Classification of normal and cancerous lung tissues by electrical impedance tomography

Jianling Gao a, Shihong Yue a*, Jun Chen b and Huaxiang Wang a

aSchool of Electrical Engineering and Automation, Tianjin University, Tianjin 300072, China
bGeneral Hospital of Tianjin Medical University, Tianjin 300052, China

Abstract. Biological tissue impedance spectroscopy can provide rich physiological and pathological information by measuring the variation of the complex impedance of biological tissues under various frequencies of driven current. Electrical Impedance Tomography (EIT) technique can measure the impedance spectroscopy of biological tissue in medical field. Before application, a key problem must be solved on how to generally distinguish normal tissues from the cancerous in terms of measurable EIT data. In this paper, the impedance spectroscopy characteristics of human lung tissue are studied. On the basis of the measured data of 109 lung cancer patients, Cole-Cole Circle radius (CCCR) and the complex modulus are extracted. In terms of the two characteristics, 71.6% and 66.4% samples of cancerous and normal tissues can be correctly classified, respectively. Furthermore, two characteristics of the measured EIT data of each patient consist of a two-dimensional vector and all such vectors comprise a set of vectors. When classifying the vector set, the rate of correctly partitioning normal and cancerous tissues can be raised to 78.2%. The main factors to affect the classification results on normal and cancerous tissues are generally analyzed. The proposed method will play an important role in further working out an efficient and feasible diagnostic method for potential lung cancer patients, and provide theoretical basis and reference data for electrical impedance tomography technology in monitoring pulmonary function.

Keywords: Cancerous tissue, normal tissue, classification, electrical impedance spectroscopy

1. Introduction

In recent years, lung cancer has been one of the leading causes for human death and the most commonly diagnosed types among all cancerous patients. In terms of averagely surviving time after operation, the early-term patients are almost three times as long as the late-term ones. Therefore, the early diagnosis of lung cancers plays an important role in saving lung cancer patients. In order to achieve early diagnosis, many techniques including Electrical Impedance Tomography (EIT) have been developed in the past decades [1,2]. EIT is fast-response, non-invasive and low-cost, so it becomes a widely accepted tool in obtaining rich physiological and pathological information. Compared with other existing techniques, EIT realizes the early diagnoses of potential lung cancers by measuring the variation of the complex impedance of biological tissues under various frequencies of driven current. Biological tissue impedance spectroscopy can show intrinsic biophysical properties of the normal and cancerous tissues that correspond to healthy and unhealthy patients, respectively.

*Corresponding author: Shihong Yue, School of Electrical Engineering and Automation, Tianjin University, Tianjin 300072, China. Tel.: 13102161306; Fax: +86-22-27405477; E-mail: shyue1999@tju.edu.cn.
However, without the real measurement data, a key unsolved problem is to generally identify normal and cancerous tissues in terms of measurable EIT data [3–5]. In the past three years, the Medical Research Group of Tianjin University has been working on lung cancer sample collection of various patients. On the individually operational spot of 109 patients with lung cancer, EIT data of cancerous and normal tissues have been measured at different exciting frequency. These measured data consist of a set of samples. Depending on the data set, two main characteristics of human lung tissue are extracted. Firstly, according to Cole-Cole mathematical model that the impedance spectroscopy of human tissue nearly satisfies, the CCCR [6–8] of each patient’s two classes of tissues is calculated as it is very sensitive to normal and cancerous tissues. The second characteristic is the complex modulus of each measured data with its own applicable range, whose function is different from CCCR. In terms of these two characteristics, 71.6% and 66.4% samples of cancerous and normal tissues can be correctly classified, respectively. Furthermore, the two characteristics of the measured EIT data of each patient consist of a two-dimensional vector and all such vectors comprise a set of vectors. When classifying the vector set, the rate of correctly classifying all normal and cancerous tissues can be raised to 78.2%. Specifically, three key factors are evaluated which affect the correctness of the two characteristics when generally classifying normal and cancerous tissues. The proposed method is expected to play an important role in further working out an efficient and feasible diagnosing system for potential lung cancer patients, and provide a valuable proof on whether a diagnosed patient has suffered from the lung cancer.

