Multimodal Imaging and BCI

Klaus-Robert Müller, Siamac Fazli, Jan Mehnert, Stefan Haufe, Frank Meinecke, Wojciech Samek, Felix Biessmann, Sven Dähne, Gabriel Curio, Benjamin Blankertz et al.
Multimodal Data

- Combining **multiple** physiological **features** (oscillations, SCPs, maps)  
  (cf. Dornhege et al 2006)

- Combining **multiple subjects** data (cf. nonstationarity)  
  (cf. Fazli et al 2009 and 2011, Samek, Meinecke & Müller submitted)

- **Correlating apples & oranges** i.e. computing correlations between multiple measuring modalities (EEG & EMG, EEG & NIRS, EEG & fMRI, LFP & fMRI)  

- Combining **multiple measuring modalities** (EEG & EMG, EEG & NIRS, EEG & fMRI)  
  (see Fazli et al 2012, Biessmann et al 2011; Pfurtscheller, Müller-Putz, Calhoun, Ritter et al, Cohen, Villringer, Eichele)

- **Nonlinear correlations between modalities** NIRS & EEG  
  (see Dähne et al submitted)
Multimodal ↔ Nonstationary
Motivation: Shifting distributions within experiment

But: Is the nonstationarity different between subjects, i.e. could we learn it from other subjects?
Changes are similar!

Modalities = Other Subjects
Changes between training and test data are similar between users.
Other multi-subject methods, e.g. cov matrix shrinkage, may improve estimation quality but do not reduce non-stationarities.
Cartoon: learn from adverse nonstationary subspace across subjects

Usually discriminative information is transferred between subjects.
Algorithm

(1) For each subject $i = 1 \ldots n$, $i \neq i^*$ compute the eigenvectors $v_i^{(1)} \ldots v_i^{(d)}$ of $\Sigma_i^{\text{train}} - \Sigma_i^{\text{test}}$.

(2) For each subject $i$ select the $l$ eigenvectors with largest absolute eigenvalues.

(3) Aggregate the vectors into a matrix $P$.

(4) Apply PCA to reduce the dimensionality of the non-stationary subspace $S_P = \text{span}(P)$ to $\nu$.

(5) Compute the projection matrix $P^\perp$ to the orthogonal complement of $S_P$.

(6) Make $i^*$'s data invariant to the changes by projecting out non-stationarities $\tilde{X} = (P^\perp)^T P^\perp X$.

(7) Compute spatial filters from $\tilde{X}$ using CSP.
Results

Two data sets with different stimulus cues in training and test

1. visual cue in training & auditory cue in test
2. letters in training & moving objects in test

The size of the non-stationary subspace is determined by CV in a leave-one-subject-out manner on the other users.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Audio-Visual Data Set</th>
<th>BCI Competition III</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A1</td>
<td>A2</td>
<td>A3</td>
</tr>
<tr>
<td>CSP</td>
<td>79.5</td>
<td>80.0</td>
<td>65.8</td>
</tr>
<tr>
<td>ssCSP</td>
<td>87.1</td>
<td>80.8</td>
<td>67.5</td>
</tr>
</tbody>
</table>

ssCSP: stationary subspace CSP
Interpretation

The most non-stationary directions are very similar between users.

Activity in occipital and temporal areas is penalized as these regions are mainly responsible for visual and auditory processing.
Feature distribution becomes stationary
Summary Part I

- Novel “multi-modal” approach to reduce non-stationarities in data

- In contrast to other multi-subject methods it does NOT transfer discriminative information, thus is more robust if subject similarity is low.

- Non-stationary information appears physiologically interpretable and meaningful.

- The idea of transferring stationary subspaces between subjects can be applied to many other problems.
NIRS-EEG Brain Computer Interfaces

[Fazli et al. Neuroimage 2012]
Experimental Setup and Paradigm

EEG: 37 electrodes

NIRS 26 channels (frontal, parietal, occipital)

EEG-based cursor feedback (ISI = 15 s)

Executed movement vs imagery movements

Imagery movements: EEG-feedback for left and right motor imagery

Number of subjects: 14

Can a simultaneous measurement of NIRS and EEG during Brain Computer Interfacing enhance the classification accuracy?

