Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Separation of gastric electrical control activity from simultaneous MGG/EGG recordings using independent component analysis

Andrei Irimia, Michael R. Gallucci, William O. Richards and L. Alan Bradshaw

Abstract— **Spatiotemporal parameters of gastric electrical control activity such as its amplitude, direction and propagation velocity are physiological parameters of distinctive clinical interest due to their potential use for differentiating between the healthy and diseased states of the human stomach. Whereas their time evolution is relatively well behaved in the case of healthy subjects, significant deviations from normal have been observed in patients suffering from a number of gastric diseases such as gastroparesis and gastropathy. For this reason, monitoring ECA parameters noninvasively may offer a useful test for the presence of such diseases whose diagnosis remains problematic. Here, we describe a method for computing ECA direction and orientation from simultaneous, noninvasive magnetogastrographic (MGG) and electrogastrographic (EGG) recordings. We demonstrate how independent component analysis and standard frequency analysis methods can be used to predict the locations and orientations of gastric current dipoles from MGG/EGG data. We compare our MGG-based dipole parameters to analogous ones obtained from simultaneous EGG recordings within the experimental framework of a human model. We find that magnetic recordings are superior in their ability to portray the underlying physiology of the stomach.**

I. INTRODUCTION

The study of gastric motility is of clinical interest due to the relationship between gastrointestinal (GI) disorders and the characteristics of gastric electrical control activity (ECA). ECA is a slow, sinusoidal wave with a frequency of 3 cycles per minute (cpm) originating in the antral gastric region and propagated along smooth muscle cells. Abnormal ECA has been associated with many GI disorders, such as gastroparesis, diabetic gastropathy, and gastric myoelectrical dysrhythmia [12]. Two methods for characterizing ECA are electrogastrography (EGG) and magnetogastrography (MGG). EGG involves the placement of cutaneous electrodes on the abdomen to record the bioelectric fields due to ECA at the body surface. The reliability of this procedure has been questioned due to the high dependence of electrical recordings upon tissue conductivity, the thickness of the abdominal wall, and the variable propagation velocity of ECA [10]. MGG involves the positioning of a magnetometer in close proximity to the abdomen to detect the biomagnetic field generated by the electric current of ECA. The measurement of biomagnetic fields is advantageous because they are more strongly dependent on tissue permeability, which is nearly equivalent to that of free space.

In this study, we describe our application of principal and independent component analysis (PCA and ICA, respectively) for measuring ECA parameters from simultaneous, noninvasive MGG/EGG recordings. Our approach makes use of ICA to isolate the gastric signal, whereafter gastric dipole locations and orientations can be determined from ICA field maps. We compare our results based on magnetic data to analogous ones obtained using simultaneous electrode recordings.

II. DATA ACQUISITION AND ANALYSIS

MGG recordings were acquired using a Superconducting QUantum Interference Device (SQUID) biomagnetometer. This device is sensitive enough to detect GI biomagnetic signals, which are very weak ($O(10^{-12}$ T)). The multichannel 637i SQUID biomagnetometer (Tristan Technologies Inc., San Diego, CA, USA) in the Gastrointestinal SQUID Technology Laboratory at Vanderbilt University has 37 detection coils located at the bottom of an insulated dewar filled with liquid He (T = 4.5 K). The detection coils are magnetically coupled to the SQUID coils, which are able to convert magnetic flux incident upon the detection coils to voltage signals. These signals are amplified and then acquired at the rate of 3 MHz. The detection coils are arranged in gradiometer format as a horizontal grid with 19 of them recording the Cartesian component of the magnetic field that is normal with respect to the grid plane (B_z). At five of these 19 locations, the other two Cartesian components of the field (B_x and B_y) are also measured. Our study was approved by the Vanderbilt University Institutional Review Board. In our experiments, the subject is asked to lie horizontally under the SQUID in a magnetically shielded room (Amuneal Manufacturing Corp., Philadelphia, PA, USA). The magnetometer is oriented so that the coils measuring the $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ components of the signal (tangential to the body surface) are oriented in the sagittal and horizontal planes, while the coils measuring the $\hat{\mathbf{z}}$ component (normal to the body surface) are oriented in the frontal plane. The electrode equipment employed for this experiment consisted of 18 pre-gelled, disposable, pick-up/ground electrodes (Rochester Electro-Med Inc., Tampa, FL, USA) connected to an isolated bioelectric amplifier (James Long Co., Caroga Lake, NY, USA). The electrodes were manually affixed to the skin in the configuration shown in Fig. 1. The relative spatial positioning between the MGG and EGG channels is also depicted in that figure.

