(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
30 November 2017 (30.11.2017)

WIPO | PCT

(10) International Publication Number
WO 2017/205734 A1

(51) International Patent Classification:
A61B 5/0476 (2006.01) *A61B 5/05* (2006.01)
A61B 5/00 (2006.01) *A61B 5/055* (2006.01)
A61B 5/04 (2006.01)

(21) International Application Number:
PCT/US2017/034667

(22) International Filing Date:
26 May 2017 (26.05.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/342,109 26 May 2016 (26.05.2016) US

(71) Applicant: **UNIVERSITY OF WASHINGTON** [US/US]; 4545 Roosevelt Way Ne, Suite 400, Seattle, WA 98105 (US).

(72) Inventors: **TAULU, Samu**; C/o University Of Washington, 4545 Roosevelt Way Ne, Suite 400, Seattle, WA 98105 (US). **LARSON, Eric, D.**; C/o University Of Washington, 4545 Roosevelt Way Ne, Suite 400, Seattle, WA 98105 (US).

(74) Agent: **LINCICUM, Matthew, V.** et al.; Perkins Coie LLP, P.O. Box 1247, Seattle, WA 98101 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KH, KN, KP, KR,

(54) Title: REDUCING SENSOR NOISE IN MULTICHANNEL ARRAYS USING OVERSAMPLED TEMPORAL PROJECTION AND ASSOCIATED SYSTEMS AND METHODS

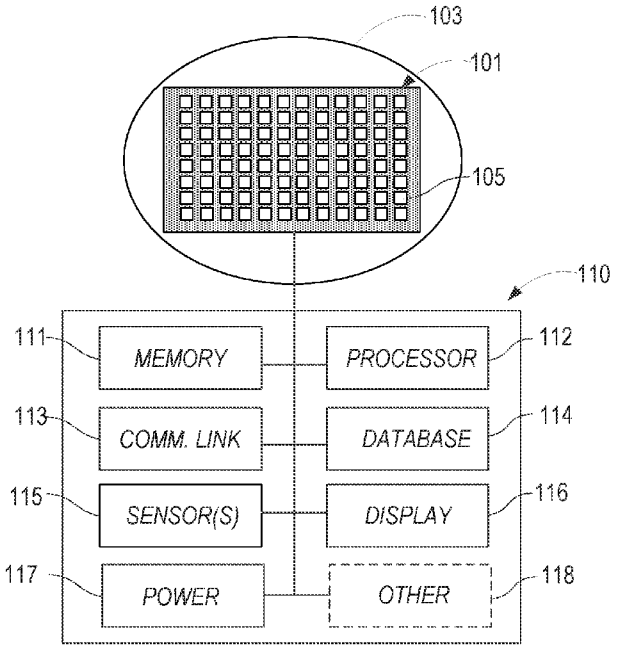


FIG. 1

(57) Abstract: A method for suppressing sensor noise in a spatially oversampled sensor array includes receiving spatially oversampled multi-channel sensor data from a region of interest and building a spatial model from the data for essential spatial degrees of freedom. The method further includes decomposing the data into the underlying spatial model to obtain associated amplitude components containing a mixture of original temporal waveforms of the data and, for each channel of the multi-channel sensor, estimating time-domain amplitude components using cross-validation. Next, for each channel, based on the estimated time-domain amplitude components, sensor noise and/or artifacts for that channel are identified. Finally, for each channel, the identified sensor noise and/or artifacts can be suppressed from the data.

WO 2017/205734 A1

[Continued on next page]

KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

REDUCING SENSOR NOISE IN MULTICHANNEL ARRAYS USING OVERSAMPLED TEMPORAL PROJECTION AND ASSOCIATED SYSTEMS AND METHODS

BACKGROUND

[0001] This application claims the benefit of U.S. Provisional Application No. 62/342,109, filed May 26, 2016, which is incorporated by reference herein in its entirety.

TECHNICAL FIELD

[0002] The present application generally relates to reducing sensor noise in multichannel measurement techniques including, for example, magnetoencephalography (MEG) and electroencephalography (EEG).

BACKGROUND

[0003] Multichannel measurement techniques are typically used to record signals emitted by some object of interest in order to obtain data on spatial or temporal structures of the object. The object may be, e.g., the human brain with electric currents flowing in the brain tissue, reflecting ongoing resting-state activity or responses to external stimuli. The corresponding multichannel measurement could be related either to the detection of the magnetic fields or electric potential outside the head. These measurement techniques are called magnetoencephalography (MEG) and electroencephalography (EEG), respectively. Other measurements of magnetic fields include, for example, geomagnetic measurements with the aim of revealing information about the electric fluctuations taking place under the surface of the earth, and magnetocardiography (MCG) recordings of the magnetic fields produced by the heart. In general, embodiments of the present technology can be applied to any measurement that provides oversampling of the signal of interest with a multitude of sensors whose intrinsic sensor noise and artifacts can be considered statistically independent between channels and also with the actual signal of interest.

[0004] As a prerequisite for the analysis of a multichannel measurement, the acquired signals must usually be processed against environmental interference, and sometimes certain signals generated by the object itself with temporal or spatial structures that mask the actual target of the analysis must be suppressed as well. The spatial processing methods act as spatial

filters that suppress certain spatial patterns by forming suitable linear combinations of all or of a subset of channels. Temporal processing methods alter the data in the temporal domain by, e.g., decomposing the signals into Fourier components and reconstructing the modified signals from a subset of frequency components so as to suppress signals from frequency bands that are not of interest in the analysis. In both of the above approaches, one exploits some *a priori* information about the spatial or temporal distribution of the signal of interest, and, if available, differences of these distributions in relation to signal components that are considered interference.

[0005] The intrinsic sensor noise, however, may also compromise the analysis results. Efficient suppression against this kind of noise is difficult because typically only some statistical information is available about its temporal and spatial structure. There is literature available on the hardware mechanisms causing device-related noise, but this information usually does not help to construct exact mathematical models for sample-by-sample suppression of the noise. In MEG, for example, researchers often know the levels of the standard deviation of intrinsic noise at different channels, but the noise fluctuations are random processes without a deterministic temporal or spatial pattern and, therefore, they are mixed between channels in the spatial filtering or have residual in the pass-band during temporal filtering. The aforementioned mixing that arises in spatial filtering is important because the spatial models do not have an accurate representation for the random sensor noise. Therefore, for any given moment of time, the momentary spatial distribution of the sensor noise values gets redistributed among measurement channels through a spatial model, which requires accurate *a priori* information about the spatial signal structures. This may be especially harmful if a specific artifact with temporal pattern resembling a potential signal of interest, but appearing in one channel only, spreads among other channels and may thus appear as a signal of interest, both spatially and temporally. Traditionally, the only reliable procedure to avoid such channel artifact spreading is visual inspection of the original data and manual rejection of bad channels or segments of time in signal processing and analysis. With modern MEG systems containing hundreds of channels, visual inspection of raw data is cumbersome and may significantly hamper clinical work, which is time-sensitive in nature. Accordingly, automated sensor noise suppression is desirable.

[0006] Suppression of sensor noise is most beneficial in the analysis of weak signal components of interest that correspond to a frequency band initially dominated by uncorrelated

sensor noise. An example of such signals in biomagnetic measurements is the high-frequency oscillations (HFO) related to epilepsy. HFO signals are hypothesized to contain significant clinical value but their detection with MEG or EEG is difficult. Broadband forms of high-frequency activity, e.g., over 60 Hz, which are often observed in intracranial recordings, are also difficult to isolate using MEG or EEG. On the other hand, high frequencies are usually not significantly contaminated by external interference signals, so the main component compromising the signal-to-noise ratio is random sensor noise. Thus, sensor noise suppression could play a major role in improving the detectability of HFO and broadband signals in EEG and MEG.

[0007] MEG and EEG are commonly used to detect brain responses to events such as auditory, visual, or somatosensory stimuli. Even in favorable measurement conditions, individual responses are often too weak to be detected under sensor noise. Therefore, several repetitions of the evoked responses are collected and averaged in order to suppress the contribution of noise that is independent of the stimulus presentation. In such a case, the signal-to-noise ratio improves proportionally to $\sqrt{N_t}$, where N_t is the number of trials. With a lower initial sensor noise level, a satisfactory signal-to-noise ratio could be reached with a smaller N_t, which would enable shorter measurements and reduce the effects of habituation in the analysis assuming accurate repeatability of the evoked response.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] Figure 1 is a schematic diagram of a multi-channel sensor system configured in accordance with embodiments of the present technology.

[0009] Figure 2 is a schematic diagram of the process flow for suppressing sensor noise for a multi-channel sensor system in accordance with embodiments of the present technology.

[0010] Figure 3 illustrates examples of various noise-suppression techniques operating on exemplary original data.

[0011] Figure 4 illustrates examples of suppressing non-correlated sensor artifacts using various noise-suppression techniques.

[0012] Figure 5 illustrates an example of results of an oversampled temporal projection (OTP) noise-suppression operation on raw data.

[0013] Figure 6 illustrates various examples of source space and sensor space suppression with different noise-suppression techniques.

[0014] Figure 7 illustrates various examples of reducing localization bias using different noise-suppression techniques.

DETAILED DESCRIPTION

[0015] The present technology is generally related to methods for sensor noise and artifact suppression that do not require user intervention or assumptions about the spatial and temporal structures of the noise, other than statistical independence between the noise and signals of interest and spatial oversampling of the signals of interest. Spatial oversampling is a reasonable assumption for MEG and EEG data, but applies equally well in many other potential domains. The disclosed techniques, referred to herein as "oversampled temporal projection" (OTP), operate on overlapping windowed segments of data. Within each window, each channel is projected onto the temporal basis formed by all other channels, and the residual is taken as an estimate of the sensor noise signal of the channel that can be suppressed for efficient sensor noise suppression. As described below, this procedure can be automatically performed on each channel at a time without any user intervention. This method is expected to minimize spreading of channel artifacts to other channels, a feature that may be of particular relevance in clinical work. In addition to facilitating signal analysis workflows by reducing the burden of visual data inspection with respect to sensor artifacts, the disclosed technique is expected to make signal processing and analysis of multichannel data more robust. OTP can be used as a first step in a chain of signal processing procedures, as any other spatial filters used in signal processing or analysis will have significantly reduced spreading of channel artifacts when the signals have first been processed with OTP.

State-of-the-Art Noise Suppression Techniques

[0016] To address problems caused by sensor noise, multiple algorithms have been developed. Some of the most relevant state-of-the art sensor noise suppression methods include a method developed primarily for EEG that uses Kalman filtering targeted to remove electromyography (EMG) artifacts, but requires making *a priori* assumptions about the nature of noise – in particular that it consists of spikes with a Poisson distribution – which make this method less well suited in general for reducing sensor noise. Adaptive auto-regressive (AR) modeling of noise has also been proposed, but requires an assumption of time dependence of

samples and Gaussian distribution of noise sources. More recently, a temporal artifact removal method was developed specifically for MEG, but assumes a parametric generative model for MEG data, which might not generalize to all forms of sensor-specific noise and artifacts.

[0017] Cross-validation (CV) based algorithms have recently been developed to address sensor noise. In CV methods, one channel at a time is left out from the acquired data and the associated data models, and the signal of the channel under investigation is then reconstructed based on all other channels or a suitable subset of channels. More formally, the procedure assumes spatial oversampling of the signal of interest and a spatial model that only contains the essential spatial features capable of explaining the data. Thus, the measured signals can be projected onto the parameters of the model and these parameters can subsequently be used to reconstruct the signal of the channel that was left out. It is evident that in such a procedure the individual noise of the channel that was left out will be absent in the reconstructed signal, but the challenge remains that part of the noise on the other channels will distort the reconstructed signal. The "sensor noise suppression" (SNS) algorithm, which uses the correlation between channels to build a spatial denoising operator, is one example of such a CV algorithm. More recently, another CV-based algorithm utilizing vector spherical harmonics as the spatial basis was introduced.

[0018] Denoising algorithms can also be performed in multiple (time) segments, which has been suggested as one way to overcome computational challenges. However, such window-based processing can effectively change the process from linear time invariant to linear time varying. Window-based processing could thus also provide potential benefits in terms of denoising quality, especially in cases where the set of sensors operating optimally changes gradually over time. A variant of this time-varying processing idea is utilized by the sparse-time artifact removal (STAR) algorithm, which uses statistical thresholding to identify temporally sparse artifacts and subsequently repair them individually with a SNS-like operation.

Suitable System(s)

[0019] The following discussion provides a brief, general description of a suitable environment in which the present technology may be implemented. Although not required, aspects of the technology are described in the general context of computer-executable instructions, such as routines executed by a general-purpose computer. Aspects of the

technology can be embodied in a special purpose computer or data processor that is specifically programmed, configured, or constructed to perform one or more of the computer-executable instructions explained in detail herein. Aspects of the technology can also be practiced in distributed computing environments where tasks or modules are performed by remote processing devices, which are linked through a communication network (e.g., a wireless communication network, a wired communication network, a cellular communication network, the Internet, a short-range radio network (e.g., via Bluetooth)). In a distributed computing environment, program modules may be located in both local and remote memory storage devices.