Are the results physiologically reliable?
Temporal Dependency of Classification in Executed Movements

EEG peaks earlier as compared to HbO and HbR

Physiological reliability: HRF shaped classification accuracies over time

Classification accuracy higher for EEG

Fazli et al. 2012
Temporal Dependency of Classification in Motor Imagery

EEG peaks earlier as compared to HbO and HbR

Physiological reliability: HRF shaped classification accuracies over time

Classification accuracy higher for EEG

Classification accuracy lower than in executed movements
Combination of EEG and NIRS

LDA classifier estimated for EEG, HbO and HbR (individually)

Meta-classifier estimated for combination in each subject

All within cross-validation (8 chronological splits)

Fazli et al. 2012
Feature Combination

NIRS-EEG combinations have higher classification accuracies for vast majority of subjects

Fazli et al. 2012
t-tests reveal a significant increase of classification accuracy for combination.

Fazli et al. 2012
Some subjects, which were not classifiable with EEG become classifiable by a meta-classifier in combination with NIRS.
Mutual Information

NIRS features for all correct EEG trials (EEG+) and incorrect EEG trials (EEG-)

Pattern is similar although the significance drops

NIRS can complement the EEG with physiological meaningful information
Discussion Part II

Problems

• Different temporal properties of the measurement devices (e.g. EEG: 1000 Hz, NIRS: max. 10 Hz)

• Temporal lag between parameters

• Different signal qualities

Ideas to Overcome the Temporal Lag

• NIRS as a measure of subjects’ attention to predict EEG-based performance

• NIRS as a localizer of the source of EEG signals

• NIRS as a ‘stop’, e.g. to discard a EEG-based classified trial when not confirmed by NIRS
Finding nonlinear correlations between NIRS & EEG

[Dähne et al. submitted]
Generative Model

Mapping to sensor space:

\[
\begin{align*}
\mathbf{x}(t) &= \mathbf{A}_x \mathbf{s}(t) + \epsilon_x(t) \\
\mathbf{y}(t) &= \mathbf{A}_y \Phi(s(t)) + \epsilon_y(t)
\end{align*}
\]

\(\Phi(s_i(t)) := (h * p_{s_i})(t) = \sum_{\tau>0} h(\tau) p_{s_i}(t - \tau)\)

"Convolution of source band power with hemodynamic response function (HRF)"
Generative Model

These steps do not commute! Nonlinearity is applied on source level.
Approaches to multimodal data analysis

- order of processing steps in line with generative model, i.e. first transformation into source space, then computation of spectral power

- modality-specific unmixing → does not take information from other modality into account to guide the unmixing

- post-hoc matching of components, thus not truly multimodal
Approaches to multimodal data analysis

- multi-modal unmixing $\rightarrow$ optimizes the coupling between components

- order of processing steps **not** in line with generative model, i.e. nonlinearity is applied in sensor space instead of source space, **!WRONG!**

- resulting "EEG/MEG power patterns" cannot be subjected to standard source localization techniques, because these methods are designed to localize time-domain patterns, not spectral-domain patterns
Approaches to multimodal data analysis

PCA / ICA

- EEG / MEG
- fMRI / NIRS

- Unimodal unmixing
- Spectral features
- Convolution

PLS / CCA

- EEG / MEG
- fMRI / NIRS

- Spectral features
- Convolution
- Unimodal unmixing

mSPoC

- EEG / MEG
- fMRI / NIRS

- Multimodal unmixing
- Spectral features
- Convolution
Multimodal source power correlation analysis (mSPoC)

mSPoC objective function:

\[ f_{\text{obj}}(w_x, w_y, w_\tau) := \text{Cov} \left( \hat{h}(\hat{p}_{s_x}), \hat{s}_y \right) \]

norm constraints:

\[
\begin{align*}
\| w_x \|_{C_{xx}} & := w_x^T C_{xx} w_x = 1 \\
\| w_y \|_{C_{yy}} & := w_y^T C_{yy} w_y = 1 \\
\| w_\tau \|_{B} & := w_\tau^T B w_\tau = 1
\end{align*}
\]