This work was supported by NIH grants RO1 DK 58197 and 58697.

A. Irimia is with the Living State Physics Laboratories, Department of Physics and Astronomy, Vanderbilt University, Nashville, TN 37235-1807, USA andrei.irimia@vanderbilt.edu

M.R. Gallucci, W.O. Richards and L.A. Bradshaw are with the Department of Surgery, Vanderbilt University School of Medicine, Nashville, TN 37235 USA m.gallucci@vanderbilt.edu, bill.richards@vanderbilt.edu, alan.bradshaw@vanderbilt.edu

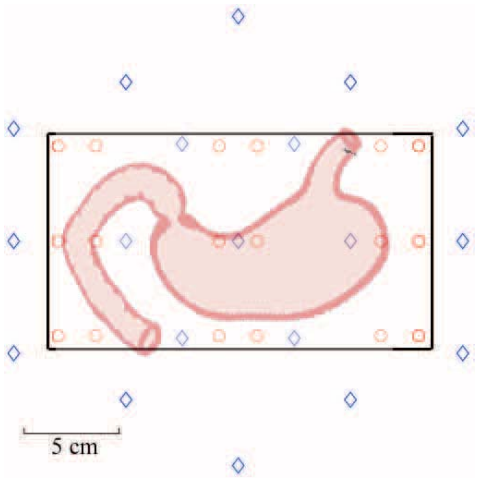

Fig. 1. Data acquisition grid for simultaneous MGG/EGG experiments. Bipolar electrode channels are indicated by red circles and SQUID magnetometer channels by diamonds. The black rectangle enclosing the electrode channels indicates the extent of the area shown in the maps of Fig. 3. The outline of the stomach is approximate and is meant to serve as a guide to the eye.

A common problem in MGG/EGG signal analysis is that the recorded signals are usually superpositions of the waveforms due to the underlying sources in the body. For this reason, there is usually not a one-to-one correspondance between the number of measured signals and that of the sources. The recorded signals are thus due to an unknown mixing of the true system variables [13]. To address this problem known as *blind signal separation*, we employed a signal processing technique called fast independent component analysis (fast ICA). This is a data analysis method that attempts to recover the unobserved signals or 'sources' from observed mixtures, i.e. from the output of the SQUID sensor array [3]. One of our previous studies [9] already demonstrated the use of ICA for noninvasive MGG signal processing, although the method has been widely used in related areas, including electrogastrography [11].

Because fast ICA requires the calculation of the data covariance matrix−which can be (near)-singular− dimensionality reduction is applied in our approach using principal component analysis (PCA), whereafter the fast ICA algorithm is applied. PCA is a technique that attempts to describe the variation in a set of multivariate data in terms of a set of uncorrelated variables, each of which is a distinct linear combination of the original variables [6]. The principal components (PCs) of the observation set are linear combinations of the underlying variables in that set. In particular, they are those specific combinations that maximize the variances of each PC subject to the constraint of orthonormality. After dimensionality reduction, the number of PCs that was retained in our study was set to the number of input signals, which account for most of the explained data variance.

The fast ICA algorithm of Hyvärinen and Oja [8] was employed in our study. Their method involves the minimization of the mutual information between the n scalar random variables that define the separated signals of interest. A detailed description of our approach to implementing fast ICA is available in [9].

