The Laboratory for Advanced Brain Signal Processing -RIKEN BSI
Why It Is, and How It Came to Be
Andrzej Cichocki

1. Structure, Vision and Philosophy of the Laboratory

The Laboratory for Advanced Brain Signal Processing (LAbBS) has been established in 1998 after the reorganization of its predecessor the Laboratory for Open Information Systems. It belongs to the Brain-Style Information Systems Research Group directed by Professor Shun-ichi Amari. The laboratory employs 15 researchers from 7 different countries. More than 30% of the researchers are from Japan. We are representing different fields and areas of expertise: Electrical engineering and computers science, mathematics, physics and recently also neurophysiology. Our policy is to work together with the best talented and highly motivated researchers. The most important criteria for recruitment and extension of contracts are: Scientific curiosity, passion, self-motivation, efficiency, scientific productivity, discipline, responsibility and sometimes also a wish to take some risk by involving in new areas of research. Of course, the members of the laboratory strongly influence the direction and the structure of the laboratory. On the other hand, the structure and shape of the laboratory are determined by carefully planned and executed decisions of the Director, Steering Committee and Brain Science Promotion Office.

Most of our research results and laboratory facilities itself are open. Riken, BSI generously promotes and allow us to invite top researchers from other research institutes and universities worldwide to work together. These scientific interactions are very important for research progress and they are mutually beneficial and allows us to focus on important scientific problems and methodological developments.

I believe that the variety of expertise and size of the laboratory have worked out for our advantage and can increase our cross fertilization and efficiency.

The signal and image processing are now very


---

Dr. A. Cichocki is with Riken, Brain Science Institute, Wako-shi, Saitama 351-0198, Japan on leave from Institute of the Theory of Electrical Engineering and Electrical Measurements, Warsaw University of Technology, Poland

advanced tools and methods with possibly wide potential applications.

Perhaps one of the most distinguished efforts in the laboratory is the continued effort to create new representations for the brain signals and develop new algorithms, for example, BSS, Multitask Subband Decomposition ICA (MSDICA), Sparse Component Analysis (SCA), Non-negative Matrix Factorization (NMF), Smooth Component Analysis (SmICA), or Dynamic Dependent Component Analysis (DCA) with some imposed constraints and a priori knowledge.

Moreover, the concept to combine power of sophisticated mathematical results with smart experiments design and advanced computerized representation and visualization of data provides us necessary integrated tools to extract hidden and dynamic distributed brain patterns which are often subtle and elusive.

We attempt to mix and compare different techniques and methodologies and try to apply them, sometimes to translated problems.

My objective, as much as possible, is to drive the laboratory in directions which I believe are important and at the same time interesting, and to make long-term perspectives with concrete or potential practical applications not necessarily only in brain science but also in other areas such as speech processing, communication or data mining.

Since 1998 the members of the laboratory published over 100 papers in refereed scientific journals and more than 100 conference papers some of them in collaboration with researchers from other laboratories.

2. Research Topics

The LAbSP investigates and develops tools and software for analysis (extraction, enhancement, de-noising, detection, localization, recognition and classification tasks) of brain signals and patterns. One of the challenging tasks is how to detect, enhance and localize reliably very weak, nonstationary brain source signals corrupted by noise (e.g., evoked and event related potentials ERP). To understand human neurophysiology, we currently rely on several types of non-invasive neuroimaging techniques. These techniques include electroencephalography (EEG) and magnetoencephalography (MEG) (see Fig. 1). In recent years, our primary interest has been in applying high density array EEG systems to analyze patterns and imaging of the human brain, where EEG has desirable properties of excellent time resolution. This property combined with other systems such as eye tracking and EMG (electromyography) systems with relatively low cost of instrumentation makes it attractive for investigating the higher cognitive mechanisms in the brain and opens a unique window to investigate the dynamics of human brain functions as they are able to follow changes in neural activity on a millisecond time-scale. In comparison, the other functional imaging modalities (positron emission tomography PET and functional magnetic resonance imaging fMRI) are limited in temporal resolution to time scales on the order of, at best, one second by physiological and signal-to-noise considerations.

