Information Content of EIT Measurements

Andy Adler[1] and William R.B. Lionheart[2]

[1] Systems and Computer Engineering, Carleton University, Ottawa, Canada
[2] School of Mathematics, University of Manchester, U.K.

Abstract—**Electrical Impedance Tomography (EIT) calculates internal conductivity from surface measurements; image reconstruction is most commonly formulated as an inverse problem using regularization techniques. Regularization adds "prior information" to address the solution ill-conditioning. This paper presents a novel approach to understand and quantify this *information*. We ask: how many bits of information (in the Shannon sense) do we get from an EIT data frame. We define the term *information in measurements* (IM) as the: decrease in uncertainty about the contents of a medium, due to a set of measurements. Before the measurement, we know the prior information (inter-class model, q). The measured data tell us about the medium (which, corrupted by noise, gives the intra-class model, p). The measurement information is given by the relative entropy (or Kullback-Leibler divergence). Based on this expression, and given a noise covariance Σ_n and a prior model of the element covariances Σ_x, IM$= \frac{1}{2}log_2|J\Sigma_x J^T \Sigma_n^{-1} + \mathbf{I}|$. Under the simplification that measurement and noise covarianes are uncorrelated, IM may be approximated as a function of the signal to noise ratio and the Jacobian and prior matrices. For an example 16 electrode EIT system, IM was calculated to be 245.1 bits. Finally, several applications of an information measure for EIT are given.**

Keywords—**Measurement Information, Kullback Leibler Divergence, Electrical Impedance Tomography**

I. INTRODUCTION

In this paper we ask: 1) how much information is there in a set of measurements? and 2) how much information do we add when we "add" prior information via regularization? To address these questions, we develop an information-theoretic formulation of the image reconstruction inverse problem.

Electrical Impedance Tomography (EIT) calculates an estimate of the conductivity distribution within a body based on current stimulations and voltage measurements on the body surface. EIT has fairly low spatial resolution, limited by the low sensitivity of surface measurements to conductivity changes deep within the body. EIT images are most commonly reconstructed by formulating the inverse problem using regularization techniques (eg. [1]) in order to calculate a conductivity change image estimate, $\hat{\mathbf{x}}$, which is both faithful to the measurements, $\mathbf{y}$, and to *a priori* constraints on a "reasonable" image. Regularization adds

"prior information" to address the solution ill-conditioning. In order to quantify this *information*, we ask: how many bits of information (in the Shannon sense) do we get from an EIT data frame. We define the term *information in measurements* (IM) to represent this information content.

> *Information in Measurements:* the decrease in uncertainty about the contents of a medium, due to a set of measurements.

In order to interpret this definition, we refer to two instants: *Instant t_1:* before an EIT measurement, at which time we only know that the medium is part of a general class p. This is the prior information; if we know that the EIT system is connected to a patient's thorax, then the class p is smaller (ie. conductivities will be in a narrow range, and the most likely number of lungs is two). If we have little information on the nature of the medium the EIT system is connected to, then the class p will be broader and the prior information less (ie. conductivities are positive). *Instant t_2:* after receiving a frame of data (with added noise) we have specific information on this medium q as well as prior information on the class p. The uncertainly on the contents of the medium will thus be less, reflecting the *information in measurements* gained.

Based on this formulation, we develop a strategy to calculate optimal stimulation patterns in an information-theoretic sense. We then show two examples: 1) for a single measurement impedance plethysmography equipment, and 2) for a 16 electrode EIT system.

II. METHODS

A. Information Theory

In Bayesian statistics the Kullback-Leibler divergence (KLD) can be used as a measure of the information gain in moving from a prior distribution, $q(\mathbf{y})$, to a posterior distribution, $p(\mathbf{y})$. It can also be understood to be the "extra bits" of information needed to represent $p(\mathbf{y})$ with respect

to $q(\mathbf{y})$. The KLD is defined as [3]:

$$D(p\|q) = \int\limits_{\mathbf{y}} p(\mathbf{y})log_2\frac{p(\mathbf{y})}{q(\mathbf{y})}d\mathbf{y} \qquad (1)$$

where the integral is over all measurement channels. A comment on notation: we use p to refer to both a specific realization of an EIT image, and the distribution of measurements from that image (including measurement noise), while q represents the *prior* distribution and the distribution of its features.