Let $x_1, x_2, \ldots, x_n$ be the set of n observed MGG or EGG signals expressed as linear combinations of m random variables $s_1, s_2, \ldots, s_m$, that are the m underlying anatomical sources i.e. $x_i = \sum_{j=1}^{m} a_{ij} s_j$, where $i = 1, \ldots, n$, $a_{ij} \in \mathbb{R}$ and the sources s_j are assumed to be statistically mutually independent. Let $\mathbf{x}$ and $\mathbf{s}$ be the random vectors containing the mixtures $x_1, x_2, \ldots, x_n$ and $s_1, s_2, \ldots, s_n$, respectively, and let $\mathbf{A}$ denote the matrix with entries $A_{ij} = a_{ij}$. The mixing model of ICA can then be written simply as $\mathbf{x} = \mathbf{As}$, where $\mathbf{A}$ is known as the *attenuation matrix*. In terms of this formalism, the task of ICA is to find $\mathbf{s}$ (in our case, the underlying gastric, cardiac, etc. signals) in terms of the given $\mathbf{x}$ (i.e. the MGG or EGG signals). In mathematical language, this statement is equivalent to the task of finding a matrix $\mathbf{W}$ such that $\hat{\mathbf{s}} = \mathbf{Wx}$. On a side note, it should be mentioned that a distinction is made here between the true sources $\mathbf{s}$ and their estimators $\hat{\mathbf{s}}$ obtained via ICA. The task of ICA is carried out by identifying a suitable choice of the matrix elements of $\mathbf{A}$. The resulting estimator of $\mathbf{A}$ as found by ICA can be denoted by $\hat{\mathbf{A}} = \mathbf{W}^{-1}$, such that the relationship $\mathbf{x} = \hat{\mathbf{A}}\hat{\mathbf{s}}$ is satisfied.

The fast ICA method was separately implemented for each of the simultaneous MGG and EGG data sets to determine the independent components (ICs) derived from magnetic field and electric potential recordings, respectively. After computing the ICs of each MGG/EGG data set, the *labeling problem* of ICA was addressed. This problem is defined as the "task of finding substantive interpretations of some set of hypothetical latent variables which have been derived" [7]. In our case, this leads to the challenge of associating one of the computed ICs with the gastric signal (which is the latent variable of interest in our study). The labeling problem was addressed using an automated analysis of the ICs, whose purpose was to identify the IC whose waveform best matched the gastric ECA waveform (which has a distinct frequency of 3 cpm in humans). In most cases, the gastric IC was found to be among the first few ICs that accounted for the highest percentage of the variance in the recorded signals. We found this to be the case because the gastric signal is relatively strong compared to other components of biological origin and because it is superseded, in most cases, only by motion artifact signals.

Although there are advantages to performing a visual, offline analysis of ICs to determine the optimal ones to be used for analysis, an automated algorithm was developed instead. As an aid to the decision process, the power spectral density (PSD) of each IC (denoted henceforth by $\rho(f)$, where f is the frequency) was first computed using the classical fast Fourier transform, thus providing a time-frequency representation of MGG/EGG data. Both signal- and zero-padding were applied to reduce picket fence effects. In addition, appropriate Hamming windows were used. To investigate frequency changes in the gastric signal as a function of time, the recorded signals were split into segments denoted by

$j = 1, \ldots, n$ and fast ICA was applied separately to each of them. Suppose that m ICs were computed for the i-th segment and let k represent the index of each IC, i.e. $k = 1, \ldots, m$ (it is assumed that the ICs are ordered in decreasing order of the variance in the data for which they account). If we let the symbol I represent some given IC, the notation I_{jk} denotes, in this convention, the k-th IC in data segment j. In the first step of our approach, we computed $\rho(I_{1k}; f)$, i.e. the PSD for each IC in the first time segment of the data set. To assess the similarity between the PSDs of different ICs (as described below), ρ values were normalized to $\tilde{\rho}$:

$$\tilde{\rho}(f) = \frac{\rho(f)}{\int_0^\infty \rho(f)\mathrm{d}f}, \quad (1)$$

where in practice the upper limit of integration is constrained by the Nyquist Theorem. Next, the energy integral

$$\mathscr{E}(\alpha,\beta) = \int_\alpha^\beta \tilde{\rho}(f)\mathrm{d}f \quad (2)$$

was calculated, where α and β are the cutoff frequencies for the signal of interest. In our case, the choice made was [α, β] = [2, 5] cpm, which is the frequency interval where most of the energy in the human gastric signal is contained. One can see that the quantity