In order to understand the higher order functioning of the brain, it is necessary to develop new methods of signal processing for brain data. On the other hand, the brain itself performs excellent information processing, by using diversified representations of information.

We are interested in the development of new efficient algorithms and the compare science aspects which are important to biomedical signals and neuroimaging. High density array EEG together with new processing techniques are employed in the study of human brain patterns and their variability. The current focus in this work is the examination of the visual, auditory and somatosensory cortex.

Recently the laboratory has started to work on bioelectromagnetic inverse problem to improve reliability of estimation of temporal patterns, spatial resolution, detection and classification of various mental tasks and for the development a Human Computer Interface HCI.

Determining active regions of the brain, given EEG/MEG measurements on the scalp is an important problem. A more accurate and reliable solution to such a problem can give information about higher brain functions and patient-specific cortical activity. However, estimating the location and distribution of electric current sources within the brain from EEG/MEG recording is an ill-posed problem, since there is no unique solution and the solution does not depend continuously on the data. The ill-posedness of the problem and distortion of sensor signals by large noise sources makes finding a correct solution a challenging analytic and computational problem.
2.1 What is Blind Signal Processing?

Related extensive activities in the laboratory include the development of new BSS or more generally Blind Signal Processing (BSP) methods and algorithms [1]. A fairly general blind signal separation (BSS) problem often referred as blind signal decomposition or blind signal extraction can be formulated as follows (see Fig. 2). We observe records of sensor signals \( x(t) = [x_1(t), \ldots, x_m(t)]^T \) from a MIMO (multiple-input/multiple-output) nonlinear dynamical system. The objective is to find an inverse system, termed a reconstruction system, neural network or an adaptive inverse system, if it exists and is stable, in order to estimate the primary source signals \( s(t) = [s_1(t), \ldots, s_N(t)]^T \). This estimation is performed on the basis of the output signals \( y(t) = [y_1(t), \ldots, y_p(t)]^T \) and sensor signals as well as some a priori knowledge of the mixing system. Preferably, the inverse system should be adaptive in such a way that it has some tracking capability in nonstationary environments. Instead of estimating the source signals directly, it is sometimes more convenient to identify an unknown mixing and filtering dynamical system first (e.g., when the inverse system does not exist or the number of observations is less than the number of source signals) and then estimate source signals implicitly by exploiting some a priori information about the system and another, a suitable specific temporal structure. The problems of separating or extracting the original source waveforms from the sensor array, without knowing the transmission channel characteristics and the sources can be considered as a number of related BSS or blind signal decomposition problems such SCA, SmoCA, ICA and its extensions, topographic ICA (TICA), pICA, SSD-ICA, DCA or Multichannel Blind Deconvolution (MBD) [1, 2].

The mixing and filtering processes of the unknown input sources \( x(k) \) may have different mathematical or physical models, depending on the specific applications. Most of linear BSS models in the simplest forms can be expressed algebraically as some specific problems of matrix factorization: Given observation (often called sensor or data) matrix \( X = [x_1, \ldots, x_N] \in \mathbb{R}^{m \times N} \) perform the matrix factorization: \( X = AS + E \) where \( A \in \mathbb{R}^{m \times N} \) represents the basis data matrix or mixing matrix (depending on applications), \( E \in \mathbb{R}^{m \times N} \) is a matrix representing errors or noise and matrix \( S = [s_1, \ldots, s_N] \in \mathbb{R}^{N \times N} \) contains the corresponding hidden components that give the contribution of each basis vector. Often these components represent unknown signals with specific temporal structures, features or properties. The matrices have usually clear physical meanings and some constraints such as nonnegativity or normalization are imposed to them (for example, the matrix \( A \) has often columns normalized to unit length and/or estimated sources have normalized variance). Often it is required that components do not have negative log-likelihoods (SmoCA) or take only nonnegative values (NMF) or values with specific constraints [1]. For example, the rows of matrix \( S \) should be sparse as possible for SCA or independent as possible for ICA.