[0020] Computer-implemented instructions, data structures, screen displays, and other data under aspects of the technology may be stored or distributed on computer-readable storage media, including magnetically or optically readable computer disks, as microcode on semiconductor memory, nanotechnology memory, organic or optical memory, or other portable and/or non-transitory data storage media. In some embodiments, aspects of the technology may be distributed over the Internet or over other networks (e.g. a Bluetooth network) on a propagated signal on a propagation medium (e.g., an electromagnetic wave(s), a sound wave) over a period of time, or may be provided on any analog or digital network (packet switched, circuit switched, or other scheme).

[0021] Figure 1 is a schematic diagram of a system 100 configured in accordance with an embodiment of the disclosed technology. The system 100 includes a multi-channel sensor module 101 disposed over a region of interest 103. The multi-channel sensor module 101 can include a plurality of individual sensors 105. The multi-channel sensor module 101 can be any type of sensor, for example, a MEG sensor array, an EEG sensor array, geomagnetic sensor array, magnetocardiography (MCG) sensor array, etc. The region of interest 103 can be, for example, a patient's head in the case of EEG or MEG sensors, a region of earth in the case of geomagnetic sensors, or a patient's heart in the case of MCG sensors. It will also be appreciated that the region of interest 103 may comprise other suitable anatomical and/or geographic locations.

[0022] The multi-channel sensor module 101 can be configured such that the combination of individual sensors 105 provides spatial oversampling of the signal of interest (i.e., the multi-channel sensor module 101 contains more individual sensors 105 or measurement channels than degrees of freedom in the detectable signal with a spatially correlated structure). The

multi-channel sensor module 101 can be communicatively coupled to a processing subsystem 110 as described in more detail below.

[0023] The processing subsystem 110 comprises several components including memory 111 (e.g., one or more computer readable storage modules, components, devices) and one or more processors 112. The memory 111 can be configured to store information (e.g., signal data, subject information or profiles, environmental data, data collected from one or more sensors, media files) and/or executable instructions that can be executed by the one or more processors 112. The memory 111 can include, for example, instructions for processing multi-channel sensor data and suppressing noise and other artifacts from individual channels.

[0024] The processing subsystem 110 also includes communication components 113 (e.g., a wired communication link and/or a wireless communication link (e.g., Bluetooth, Wi-Fi, infrared and/or another wireless radio transmission network)) and a database 114 configured to store data (e.g., signal data acquired from the region of interest, equations, filters) used in the noise-suppression techniques disclosed herein. One or more sensors 115 can provide additional data for use in noise-suppression and/or other signal processing and analysis. The one or more sensors 115 may include, for example, one or more ECG sensors, blood pressure monitors, galvanometers, accelerometers, thermometers, hygrometers, blood pressure sensors, altimeters, gyroscopes, magnetometers, proximity sensors, barometers and/or hall effect sensors. One or more displays 116 can provide video output and/or graphical representations of data obtained by the system 100. A power supply 117 (e.g., a power cable connected to a building power system, one or more batteries and/or capacitors) can provide electrical power to components of the processing subsystem 110 and/or the system 100. In embodiments that include one or more batteries, the power supply 117 can be configured to recharge, for example, via a power cable, inductive charging, and/or another suitable recharging method. Furthermore, in some embodiments, the processing subsystem 110 may one or more additional components 118 (e.g., one or more microphones, cameras, Global Positioning System (GPS) sensors, Near Field Communication (NFC) sensors).

[0025] In some embodiments, the processing subsystem 110 may include one or more components partially or wholly incorporated into the sensor module 101. In other embodiments, however, the processing subsystem 110 may include components remote from the sensor module 101 and connected thereto by a communication network (e.g., the Internet and/or another network). In some embodiments, for example, at least a portion of the

processing subsystem 110 may reside on a mobile device (e.g., a mobile phone, a tablet, a personal digital assistant) and/or a computer (e.g., a desktop computer, a laptop) communicatively coupled to the sensor module 103.

Suitable Methods

[0026] Figure 2 is a flow diagram of a process 200 configured in accordance with an embodiment of the present technology. The process 200 can include instructions stored, for example, in the memory 111 of the system 100 (Figure 1) that are executable by the one or more processors 112 (Figure 1). In some embodiments, portions of the process 200 are performed by one or more hardware components (e.g., the sensor module 101 of Figure 1). In certain embodiments, portions of the process 200 are performed by a device external to the system 100 of Figure 1.

[0027] At block 202, the process 200 receives spatially oversampled data from the multi-channel sensor (e.g., the sensor module 101 of Figure 1). In one example, the data can be MEG data recorded over a period of time from an array of MEG sensors disposed over a patient's head, for example 300 or more individual MEG sensors. The data includes a plurality of individual channels, for example in some embodiments one channel per individual sensor.

[0028] At block 204, the process 200 constructs a spatial model for essential degrees of freedom for the data obtained in block 202. At block 206, the process 200 continues by decomposing the data into the spatial model. The spatial model has amplitude components that contain temporal waveforms from the original data obtained in block 202. In some embodiments, for example, the spatial model can be built by implementing a combination of a basic solution of Laplace's equation for the harmonic scalar potential, taking the corresponding gradient, and thus determining the magnetic field in the measurement area to calculate a spatial basis. In certain embodiments, the spatial model is built using dominating spatial patterns found in a statistical analysis of the original data to calculate a spatial basis, for example using dominating spatial patterns found in a principle component analysis.

[0029] At block 208, the process 200 estimates time-domain amplitude components using cross-validation for each channel, and at block 210, the process 200 determines the sensor noise and/or artifacts for each channel. Once identified, the process 200 suppresses the sensor noise and/or artifacts for each channel.

[0030] In some embodiments, sensor noise and/or artifacts are suppressed by applying an orthogonal projection operator, while in other embodiments a parallel projection operator is used. In one example, identifying and suppressing sensor noise and/or artifacts can include decomposing the data in two situations for each channel of the multi-channel device. In the first situation, the amplitude components are estimated when applying the spatial model to all channels, while in the second situation, the amplitude components are estimated when applying the spatial model with the channel under investigation excluded from the data and the spatial model. Next, the results of the first situation and the results of the second situation can be compared, with the difference indicating the sensor noise and/or artifacts for the channel under investigation. This process can be repeated for each channel to identify sensor noise and/or artifacts for each individual channel. In some embodiments, the results are compared using an orthogonal projection operation. In some embodiments, the results are compared using a singular value decomposition of the temporal domain signals of the amplitude coefficients corresponding to the spatial model to find a temporal basis.

[0031] In another example, the sensor noise and/or artifacts are identified and suppressed by applying the spatial model with the channel under investigation excluded from the data and the spatial model. The difference in the temporal domain between the decomposition and the original data for the channel under investigation can indicate the sensor noise and/or artifacts for the channel under investigation. To suppress the identified sensor noise and/or artifacts, a temporal waveform common to the channel under investigation and the decomposition can be found. In some embodiments, the common temporal waveform can be found by projecting the original channel waveform to a temporal basis spanned by the decomposed amplitude coefficients by a parallel projection operator. The parallel projection operator can be a singular value decomposition of the temporal domain signals of the amplitude coefficients corresponding to the spatial model to find the temporal basis.

[0032] In some embodiments, the spatial model can be constructed as outlined below. For an instance of a multichannel measurement instrument with N independent measurement channels that essentially detects signals belonging to a signal subspace with $n < N$ degrees of freedom (*i.e.*, a spatially oversampled measurement system), a multichannel data matrix D of N channels (rows) and n_s temporal samples (columns) can be defined as:

$$D = M_a X_a + M_b X_b + E_u + E_c \qquad (1)$$

Here M_a and M_b are $N \times na$- *and* $N \times nb$ -dimensional matrices defining the detectable subspaces for signals of interest and externally produced interference, respectively, such that $n_a + n_b = n$. X_a and X_b are the corresponding $n_a \times n_s$- and $n_b \times n_s$ dimensional coordinate matrices. For example, M_a could be a brain-to-sensor physical mixing model in MEG and M_b could span magnetic fields produced by environmental or biological, such as heart, interference sources. However, embodiments of the present technology are particularly useful in addressing the role of uncorrelated sensor noise without considering external interference or the existing algorithms that effectively suppress it, e.g., signal-space projection (SSP), signal space separation (SSS), or time-shift principal component analysis (tsPCA). Therefore eq. 1 can be simplified by making the substitutions $M = [M_a\ M_b]$ and $X = [X_a^T\ X_b^T]^T$, which leads to:

$$D = MX + E_u + E_c \quad (2)$$

[0033] Thus, $D_0 = MX$ is the noiseless signal that can be modeled with an overdetermined spatial model. In some embodiments, the assumption can be made that cross-talk between channels is negligible or has been compensated for mathematically using other processes.

[0034] There are two $N \times n_s$ -dimensional noise matrices:

1) E_u contains spatially uncorrelated artifacts and individual sensor noise. The corresponding covariance matrix $C_u = E[E_u E_u^T]$ is diagonal (E indicates expectation).

2) E_c contains spatially correlated artifacts that are not spanned by matrix M. For example, radiation shield noise or common-mode artifacts caused by the measuring instrument itself typically belong to this category.

[0035] For the sake of simplicity, E_c can be set to 0 in the sensor noise investigations addressed herein. Regarding modeling error, if M is composed of basis functions spanning the signal distributions of interest, it is often possible to construct M in such a way that the modeling error is negligible. The quasistatic magnetic field, for example, can be expressed with a rapidly converging series of spherical harmonic basis functions that allows one to express the measured signals with a relatively small number of degrees of freedom, that is, number of columns in M. Thus, model in eq. 1 reduces to

$$D = MX + E_u \quad (3)$$

[0036] As noted previously, spatial operators for noise-suppression have been identified, and examples of such spatial operators are outlined in more detail in Appendix A. However, there are several benefits to employing a temporal operator for sensor noise suppression using cross-validation.

[0037] Traditionally, the approaches for sensor noise suppression are formulated with the spatial domain as the starting point. However, CV operators can also be constructed by considering signals in the temporal domain. The simplest formulation is perhaps the channel-wise time domain-parallel projection

$$\tilde{\mathrm{d}}_j = \mathrm{d}_j \mathrm{V}_{\notin j} \mathrm{V}_{\notin j}^T \tag{4}$$

where d_j and $\tilde{\mathrm{d}}_j$ are the original and sensor-noise-suppressed $1 \times n_s$ signal vectors, respectively, of channel j, and $\mathrm{V}_{\notin j}$ is the orthonormal *temporal* subspace derived from all channels excluding channel j itself. The set denoted as $\notin j$ could also correspond to a smaller subset of channels excluding j, as long as the number of channels in this subset would exceed n and properly sample the temporal subspace.

[0038] The main question in the temporal-domain cross validation is how to determine the signal subspace $\mathrm{V}_{\notin j}$. There are many such potential formulations, but the simplest is perhaps given by right singular vectors of $\mathrm{D}_{\notin j}$:

$$\mathrm{D}_{\notin j} = \mathrm{U}_{\notin j} \mathrm{S}_{\notin j} \mathrm{V}_{\notin j}^T \tag{5}$$

with the orthonormalized temporal waveforms of $\mathrm{D}_{\notin j}$ being represented in the rows of matrix $\mathrm{V}_{\notin j}^T$.

[0039] This simple formulation results in a relationship between the temporal formulation of sensor noise suppression via OTP to the SNS method's $\mathrm{K}_{\mathrm{SNS}}$ operator matrix. Coefficients for $\mathrm{K}_{\mathrm{SNS}}$ are derived from the principle component analysis (PCA) of the data. The eigenvectors A and eigenvalues λ from PCA are related to the left singular vectors and singular values as $\mathrm{A} = \mathrm{U}^{\mathrm{T}}$ and $\lambda = \mathrm{diag}(\mathrm{S})^2$, respectively. In other words, SNS uses the singular values alongside the left singular vectors of the original data matrix D to create its operator, whereas OTP uses the right singular vectors.

[0040] However, this basic temporal formulation based on right singular vectors actually gives rise to an equivalent spatial operation. A closer look at eq. 5 shows that $V_{\notin j}^{T} = S_{\notin j}^{-1} U^{T} D_{\notin j}$, so combined with eq. 4, it can be seen that

$$\tilde{d}_j = (d_j D_{\notin j}^{T} U_{\notin j} S_{\notin j}^{-1} S_{\notin j}^{-1} U_{\notin j}^{T})\, D_{\notin j} = {}_{K_j} D_{\notin j}$$

The matrix multiplications can thus be collapsed to a single $1 \times (N-1)$ matrix of coefficients ${}_{K_j}$ that are used to multiply $D_{\notin j}$ to obtain $\tilde{d}_j$. These coefficients ${}_{K_j}$ can be trivially inserted row-by-row (leaving zeros on the diagonal) to itself form a spatial operator K_{OTP} where $\tilde{D} = K_{OTP} D$. Thus, this basic time-domain investigation leads to a spatial cross-validation matrix when singular value decomposition (SVD) is used to determine the waveforms. In fact, any linear decomposition method operating on the data matrix D will similarly result in a spatial operator.