$$\frac{\mathscr{E}(\alpha,\beta)}{\mathscr{E}(0,\infty)} = \frac{\int_\alpha^\beta \tilde{\rho}(f)\mathrm{d}f}{\int_0^\infty \tilde{\rho}(f)\mathrm{d}f} = \int_\alpha^\beta \tilde{\rho}(f)\mathrm{d}f \quad (3)$$

is maximized by the IC that has the largest proportion of energy in the interval [α, β]. For this reason, the IC selected for analysis from the first segment of data (i = 1) was chosen based on this criterion. Let this IC be denoted by I_{1c} (c for 'chosen'). In the second stage of our selection algorithm, the subsequent time segments of data $j = 2, \ldots, n$ were processed. For each of these, the squared energy residuals

$$\Delta\mathscr{E}_{jk}^2 = \left[\int_0^\infty \tilde{\rho}(I_{1c}; f)\mathrm{d}f - \int_0^\infty \tilde{\rho}(I_{jk}; f)\mathrm{d}f\right]^2 \quad (4)$$

were evaluated. What $\Delta\mathscr{E}_{jk}^2$ provides is a measure of the 'similarity' between the first chosen independent component (I_{1c}) and each of the components in the subsequent time segments (I_{jk}). Naturally, this measure is minimized when the similarity is highest. For this reason, the IC with the lowest value of the squared energy residual was selected for analysis. A legitimate question can be raised by the reader as to why different selection criteria were chosen for the two steps of our algorithm. The reason lies in the fact that fast ICA can very possibly separate more than one gastric component from the data. Naturally, since these components may account for distinct dipolar patterns with distinct locations and rhythms, it is important to distinguish between the IC that was chosen initially (I_{1c}) and other ICs. This is because, since one wishes to determine the spatiotemporal evolution of I_{1c}, it is more useful to identify ICs in subsequent data segments that are most similar to I_{1c} (and therefore most likely to represent the signal due to the source being sought) rather than other dipoles−whether gastric or not−whose evolution is not under focus. Thus, to identify $I_{2c}, \ldots, I_{nc}$, the use of the quantitative measure $\Delta\mathscr{E}_{jk}^2$ is preferable. After selecting suitable ICs for each time segment in the data set, the signals provided by the chosen ICs were analyzed. First, each data set was detrended using a linear fit to eliminate short-lived trends that are due to extraneous causes, such as low-frequency noise. Then, a number of bandpass, second-order Butterworth filters were applied.

Our source model consists of a current dipole $\mathbf{Q}_1$ that approximates a current density $\mathbf{J}(\mathbf{r}')$ with spatial extent at $\mathbf{r}'$ in an ellongated cylinder with an inhomogeneous conductivity profile. Since the ECA is a depolarization/repolarization wavefront advancing along the corpus, the current dipole approximation is appropriate in our context. For our purposes, the inhomogeneity of the spatial conductivity function σ can be approximated through the use of small regions of piecewise homogeneous conductivity. The sources $\mathbf{Q}_i$ are assumed to be buried beneath the skin surface, which is taken to be the plane z = 0 cm. It is relatively straightforward [15] to show that, if the distance between the two extrema of $\mathbf{B}(\mathbf{Q}_i)$ on the surface is taken to be γ and the dipole is approximately parallel to the surface, then the dipole lies at a depth $z = -\gamma/\sqrt{2}$ on the vertical plane that passes through the extrema.

The process of identifying the location and orientation of bioelectric sources from information provided by ICA has already been successfully implemented in magnetoencephalography [5] and magnetocardiography [4]. Our particular approach closely follows the method proposed by Tang and colleagues [14]. Spatial information regarding a separated IC is provided by its sensor projection (field map), which represents the measured sensor response to the activation of the component $\hat{\mathbf{s}}_{jk}$. The information required to construct the map of component k in time segment j is contained in the k-th column of the estimated attenuation matrix $\hat{\mathbf{A}}$ for that particular IC, which we denote by $\hat{\mathbf{A}}_k$. The values in $\hat{\mathbf{A}}_k$ were used to create a two-dimensional field/potential map using adaptive cubic spline interpolation. This method was then employed to generate 2D spatial maps that show, in addition to the underlying field/potential maps, the locations of the current dipoles as inferred from the field/potential map data.