Although some decompositions or matrix factorizations provide an exact reconstruction data, we usually consider decompositions which are approximate in nature due to noise and interference. In fact, many problems in signal and image processing can be expressed in such terms of matrix factorization. Different cost functions and imposed constraints may lead to different types of matrix factorization.

The above formulated problems are related closely to the linear inverse problem or more generally, to solving a large ill-conditioned system of linear equations (overdetermined or underdetermined depending on applications) where it is necessary to estimate reliably source \( s(k) \) and also to identify a matrix \( A \) for noisy data. Such systems of equations are often contaminated by noise or errors, thus the problem of finding an
optimal and robust with respect noise solution
arises.

There appears to be something magical about blind signal processing; we are estimating the
original source signals without knowing the pa-
rameters of mixing and/or filtering processes. It
is difficult to imagine that one can estimate this
at all. In fact, without some a priori knowledge,
it is not possible to uniquely estimate the original
source signals. However, one can usually estimate
them up to certain indeterminacies. In mathe-
matical terms these indeterminacies and ambigu-
ities can be expressed as arbitrary scaling, per-
mutation and delay of estimated source signals.
These indeterminacies preserve, however, usually
the waveforms of original sources. Although these
indeterminacies seem to be rather severe limita-
tions, in a great number of applications these limi-
tations are not essential, since the most relevant
information about the source signals is contained
in the waveforms of the source signals and not in
their amplitudes or order in which they are ar-
ranged in the output of the system.

Fig. 4 EEG Lab (Shield room)

The most promising and known BSS techniques
are ICA and SCA, although new concepts such as
DCA, SmoICA, Sparse Principal Component
Analysis (SPCA) and Parallel Factor Analysis
(PARAFAC) are also emerging.

2.2 Beyond ICA

Despite the success of using standard ICA in
many applications, the basic assumptions of ICA
may not hold for some kind of signals. Hence,
some caution should be taken when using stan-
dard ICA to analyze real-world problems, espe-
cially in biomedical signal processing. In fact, by
definition, the standard ICA algorithms are not
able to estimate statistically dependent original
sources, that is, when the independence assump-
tion is violated. Recently several extensions of
ICA has been proposed, e.g., Topographic ICA,
Multidimensional ICA and Multiresolution Sub-
band Decomposition ICA (MSD-ICA) which re-
laxes considerably the assumption regarding mu-
tual independence of primarily sources [1]. For
example, for some applications the wide-band
source signals can be strongly dependent, how-
ever some their narrow band subcomponents are
independent. The key idea in the Multiresolu-
tion Subband Decomposition ICA is to exploit
this assumption and to divide the source signal
spectra into their subbands or subcomponents
and then to treat those subspectra individually for
the purpose at hand. The subband signals can be
ranked and processed independently. Provided
that for some of the frequency subbands (at least
one) all subcomponents are mutually indepen-
dent or spatio-temporally decorrelated, then we
can easily estimate the mixing or separating sys-
tem under condition that these subbands can be
identified by some a priori knowledge or detected
by some self-adaptive process. For this purpose,
we simply apply suitable BSS algorithms, how-
ever, not for all available raw sensor data but only
for suitably pre-processed (e.g., subband filtered)
sensor signals.

We have implemented these concepts in our
ICALAB software and extensively tested for some
experimental data [1], [4]. Promising results have
been obtained.

Summarizing, The MSD-ICA (Multiresolution
Subband Decomposition ICA) can be formulated
as a task of estimation of the separating matrix
W and/or the mixing matrix A on the basis of
suitable multiresolution subband decomposi-
tion of sensor signals and by applying a classi-
cal ICA/BISS (instead for raw sensor data) for
one or several presented subbands for which
source sub-components are independent. In the
preprocessing stage, more sophisticated methods,
such as block transforms, multirate subband fil-
ter bank or wavelet transforms, can be applied.
We have extended and generalized further this
concept by performing the decomposition of sen-
sor signals in a composite time-frequency domain
rather than in frequency domain. This naturally
leads to the concept of wavelets packets
(subband hierarchical trees) and to block trans-
form packets.