[0041] The temporal operator can be improved with an appropriate estimate of the model parameters M. For example, for MEG data, the physical relationship between the sources of interest and sensors can be used to create the regularized multipole moment basis also used by SSS. This can then be used to reduce the matrix as $M_{\notin j} D_{\notin j}$ prior to computing the right singular vectors for d_j projection. This is referred to herein as variant OTP_{SSS}.

[0042] Both SNS and OTP can be formulated to operate on a given data matrix D in a single pass, or by processing data in chunks of n_{seg} samples (with $N < n_{seg} < n_s$). A simple constant overlap-add (COLA) approach can be used to minimize edge artifacts using a Hann window with 50% overlap. This gives rise to a single algorithm parameter, namely the processing window duration. Within each window, the algorithm proceeds according to eq. 4. For numerical reasons, in practice it can be useful to rescale channels to a similar scale before processing, and also demean the data before computing the right singular vectors $V_{\notin j}$.

[0043] When applied to Elekta Neuromag® MEG data (306 channels, 1000 Hz sample rate) this type of algorithm can operate roughly 3 times slower than realtime. Using a non-identity basis M, such as in OTP_{SSS}, can reduce computation time by a factor roughly equivalent to the rank reduction of the data (e.g., 306 channels to 88) would yield roughly a 3x speedup over the identity version). Note that the rank of 88 is just an example related to the initial rank of the full SSS basis, typically 95, followed by regularization to optimize the numerical stability of the basis.

[0044] As noted above, multichannel data can be roughly divided into contributions from the object of interest, interference sources common to the sensor array, and intrinsic sensor noise and artifacts arising from the sensors themselves. While the spatially correlated part of the signal can often be processed with deterministic signal models, the momentary contributions of intrinsic sensor noise and artifacts are unpredictable even though one may have some statistical information about them. In case of oversampling, that is, with a sensor array containing more measurement channels than degrees of freedom in the detectable signal with spatially correlated structure, one can use CV (leave one channel out at a time) techniques for estimating the intrinsic sensor noise and artifact contribution. The methods outlined herein are based on CV with the components of the sensor noise suppression methods derived in the temporal rather than in the spatial domain. The spatial domain approach has been demonstrated before, but the temporal approach has certain advantages that may be useful, e.g., in clinical MEG work. Another way to compensate for the effects of sensor noise and artifacts is to take their contribution into account in regularized projections of the measured data onto signal models. Such methods include, e.g., Wiener filtering, signal space projection, or matrix inversions with incorporated noise covariance estimates. However, sensor-level noise suppression facilitates visual inspection of raw data while suppressing the noise of the associated noise estimates. In addition to introducing the specific framework of sensor noise suppression derived in the temporal domain, the mathematics of CV-based sensor noise suppression are described with sufficient depth to facilitate understanding of the subsequent characteristics of the processed data.

[0045] The benefit of spatially derived CV is that in the process of estimating each channel signal exclusively from all other channels but the channel itself, it is evident that each channel has its own sensor noise and artifact contribution completely removed. The disadvantage of this approach is, however, that it is difficult to control for inter-channel spreading of sensor noise and artifacts in this framework. As a consequence, while decreased in amplitude, the intrinsic sensor noise is redistributed among channels in this operation. Another potential complication of the spatial model is that, depending on the implementation, the subsequent source modeling may require compensation against the spatial bias produced by the spatial sensor noise suppression operator. Testing has established that such compensation cancels the effect of sensor noise suppression in the source space.

[0046] Testing has demonstrated that the temporal domain projection provided by the OTP method is an efficient way to suppress intrinsic sensor noise without causing significant spreading of sensor noise and artifact contributions between channels, which is an important practical benefit. The main consequences are that the users of the technique do not manually have to specify any bad channels to be excluded from the data, the risk of misinterpretation of data is mitigated due to the insignificant spreading of intrinsic noise, and there is no need to compensate for the sensor noise suppression operation in source models. As noted above, certain embodiments of the present technology do assume statistical independence and linear algebraic orthogonality. Whether these assumptions hold true on a case-by-case basis is a statistical question that requires practical evidence and experience. As described below, data has been obtained from seven different MEG sites, recorded at a large range of amplitudes (20 nAm to 500 nAm), and the results consistently show improvements using OTP.

[0047] OTP provides the best combination of simultaneously providing satisfactory sensor level noise suppression and ensuring minimal spatial bias of the signal of interest. Therefore, OTP does not require lead field level compensation, which would leave to compromised suppression of random noise in the source space. Moreover, using a non-identity basis M, such as the SSS-based matrix used here, can yield additional improvements. Additionally, the window duration parameter of OTP provides a tradeoff that can be set based on the data of interest. At the long-window extreme, the suppression factors are the largest. At the short-window extreme, the OTP technique adapts well based on time-varying sensor structure and quality, better compensating for non-stationarity in the data. In practice for MEG data, 10 second windows appear to provide the best compromise.

[0048] The simplest formulation of the OTP operator can be expressed as a matrix multiplication of an N x N-dimensional spatial matrix and the original data if the temporal subspace of interest is derived by a simple SVD algorithm. Under this formulation, OTP could also result in some inter-channel artifact spreading in some particularly difficult cases, even if it is designed to be minimized by the right singular vector computation and subsequent projection. However, OTP as a general framework provides opportunities for extension to address these sorts of issues. For example, within a given window, the uncorrelated sensor noise and artifact patterns can be estimated, yielding an N x n_s-dimensional "random noise" matrix. A small number of dominating waveforms of this matrix may then be selected and projected out from the original data. The most significant channel artifacts are thus suppressed

from the data by an overall temporal domain projection, *i.e.*, by applying the same projection, which contains the worst channel artifacts, to all channels at once. As a consequence, the artifact spreading introduced by the aforementioned spatial matrix multiplication, which is already designed to be minimized by OTP, is even less significant with the data that have been pre-cleaned with respect to the worst artifacts.

[0049] Suppression of sensor noise also opens up the opportunity for applications in areas where the signal of interest is known to be low with respect to the intrinsic sensor noise. An example of such an application area is the analysis of high-frequency oscillations (HFO) in MEG. The HFO signals may be higher in frequency than most of the environmental interference, rendering the intrinsic sensor noise the dominating noise contribution in the HFO frequency band. This topic is especially relevant in epilepsy where the HFO contribution is believed to convey very important information about the seizure onset area and while HFO is visible in intracranial recordings, it is difficult to detect in a non-invasive fashion with present SNR levels.

Test Cases

[0050] Several illustrative cases are provided below to demonstrate the usefulness of the noise-suppression techniques enabled by the present technology. Performance of OTP against SNS was tested using the PCA-based formulation described above. Processed data was also tested with the STAR algorithm (using the software package NoiseTools, version 18-Nov-2016), which applies a statistical thresholding procedure to correct artifacts timepoint-by-timepoint using an SNS-like procedure.

[0051] For illustrative purposes, signal space separation (SSS) as implemented in MNE was also used, which acts as a linear spatial operator to suppress external interference artifacts. Note that SSS is not a sensor-noise removing algorithm, but is rather designed to suppress environmental artifacts. It is shown here only as an example of preprocessing spatial operators that could be applied to data, and without loss of generality.

1. Phantom and empty-room data

[0052] Phantom MEG data were recorded at six sites equipped with an Elekta Neuromag® or Triux®MEG system: the University of Washington (UW) at 500 nAm, Massachusetts General Hospital (MGH) at 500 nAm, the Epilepsy Center at the Cleveland

Clinic Neurological Institute at 20 and 200 nAm, Minamata at 500 nAm, BioMag Laboratory at Helsinki University Central Hospital at 500 nAm, and at the Otaniemi site of Aalto University at 20 and 100 nAm. The data from Cleveland Clinic were obtained as part of the informative and publicly available BrainStorm tutorials. Elekta Neuromag® MEG systems contain 306 MEG sensors, consisting of 102 magnetometers and 204 planar gradiometers (in orthogonal pairs). Phantom data were acquired with an MEG phantom with 32 controllable dipoles at fixed, known locations and orientations. All 32 dipoles were used for UW, Cleveland, BioMag, and Minamata, with 27 used for MGH and 28 for Otaniemi based on reliable localization. Only the 200 nAm Cleveland data are shown here because the 20 nAm data did not produce reliable single-trial localization. For each recording, before placing the phantom in the MEG room, head position indicator (HPI) coils were attached, and the HPI locations and cardinal landmarks (nasion, left/right periauriculars) were digitized with a 3Space Fastrak (Polhemus). The phantom was placed inside the MEG helmet and coregistered to the MEG sensors using the HPI coils. The phantom then repeated brief two-cycle sinusoidal pulses from each dipole.

[0053] Data were also recorded at Boston Children's Hospital (BCH) on the Tristan Systems BabyMEG system with 270 magnetometers, 105 compensation magnetometers, 9 reference magnetometers, and a custom phantom with 6 usable dipoles at 200 nAm. A separate set of over 300 trials of 20 nAm were used to empirically determine veridical dipole locations in head coordinates for use with the 200 nAm data. The device-to-head coordinate frame transformation for the BabyMEG system was determined by doing a nonlinear quaternion fit of a transformation matrix between the veridical dipole locations and the empricial locations also using many trials (over 300 per dipole).

[0054] To emulate a strong jump artifact, a 4000 fT/cm amplitude 15 ms rectangular jump artifact was added to a single MEG gradiometer (2422) which exhibited a strong response (based on visual inspection) to the first epoch of the first phantom dipole in the UW phantom recording. This data had a high SNR (25 dB) 20 Hz sinusoid power on gradiometers compared to baseline. This artifact was present only for visual demonstration purposes (Figure 3), and was not present during source localization.

[0055] Phantom data were processed with OTP, STAR, and/or SNS. STAR processing used NoiseTools version "18-Nov-2016" and default options. SNS processing used the version implemented in the MNE sandbox (https://github.com/mne-tools/mne-sandbox; commit

1d050dc) as it consistently outperformed the NoiseTools implementation for these data. For each recording, data were processed by these techniques without excluding bad channels, except for BCH data, where 247 of the 270 available magnetometers were used (and all 114 reference magnetometers were ignored).

[0056] Phantom dipoles were localized for these processed (and raw) data by using a fairly simple processing pipeline: 1) zero-phase FIR notch filtering at line frequencies (50 or 60 Hz plus harmonics) if necessary (plus band-pass filtering between 2 and 50 Hz for BCH); 2) excluding bad MEG channels based on visual inspection for Aalto (2233, 1211, 0511, 2142, 2422, 2231), BioMag (1013), UW (0441, 1913, 2611), Cleveland (1423, 1522, 2413, 2421), MGH (none), Minamata (none), and BCH (150); 3) epoching the data at the onset of each dipole activation; 4) throwing out epochs that exceeded lenient thresholds (10000 fT/cm for gradiometers, 40000 fT for magnetometers); 5) averaging across epochs; 6) computing sensor covariances using the baseline period; 7) regularizing covariances to deal with rank deficiency (e.g. SNS at low neighbor counts); 7) performing a single dipole fit at the peak of the phantom dipole activation using a spherical conductor model with center at the head origin.

[0057] Note that this pipeline involves very little effective environmental denoising; only the filtering and whitening via covariance during dipole fitting reject some of the environmental noise. So the dipole fits recorded here for each site should not be considered an optimal representation of data quality for any given site.

[0058] For SNS, optional compensation of dipole fits (SNS_{COMP}) was performed by multiplying the lead fields during fitting by the same spatial operator that was used to denoise the channels. Although in principle SNS and OTP are designed to correct the bad channels above, here they are omitted so that the original, SNS, and OTP localization procedures operate on equivalent datasets. For each site, the average across 10 single-dipole fits for each dipole was used to assess the localization bias.

[0059] Empty-room data were acquired with no subject present in the MEG room for a duration of approximately 121 seconds at UW. After the recording, data were visually inspected and six MEG channels (0142, 0413, 0442, 1442, 1722, and 2643) were identified as having clear channel-specific artifacts. This is more than is typical in an MEG recording, but this recording was selected to demonstrate various kinds of MEG channel artifacts.

[0060] Single-subject MEG data during an audio-visual experiment were also analyzed from the publicly available MNE-Python "multimodal" dataset, which was recorded at Otaniemi using left and right auditory stimuli, somatotopic stimulation, and visual upper/lower left/right stimuli.

