NeuroImage 110 (2015) 136–148

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Optimizing methods for linking cinematic features to fMRI data

Janne Kauttonen [a,c,*,1], Yevhen Hlushchuk [a,b,c,d,1], Pia Tikka [a,c]

[a] Department of Film, Television and Scenography, Aalto University School of Arts, Design and Architecture, FI-00076 AALTO, Finland
[b] Brain Research Unit, O.V. Lounasmaa Laboratory, Aalto University, FI-00076 AALTO, Finland
[c] Aalto NeuroImaging, Aalto University, FI-00076 AALTO, Finland
[d] Department of Radiology, Hospital District of Helsinki and Uusimaa (HUS), HUS Medical Imaging Center, Helsinki University Central Hospital (HUCH), University of Helsinki, Helsinki, Finland

ARTICLE INFO

Article history:
Accepted 30 January 2015
Available online 7 February 2015

Keywords:
fMRI
Neurocinematics
Elastic-net regularization
Linear regression
Independent component analysis
Naturalistic stimuli
Annotation

ABSTRACT

One of the challenges of naturalistic neurosciences using movie-viewing experiments is how to interpret observed brain activations in relation to the multiplicity of time-locked stimulus features. As previous studies have shown less inter-subject synchronization across viewers of random video footage than story-driven films, new methods need to be developed for analysis of less story-driven contents. To optimize the linkage between our fMRI data collected during viewing of a deliberately *non-narrative* silent film 'At Land' by Maya Deren (1944) and its annotated content, we combined the method of elastic-net regularization with the model-driven linear regression and the well-established data-driven independent component analysis (ICA) and inter-subject correlation (ISC) methods. In the linear regression analysis, both IC and region-of-interest (ROI) time-series were fitted with time-series of a total of 36 binary-valued and one real-valued tactile annotation of film features. The elastic-net regularization and cross-validation were applied in the ordinary least-squares linear regression in order to avoid over-fitting due to the multicollinearity of regressors, the results were compared against both the partial least-squares (PLS) regression and the un-regularized full-model regression. Non-parametric permutation testing scheme was applied to evaluate the statistical significance of regression. We found statistically significant correlation between the annotation model and 9 ICs out of 40 ICs. Regression analysis was also repeated for a large set of cubic ROIs covering the grey matter. Both IC- and ROI-based regression analyses revealed activations in parietal and occipital regions, with additional smaller clusters in the frontal lobe. Furthermore, we found elastic-net based regression more sensitive than PLS and un-regularized regression since it detected a larger number of significant ICs and ROIs. Along with the ISC ranking methods, our regression analysis proved a feasible method for ordering the ICs based on their functional relevance to the annotated cinematic features. The novelty of our method is – in comparison to the hypothesis-driven manual pre-selection and observation of some individual regressors biased by choice – in applying data-driven approach to all content features simultaneously. We found especially the combination of regularized regression and ICA useful when analyzing fMRI data obtained using non-narrative movie stimulus with a large set of complex and correlated features.

Introduction

The neuroimaging experiment settings using free-viewing of movies as stimuli, also referred to as *naturalistic* neuroscience, have drawn increasing interest among cognitive neuroscientists. Conventional neuroimaging experiments that use artificial stimuli such as still images, beep sounds, or check boards in relatively isolated conditions, have greatly accumulated our understanding on specific brain phenomena. In turn, showing movies of full length in the brain scanner has significantly advanced the possibilities to study human brain in relation to more socially valid life-events.

The first study that used narrative film in functional magnetic resonance imaging (fMRI), helped to understand how brain responds to event-boundaries in a continuous film stimulus (Zacks et al., 2001). Since then, a range of studies have paved the way for relating complex naturalistic stimulus events to equally complex neural events detected in the brain using fMRI (Bartels and Zeki, 2004a,b; Bartels et al., 2008; Hasson et al., 2004; Jääskeläinen et al., 2008; Lahnakoski et al., 2012a, b; Zacks et al., 2010), and, more recently, magnetoencephalography (MEG; Lankinen et al., 2014).