2.3 Sparse Component Analysis and Beyond

Sparse Component Analysis (SCA) and related
problems Sparse Signal Representations (SSR)
and Sparse Coding (SC) arise in many scientific
problems, especially, when we wish to represent
signals of interest by using a small (or sparse)
number of basic signals from a much larger set
of signals, often called a dictionary. Such prob-
lems arise also in many applications such as
electro-magnetic and biomagnetic inverse prob-
lems.
lens (EEG/MEG), feature extraction, filtering, wavelet denoising, time-frequency representation, neural and speech coding, spatiotemporal description of arrival estimation, culture diagnosis and speed-up processing.

A closely related problem is a sparse coding in which is to seek a linear basis representation such that each signal or image is represented by a small number of active coefficients. The learning algorithm involves adapting a basis vector set while imposing a low-sparsity, or sparse, prior on the output coefficients. The crucial assumption in the sparse coding framework is that these hidden variables exhibit sparseness. Sparse coding applied on natural images has been shown to extract wavelet-like structure. The basis vectors appear self-similar, localized, oriented and bandpass.

We believe that in order to process and extract efficiently information the sparse coding should involve at least four basic properties and features:

- Use an overcomplete representation.
- Project data into a low-dimensional subspace before attempting to resolve the sparse structure.
- Apply a sparsity constraint on the basis vectors $s_n$ and/or on the latent variables $s_p$.
- Apply a nonnegativity and possibly other constraints like smoothness, independence, direct or specific temporal structures. Much of the neurophysiological data from high-level visual cortex support the hypothesis that the neural code is not only sparse, but also that it exhibits some spatio-temporal structures. Moreover, neural sparse coding also consists some kind of sparse time-frequency representation with sub-band multiscale selective filtering in which the elements of the code stand for meaningful features of the world, such as complex shapes, object-components and faces and even biomechanical stages of the visual visual pathway some selective responses to interpretable shape primitives and contour features.

While the sparse coding and the sparse representation of signals are interesting issues both experimentally and theoretically from the neuro-computation point of view, these issues only cover a some limited aspect of neural representation. A much more significant question is what auxiliary constraints or properties have sparse components such as some predictability or a temporal structure, smoothness and/or nonnegativity or box constraints. What is optimal sparsity profile or degree of sparseness? Is always the sparest representation the best and the most robust? Other important questions are: What kind of signal representations and related fundamental component analysis (if any) are essential in multidimensional, multilayer, hierarchical brain information processing, for example. Smooth Component Analysis (SMCA), Multidimensional or

Multiresolution Independent Component Analysis (MICA/MISO-ICA), Sparse Principal Component Analysis (SPCA), or Dense Component Analysis (DentCA)? Another issue to investigate is how useful are these relatively new techniques in pre-processing of brain signals. We are systemically working in these directions in the hope to give at least some partial answers.

2.4 Applications

BSS and its related methods like PARAFAC or SPICA are promising approaches for the elimination of artifacts and noise from EEG/MEG data. In fact, for these applications, ICA/BSS techniques have been successfully applied to remove artifacts and noise including background brain activity, electrical activity of the heart, eye-blink and other muscle activity, and environmental noise efficiently.

One important problem is how to automatically detect, extract and eliminate noise and artifacts. Another important problem is how to construct and classify the “brain sources”.