2. Synthetic data

[0061] Independent identically distributed (IID) zero-mean, unit standard deviation Gaussian sensor noise data were generated (using a standard Mersenne Twister from NumPy) for each of the 306 MEG channels for 60 seconds of data. The temporally white signals were then also processed them with an autoregressive IIR filter (numerator [1], denominator with coefficients [1, —1, 0.2]) to approximate temporally pink noise. Shot-like noise was simulated by scaling the white noise by 0.001 and randomly adding unit impulses to channels (at most one per channel at any given time) at an approximate rate of 1 Hz per channel.

[0062] Suppression levels for OTP and SNS in decibels (dB) were then calculated as

$$20\log_{10}\left(\frac{\|\mathbf{D}\|}{\|\tilde{\mathbf{D}}\|}\right). \qquad (6)$$

[0063] The suppression factors were calculated in source space by using a simplified dipole fitting routine to estimate "neural currents" (in nAm) in source space. To estimate these, for each of the 32 Neuromag phantom dipole locations and orientations (which have a reasonable spatial sampling), the inner product between the lead field vector for that dipole and the sensor space data was taken. This is mathematically equivalent to standard fixed-position, fixed-orientation single equivalent-current dipole (ECD) amplitude estimation, given that the sensor noise covariance is a diagonal with equivalent values per channel. Suppression could then be measured in this 32-dipole source space by using the same metric as in sensor space. By estimating the neural currents for this simulated data, which by construction contains only IID sensor noise, it can be determined how well such sensor noise would be suppressed during the source estimation process if there were actually neural signals present in the data.

3. Results

[0064] First, a condition was simulated that could be problematic especially with the interpretation of clinical MEG recordings: a single channel contains a temporal artifact (e.g., from flux trapping) that spreads to other channels when processed using spatial methods. A

brief 15 ms rectangular pulse was added into a single gradiometer channel of a recording of an MEG phantom, which repeatedly activates a dipole, with results shown in Figure 3. The top row shows phantom dipole activation (2-cycles of 20 Hz sinusoid) across channels combined with a simulated channel artifact (square pulse) on a single gradiometer channel of a similar magnitude to the phantom activation. Spatial methods such as Maxwell filtering (SSS; second row) and SNS (third row; $N_{neighbors}$ = 305) spread the artifact across channels, conferring a spatial profile to the artifact, whereas STAR (fourth row) and OTP (bottom row; 30 sec duration) remove the artifact from the data without compromising the other channels. Although SSS is not designed to remove individual sensor noise, it is a representative example of a spatial filter commonly used on MEG data to suppress environmental noise.

[0065] Next, the examined data was recorded with no subject present, which is often called "empty-room" data. This recording contained a handful of channels with clearly visible sensor-noise artifacts. In each case, the visible temporal artifacts in the data were greatly suppressed by both OTP and SNS (as shown in Figure 4), while only a subset were suppressed by STAR. Figure 4, for example, illustrates how uncorrelated sensor noise artifacts are suppressed by OTP and SNS. Temporal artifacts that are restricted to individual channels (top) are partially removed by STAR (second row) and effectively suppressed by SNS (third row; $N_{neighbors}$ = 305) and OTP (fourth row; 30 second duration).

[0066] In addition to correcting artifacts, OTP can also be used to improve visual inspection of evoked data. An example of 10 seconds of raw data from a single-subject, multi-sensory MEG experiment is shown in Figure 5. As shown in Figure 5, data from an example multi-sensory experiment show examples of how OTP can improve single-trial evoked response visibility, with trial types auditory left (AL), visual lower left (VLL), visual upper right (VUR), visual upper left (VUL), and somatotopic stimulation right (SR), before (top) and after (bottom) denoising with OTP. This optimal window-size tradeoff might depend on the given sensor configuration and sensor noise characteristics, and thus could be domain or system dependent. The sensor denoising makes single-trial responses (e.g., to somatosensory stimulation) more prominent when browsing raw data.

[0067] Using simulated data, the suppression of idealized sensor noise for data both in sensor space, and in source space (i.e., during source localization) was quantified. Independent noise was generated for 306 signals, analogous to the 306 sensors in the Elekta Neuromag® MEG acquisition system. Next, signals with white (flat) and approximately pink ($1/f$ frequency

falloff) temporal spectra were used, as well as data with shot noise processes. SNS and OTP were applied with varying $N_{neighbors}$ and duration, respectively. In sensor space, OTP and SNS have matching suppression levels when OTP uses the entire duration (here, 60 seconds) and SNS uses the maximum $N_{neighbors} = 305$. For example, Figure 6 illustrates how OTP suppresses simulated independent noise in sensor and source space. As shown, OTP using an identity matrix (first column) shows slightly better performance than SNS applied in an overlap-add scheme ($N_{neighbors} = 305$; second column) at short durations, but otherwise has similar performance in both sensor (top row) and source space (bottom row) for three noise types: shot noise (solid lines), white noise (dashed lines), and pink noise (dotted lines). A reduced-rank SSS basis (third column) shows slightly better suppression. Although STAR followed by standard SNS (fourth column) has good noise suppression in sensor space, applying compensation in source space eliminates the noise reduction for pink and white noise, and reduces it for shot noise.

[0068] The suppression factor of SNS drops effectively to zero for white and pink noise in source space during localization when the lead fields are properly compensated, whereas OTP suppression remains at the same levels as in sensor space. However, when SNS is used to compute coefficients for overlap-add operation ($SNS_{305,OLA}$), suppression is similar to that of OTP, and is maintained in source space. OTP can also be modified to use a rank-reduced basis for projection, termed OTPsss, to obtain modest gains in suppression factors in sensor and source space.

[0069] The tests also examined the effect of OTP on source localization using experimental data from MEG phantoms from seven sites by measuring single-trial equivalent-current dipole (ECD) localization error. For example, Figure 7 illustrates how OTP reduces localization bias in recorded data. As shown, localization bias (y-axis, log scale) averaged across 10 single-trial localizations of phantom dipoles for five different sites (with 20 nAm and 200 nAm recordings from Otaniemi), with unprocessed data localization shown for each site (dashed black lines). Two OTP variants are shown using an identity matrix (left) and reduced-basis (right; SSS), with minor variations in performance across durations (dots: 1, 2, 5, 10, 30, and 60 seconds). The improvement for OTP also changes as a function of window duration, with an optimal value generally around 10 seconds. STAR alone had no discernable effect on localization. STAR followed by SNS and SNS alone performed very similarly, so only SNS is shown ($N_{neighbors}$: 1, 2, 5, 10, 20, 50, 100, and 246 [BCH] or 305 [all others] as sequential dots).

SNS usually biases localization, but lead-field compensated SNS processing approximately restores bias to the level of unprocessed data. When applied in an overlap-add windowed processing scheme, SNS approaches standard OTP performance for some but not all sites (e.g., not Otaniemi at 20 or 100 nAm). Overall, the SSS variant of OTP shows the best performance.

[0070] OTP, which required no compensation of the lead fields during fitting, typically reduced the localization bias relative to unprocessed data (especially for the 10 second duration), with long durations sometimes leading to a slight bias increase (e.g., 60 seconds for BCH). STAR had no discernable effect on localization for any site. SNS, if no lead field compensation was applied, generally increased the localization bias, particularly when $N_{neighbors}$ was small (in some cases increasing errors by an order of magnitude). When the operator was applied to the lead field during fitting, SNS typically converged to the same localization error as the unprocessed data. One interesting exception is the Cleveland data, where the compensated SNS showed a reduction in localization bias, although it still had larger bias than OTP; for this data in the $N_{neighbors} = 20$ condition, the SNS operator matrix discarded 16.6% (51 of 306) of the channels. When SNS was applied in an overlap-add scheme, localization was usually similar to OTP, but worse for some sites (e.g., Otaniemi). The reduced-rank version OTPsss had the lowest localization bias overall.

[0071] As explained above, and as demonstrated by the above test cases, sensor noise in multichannel arrays that provide spatial oversampling can be effectively suppressed using OTP, a leave-one-out cross-validation method with temporal projection using overlapping windows. The relevant mathematical properties of sensor noise suppression operators have been derived in both sensor space and during source localization. Finally, illustrative studies have demonstrated that OTP is expected to provide superior performance compared to existing methods, especially during localization of the generators that give rise to the sensor-level signals.

Examples

1. A multi-channel sensor system, comprising:

a multi-channel sensor configured to obtain spatially oversampled data; and

a computing device communicatively coupled to the multi-channel sensor, the computing device configured to—

(a) receive the spatially oversampled data from the multi-channel sensor;

(b) build a spatial model from the data for essential spatial degrees of freedom;

(c) decompose the data into the underlying spatial model to obtain associated amplitude components containing a mixture of original temporal waveforms of the data;

(d) for each channel of the multi-channel sensor, estimate time-domain amplitude components using cross-validation;

(e) for each channel, based on the estimated time-domain amplitude components, identify sensor noise and/or artifacts for that channel; and

(f) for each channel, suppress the identified sensor noise and/or artifacts from the data.

2. The system of example 1 wherein the multi-channel sensor comprises at least one of: a magnetoencephalographic (MEG) sensor array or an electroencephalographic (EEG) sensor array.

3. The system of any one of examples 1–2 wherein the computing device is configured to decompose the data into the underlying spatial model by decomposing the data in two situations for each channel of the multi-channel sensor device, the two situations including:

a first situation in which the amplitude components are estimated when applying the spatial model to all channels; and

a second situation in which the amplitude components are estimated when applying the spatial model with the channel under investigation excluded from the data and the spatial model; and

wherein the computing device is further configured to estimate the time-domain amplitude components using cross-validation by comparing, for each channel of the multi-channel sensor device, results of the first situation to results of the second situation in the time domain wherein, for each channel, the difference between the results of the first situation and the results of the second situation indicate the sensor noise and/or artifacts for that channel.

4. The system of any one of examples 1–3 wherein the computing device is configured to estimate the time-domain amplitude components using cross-validation by

applying the spatial model with the channel under investigation excluded from the data and the spatial model, wherein the difference in the temporal domain between the decomposition and the original data for the channel under investigation indicates the sensor noise and/artifacts of the channel under investigation, and

wherein the computing device is further configured to suppress the identified sensor noise and/or artifacts from the data by finding the temporal waveform common to the channel under investigation and the decomposition.

5. The system of any one of examples 1–4 wherein the data spans a first length of time, and wherein the computing device is further configured to:

divide the data into a plurality of time segments that each span lengths of time shorter than the first length of time; and

perform steps (b)–(f) separately for each time segment.

6. The system of example 5 wherein at least some of the time segments are overlapping in time.

7. A method, comprising:

(a) receiving spatially oversampled multi-channel sensor data from a region of interest;

(b) building a spatial model from the data for essential spatial degrees of freedom;

(c) decomposing the data into the underlying spatial model to obtain associated amplitude components containing a mixture of original temporal waveforms of the data; and

(d) for each channel of the multi-channel sensor,

estimating time-domain amplitude components using cross-validation;

based on the estimated time-domain amplitude components, identifying sensor noise and/or artifacts for that channel; and

suppressing the identified sensor noise and/or artifacts from the data.

8. The method of example 7 wherein decomposing the data into the underlying spatial model comprises decomposing the data in two situations for each channel of the multi-channel sensor device, the two situations including:

a first situation in which the amplitude components are estimated when applying the spatial model to all channels; and

a second situation in which the amplitude components are estimated when applying the spatial model with the channel under investigation excluded from the data and the spatial model;

wherein estimating the time-domain amplitude components using cross-validation comprises comparing, for each channel of the multi-channel sensor device, results of the first situation to results of the second situation in the time domain wherein, for each channel, the difference between the results of the first situation and the results of the second situation indicate the sensor noise and/or artifacts for that channel.

9. The method of example 8 wherein comparing, for each channel of the multichannel device, results of the first situation to results of the second situation in the time domain comprises implementing an orthogonal projection operation.

10. The method of example 8 wherein comparing, for each channel of the multichannel device, results of the first situation to results of the second situation in the time domain comprises implementing a singular value decomposition of the temporal domain signals of the amplitude coefficients corresponding to the spatial model to find a temporal basis.

11. The method of any one of examples 7–10 wherein estimating the time-domain amplitude components using cross-validation comprises applying the spatial model with the channel under investigation excluded from the data and the spatial model, wherein the difference in the temporal domain between the decomposition and the original data for the channel under investigation indicates the sensor noise and/artifacts of the channel under investigation, and

wherein suppressing the identified sensor noise and/or artifacts from the data comprises finding the temporal waveform common to the channel under investigation and the decomposition.

12. The method of example 11 wherein suppressing, for each channel of the multichannel device, the sensor noise and/or artifacts by finding the temporal waveform common to the channel under investigation and the decomposition comprises projecting the original channel waveform to a temporal basis spanned by the decomposed amplitude coefficients by a parallel projection operator.

13. The method of example 12 wherein projecting the original channel waveform to a temporal basis spanned by the decomposed amplitude coefficients by a parallel projection operator comprises implementing a singular value decomposition of the temporal domain signals of the amplitude coefficients corresponding to the spatial model to find the temporal basis.