2. Related work

In this section, two parts of related researches will be introduced, including the Cole-Cole mathematical model and the impedance spectroscopy of lung cancer.

2.1. Cole-Cole equation

One of the most important behaviors of electrical conductivity is frequency-dependent behaviors which demonstrate the induced polarization, so-called ‘low-frequency dispersion’. Up to now, no general physical-chemical pattern which can depict frequency dispersion has been found; hence, researchers always choose the way in which the experimental result can be met on the strength of phenomenological pattern [9,10]. The Cole-Cole equation is one of the most favorite patterns which was initially used in complex dielectric constant, as shown in Eq. (1).

\[
\varepsilon^* = \varepsilon_s + \frac{\varepsilon_c - \varepsilon_s}{1 + (i\omega\tau)\alpha}
\]

(1)

where \(\varepsilon_0\) and \(\varepsilon_c\) are the low-frequency and the high-frequency dielectric constant values, respectively, \(\tau\) is the central relaxation time and \(\alpha = 1 - \alpha\) is so-called Cole–Cole equation index, which describes the broadness of the relaxation time distribution. On the strength of the electrostatic analogy, for dielectric materials with losses, Eq. (1) can be evolved to the complex electrical conductivity as follows,
where $\sigma_0$ and $\sigma_e$ are the low and high-frequency electrical conductivity values, respectively. Eq. (2) is always applied in actual practice as it is computed firsthand by using the collected electrical impedance. In the 1970s, Pelton found that the Cole-Cole equation is very suitable for describing the behavior of the complex resistivity by a lot of experiments [11], as expressed by

$$\sigma^* = \sigma_0 + \frac{\sigma_e - \sigma_0}{1 + (i\omega\tau)}.$$  

(2)

where $\rho_0$ and $\rho_e$ refer to the low- and the high-frequency electrical resistivity values, respectively. Both Eqs. (2) and (3) are used to estimate the experiments data representatively [12]. In the past two decades, the data obtained by different research groups have been compared on the strength of the Cole–Cole equation parameters. Their results validate the Cole-Cole equation [13–15].

2.2. Cancerous cancers and their electrical characteristics

According to different histopathology characteristics, lung cancers mainly consist of 4 classes, as explained below.

1) Squamous cell carcinoma. Figure 1(a) shows the visual characteristics of squamous epithelial cells in a high-power visual microscope, where the pathological section consists of sphere-distributed tissue cells with keratinocytes inside. Squamous cell carcinoma is the most common type and accounts for 40%-50% of primary lung cancer. Squamous cell carcinoma has a slower growth process, and has a higher probability of successful surgery removal, thus the patients can have a longer survival time [16].

2) Adenocarcinoma. Figure 1(b) shows the adenocarcinoma that accounts for 25% of primary lung cancer. The characteristics of adenocarcinoma pathological section are that cancer cells have developed into adenoid structure and keratinocytes with various sizes, shapes and arrangements. Coenothecalia shape is observable among normal gland, and pathological karyokinesis phenomenon also exists among cancer cells [17,18].

3) Large cell carcinoma. Figure 1(c) shows the undifferentiated lung cancer cell distributions. The

(a) Squamous cell carcinoma   (b) Adenocarcinoma     (c) Large cell carcinoma   (d) Small cell carcinoma

Fig. 1. Pathological section of four classes of typical lung cancers.
characteristic of pathological section in large cell carcinoma is that cancer nest is formed by cancer cells. These neoplastic cells are pleomorphic and appear clear cytoplasm [19].

4) Small cell carcinoma. Figure 1(d) shows the small undifferentiated cells in small cell carcinoma for short which is the most malignant type of lung cancer and accounts for 20% of primary lung cancer. The characteristics of pathological section are that cancer cells are small, rounded or oat-shaped, and have various pathological karyokinesis [20].