Movies typically depict people in their natural situations and contexts. With some confidence, their goal-directed actions can be assumed to engage action observation and imitation (loosely, 'mirroring')

* Corresponding author at: Aalto School of Science, AMI centre, P.O. Box 13000, 00076 AALTO, Finland.
E-mail address: janne.kauttonen@gmail.com (J. Kauttonen).
[1] Equal contribution.

http://dx.doi.org/10.1016/j.neuroimage.2015.01.063
1053-8119/

networks (Caspers et al., 2010; Rizzolatti et al., 2014). Most of the previous film-viewing studies have synchronized annotated appearance of particular features, such as faces or body parts, to the viewer's observed brain activation, thus adding to understanding of how the face- and body-specific brain functions intertwine with motion specific functions (Abdollahi et al., 2013; Jastorff and Orban, 2009; Kanwisher and Yovel, 2006; Large et al., 2008; Peelen and Downing, 2005; Saxe et al., 2004) or, with brain functions associated with movements, either in the fixed camera frame (local motion), or in mobile camera frame (global motion) (Bartels et al., 2008). Taken the complexity of moving images, creating linkages between simultaneous interdependent features and the viewer's brain functions is one of the most critical issues. In our view, the naturalistic paradigm is still lacking adequate tools for the interpretation of observed brain activations in relation to the complex multiplicity of time-locked stimulus features.

Regarding the methods of linking stimulus features with the related fMRI data, conventional methods used with more artificial experiment conditions have considerable limitations. For instance, the general linear model (GLM; Friston et al., 2006) analysis works best with controlled conditions, such as event- or block designs, in which the parametric model for the BOLD signal changes related to the activation is defined *a priori* (Pajula et al., 2012). However, when applied to naturalistic stimuli conditions in fMRI, GLM seems less sensitive in detecting brain activation related to annotated stimulus content when compared to model-free analysis methods, such as independent component analysis (ICA) (Malinen et al., 2007). ICA is an efficient dimension reduction method for distinguishing a set of independent functional brain networks (ICs, i.e., groups of voxels; Calhoun and Adali, 2012; Calhoun et al., 2001; Chen and Yao, 2004; Mckeown et al., 1998). Yet, further analysis is needed to pinpoint those components (ICs) that relate to the particular features of the stimulus (Bartels and Zeki, 2004a; Lahnakoski et al., 2012b). Often considered as an alternative method to ICA, intersubject correlation (ISC), in turn, has been proven effective in identifying brain areas that are activated in similar manner across subjects, as they view the same narrative film (Hasson et al., 2004; Jääskeläinen et al., 2008). Interestingly, story-driven drama films seem to elicit significantly higher correlations across different viewers compared to more random non-narrative video recordings (Hasson et al., 2008b).

Neither ICA nor ISC alone, however, allows straightforward identification of the functional tasks of particular brain regions. When combined, the inter-subject correlation method can be used for sorting a set of ICs in two groups. The first includes inter-subjectively shared brain networks that can be related to *extrinsic* stimulus-induced activation, and the other relates to *intrinsic* mental processes that are not time-locked with the stimulus (Bartels and Zeki, 2004a; Malinen and Hari, 2011). The next question is how to identify the time-locked linkages between the detected brain activation and complex stimulus features. Previous studies have mainly used story-driven films, which typically follow established storytelling conventions that cue the viewers' shared intersubjective synchronization in an expected manner, thus providing some basis for linking of the story content with the brain data. A recent study revealed high synchronization in the executive parieto-frontal network, when data collected during the viewing a Hitchcock suspense movie was linked to the viewer's subjective time-locked annotations of “suspense” (Naci et al., 2014). The study exemplifies once again the catchy nature of story-driven films, and particularly that of thrillers.

In our study, we aimed to go beyond story-driven narratives. What methods could improve the analysis of time-locked interdependence between the brain data and the content of more ambiguous non-narrative films, or video recordings of non-structured events, such as improvised conversation? This question motivated us to study the linkage between our fMRI data collected during viewing of a *non-narrative* silent film ‘At Land’ directed by Maya Deren (1944, 14′40″) and its annotated content. The film ‘At Land’ shows an expressionless young woman wandering in her surroundings without any explicit motivation for her behavior, such as collecting stones in the beach, or jumping down from a rock. In addition, according to the director-actress herself, she deliberately avoided emotional expressions, while the cinematographic aspects, such as camera movements and framing, have been carefully composed (Deren, 2005). How to link the cinematic features of such an ambiguous film with the viewers' brain activation detected while they are trying to make sense of it? As the film ‘At Land’ does not give any tools for inferring the character's mental state, goals of her actions, or inner reasons, story-based film structure analysis methods do not necessarily allow adequate interpretation for the resulting linkages. Instead, by annotating all bodily actions and camera-related features specifically pointed out by the director of the film, one might find meaningful linkages between the film content and collected fMRI data.