14. The method of any one of examples 7–13 wherein suppressing, for each channel of the multichannel device, the sensor noise and/or artifacts from the signal of interest comprises implementing an orthogonal projection operator.

15. The method of any one of examples 7–14 wherein building the spatial model for the essential spatial degrees of freedom comprises implementing a combination of a basic solution of Laplace's equation for the harmonic scalar potential determining the magnetic field in the measurement area to calculate a spatial basis.

16. The method of any one of examples 7–15 wherein building the spatial model for the essential spatial degrees of freedom comprises using dominating spatial patterns found in a statistical analysis of the original data to calculate a spatial basis.

17. The method of example 16 wherein using dominating spatial patterns found in a statistical analysis of the original data to calculate a spatial basis comprises using dominating spatial patterns found in a principle component analysis.

18. The method of any one of examples 7–17 wherein the data spans a first length of time, the method further comprising:

dividing the data into a plurality of time segments that each span lengths of time shorter than the first length of time; and

performing steps (b)–(d) separately for each time segment.

19. The method of example 18, wherein at least some of the time segments are overlapping in time.

20. A computer-readable medium storing instructions that, when executed by one or more processors, cause the processors to perform operations comprising:

(a) receiving spatially oversampled multi-channel sensor data from a region of interest;

(b) building a spatial model from the data for essential spatial degrees of freedom;

(c) decomposing the data into the underlying spatial model to obtain associated amplitude components containing a mixture of original temporal waveforms of the data;

(d) for each channel of the multi-channel sensor, estimating time-domain amplitude components using cross-validation;

(e) for each channel, based on the estimated time-domain amplitude components, identifying sensor noise and/or artifacts for that channel; and

(f) for each channel, suppressing the identified sensor noise and/or artifacts from the data.

21. The computer-readable medium of example 20 wherein decomposing the data into the underlying spatial model comprises decomposing the data in two situations for each channel of the multi-channel sensor device, the two situations including:

a first situation in which the amplitude components are estimated when applying the spatial model to all channels; and

a second situation where the amplitude components are estimated when applying the spatial model with the channel under investigation excluded from the data and the spatial model;

wherein estimating the time-domain amplitude components using cross-validation comprises comparing, for each channel of the multi-channel sensor device, results of the first situation to results of the second situation in the time domain wherein, for each channel, the difference between the results of the first

situation and the results of the second situation indicate the sensor noise and/or artifacts for that channel.

22. The computer-readable medium of example 21 wherein comparing, for each channel of the multichannel device, results of the first situation to results of the second situation in the time domain comprises implementing an orthogonal projection operation.

23. The computer-readable medium of any one of examples 20–22 wherein estimating the time-domain amplitude components using cross-validation comprises applying the spatial model with the channel under investigation excluded from the data and the spatial model, wherein the difference in the temporal domain between the decomposition and the original data for the channel under investigation indicates the sensor noise and/artifacts of the channel under investigation, and

wherein suppressing the identified sensor noise and/or artifacts from the data comprises finding the temporal waveform common to the channel under investigation and the decomposition.

24. The computer-readable medium of example 23 wherein suppressing, for each channel of the multichannel device, the sensor noise and/or artifacts by finding the temporal waveform common to the channel under investigation and the decomposition comprises projecting the original channel waveform to a temporal basis spanned by the decomposed amplitude coefficients by a parallel projection operator.

25. The computer-readable medium of any one of examples 20–24 wherein the data spans a first length of time, and wherein the operations further comprise:

dividing the data into a plurality of time segments that each span lengths of time shorter than the first length of time; and

performing steps (b)–(f) separately for each time segment.

26. The computer-readable medium of example 25 wherein at least some of the plurality of segments are overlapping in time.

[0072] The above detailed descriptions of embodiments of the technology are not intended to be exhaustive or to limit the technology to the precise form disclosed above. Although specific embodiments of, and examples for, the technology are described above for illustrative purposes, various equivalent modifications are possible within the scope of the technology, as those skilled in the relevant art will recognize. For example, while steps are presented in a given order, alternative embodiments may perform steps in a different order. Moreover, in some embodiments, the technology can be used, for example, to reduce noise in data other than electrical or magnetic signals, for example economic or other data. The various embodiments described herein may also be combined to provide further embodiments.

[0073] Moreover, unless the word "or" is expressly limited to mean only a single item exclusive from the other items in reference to a list of two or more items, then the use of "or" in such a list is to be interpreted as including (a) any single item in the list, (b) all of the items in the list, or (c) any combination of the items in the list. Where the context permits, singular or plural terms may also include the plural or singular term, respectively. Additionally, the term "comprising" is used throughout to mean including at least the recited feature(s) such that any greater number of the same feature and/or additional types of other features are not precluded. It will also be appreciated that specific embodiments have been described herein for purposes of illustration, but that various modifications may be made without deviating from the technology. Further, while advantages associated with certain embodiments of the technology have been described in the context of those embodiments, other embodiments may also exhibit such advantages, and not all embodiments need necessarily exhibit such advantages to fall within the scope of the technology. Accordingly, the disclosure and associated technology can encompass other embodiments not expressly shown or described herein.

APPENDIX A
SPATIAL SENSOR NOISE SUPPRESSION

[0074] Under the assumption of spatial oversampling, the contribution of E_u can be suppressed by cross-validation. A spatial CV operator K can be constructed to compute denoised data matrix $\tilde{D}$ as

$$\tilde{D} = KD \tag{7}$$

In this CV formulation, K is an $N \times N$ matrix with zeros on the diagonal. Because this operator reconstructs each channel d_j using only a weighted combination of the other channels $d_{j' \neq j}$, it eliminates the contribution of individual noise of d_j from the denoised $\tilde{d}_j$. Here, the d-vectors consist of the sampled signals corresponding to the channel under investigation and are thus n_s-dimensional.

[0075] However, this benefit comes at the expense of also contaminating $\tilde{d}_j$ with the sensor noise from $d_{j' \neq j}$. It is thus critical to minimize spreading of artifacts and noise through appropriate weights in the K matrix off-diagonals. Therefore, the theoretical effects of such operators on this data must be examined.

[0076] 1) *Operator properties*: In principle, any $N \times N$ -dimensional matrix K with a zero diagonal could be a CV matrix, providing a cross-validated estimate of the data

$$\tilde{D} = KD = (KM)\,X + KE_u \tag{8}$$

with $K(j, j) = 0 \ \forall j$. Any such operator will minimize the diagonal elements of the initially diagonal noise covariance, as the covariance $C_u = E[E_u E_u^T]$ will be transformed to

$$\tilde{C}_u = E\left[KE_u E_u^T K^T\right] = KC_u K^T. \tag{9}$$

It is not guaranteed, however, that any K would result in small off-diagonal elements in covariance $\tilde{C}_u$, which needs to be taken into account in the design of the operator K. The trivial example $K = 0$ satisfies the condition $K(j, j) = 0$ and minimizes the off-diagonal elements of $\tilde{C}_u$, but it would lead to $KM = 0$, and, consequently, all signal of interest would be lost as well.

[0077] A good sensor noise suppression operator K should:

[0078] 1) Minimize spatial bias of the signal of interest.

[0079] 2) Minimize spatial spreading of channel artifacts.

[0080] 3) Maximize the suppression of sensor noise (ratio of diagonal elements of C_u and $\tilde{C}_u$).

[0081] Comparing eqs. 3 and 8, the requirement of no spatial bias becomes KMX = MX, or (K — I) M where I is the N x N identity matrix. For no spatial bias to occur, matrix K — I thus has to belong to the left null space of the data-dependent matrix MX, which is spanned by the rows of U^T that correspond to zero-valued singular values of the SVD decomposition MX = $U\Sigma V^T$. See Appendix A for a formulation based on this principle.

[0082] If one uses a matrix K that distorts the spatial profile of the signal of interest, then for many subsequent analyses this must be taken into account. For example, in MEG the aim is often to estimate the underlying neural current distribution y. The estimation process requires that the N-dimensional sensor-level signal distribution b produced by a set of n_q neural current elements be calculated, each charaterized with an N-dimensional lead field l_k. Given an n_q-dimensional amplitude distribution y, the model is b = Ly, where $L = [l_1, l_2, ..., l_{n_q}]$ is the N x n_q-dimensional model matrix. By fitting this model to the acquired data, the amplitudes are estimated as $\hat{y} = L^{\dagger}D$, which is associated with a source-level noise covariance

$$C_{uy} = L^{\dagger}\tilde{C}_u L^{\dagger T}. \quad (10)$$

[0083] In case of a non-biasing sensor noise suppression operator K_M, the source-space noise covariance is

$$\tilde{C}_{uy} = L^{\dagger}\tilde{C}_u L^{\dagger T}, \quad (11)$$

Where $\tilde{C}_u = KMC_uK_M^T$, and the noise suppression performance is preserved in the source space. If spatial bias is expected, however, one has to compensate for the bias of K at the lead field level during localization to ensure that the model matches the data as

$$KD = KLy. \quad (12)$$

[0084] This leads to $\hat{y} = (KL)^{\dagger}KD$, corresponding to source space noise covariance

$$\begin{aligned}\tilde{C}_{uy} &= (KL)^{\dagger}\tilde{C}_{u}(KL)^{\dagger T} \\ &= (KL)^{\dagger}KC_{u}K^{T}(KL)^{\dagger T} \\ &= L^{\dagger}C_{u}L^{\dagger T} \\ &= C_{uy},\end{aligned} \tag{13}$$

where it is assumed that K and L are full-rank matrices ($(KL)^{\dagger} = L^{\dagger} K^{\dagger}$). Thus, the sensor noise suppression performance is lost in the source space in case lead field compensation is required.

[0085] In prior work, multiple potential formulations for spatial operators were provided, e.g., using a weighted average of sensors based on physical proximity. However, the most general formulation uses the temporal correlations between channels to construct K_{SNS}. Specifically, assuming that D has zero sample mean, to get the non-*j* coefficients k_j for row *j* of K_{SNS} (where the subscript $\notin j$ indicate the exclusion of channel *j*):

[0086] 1) Compute the PCA of $D_{\notin j}$ (eigenvalues λ and eigenvectors A of the covariance matrix $C_{\notin j}$).

[0087] 2) Compute weighted eigenvectors $A_w = \mathrm{diag}\left(\sqrt{\lambda}\right)A$ (where the square root is taken element-wise).

[0088] 3) Project j^{th} row (omitting column j) of the full covariance C onto this subspace to obtain $k_j = A_w c_j A_w$.

[0089] This formulation can be modified to use only a subset of channels $1 < N_{neighbors} < N$ by subselecting $N_{neighbors}$ rows and columns of C according to some criteria, e.g., the *j*:th row of the correlation matrix C_N, obtained by dividing the rows and columns of C by the square root of the diagonal elements of C.

APPENDIX B
CONTROLLING BIAS

[0090] In the spatial-domain formulation, properly controlling of the spatial bias requires access to MX. This is facilitated if the data-independent matrix M or any other matrix that spans the same signal subspace as M is known because then, under the assumption of spatial oversampling, eq. 2 indicates that cross-validation for any channel j can be expressed as

$$\tilde{\mathrm{d}}_j^T = \mathrm{m}_j^T \mathrm{M}_{\notin j}^{\dagger} \mathrm{D}_{\notin j} = \mathrm{d}_0^T + \mathrm{m}_j^T \mathrm{M}_{-j}^{\dagger} \mathrm{E}_{u,\notin j}, \qquad (14)$$

where m_j^T is the j:th row in matrix M, subscript $\notin j$ indicates exclusion of channel j, and $\dagger$ means pseudoinverse. Now, the j:th row of matrix K can be constructed by setting the values of the 1 x (N — 1) -dimensional row vector $\mathrm{m}_j^T \mathrm{M}_{-j}^{\dagger}$ on off-diagonal entries of the j:th row of K. To emphasize that information about matrix M was used in the derivation of the CV matrix, the introduce subscript M is introduced as follows:

$$\tilde{\mathrm{D}} = \mathrm{K}_{\mathrm{M}} \mathrm{D} = \mathrm{D}_0 + \mathrm{K}_{\mathrm{M}} \mathrm{E}_u \qquad (15)$$

[0091] In this formulation, no spatial bias is introduced, as demonstrated by eq. 14, but the random noise contribution changes from E_u to $\mathrm{K}_{\mathrm{M}}\mathrm{E}_u$. Note that K_{M} has zeros at least on the diagonal, so each channel's own noise contribution is completely eliminated from the channel itself.