III. RESULTS AND DISCUSSION

Sample simultaneous MGG/EGG waveforms are shown in Fig. 2. In Fig. 3, examples of field/potential maps from three non-consecutive time points (A, B, C) are shown. The quantity visualized in each spatial map is the *normalized* amplitudes of the magnetic field ($\tilde{B}_i$) and electric potential ($\tilde{V}_i$) due to three identified gastric sources, where $\tilde{B}_i = B_i/\max\{B_i\}$ and $\tilde{V}_i = V_i/\max\{V_i\}$. Also depicted are the inferred locations of the associated current dipole producing the fields/potentials shown in the map. The decision to plot $\tilde{B}_i$ and $\tilde{V}_i$ rather than B_i and V was made so as to simplify the visual comparison of different maps. Because only one colorbar and colormap were used for *all* maps (although $\max\{B_i\}$ and $\max\{V_i\}$ are different for each map), plotting

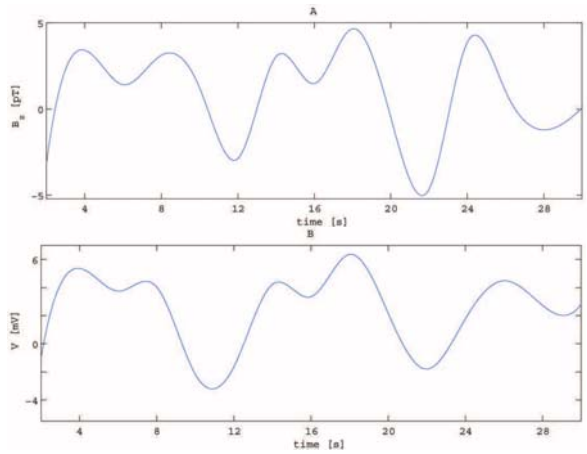

Fig. 2. Sample ECA waveform as recorded using simultaneous MGG (A) and EGG (B). Signal extraction was performed using fast ICA and appropriate Butterworth filters and Hamming windows were applied to the selected ICs to generate these plots. Both waveforms display the characteristic gastric frequency of 3 cpm.

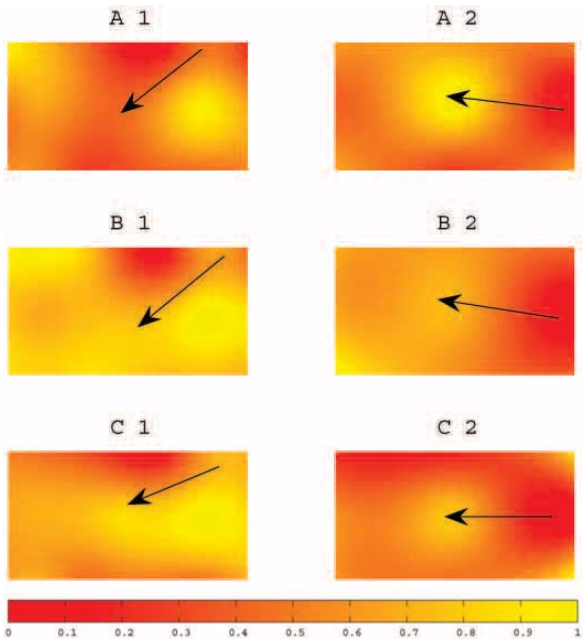

Fig. 3. Examples of field/potential maps from three non-consecutive time points (A, B, C). Maps of magnetic field (B_z, see A1, B1, C1) and electric potential (V, see A2, B2, C2) patterns were extracted from simultaneous MGG/EGG recordings using fast ICA. Pictured are spatial maps of *normalized* (see text) B and V values as inferred via fast ICA (the first column depicts $\tilde{B}_i$, the second column depicts $\tilde{V}$). Grid dimensions are 17 $\times$ 10 cm. Black arrows are shown as guides to the eye to indicate the locations and orientations of the underlying current dipoles as inferred from the spatial maps using the right hand rule.