One important general difficulty with direct information-theoretic measures is that of data availability. Distributions are difficult to estimate accurately, especially at the tails; and yet $log_2\left(p(\mathbf{y})/q(\mathbf{y})\right)$ will give large absolute values for small $p(\mathbf{y})$ or $q(\mathbf{y})$. Instead, it is typical to fit data to a model with a small number of parameters. The Gaussian distribution is the most common model; it is often a good reflection of the real world distributions, and is analytically solvable in entropy integrals. Another important property of the Gaussian is that it gives the maximum entropy for a given standard deviation, allowing such models to be used to give an upper bound to entropy values. Thus, we model the distributions p and q as Guassian with vector means $\boldsymbol{\mu}_p, \boldsymbol{\mu}_q$ and covariances $\boldsymbol{\Sigma}_p, \boldsymbol{\Sigma}_q$. For Gaussian distributions the KLD is calculated:

$$\begin{aligned} D(p\|q) = \underset{p}{E}&\left[(\mathbf{y} - \boldsymbol{\mu}_p)^t\boldsymbol{\Sigma}_p^{-1}(\mathbf{y} - \boldsymbol{\mu}_p)\right. \\ &\left. + (\mathbf{y} - \boldsymbol{\mu}_q)^t\boldsymbol{\Sigma}_q^{-1}(\mathbf{y} - \boldsymbol{\mu}_q) + log_2\sqrt{\frac{|2\pi\boldsymbol{\Sigma}_q|}{|2\pi\boldsymbol{\Sigma}_p|}}\right] \\ = \frac{1}{2}&log_2|\boldsymbol{\Sigma}_q\boldsymbol{\Sigma}_p^{-1}| + tr\left(\boldsymbol{\Sigma}_p\boldsymbol{\Sigma}_q^{-1} - \mathbf{I}\right) + \mathbf{t}^T\boldsymbol{\Sigma}_q\mathbf{t} \quad (2) \end{aligned}$$

where $\mathbf{t} = \boldsymbol{\mu}_p - \boldsymbol{\mu}_q$ is the difference between the mean measurement for the instance and the prior mean. As $\mathbf{t}$ becomes large, the KLD increases dramatically as the current realization moves to the tails of the prior.

Equation 2 represents the information for a specific p. Instead, we are interested in the average information from measurements, weighted over the prior probability of each measurement configuration. Thus, we define the *information in measurements* (IM) to be

$$\begin{aligned} \text{IM} &= \underset{q}{E}\left[D(p\|q)\right] \\ &= \frac{1}{2}log_2|\boldsymbol{\Sigma}_q\boldsymbol{\Sigma}_p^{-1}| + tr\left(\boldsymbol{\Sigma}_p\boldsymbol{\Sigma}_q^{-1} - \mathbf{I}\right) + \underset{q}{E}\left[\mathbf{t}^T\boldsymbol{\Sigma}_q\mathbf{t}\right] \\ &= \frac{1}{2}log_2|\boldsymbol{\Sigma}_q\boldsymbol{\Sigma}_p^{-1}| + tr\left(\boldsymbol{\Sigma}_p\boldsymbol{\Sigma}_q^{-1}\right) \qquad (3) \end{aligned}$$

In normal applications, the variability due to noise is much less than the prior variability, and $|\boldsymbol{\Sigma}_p| \ll |\boldsymbol{\Sigma}_q|$, giving

$$\text{IM} = \frac{1}{2}log_2(|\boldsymbol{\Sigma}_q|/|\boldsymbol{\Sigma}_p|) \qquad (4)$$

The tr (trace) term will only have an effect in the very unusual case that the measurement distribution, p, is larger than that of the prior, q.

B. Information in Impedance Plethysmography

To clarify the application of IM, we initially consider an impedance plethymography system which measures a single impedance value for the change in impedance, $\mathbf{y}$, across the thorax, which varies with breathing, posture and heart activity.

In order to estimate the prior distribution, data acquisitions would be performed on a number of patients. From these data the inter-class mean (μ_q) and variance ($\boldsymbol{\Sigma}_q = \sigma_q^2$) would be calculated. Since we are calculating difference measurements, μ_q would be zero. For illustration purposes, consider that $\sigma_q = 800 \ mV$.

When using the measurement system on a specific patient at a given time, the measurements, $\mathbf{y}$, will be due to the current physiological condition and the noise. Consider measurement noise $\sigma_n = 10 \ mV$, and at a specific time, the difference measurement is $1 \ V$. The information via the KL divergence is

$$log_2\frac{\sigma_q}{\sigma_p} + \left(\frac{\mu_q - \mu_p}{\sigma_q}\right)^2 + \left(\frac{\sigma_p}{\sigma_q}\right)^2 - 1 \qquad (5)$$

which is 7.9 bits. In order to calculate the *information in measurements*, we must average over measurements on the specific patient, giving

$$\text{IM} = log_2\frac{\sigma_q}{\sigma_p} + \frac{1}{2}\left(\frac{\sigma_p}{\sigma_q}\right) \approx log_2\frac{\sigma_q}{\sigma_p} \qquad (6)$$

which is 6.3 bits.