[0092] According to eq. 9, the noise covariance is now

$$\tilde{\mathrm{C}}_{u,\mathrm{M}} = \mathrm{E}\left[\mathrm{K}_{\mathrm{M}} \mathrm{E}_u \mathrm{E}_u^T \mathrm{K}\right] = \mathrm{K}_{\mathrm{M}} \mathrm{C} \mathrm{K}_{\mathrm{M}}^{\mathrm{T}}. \qquad (16)$$

[0093] It is to be noted that in such a purely spatially driven sensor noise suppression method, the inter-channel artifact spreading, manifested by non-zero off diagonal elements in $\tilde{\mathrm{C}}_{u,\mathrm{M}}$ depends on the physical and geometrical condition of the measurement arrangement. For example, in MEG the K_{M} would change with changing position of the subject's head with respect to the sensors, leading to sensor noise suppression performance dependent on the head position.

APPENDIX C
TEMPORAL ORTHOGONALITY AND SPATIAL RANK -REDUCTION

[0094] The temporal information is preserved in the estimate of the model parameters:

$$\hat{\mathrm{X}} = \mathrm{M}^{\dagger}\mathrm{D} = \mathrm{X} + \mathrm{M}^{\dagger}\mathrm{E}_{u}, \tag{17}$$

where the last term $\mathrm{M}^{\dagger}\mathrm{E}_{u}$ represents the projections of the 1 x n_s-dimensional random noise vectors of different channels onto an n-dimensional subspace. Now, let us consider a pair of random vectors v_1 and v_2 with elements that are statistically independent between the two vectors. Statistical independence means that at the statistical limit information about vector v_1 will not yield any information that could be used to infer properties of vector v_2, and vice versa. Therefore, not only the amplitude but also direction of these two vectors are uncorrelated at the statistical limit. That is, the larger the number of elements n_s is in the two vectors, the closer the vectors are to orthogonality, and so:

$$\lim_{n_s \to \infty} \left(\mathrm{v}_1 \mathrm{v}_2^{T} \right) = 0. \tag{18}$$

[0095] Thus, assuming a large number of samples, for individual noise vectors corresponding to channels j and k, $\in_{uj} \in_{uk}^{T} = 0$ and therefore $\in_{uj} \mathrm{E}_{u,\notin j}^{T} = 0$. Here the j:th row of E_{u} is denoted as $\in_{j}$. Comparing this with the projected space yields $\mathrm{M}_{\notin j}^{\dagger} \mathrm{E}_{,\notin j} \in_{uk}^{T} = 0$, which means that the original noise contribution of channel j is orthogonal to the noise contributions of other channels also in the projected space. Cross validation can be performed in the temporal space by estimating the waveforms from $N-1$ channels:

$$\hat{\mathrm{X}}_{\notin j} = \mathrm{M}_{\notin j}^{T} \mathrm{D}_{\notin j} = \mathrm{X} + \mathrm{M}_{\notin j}^{\dagger} \mathrm{E}_{\notin j}, \tag{19}$$

assuming spatial oversampling. By performing a temporal domain analysis on $\hat{\mathrm{X}}_{\notin j}$ that emphasizes the common waveforms, rows of X, one can find an $n_s \times n_t$ -dimensional ($n_t \leq n$) orthonormal temporal subspace $\mathrm{V}_{\notin j}$ that spans X but has no contribution of $\in_{uj}$. Then

$$\mathrm{d}_{j} = \mathrm{d}_{j} \mathrm{V}_{\notin j} \mathrm{V}_{\notin j}^{T} = \left(\mathrm{d}_{0j} + \in_{uj} \right) \mathrm{V}_{\notin j} \mathrm{V}_{\notin j}^{T}, \tag{20}$$

where $\in_{uj} V_{\notin j} V_{\notin j}^T$ represents the parallel projection of the individual noise vector of channel j onto the temporal subspace derived from all other channels. With large n_s, the subspace $V_{\notin j} V_{\notin j}^T$ and vector $\in_{uj}$ are very close to orthogonality due to statistical independence and therefore $\in_{uj} V_{\notin j} V_{\notin j}^T \approx 0$. On the other hand, by determining the temporal subspace $V_{\notin j}$ in a way that emphasizes the common waveforms, X, over the random vectors in $E_{u,\notin j}$, the angle between $d_{0j} = m_j X$ and $V_{\notin j}^T$ is small. Thus, $d_{0j} V_{\notin j} V_{\notin j}^T \approx d_{0j}$ and eq. 20 becomes

$$d_j \approx d_{0j}. \tag{21}$$

[0096] If the statistical assumptions hold true, the temporal cross-validation method introduces minimal spatial bias and maximal sensor noise suppression performance. Thus, the techniques disclosed herein do away with the inherent problems of spatial bias and inter-channel artifact spreading of the spatial models. Now the limitations are posed by statistical properties of the data rather than inherent complications of spatial-domain operations.

CLAIMS

I/We claim:

1. A multi-channel sensor system, comprising:

a multi-channel sensor configured to obtain spatially oversampled data; and

a computing device communicatively coupled to the multi-channel sensor, the computing device configured to—

(a) receive the spatially oversampled data from the multi-channel sensor;

(b) build a spatial model from the data for essential spatial degrees of freedom;

(c) decompose the data into the underlying spatial model to obtain associated amplitude components containing a mixture of original temporal waveforms of the data;

(d) for each channel of the multi-channel sensor, estimate time-domain amplitude components using cross-validation;

(e) for each channel, based on the estimated time-domain amplitude components, identify sensor noise and/or artifacts for that channel; and

(f) for each channel, suppress the identified sensor noise and/or artifacts from the data.

2. The system of claim 1 wherein the multi-channel sensor comprises at least one of: a magnetoencephalographic (MEG) sensor array or an electroencephalographic (EEG) sensor array.

3. The system of claim 1 wherein the computing device is configured to decompose the data into the underlying spatial model by decomposing the data in two situations for each channel of the multi-channel sensor device, the two situations including:

a first situation in which the amplitude components are estimated when applying the spatial model to all channels; and

a second situation in which the amplitude components are estimated when applying the spatial model with the channel under investigation excluded from the data and the spatial model; and

wherein the computing device is further configured to estimate the time-domain amplitude components using cross-validation by comparing, for each channel of the multi-channel sensor device, results of the first situation to results of the second situation in the time domain wherein, for each channel, the difference between the results of the first situation and the results of the second situation indicate the sensor noise and/or artifacts for that channel.

4. The system of claim 1 wherein the computing device is configured to estimate the time-domain amplitude components using cross-validation by applying the spatial model with the channel under investigation excluded from the data and the spatial model, wherein the difference in the temporal domain between the decomposition and the original data for the channel under investigation indicates the sensor noise and/artifacts of the channel under investigation, and

wherein the computing device is further configured to suppress the identified sensor noise and/or artifacts from the data by finding the temporal waveform common to the channel under investigation and the decomposition.

5. The system of claim 1 wherein the data spans a first length of time, and wherein the computing device is further configured to:

divide the data into a plurality of time segments that each span lengths of time shorter than the first length of time; and

perform steps (b)–(f) separately for each time segment.

6. The system of claim 5, wherein at least some of the time segments are overlapping in time.

7. A method, comprising:

(a) receiving spatially oversampled multi-channel sensor data from a region of interest;

(b) building a spatial model from the data for essential spatial degrees of freedom;

(c) decomposing the data into the underlying spatial model to obtain associated amplitude components containing a mixture of original temporal waveforms of the data;

(d) for each channel of the multi-channel sensor,

estimating time-domain amplitude components using cross-validation;

based on the estimated time-domain amplitude components, identifying sensor noise and/or artifacts for that channel; and

suppressing the identified sensor noise and/or artifacts from the data.

8. The method of claim 7 wherein decomposing the data into the underlying spatial model comprises decomposing the data in two situations for each channel of the multi-channel sensor device, the two situations including:

a first situation in which the amplitude components are estimated when applying the spatial model to all channels; and

a second situation in which the amplitude components are estimated when applying the spatial model with the channel under investigation excluded from the data and the spatial model;

wherein estimating the time-domain amplitude components using cross-validation comprises comparing, for each channel of the multi-channel sensor device, results of the first situation to results of the second situation in the time domain wherein, for each channel, the difference between the results of the first situation and the results of the second situation indicate the sensor noise and/or artifacts for that channel.

9. The method of claim 8 wherein comparing, for each channel of the multichannel device, results of the first situation to results of the second situation in the time domain comprises implementing an orthogonal projection operation.

10. The method of claim 8 wherein comparing, for each channel of the multichannel device, results of the first situation to results of the second situation in the time domain comprises implementing a singular value decomposition of the temporal domain signals of the amplitude coefficients corresponding to the spatial model to find a temporal basis.

11. The method of claim 7 wherein estimating the time-domain amplitude components using cross-validation comprises applying the spatial model with the channel under investigation excluded from the data and the spatial model, wherein the difference in the

temporal domain between the decomposition and the original data for the channel under investigation indicates the sensor noise and/artifacts of the channel under investigation, and

wherein suppressing the identified sensor noise and/or artifacts from the data comprises finding the temporal waveform common to the channel under investigation and the decomposition.

12. The method of claim 11 wherein suppressing, for each channel of the multichannel device, the sensor noise and/or artifacts by finding the temporal waveform common to the channel under investigation and the decomposition comprises projecting the original channel waveform to a temporal basis spanned by the decomposed amplitude coefficients by a parallel projection operator.

13. The method of claim 12 wherein projecting the original channel waveform to a temporal basis spanned by the decomposed amplitude coefficients by a parallel projection operator comprises implementing a singular value decomposition of the temporal domain signals of the amplitude coefficients corresponding to the spatial model to find the temporal basis.

14. The method of claim 7 wherein suppressing, for each channel of the multichannel device, the sensor noise and/or artifacts from the signal of interest comprises implementing an orthogonal projection operator.

15. The method of claim 7 wherein building the spatial model for the essential spatial degrees of freedom comprises implementing a combination of a basic solution of Laplace's equation for the harmonic scalar potential determining the magnetic field in the measurement area to calculate a spatial basis.

16. The method of claim 7 wherein building the spatial model for the essential spatial degrees of freedom comprises using dominating spatial patterns found in a statistical analysis of the original data to calculate a spatial basis.

17. The method of claim 16 wherein using dominating spatial patterns found in a statistical analysis of the original data to calculate a spatial basis comprises using dominating spatial patterns found in a principle component analysis.

18. The method of claim 7 wherein the data spans a first length of time, the method further comprising:

dividing the data into a plurality of time segments that each span lengths of time shorter than the first length of time; and

performing steps (b)–(d) separately for each time segment.

19. The method of claim 18 wherein at least some of the time segments are overlapping in time.

20. A computer-readable medium storing instructions that, when executed by one or more processors, cause the processors to perform operations comprising:

(a) receiving spatially oversampled multi-channel sensor data from a region of interest;

(b) building a spatial model from the data for essential spatial degrees of freedom;

(c) decomposing the data into the underlying spatial model to obtain associated amplitude components containing a mixture of original temporal waveforms of the data;

(d) for each channel of the multi-channel sensor, estimating time-domain amplitude components using cross-validation;

(e) for each channel, based on the estimated time-domain amplitude components, identifying sensor noise and/or artifacts for that channel; and

(f) for each channel, suppressing the identified sensor noise and/or artifacts from the data.

21. The computer-readable medium of claim 20 wherein decomposing the data into the underlying spatial model comprises decomposing the data in two situations for each channel of the multi-channel sensor device, the two situations including:

a first situation in which the amplitude components are estimated when applying the spatial model to all channels; and

a second situation where the amplitude components are estimated when applying the spatial model with the channel under investigation excluded from the data and the spatial model;

wherein estimating the time-domain amplitude components using cross-validation comprises comparing, for each channel of the multi-channel sensor device, results of the first situation to results of the second situation in the time domain wherein, for each channel, the difference between the results of the first situation and the results of the second situation indicate the sensor noise and/or artifacts for that channel.

22. The computer-readable medium of claim 21 wherein comparing, for each channel of the multichannel device, results of the first situation to results of the second situation in the time domain comprises implementing an orthogonal projection operation.

23. The computer-readable medium of claim 20 wherein estimating the time-domain amplitude components using cross-validation comprises applying the spatial model with the channel under investigation excluded from the data and the spatial model, wherein the difference in the temporal domain between the decomposition and the original data for the channel under investigation indicates the sensor noise and/artifacts of the channel under investigation, and

wherein suppressing the identified sensor noise and/or artifacts from the data comprises finding the temporal waveform common to the channel under investigation and the decomposition.

24. The computer-readable medium of claim 23 wherein suppressing, for each channel of the multichannel device, the sensor noise and/or artifacts by finding the temporal waveform common to the channel under investigation and the decomposition comprises projecting the original channel waveform to a temporal basis spanned by the decomposed amplitude coefficients by a parallel projection operator.

25. The computer-readable medium of claim 20 wherein the data spans a first length of time, and wherein the operations further comprise:

dividing the data into a plurality of time segments that each span lengths of time shorter than the first length of time; and

performing steps (b)–(f) separately for each time segment.