$\tilde{B}_i$ and $\tilde{V}_i$ ensured that the same proportional scaling was used for all maps. Thus, comparison between different maps is simplified.

All maps have a spatial extent of 17 $\times$ 10 cm. Because simultaneous MGG/EGG recordings were acquired, maps were computed from both magnetic field (B) and electric potential (V) signals. This procedure allowed for the comparison of noninvasive biomagnetic recordings to cutaneous electrode data. In Fig. 3, the orientations of MGG-based dipoles are more realistic and are similar, in fact, to those obtained from MGG data in a previous study using the surface current density method [1].

One important observation to be made from the analysis of our figure is that, due to its higher spatial resolution, the magnetic sensor grid is superior to the electrode grid from the standpoint of its ability to capture the configuration of the gastric biomagnetic field. In spite of grid coarseness issues, electric maps are nevertheless valuable because they offer a method of comparison between the dipole locations and orientations obtained using the two protocols. Quite possibly, differences in location and orientation between MGG- and EGG-based dipoles are due to the different signal-to-noise ratios of the two procedures as well as to the difference in the finesse and configurations of the two grids. In conclusion, there is agreement between the two although magnetic recordings are superior in their ability to portray the underlying physiology.

REFERENCES

[1] LA Bradshaw, A Irimia, JA Sims, RL Palmer, WO Richards (2006) Biomagnetic characterization of spatiotemporal parameters of the gastric slow wave *Neurogastroenterol Motil* in press

[2] RJ Brzana, KL Koch, S Bingaman (1998) Gastric myoelectrical activity in patients with gastric outlet obstruction and idiopathic gastroparesis *Am J Gastroenterol* **93** 1803-1809

[3] J-F Cardoso (1998) Blind signal separation: statistical principles *Proc IEEE* **86** 2009-2025

[4] S Comani, D Mantini, G Alleva, S di Luzio, GL Romani (2004) Fetal magnetocardiographic mapping using independent component analysis *Physiol Meas* **25** 1459-1472

[5] A Delorme, S Makeig, M Fabre-Thorpe, TJ Sejnowski (2002) From single-trial EEG to brain area dynamics *Neurocomput* **44-46** 1057-1064

[6] BS Everitt and G Dunn (1992) Applied multivariate data analysis *Oxford University Press*: New York, NY 45-64

[7] RJ Harris (1975) A primer of multivariate statistics *Academic*: New York, NY 156-204

[8] A Hyvärinen (1999) Fast and robust fixed-point algorithms for independent component analysis *IEEE Trans Neural Networks* **10** 626-634

[9] A Irimia, LA Bradshaw (2005) Artifact reduction in magnetogastrography using fast independent component analysis *Physiol Meas* **26** 1059-1073

[10] J Liang and JDZ Chen (1997) What can be measured from surface electrogastrography - computer simulations *Digest Dis Sci* **42** 1331-1343

[11] H Liang (2005) Extraction of gastric slow waves from electrogastrograms: combining independent component analysis and adaptive signal enhancement *Med Biol Eng Comput* **43** 245-251

[12] LW Qian, PJ Pasricha and JDZ Chen (2003) Origins and patterns of spontaneous and drug-induced canine gastric myoelectrical dysrhythmia *Digest Dis Sci* **48** 508-515

[13] CJ Stam (2005) Nonlinear dynamical analysis of EEG and MEG: review of an emerging field *Clin Neurophysiol* **116** 2266-2301

[14] AC Tang, BA Pearlmutter, NA Malaszenko, DB Phung and BC Reeb (2002) Independent components of magnetoencephalography: single-trial response onset times *NeuroImage* **17** 1773-1789

[15] Tripp JH (1983) Biomagnetism: an interdiscipliary approach *SJ Williamson, GL Romani, L Kaufman and I Modena, eds.* NATO ASI Series A: Life Sciences **66** pp. 101-140

[16] MAMT Verhagen, M Samson and AJPM Smout (1998) Gastric myoelectrical and antroduodenal motor activity in patients with achalasia *Neurogastroenterol Mot* **10** 211-218