Electrical impedance tomography of biological tissues is tightly related to tissue structures, physiological and pathological states. Biological tissues consist of a great number of plesiomorphous cells and intercellular substances including solid and liquid electrolytes with different conductivities. When exciting the biological tissues by low-frequency currents, the membrane of each cell has strong dielectric effect and stops currents from its inner. Thus the current mainly goes through the intercellular substances. As the exciting current frequency increases, the dielectric effect of cell membrane decreases since partial currents go through the inner of most cells. Thus the high-frequency electrical impedances of biological tissue are larger than the low-frequency ones. Figure 2 shows the current paths in the biological tissues.

The low-frequency and high-frequency currents lead to very different characteristics of electrical impedances. As usual, when the external drive signal frequency ranges among 20 kHz-100 kHz, $65\Omega \cdot \text{m} - 166\Omega \cdot \text{m}$ electrical impedances can be obtained in data measuring system, and the ratio of bones to soft tissues may reach 1:250. Moreover, the maximal ratio between the pump and lung tissues can attain 35:1, and different soft tissues have different electrical impedances in different physiological status. Specifically, electrical impedances between normal and cancerous lung tissues for various cancerous patients also vary with the expiratory and inspiratory process. From comparison, the ray absorption rate in the X-ray computed tomography has much smaller difference. For example, the attenuation coefficient of bone is only 1.5 times that of soft tissue. Consequently, EIT technology can bring higher contrast to different tissue diagnoses, and have a potential to be applied in medicine clinic [21]. When the characteristics of electrical impedance are applied to distinguish the normal tissues from the cancerous or the lung cancers, the electrical impedance is expected to serve as a valuable tool to diagnose potential patients with lung cancer.

2.3. EIT principle and prior electrical information effect

EIT imaging technology [22] is one nondestructive visualization measurement technique. Through exciting a group of electromagnetic sources in turns into the field containing the investigated objects, the grey-based distributions of the investigated objects by the boundary measures can be visually reconstructed. While the existing algorithms show promise in a number of applications, some issues...
remain unsolved as explained as follows:

1) Ill-posed problem. A high-resolution ET image necessarily has full and high signal-noise-ratio measurements, while the measurements in the existing ET techniques are much less than the number of solved or determined pixels. Moreover, the ET measurements usually are low signal-noise-ratio. Consequently, the existing ET techniques are of low resolution. Along with the existing algorithms, it is undesired that this problem can be solved. Therefore the reconstructed images usually have low resolution.

2) “Soft-field” effect. Owing to the use of weak current excitation to produce the measured data in ET that must be a weak signal, any small measured errors or noise may lead to large variance of the investigated object image. Essentially, any measure data is associated with the material distributions of the whole investigated field. This is called "soft-field" effect which is a much undesired case in practice [23]. The existing ET imaging algorithms are limited for this problem as the algorithms result in the obtained ET image being unstable and often unacceptable in noisy conditions.

An effective way to overcome the above problem is to apply the prior information on the investigated objects. The lung cancer type is very useful to obtain the approximate electrical characteristic ratio between the normal and cancerous tissues. Figure 3 shows the reconstructed ET images and X-ray computed tomography (XCT) for four typical cancerous tissues, where the cancerous tissues can be approximately obtained after adding some simple and generally available prior information. Thus the prior information on different cancerous types is expected to be obtained through comparing the real normal and cancerous tissues.

3. Sample measurement and collection

Experimental samples of lung tissue are taken from 109 patients in the lung surgery department in the General Hospital of Tianjin Medical University. Experimental samples are selected by clinicians of...
tumor tissue which around normal part. In accordance with the standards of checkerboard, the cancerous tissue and normal tissue are collected and filled, respectively. Table 1 shows the statistics of 109 patients according to their gender, operational positions and ages. After removing some noisy data and error-contained data, only 91 groups of data are valid. Table 1 shows the gender of 109 patients with lung cancer, surgery cutting areas and the age distributions. Table 2 groups these samples into four classes in terms of pathological types, that is, 34 adenocarcinoma, 29 squamous carcinoma, 11 large cell carcinoma, 9 small cell carcinoma, and 9 others.