\( C_{xx}, C_{yy} \) : modality specific covariance matrices

\( B \) : auto-correlation matrix of \( \hat{p}_{s_x} \)

mSPoC model:

\[
\begin{align*}
w_x^T x & = \hat{s}_x \\
w_y^T y & = \hat{s}_y
\end{align*}
\]

\[
\hat{p}_{s_x}(e) = \left\langle (w_x^T x(t))^2 \right\rangle_{t \in T_e} = w_x^T C_{xx}(e) w_x.
\]

\[
\hat{h}(\hat{p}_{s_x})(e) = \sum_n^{N_\tau} w_{\tau n} \hat{p}_{s_x}(e - n)
\]

\[
\begin{align*}
C_{xx}^{-1} \left( C_{xy} \hat{x}_{(3)} w_y \hat{x}_{(4)} w_\tau \right) w_x & = \lambda w_x \\
C_{yy}^{-1} \left( C_{xy} \hat{x}_{(1)} w_x \hat{x}_{(2)} w_y \hat{x}_{(3)} w_\tau \right) w_\tau & = \lambda w_\tau
\end{align*}
\]

\[
\begin{align*}
B^{-1} \left( C_{xx} \hat{x}_{(1)} w_x \hat{x}_{(2)} w_x \hat{x}_{(3)} w_y \right) w_y & = \lambda w_\tau
\end{align*}
\]
Multimodal analysis of simultaneously recorded EEG and NIRS

Data from Fazli et al. 2012 → 96 trials of (left/right) hand gripping. Comparison of \textbf{mSPoC} to convolutive \textbf{CCA}.
mSPoC vs CCA

Correlations between EEG and NIRS [HbR]

Correlations between EEG and NIRS [HbO]

Fig. 5. Cross-validated correlations between EEG and NIRS (left HbR, right HbO) in the motor execution task for each subject. Results of mSPoC and CCA are compared. Each point corresponds to the correlations obtained for the first set of $w_x$, $w_y$, and $w_\tau$ from a single subject by CCA (x-axis) and mSPoC (y-axis).
mSPoC vs CCA

Fig. 7. Exemplary results for one subject (VPean) as derived by mSPoC. The scalp-plots on the left side show the EEG pattern that corresponds to the obtained filter $w_x$. In the middle plot we show the temporal filter for the EEG power of the component shown left. The rightmost scalp-plots depict the spatial pattern that corresponds to the filter $w_x$, i.e. the NIRS patterns. The top row shows the results for applying mSPoC to left hand movement trials, while in the bottom row results for right hand movement trials are shown.
Conclusion

- Information from multimodal measurements increases the understanding of physiology in neuroscience

- Multimodal Imaging is of interest for numerous research questions and clinical application

- The specific fusion of data depends on the research question and the used instruments

- Numerous algorithms have been developed to merge the data

FOR INFORMATION SEE:
www.bbci.de
Thanks to BBCI core team:

Gabriel Curio
Florian Losch
Volker Kunzmann
Frederike Holefeld
Jan Mehnert
Vadim Nikulin@Charite

Florin Popescu
Motoaki Kawanabe
Guido Nolte
Thorsten Dickhaus

Benjamin Blankertz
Michael Tangermann
Claudia Sannelli
Carmen Vidaurre
Siamac Fazli
Martijn Schreuter
Stefan Haufe
Andreas Ziehe
Laura Acqualagna
Frank Meinecke
Wojciech Samek
Sven Dähne
Paul von Bünau
Felix Biessmann@TUB

Matthias Krauledat
Guido Dornhege
Steven Lemm
Yakob Badower
Marton Danozci
Roman Krepki@industry

Collaboration with: U Tübingen, Bremen, Albany, KU, TU Graz, EPFL, Daimler, Siemens, MES, MPIs, U Tokyo, TIT, RIKEN, Bernstein Center for Computational Neuroscience Berlin, Columbia, CUNY
Funding by: EU, BMBF and DFG