C. Information in EIT

We consider a difference EIT system with n_E electrodes applied to a body using sequential current stimulation with parallel voltage measurement. Using these electrodes, n_E current stimulation patterns are sequentially applied and n_V differential measurements are made for each stimulation. For an adjacent drive EIT system, voltages are typically not measured at driven electrodes, and $n_V = n_E - 3$. Each data frame measures a vector, $\mathbf{v} \in \mathbb{R}^{n_M}$, of $n_M = n_E n_V$ data points (some of which are redundant if the medium is not changing). Difference EIT calculates difference data

$\mathbf{y} = \mathbf{v}_2 - \mathbf{v}_1$. To improve its precision, $\mathbf{v}_1$ is typically averaged over many data frames, at a time when the conductivity distribution may be assumed to be stable; we thus assume that $\mathbf{v}_1$ is noise free.

The body under investigation is modelled using a finite element model (FEM) which discretizes the conductivity onto n_N piecewise smooth elements, represented by a vector $\boldsymbol{\sigma} \in \mathbb{R}^{n_N}$ (In this paragraph, $\boldsymbol{\sigma}$ represents conductivity; elsewhere in this paper, σ is the standard deviation). Difference EIT calculates a vector of conductivity change, $\mathbf{x} = \boldsymbol{\sigma}_2 - \boldsymbol{\sigma}_1$, between the present conductivity distribution, $\boldsymbol{\sigma}_2$, and that at the reference measurement, $\boldsymbol{\sigma}_1$. For small variations around the reference conductivity $\boldsymbol{\sigma}_1$, the relationship between $\mathbf{x}$ and $\mathbf{y}$ can be linearized (giving the difference EIT forward model):

$$\mathbf{y} = \mathbf{J}\mathbf{x} + \mathbf{n} \tag{7}$$

where $\mathbf{J} \in \mathbb{R}^{n_M \times n_N}$ is the Jacobian or sensitivity matrix and $\mathbf{n} \in \mathbb{R}^{n_M}$ is the measurement noise which is assumed to be uncorrelated white Gaussian. $\mathbf{J}$ is calculated from the FEM as $\mathbf{J}_{ij} = \left.\frac{\partial \mathbf{y}_i}{\partial \mathbf{x}_j}\right|_{\boldsymbol{\sigma}_1}$, and depends on the FEM, current injection patterns, the reference conductivity, and the electrode models. This system is typically ill-posed, since $n_N > n_M$.

In order to calculate the class statistics from the forward model (eqn. 7), we calculate

$$\begin{aligned}\boldsymbol{\mu}_{\mathbf{y}} =&E\left[\mathbf{y}\right] = E\left[\mathbf{J}\mathbf{x} + \mathbf{n}\right] = \mathbf{J}\boldsymbol{\mu}_{\mathbf{X}} + E[\mathbf{n}]\\ \boldsymbol{\Sigma}_{\mathbf{y}} =&E\left[(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})^T\right]\\ =&\mathbf{J}E\left[(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{X}})(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{X}})^T\right]\mathbf{J}^T + E\left[\mathbf{n}\mathbf{n}^T\right]\end{aligned} \tag{8}$$

For distribution q, $\boldsymbol{\mu}_{\mathbf{X}} = \boldsymbol{\mu}_{\mathbf{y}} = 0$, since for difference EIT, positive and negative changes are equally likely. For distribution p, $\mathbf{x}$ is the realization of the image in the medium, and $\boldsymbol{\mu}_{\mathbf{y}} = \mathbf{J}\mathbf{x}$. Assuming a Gaussian model, the distribution covariances may be calculated:

$$\begin{aligned}\boldsymbol{\Sigma}_q =&\mathbf{J}\boldsymbol{\Sigma}_x\mathbf{J}^T + \boldsymbol{\Sigma}_n\\ \boldsymbol{\Sigma}_p =&\boldsymbol{\Sigma}_n\end{aligned} \tag{9}$$

giving the *information in measurements*

$$\begin{aligned}\text{IM} =&\frac{1}{2}log_2|\boldsymbol{\Sigma}_q\boldsymbol{\Sigma}_p^{-1}|\\ =&\frac{1}{2}log_2\left|\mathbf{J}\boldsymbol{\Sigma}_x\mathbf{J}^T\boldsymbol{\Sigma}_n^{-1} + \mathbf{I}\right|\end{aligned} \tag{10}$$