The measurement system consists of measuring electrodes, specimen test box, test fixture, adapter, impedance analyzer, data acquisition interface and PC. Before measurements, the tissue will be put into the cylindrical cavity of the test box. The test box is connected to 4294A impedance analyzer by 42942A adapter and 16092A fixture, measuring the series resistance R and reactance X of tissue samples from 31 frequency points of logarithmic increments within 100Hz-100MHz. The initialization procedures and data acquisition procedures of the impedance analyzer are programmed through standard commands for programmable instruments (SCPI) on LabView 8.0 development platform. Data acquisition is completed through the GPIB data interface and measurement results are saved in

<table>
<thead>
<tr>
<th>Gender</th>
<th>Operation removal part</th>
<th>Age group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>Left lung</td>
<td>≤50 50–60 60–70 ≥70</td>
</tr>
<tr>
<td>Female</td>
<td>Right lung</td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>Superior lobe</td>
<td>13</td>
</tr>
<tr>
<td>38</td>
<td>Inferior lobe</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Superior lobe</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Middle and inferior lobe</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Inferior lobe</td>
<td>16</td>
</tr>
</tbody>
</table>

Table 2

<table>
<thead>
<tr>
<th>Tissue type</th>
<th>Lesion type</th>
<th>Number</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal tissue</td>
<td>Adenocarcinoma</td>
<td>34</td>
<td>109</td>
</tr>
<tr>
<td></td>
<td>Squamous carcinoma</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>Lesion tissue</td>
<td>Large cell carcinoma</td>
<td>11</td>
<td>91</td>
</tr>
<tr>
<td></td>
<td>Small cell carcinoma</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

(a) Surgical removal tissues. (b) Normal and cancerous tissues. (c) 4294A impedance analyzer.

Fig. 4. Sample and measurement of normal and cancerous tissues.
In order to ensure the activity of the measured tissues and minimize measurement errors, human lung tissue dielectric measurement environment control platform is set up by the test chamber, temperature control system, the humidity measuring instrument and the ultrasonic atomizer. Human lung tissue sampling steps are shown in Figure 4.

Our research group tests dielectric spectrum characteristics of normal and cancerous tissues, respectively. But before pathology confirmation errors may be included, for example, normal tissues are classified into lung cancer tissues, vice versa. So after the completion of the test, pathological analysis is done for all the tissue samples with formalin soaked, paraffin embedded, and stained. After performing pathological analysis, the measurement error of data can be decreased to some extent.

4. Classification of normal and cancerous tissues

Two characteristics in electrical impedances are extracted including Cole-Cole circle radius and the complex modulus. The motivation of choosing these characteristics is due to their sensitiveness and classification function for various normal and cancerous tissues. In order to test the application in practice, the two characteristics are evaluated under different conditions, as explained below.

4.1. Comparison of electrical impedance spectroscopy of the biological tissue

The averages of electrical impedances of normal and cancerous tissues for the four classes of lung cancers are listed in the same coordinate system, and these data from 31 exciting frequencies comprise an electrical impedance spectroscopy, as shown in Figure 5.

Figure 5 shows that normal and cancerous tissues decrease as the exciting frequency increases. In addition, the average values of electrical impedances of four classes of lung cancers according to their real parts and imaginary parts are shown in Figures 5(a)-5(d), respectively. For each class of lung cancer, its differences between normal and cancerous tissues are clear, and these averages of four classes of lung cancers are very different. As usual, there is very large difference between the normal and cancerous electrical impedance spectroscopy for the same patient. However, these electrical impedance spectroscopies of different patients are difficult to be distinguished, since some electrical impedances of cancerous tissues of some patients are larger than those of normal tissues of other patients, and the two classes of electrical impedance spectroscopy has no clear boundary. So the first task to apply electrical impedance spectroscopy for lung cancer diagnosis is to extract the characteristics of electrical impedance spectroscopy (Note: since the original measuring data of electrical impedance spectroscopy is not allowed for public application, these data fail to be shown in this paper).