A conceptual model for the elimination of noise and other undesirable components from multi-sensory data is depicted in Figure 5. First, BSS is performed using suitably chosen robust (with respect to the noise) algorithm by a linear transformation of sensory data as $y(i) = Wx(i)$, where the vector $y(i)$ represents the specific components (e.g., spatio-smooth, spatio-temporally decorrelated or statistically independent components). Then, the projection of interesting or useful components (e.g., spatio-temporal decorrelated or independent activation maps) $g(i)$ back onto the sensors (electrodes) can be done by the transformation $x(i) = W^*y(i)$, where $W^*$ is the pseudo-inverse of the unmixing matrix $W$. In addition to the denoising and artifacts removal, BSS techniques can be used to decompose EEG/MEG data into individual components, each representing a physiologically dis-
tinct process or brain source. The main idea here is to apply localization and imaging methods to each of these components in turn. The decomposition is usually based on the underlying assumption of statistical independence between the activation of different cell assemblies involved. An alternative criterion for the decomposition is temporal predictability or smoothness of components. These approaches lead to interesting and exciting new ways of investigating and analyzing brain data and develop new hypotheses how the neural assemblies communicate and process information. This is actually an extensive and potentially promising research area. However, these approaches still remain to be validated at least experimentally.

Applications of BSS show special promise in the areas of non-invasive human brain imaging techniques to delineate the neural processes that underlie human cognition and sensorimotor functions. This framework provides a methodology by which several different types of information can be combined to aid in making inferences about a problem.

The BSS or more general BSP approaches are promising methods for the blind extraction of useful signals from the EEG/MEG data. The EEG/MEG data can be first decomposed into useful signal and noise subspaces using standard techniques like local and robust PCA or Adaptive Factor Analysis (AFA) and nonlinear adaptive filtering. Next, we apply BSS algorithms to decompose the observed signals (signal subspace) into specific components. The BSS approaches enable us to project each component (independent "brain source") onto an activation map at the skull level. For each activation map, we can apply an EEG/MEG source localization procedure, looking only for a single dipole (or brain source) per map. By localizing multiple sources independently, we can reduce the computational complexity and increase the likelihood of efficiently converging to the correct and reliable solution.

The fundamental problems here are: What are the system's real properties and how can we get information about them? What is valuable information in the observed data and what is only noise or interference? How can the observed (sensor) data be transformed into features characterizing the system in a reasonable way?

3. Development of Software -ICALAB Toolboxes

The ICALAB Toolboxes are three independent MATLAB packages for images, signals and EEG/MEG data processing developed by the members and visitors of the laboratory. The first two packages are available free on the web-site of our laboratory.

The web pages of our ICALAB software have received in the last several months more than 8000 hits and more than 2500 users from more than 50 countries loaded and formally registered our software. ICALAB has proved to be very popular and its is now used throughout the world for teaching and research in fields like biomedicine, to neural computation, from electrical engineering and mechanical engineering to economics.

The ICALAB was originally designed for ICA and hence its name. The actual versions of ICALAB implement many algorithms not only for ICA but also for BSS based on various criteria such linear predictability, smoothness, or sparsity. In the future, we intend to add novel algorithms for over-complete models (more source signals than sensors) including SCA, SNoCA and NMF. One of the biggest strengths of ICALAB is that it offers a variety of powerful and efficient batch algorithms, each of which makes different tradeoffs between speed of computation, accuracy, generality, simplicity, and robustness to noise. Some of the algorithms (e.g., AMUSE and FORI) are implemented and designed for pedagogical and debugging purpose. Some algorithms are very robust in respect to noise (e.g., SOBIRO or SONS). In some cases, it is recommended to use algorithms in cascade (multiple) or parallel mode in order to extract components with various statistical properties.

BSS algorithms, e.g., ICA, SCA, NMF, SNoCA, DenCa are pure mathematical, powerful, robust, but rather mechanical procedures. There is illusion that there are not very much left for the user to do after the machinery has been optimally implemented. The successful and efficient use of such tools strongly depends on a priori knowledge, common sense and appropriate use of the preprocessing and postprocessing tools. In other words, it is preprocessing of data and postprocessing of models where an expertise is truly needed. Typi
cal preprocessing tools include: Principal Component Analysis (PCA), Adaptive Factor Anal-

ysis, (AFA), prewhitening, filtering, FFT, Time-Frequency Representation (TFR) and sparsification of data (see Fig. 6). Postprocessing tools include: Deflation and reconstruction ("cleaning") of original raw data by removing undesirable components, noise or artifacts. On the other hand, the assumed linear mixing models must be validated at least approximately and original sources signals should have specified statistical properties [1], [2], [4].