In many cases, the measurement noise is equal on all channels, and can be measured, giving $\boldsymbol{\Sigma}_n = \sigma_n^2\mathbf{I}$. In a similar way, a prior model can be constructed giving $\boldsymbol{\Sigma}_x = \sigma_x^2\mathbf{R}^{-1}$, where $\mathbf{R}^{-1}$ is the generalized Tikhonov regularization matrix. In this case, the regularization prior

is $\lambda = \frac{\sigma_n}{\sigma_x}$. In order that $(\frac{\sigma_x}{\sigma_n})^2$ represent the signal to noise ratio (SNR), $\mathbf{R}^{-1}$ should be normalized so that

$$tr(\mathbf{J}\mathbf{R}^{-1}\mathbf{J}^T) = n_N \tag{11}$$

Thus, in the case where $SNR \gg 1$, IM is as

$$\begin{aligned}\text{IM} =&\frac{1}{2}log_2\left|\frac{\sigma_x^2}{\sigma_n^2}\mathbf{J}\mathbf{R}^{-1}\mathbf{J}^T + \mathbf{I}\right|\\ =&\frac{1}{2}log_2\left|(\frac{1}{\lambda^2})(\mathbf{J}\mathbf{R}^{-1}\mathbf{J}^T + \lambda^2\mathbf{I})\right|\\ =&n_Nlog_2\left(\frac{\sigma_x}{\sigma_n}\right) + \frac{1}{2}log_2\left(\mathbf{J}\mathbf{R}^{-1}\mathbf{J}^T + \lambda^2\mathbf{I}\right)\end{aligned} \tag{12}$$

Equations for information content, such as (12) are often singular [7]; in this case IM will diverge to $-\infty$. This effect is due to the fact that measurements are not independent. To solve this problem, we wish to extract an independent set of measurements, $\mathbf{z}$, from $\mathbf{y}$. Using the singular value decomposition, we decompose the Jacobian matrix such that $\mathbf{U}\mathbf{D}\mathbf{U}^T = \mathbf{J}\mathbf{R}^{-1}\mathbf{J}^T$, with orthonormal $\mathbf{U}$ and diagonal $\mathbf{D}$. By truncating $\mathbf{U}$ after the first n_Z columns (yeilding $\mathbf{U}_{n_Z}$), we keep the n_Z most important measurements, and generate a mapping ($\mathbf{U}^T : \mathbf{y} \rightarrow \mathbf{z}$), from the original EIT measurements $\mathbf{y} \in \mathbb{R}^{n_M}$ to a new measurement space $\mathbf{z} \in \mathbb{R}^{n_Z}$. Equation (12) is then modified to give:

$$\begin{aligned}\text{IM} =&n_Zlog_2\left(\frac{\sigma_x}{\sigma_n}\right) + \frac{1}{2}log_2\left(\mathbf{U}^T\mathbf{J}\mathbf{R}^{-1}\mathbf{J}^T\mathbf{U} + \lambda^2\mathbf{I}_{n_Z}\right)\\ =&n_Zlog_2\left(\frac{\sigma_x}{\sigma_n}\right) + \frac{1}{2}log_2\left([\mathbf{D}^{\frac{1}{2}}\mathbf{R}^{-1}\mathbf{D}^{\frac{1}{2}}]_{n_Z} + \lambda^2\mathbf{I}_{n_Z}\right)\end{aligned} \tag{13}$$

where $[\cdot]_{n_Z}$ indicates selection of the upper left $n_Z \times n_Z$ submatrix.

Since λ is small, this is approximately

$$\text{IM} \approx n_Zlog_2\left(\frac{\sigma_x}{\sigma_n}\right) + \frac{1}{2}log_2\left([\mathbf{D}^{\frac{1}{2}}\mathbf{R}^{-1}\mathbf{D}^{\frac{1}{2}}]_{n_Z}\right) \tag{14}$$

This formula is intuitively appealing. The information in the measurements is proportional to the number of measurements and the log SNR, plus a term related to the Jacobian and the regularization matrix. Improvements to an EIT system design can thus be implemented by 1) improving the SNR, 2) adding measurement channels, and 3) designing better measurement schemes. That last function is the goal of the optimal current patterns, such as those of [4][5][6].