4.2. Feature extraction

The first extracted feature from electrical impedance spectroscopy is CCCR. Assume that \( f(x) \) is an arbitrary function of variable \( x \), the curvature value of \( f(x) \) at \( x_0 \) is expressed as

\[
K = \frac{|f'(x_0)|}{\sqrt{(1+f(x_0)^2)^3}}, \text{ s.t. } f''(x_0) = \lim_{x \to x_0} \frac{f'(x) - f'(x_0)}{x - x_0}
\]  

(4)
Let \( f(i,j) \) be the measuring electrical impedance of the \( i \)-th patient at the \( j \)-th excitation, where \( i = 1, 2, \ldots, 71 \) (the selected number of typical patients); \( j = 1, 2, \ldots, 30 \) (exciting times). For the \( i \)-th patient, the optimal curvature radius \( R(i) \) is solved by the following equation:

\[
R(i) = \arg\min_{K(j)} \left[ \sum_{j=1}^{30} (R - 1/K(j))^2 \right], \quad i = 1, 2, \ldots, 71
\]

(5)

where \( K(j) \) is the value of Eq. (4) when \( x \) takes 71 discrete electrical impedances. According to these curvature radii of 71 patients, their averages are shown in Figure 6.

The mean values of all cancerous and normal tissues by Eq. (5) are 423.4 and 583.6, respectively, which are regarded as the representative values to identify the normal and cancerous tissues, respectively. When the value of the measured electrical impedance is close to 423.4, the corresponding tissue is classified into the cancerous group, or else into the normal group. Consequently, 81.5% normal tissues and 73.3 % cancerous tissues can be correctly classified, respectively.

In medical research field, the experimental results are often evaluated by the statistical receiver operating characteristic (ROC) curve. ROC curve is a widely applied method to evaluate the classification (diagnosis) results. The abscissa and ordinate in a ROC curve show the computed values of specificity and sensitivity of all measuring data, respectively. The ROC curve evaluates the effectiveness of a classification result by the area below the curve. When the area value is less than 0.5, the classification result is meaningless. When the area values range between 0.5-0.7, the classification result can be adopted to some extent. When the area values range between 0.7-0.9, the classification result is valuable. Specifically, when the value is larger than 0.9, the results are perfectly believable.

By means of CCCR for classifying normal and cancerous tissues, the area value of 0.805 can be
obtained, as shown in Figure 5(b).

The second extracted feature from electrical impedance spectroscopy is the complex modulus (CM). Assume that the measuring electric impedance is expressed as

\[ f(n,i) = A(n) + B(n)i, \quad n=1,2,\ldots,30; \quad i=1,2,\ldots,109 \]  

The complex modulus (CM) of electric impedance is extracted as:

\[ R_i = \sqrt{A(n,i)^2 + B(n,i)^2}, \quad n=1,2,\ldots,30; \quad i=1,2,\ldots,109 \]  

The typical value of different lung classes is computed by the following equation:

\[ CCCR(k) = \frac{\sum_{i=1}^{109} K(n,i)}{109} \]  

where \( K(k,i) \) is the CM value of the \( k \)-th class. When all samples are partitioned into normal and cancerous classes, \( k=1,2 \); the computed values are 145.6 and 164.8, respectively. Consequently, all samples are partitioned into two types. According to the nearest neighborhoods (NN) assignment principle, 75% of normal and cancerous tissues can be correctly partitioned by CM. On the other hand, the CM area value in ROC curve is about 0.79 (see Figure 7(b)), the classification results of CM are believable as well.