26. The computer-readable medium of claim 25 wherein at least some of the plurality of segments are overlapping in time.

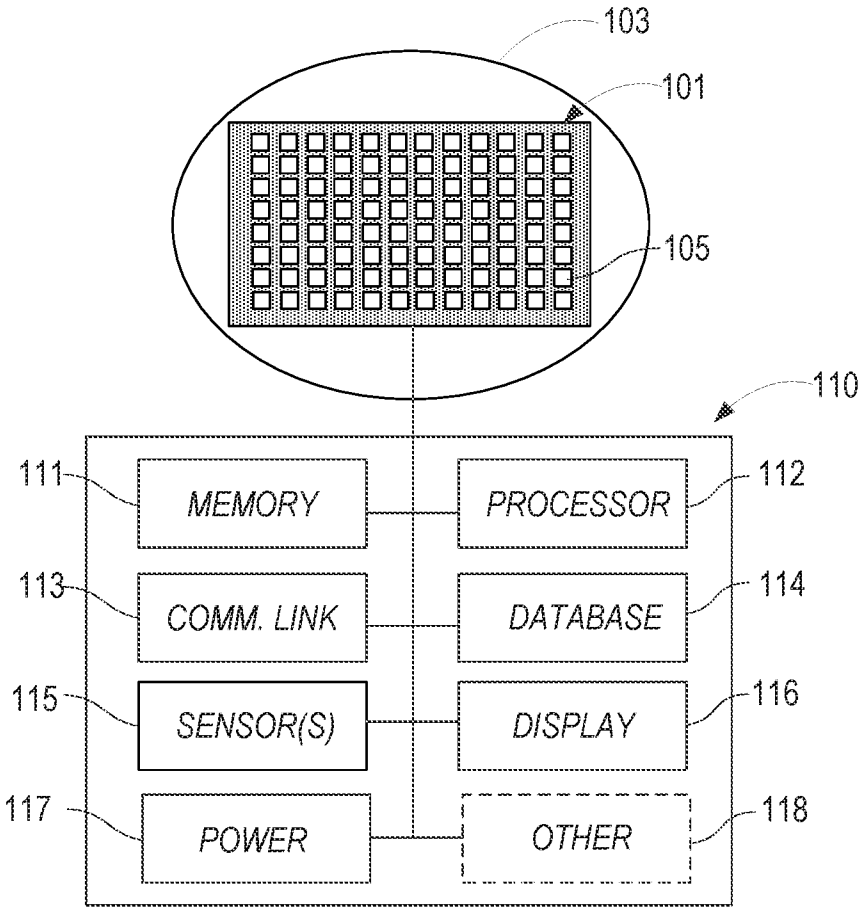

103
101
105
110
111
MEMORY
PROCESSOR
112
113
COMM. LINK
DATABASE
114
115
SENSOR(S)
DISPLAY
116
117
POWER
OTHER
118

FIG. 1

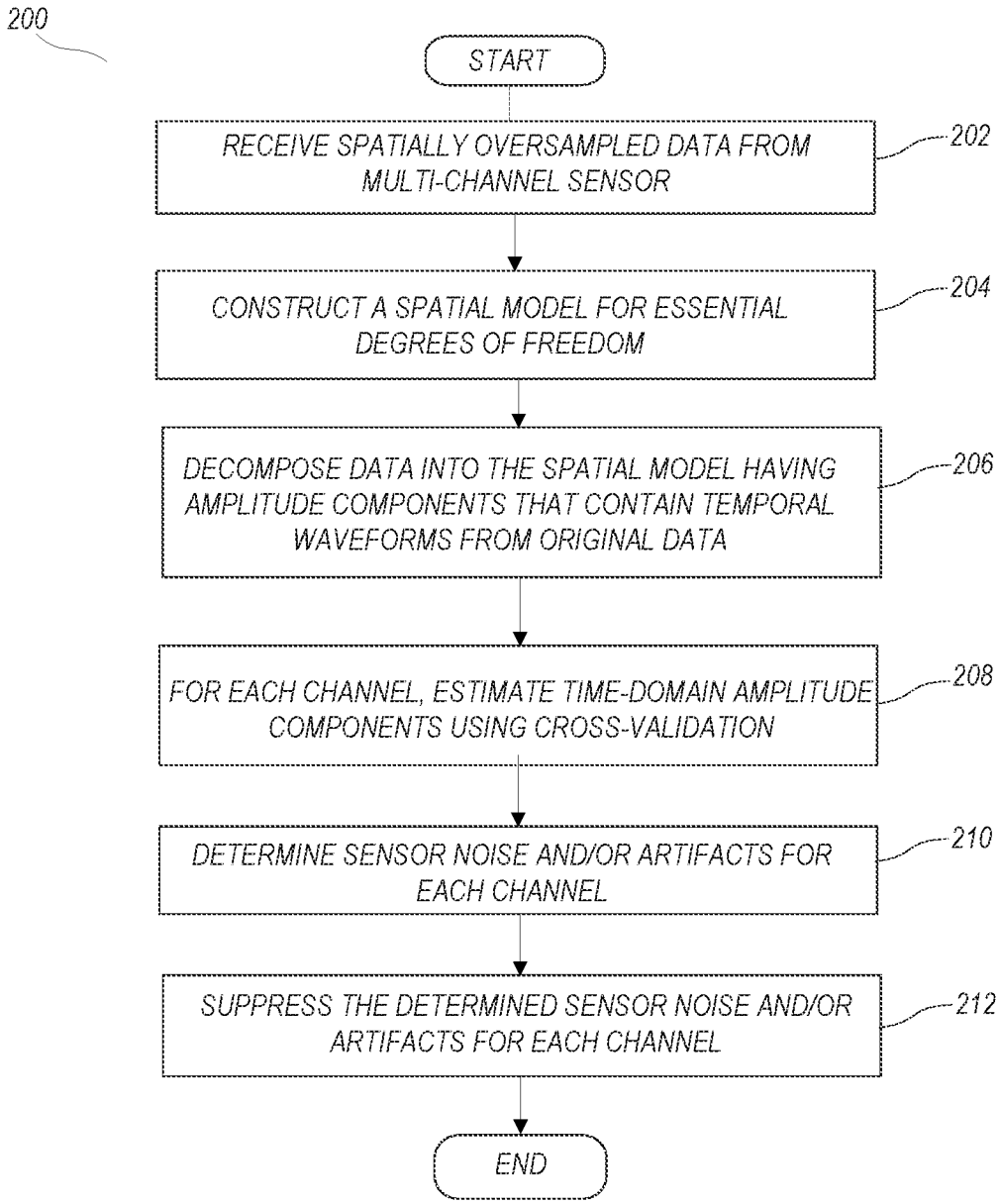

200
START
RECEIVE SPATIALLY OVERSAMPLED DATA FROM MULTI-CHANNEL SENSOR
202
CONSTRUCT A SPATIAL MODEL FOR ESSENTIAL DEGREES OF FREEDOM
204
DECOMPOSE DATA INTO THE SPATIAL MODEL HAVING AMPLITUDE COMPONENTS THAT CONTAIN TEMPORAL WAVEFORMS FROM ORIGINAL DATA
206
FOR EACH CHANNEL, ESTIMATE TIME-DOMAIN AMPLITUDE COMPONENTS USING CROSS-VALIDATION
208
DETERMINE SENSOR NOISE AND/OR ARTIFACTS FOR EACH CHANNEL
210
SUPPRESS THE DETERMINED SENSOR NOISE AND/OR ARTIFACTS FOR EACH CHANNEL
212
END

FIG. 2

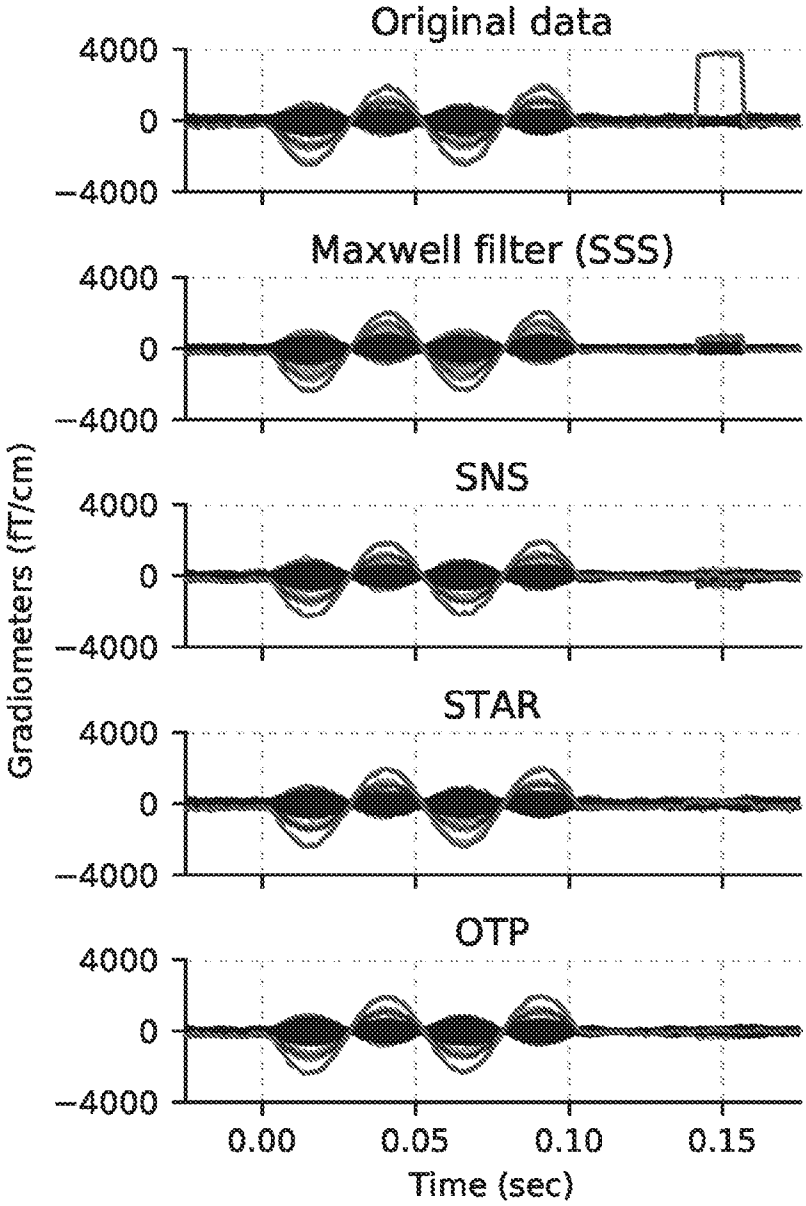

Original data
Maxwell filter (SSS)
SNS
STAR
OTP
4000
0
−4000
Gradiometers (fT/cm)
0.00
0.05
0.10
0.15
Time (sec)

FIG. 3

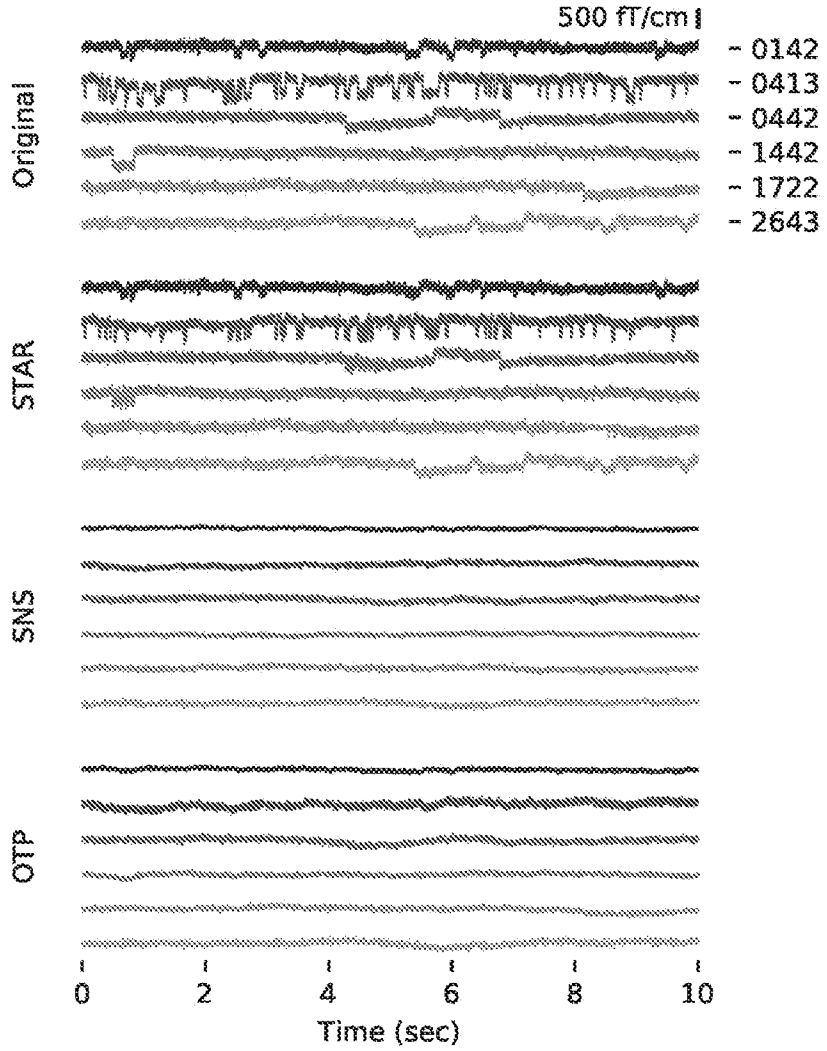

500 fT/cm
Original
STAR
SNS
OTP
- 0142
- 0413
- 0442
- 1442
- 1722
- 2643
0
2
4
6
8
10
Time (sec)