III. RESULTS

Test data were measured from a 30 cm diameter and 30 cm tall plastic cylindrical phantom filled with 0.9% saline

solution. Sixteen stainless steel electrodes were placed, equally spaced, around the circumference, and EIT data were acquired using the Goë-MF II EIT system (Viasys Healthcare, Höchberg, Germany) using an adjacent stimulation and measurement pattern. In order to measure the noise level, a sequence of 750 frames were measured from a homogeneous tank. To measure the signal level, a small non-conductive spherical object of 2 cm radius was introduced in the plane of the electrodes and moved into various positions on the horizontal plane.

To measure σ_n, the signal noise was calculated for each measurement channel by taking the standard devition for the entire frame sequence. The noise (σ_n) was taken as the average across all channels; it was somewhat consistent across measurement channels (std/mean = 0.71).

To measure σ_x the average signal (difference from homogeneous) from different placements of the spherical object was calculated in the same way as the average noise. The regularization prior $\mathbf{R}$ was set to the identity matrix, $\mathbf{I}$, scaled by the scheme of eqn 11. The Jacobian, $\mathbf{J}$ was calculated for a 2D 576 element mesh using the EIDORS software [1].

Two approaches are used to calculate IM: 1) Using eqn. 14 (unregularized); the value of IM was chosen corresponding to its maximum at $n_Z = 63$. The value calculated is IM= 239.4 bits. For n_Z above this value, the calculated IM rapidly diverges to $-\infty$. 2) Using eqn. 13 (regularized), the value of IM was calculated and shown as a function of n_Z in Fig. 1. The curve reaches a plateau of $IM = 245.1$ bits at $n_Z = 109$ and stays constant after that point. Interestingly, at $n_Z = 63$ both equations calculate exactly the same value.
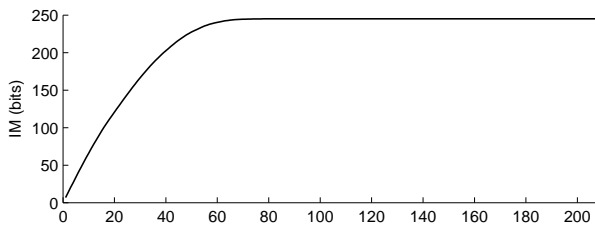

F ig. 1 *Information in measurements* (IM) (bits) vs number of independent measurements n_Z for EIT system

IV. DISCUSSION

This paper has introduced a definition of *information of measurements* and a formula to measure it for an EIT system. This calculation was implemented for a 16 electrode EIT system and a value of IM=245.1 bits calculated.

Interestingly, a naïve guess for the IM for this EIT system would be the number of independent measurements $\frac{1}{2}n_v n_e$ times $\frac{1}{2}log_2(\text{SNR})$. This calculation yeilds 406.3. The actual IM is less because EIT measurements from the medium are highly correlated, and such correlated measures are less informative.

Finally, we would like to suggest that this measure may allow novel insights into a number of questions regarding the performance of an EIT system, such as:

- Inherent limits to the compressibility of measurement data. Measured data cannot be stored in less space than the calculated IM value.
- Distinguishability limits may be defined in terms of the IM content from small contrasts [2].
- Fusion of EIT with other modalities. Measurements which are not independent will only add a small increment to the IM from EIT.
- Optimal current patterns may be defined in terms of maximizing IM [4][6].

REFERENCES

1. Adler A, Lionheart W R B (2006) Uses and abuses of EIDORS: An extensible software base for EIT. *Physiol. Meas.* 27:S25–S42
2. Cheney M, Isaacson D (1992) Distinguishability in impedance imaging, *IEEE Trans Biomed Imag*, 39:852–860
3. Cover T M, Thomas J A (1991) *Elements of Information Theory* New York: Wiley
4. Demidenko E, Hartov A, Soni N, Paulsen K D, (2005) On optimal current patterns for electrical impedance tomography *IEEE Trans Biomed Eng* 52:238–248
5. Isaacson D (1986) Distinguishability of conductivities by electric current computed tomography *IEEE Trans Med Imag* 5:91–95
6. Lionheart W R B, Kaipio J, McLeod C N, (2001) Generalized optimal current patterns and electrical safety in EIT, *Physiol Meas* 22:85–90
7. Wayman J S, "The cotton ball problem", *Biometrics Conference*, Washington DC, USA, Sep. 20-22, 2004.

Address of the corresponding author:
Author: Andy Adler
Institute: Carleton University
Street: 1125 Colonel By Drive
City: Ottawa
Country: Canada
Email: adler@sce.carleton.ca