectral audio signal analysis combined with machine learning, Comput. Biol. Med. 107 (2019) 10–17 https://doi.org/10.1016/j.compbiomed.2019.02.001.
[14] C.-H. Chen, T. Sühn, A. Illanes, I. Maldonado, H. Ahmad, C. Wex, R. Croner, A. Boese, M. Friebe, Proximally placed signal acquisition sensoric for robotic tissue tool interactions, Current Directions in Biomedical Engineering 4 (2018) 67–70 https://doi.org/10.1515/cdbme-2018-0017.
[15] February 10th, https://learn.adafruit.com/adafruit-i2s-mems-microphone-breakout/raspberry-pi-wiring-and-test.
[16] M. Basseville, A. Benveniste, Detection of Abrupt Changes in Signals and Dynamical Systems, Springer-Verlag, Berlin/Heidelberg, 1985.
[17] D. Manolakis, V. Ingle, S. Kogon, Statistical and Adaptive Signal Processing: Spectral Estimation, Signal Modeling, Adaptive Filtering and Array Processing, Artech House, Norwood, 2005.
[18] N. Enayati, E. de Momi, G. Ferrigno, Haptics in robot-assisted surgery: challenges and benefits, IEEE Rev. Biomed. Eng. 9 (2016) 49–65 https://doi.org/10.1109/RBME.2016.2538080.
[19] J. Konstantinova, A. Jiang, K. Althoefer, P. Dasgupta, T. Nanayakkara, Implementation of tactile sensing for palpation in robot-assisted minimally invasive surgery: a review, IEEE Sens. J. 14 (2014) 2490–2501 https://doi.org/10.1109/JSEN.2014.2325794.
[20] A.L. Trejos, R.V. Patel, M.D. Naish, Force sensing and its application in minimally invasive surgery and therapy: a survey, Proc. Inst. Mech. Eng. C J. Mech. Eng. Sci. 224 (2010) 1435–1454 https://doi.org/10.1243/09544062JMES1917.
[21] A.M. Okamura, Haptic feedback in robot-assisted minimally invasive surgery, Curr. Opin. Urol. 19 (2009) 102–107 https://doi.org/10.1097/MOU.0b013e32831a478c.
[22] M.I. Tiwana, S.J. Redmond, N.H. Lovell, A review of tactile sensing technologies with applications in biomedical engineering, Sens. Actuators A Phys. 179 (2012) 17–31 https://doi.org/10.1016/j.sna.2012.02.051.