4. Future Perspectives

I am not sure what will be the future and the next stages of our laboratory. However, it is within my responsibilities to create some vision, concepts and research programs which are the key factors for success and progress.

The success of any biomedical signal processing and medical imaging requires various collaboration among physicists, biologists, chemists, clinicians of many specialties, mathematicians and also engineers, especially signal processing engineers. Signal processing engineers/scientists are necessary in order to process, reconstruct, enhance and classify patterns from weak, non-stationary and usually very noisy biomedical recordings. In doing so, it is necessary that the development of signal processing algorithms is guided by the deep understanding of electrophysiological processes involved and physics behind specific imaging modality.

In the near future we plan to address several fundamental problems relevant to EEG/MEG researchers and the brain imaging studies by concentrating on using and extending the methods we have developed to date including the following tasks:

- To develop novel computational tools for inverse problems and comparing and quantifying performance of different models.
- To improve dramatically the resolution and accuracy for in estimation of temporal patterns and for the localization of multiple current sources in the brain.
- To investigate how EEG/MEG data best can be combined with functional MRI or PET activation data.
- To extract and classify brain patterns related to different mental or imaginary tasks and to develop fundamentals for a Brain Computer Interface (BCI).

Acknowledgment

I would like to express my gratitude to Professor Shin-ichi AMARI - Director of the BSI Riken for his guidance, creation of a great scientific environment for multidisciplinary research, promotion of international collaboration and giving me opportunity to conduct research in Japan. Every member of this laboratory, past and present has influenced its activities and directions to various extent. Many of them have made important and crucial contributions. I would like to warmly thank and express my appreciation for hard systematic work and scientific and technical contributions of my actual colleagues: Toru AONISHI, Hovagim BAKARDJIAN, Jianing CAO, Arato FUNASE, Pando GEORGIJEV, Wakoako HASHIMOTO, Gen HORI, Tetsuya HOYA, Junta KABANEN, Zbigniew LEONOW-ICZ, Yuanqiong LI, Yo-ichi MIYAWAKI, Sergei SHISHKIN, and Toshihisa TANAKA.

![Fig. 7 Members of the laboratory](image)
Andrzej Cichocki received the M.Sc. (with honors), Ph.D., and Habilitation degrees in electrical engineering, from Warsaw University of Technology (Poland) in 1973, 1977, and 1982, respectively. Since 1997, he has been with the Institute of Theory of Electrical Engineering and Electrical Measurements at the Warsaw University of Technology, where he became a full Professor in 1991. He is the co-author of three books: Adaptive Blind Signal and Image Processing – Learning Algorithms and Applications, Wiley, 2002, Multichannel Capacitance and Constrained Time Integrated Circuits and Systems (Springer-Verlag, 1989) and Neural Networks for Optimization and Signal Processing (Teubner-Wiese, 1991-94) and more than 150 research papers. He spent at University Erlangen-Nuremberg (Germany) a few years as Alexander Humboldt Research Fellow and Guest Professor. Since 2001 he has been working in the Brain Science Institute RIKEN (Japan), as a team leader for laboratory for Open Information Systems and currently as a head of laboratory for Arteriovenous Signal Processing. His current research interests include optimization, information, neuromarketing and signal and image processing, especially analysis and processing of multisensory biomedical data. He is currently associate editor of IEEE Transactions on Neural Networks and recently the member of core group who established a new IEEE Circuits and Systems Technical Committee for Blind Signal Processing and the member of Steering Committee of ICA workshops.