FIG. 4

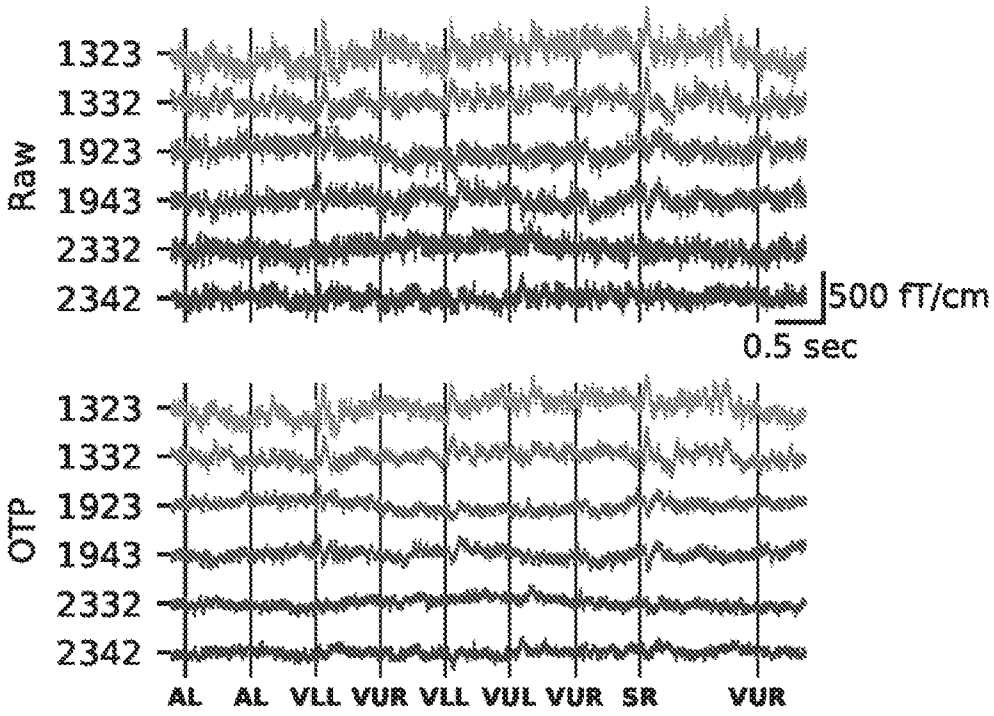

Raw
1323
1332
1923
1943
2332
2342
500 fT/cm
0.5 sec
OTP
1323
1332
1923
1943
2332
2342
AL AL VLL VUR VLL VUL VUR SR VUR

FIG. 5

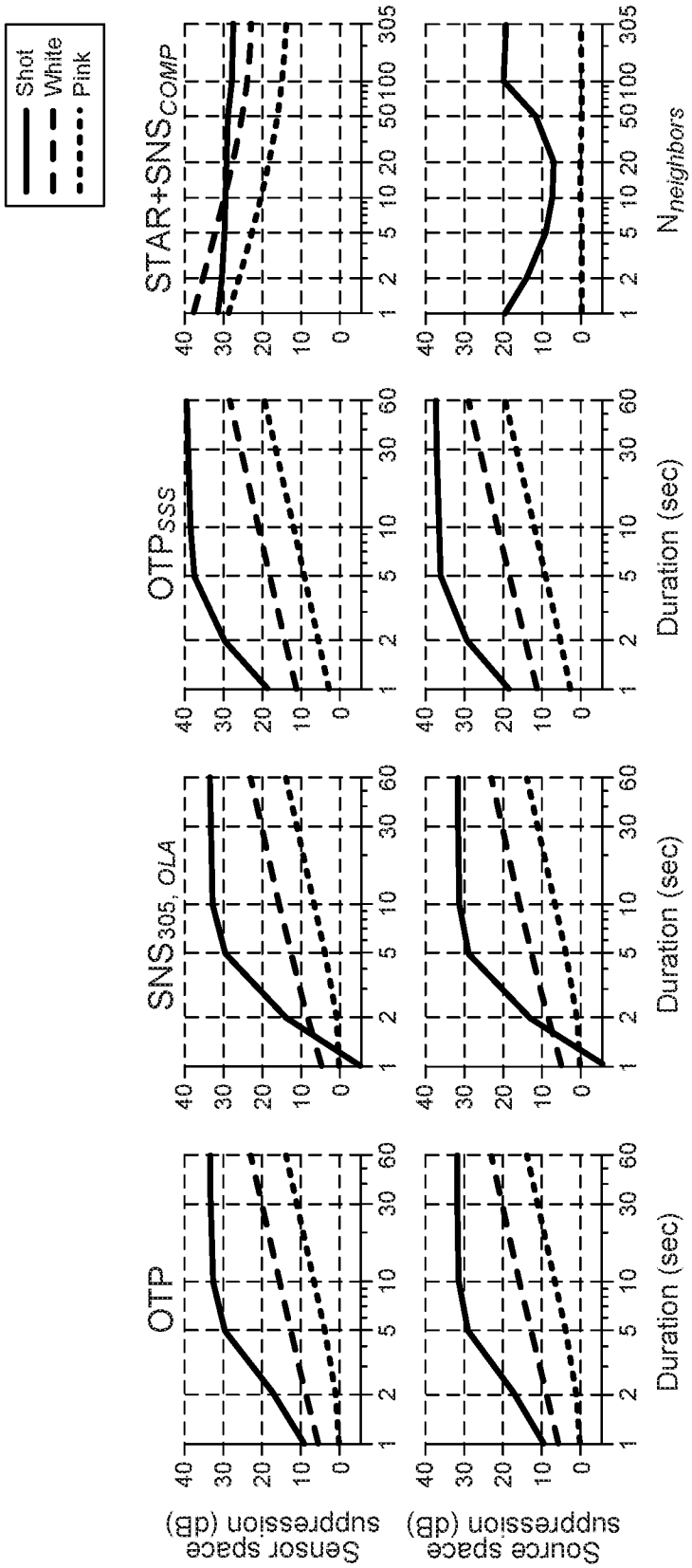

Shot
White
Pink
OTP
SNS305, OLA
OTPSSS
STAR+SNSCOMP
Nneighbors
Sensor space suppression (dB)
Source space suppression (dB)
Duration (sec)
Duration (sec)
Duration (sec)
Duration (sec)

FIG. 6

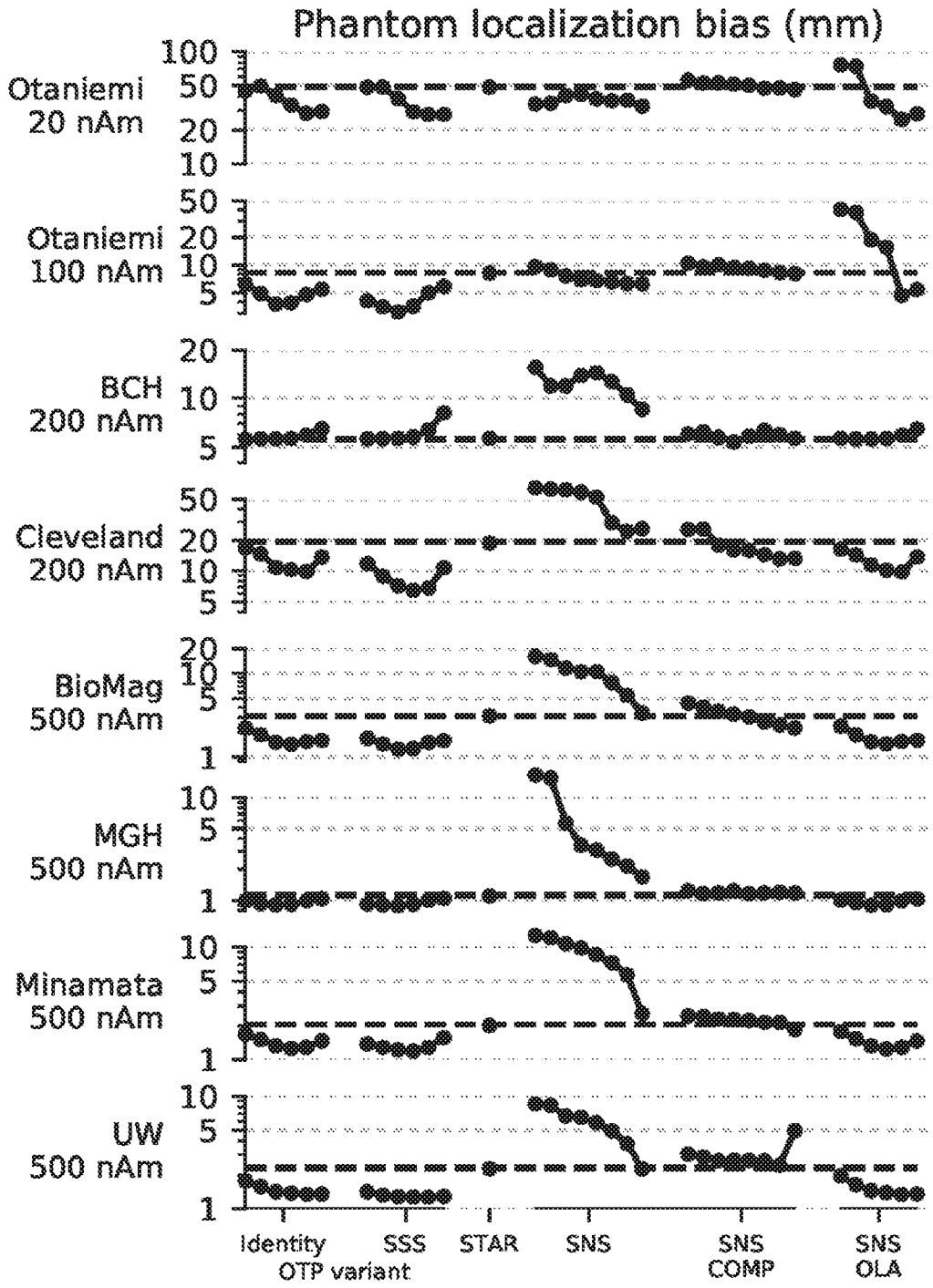

Phantom localization bias (mm)
Otaniemi 20 nAm
Otaniemi 100 nAm
BCH 200 nAm
Cleveland 200 nAm
BioMag 500 nAm
MGH 500 nAm
Minamata 500 nAm
UW 500 nAm
Identity
SSS
STAR
SNS
SNS COMP
SNS OLA
OTP variant

FIG. 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2017/034667

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - A61B 5/0476; A61B 5/00; A61B 5/04; A61B 5/05; A61B 5/055 (2017.01)

CPC - A61B 5/04008; A61B 5/04; A61B 2576/026; A61B 5/0042; A61B 5/04012; A61B 5/7203 (2017.02)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

USPC - 600/544; 382/275; 375/346; 382/131; 600/409; 702/65 (keyword delimited)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	US 2014/0343882 A1 (ELEKTA AB) 20 November 2014 (20.11.2014) entire document	1-3, 7, 8, 16, 17, 20, 21 --- 5, 6, 9, 10, 14, 15, 18, 19, 22, 25, 26
Y	US 7,933,727 B2 (TAULU et al) 26 April 2011 (26.04.2011) entire document	5, 6, 9, 10, 14, 18, 19, 22, 25, 26
Y	US 2016/0113587 A1 (THE REGENTS OF THE UNIVERSITY OF CALIFORNIA) 28 April 2016 (28.04.2016) entire document	6, 19, 26
Y	US 7,696,748 B2 (SCHLICKER et al) 13 April 2010 (13.04.2010) entire document	15
A	US 2016/0051163 A1 (BG NEGEV TECHNOLOGIES AND APPLICATIONS LTD AT BEN-GURION UNIVERSITY et al) 25 February 2016 (25.02.2016) entire document	1-26
A	US 9,008,399 B2 (WISCONSIN ALUMNI RESEARCH FOUNDATION) 14 April 2015 (14.04.2015) entire document	1-26

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 July 2017	10 AUG 2017
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, VA 22313-1450 Facsimile No. 571-273-8300	Authorized officer Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774