Combining both CCCR and CM to a vector as the following form,

\[ V = \{V_i\}, \text{s.t.} V_i = (CCCR_i, CM_i), i=1,2,\ldots,91 \]
When a widely used clustering algorithm, the K-means algorithm, is used to cluster the group of samples, their cluster centers are (423.4, 145.6) and (583.6, 164.8) for normal and cancerous tissues, respectively. According to the two values and the NN assignment principle, all vectors in $V$ are classified to two groups, and on average, 85.4% samples can be correctly partitioned. So the combined characteristics can improve the classification correctness of normal and cancerous tissues.

4.3. Major factors of affecting classification correctness

There are many factors affecting the classification results such as the patient’s age, gender, chemotherapy, and etc. These factors are analyzed as follows.

4.3.1. Test on effect of age, gender and cancerous type

In order to study the effect of patient’s age and gender, 91 samples are grouped and the classification correctness is evaluated by CCCR and CM. Table 3 shows that the classification correctness of squamous cell carcinoma by CCCR is the highest among the four classes of cancerous tissues. Moreover, CM outperforms CCCR in the three classes of cancerous tissues. But in our experiments, the number of samples of squamous cell carcinoma and adenocarcinoma is the majority, thus their statistical results are more general.

Table 4 shows that the gender has no effect on the classification correctness, where the rates for different gender are much closer to each other. However, in terms of age, a younger age results in a lower correctness.

Table 3

<table>
<thead>
<tr>
<th>Effect of different cancerous tissue on cancerous classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Squamous cell carcinoma</td>
</tr>
<tr>
<td>Rate</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>CCCR</td>
</tr>
<tr>
<td>CM</td>
</tr>
</tbody>
</table>
4.3.2. Test on chemotherapy effect

Chemotherapy is a treatment which induces the benign differentiation of tumor cells, which can inhibit the growth and reproduction of a tumor cell, or kill tumor cells by chemical drugs. Clinicians may choose chemotherapy according to the lung cancer patient’s health states. However, chemotherapy not only kills tumor cells, but also affects the systemic normal cells. Conductivity and permittivity measurements on the patients have been done before operation in the same way. The comparison of dielectric characteristics between cancer tissue and normal tissue shows that the conductivity and relative permittivity of cancer tissue are higher than those of normal tissue. The conductivity and relative permittivity of cancer tissue after preoperative chemotherapy are less than those of normal tissue, as shown in Figures 8(a) and 8(b), respectively. This may be due to the fact that chemotherapy drugs play a role in killing or inhibiting the cancer cells, destroying the structure and activity of cancer cells, which is manifested as the reduction of conductivity capacity and dielectric capacity in the biophysical characteristics. Consequently, when both CCCR and CM are applied to indentify the normal and cancerous tissues, corresponding to the healthy and unhealthy lung tissues respectively, the difference must be considered. In the collected 91 samples, about 5.5% needs to be added to the original measurable values.

5. Conclusion

Firstly, this study reports the measuring process of the electrical impedance of removing normal and cancerous tissues from 91 patients with lung cancer. These measured electrical impedances consist of a set of samples. There is an apparent distinct difference between original electrical impedance spectroscopes of the two classes of tissues. But when mixing all samples of 91 patients together, the normal and cancerous tissues can hardly be identified, since the electrical impedance spectroscopes of
partial cancerous tissues for some patients are larger than those of partial normal tissues for other patients. In order to classify normal and cancerous tissues by their electrical impedance spectroscopes, two main characteristics of Cole–Cole circle radius and complex modulus are extracted. In light of the two characteristics, most normal and cancerous tissues can be correctly classified and identified. Consequently, each class of electrical impedance spectroscopes can have interpretability and classification characteristics.

These research results can lay a solid foundation for establishing both the classification of impedance in different pathologies and the quantitative relationship between different lung illnesses, and can bring some practical experience for establishing a lung cancer diagnostic system based on electrical tomography technology in the future.

Acknowledgement

This work is supported by National Science Foundation of China (Grant No. 61174014, 2013DFAF11040 and No. 60772080) and National Science Foundation of Tianjin (Grant No. 08JCYBJC13800).

References