To tackle the topic, we decided to combine the data-driven ICA, ISC, and linear regression methods established by previous naturalistic movie-viewing studies. Also, in line with the annotation plan, we would apply a data-driven approach to the film content as well. We first ranked the 40 ICs based on the temporal and spatial ISC (Bartels and Zeki, 2004b; Malinen and Hari, 2011). Then, in order to identify which brain regions specifically respond to the annotated aspects of the ‘At Land’, we compared a set of 37 annotation time-series both against IC time-series and region-of-interest (ROI) time-series using ordinary least squares (OLS) linear regression. To further overcome two problems, the coefficient estimation instability due to correlated annotations, and the overfitting due to large annotation set, we applied the *elastic-net* (Zou and Hastie, 2005) and *partial least-squares* (PLS; Wold et al., 2001) regularization methods. The advantage of the regularization was that it allowed us to carry the analysis with a larger number of mutually dependent annotations (regressors) by prioritizing those that fit the measured data best. The results from the regularized regression analysis allowed also ordering the ICs based on the best fit. The problems related to correlated regressors (i.e., *collinearity*) are well-known in linear regression analysis literature (Hair, 2009), also acknowledged by fMRI studies that apply GLM (Friston et al., 2006). Conventional designs using artificial stimuli aim at maximizing the control over the experiment by minimizing collinearity between effects of interest. In contrast, in naturalistic free-viewing studies, which aim at simulating the real-life situations, the audiovisual and temporal richness of the stimulus unavoidably reduces the control over the multiplicity of correlated regressors. The novelty of our method is – in comparison to the hypothesis-driven manual pre-selection and observation of some individual regressors biased by choice – in applying a data-driven approach to all content features simultaneously. In addition, no other study, to our best knowledge, has applied and compared regularization methods to the analysis of naturalistic movie-viewing data.

Materials and methods

Subjects, stimuli and data acquisition

Subjects

The fMRI data from 22 healthy adults were collected, out of which 12 (7 female) of the highest quality datasets were chosen for the analysis (i.e., alert subjects and minimal motion artifacts: 9 subjects were discarded due to drowsiness and 1 due to excessive motion artifacts). All 12 subjects were right-handed and their age ranged from 20 to 50 years (mean 25.9 years and standard deviation 8.2 years). The study had a prior approval by the Ethics Committee of Helsinki and Uusimaa Hospital District and all subjects gave their informed prior consent.

Annotations

The total of 36 different binary annotation time-series that covered a wide range of cinematic features appearing during the full duration of the film 'At Land' were produced by two film experts using the ELAN annotation tool (Brugman and Russel, 2004). Each annotation time-series consisted of a sequence of body- or camera-related events defined by timestamps for start and end points of each event. They covered various features of the film ranging from the appearance of human body, faces, hands, or feet, towards more conceptual annotations of camera angles, different types of movements, camera tricks, or otherwise dramatic moments in the film. Further in the text, these categories are labeled with prefixes BODY (body), CAM (camera), FACE (face), DRAMA (drama), FRAME (framings), OBJ (objects), MOT (motion) and MOV (kinaesthetic movement). In addition, we included in the analysis one more category (TACTILE) based on the real-valued annotation on tactile experiences averaged over separate annotations of 16 subjects, conducted immediately after the neuroimaging experiment. The rating data were collected using an in-house built software that allows continuous annotation of one's experience (expressed with mouse movements) while viewing the movie on computer screen. The complete list of 37 annotations used in the analysis with their detailed descriptions is given in Appendix A.

Stimuli and data acquisition

During the fMRI scanning, the subjects were viewing a silent black-and-white film 'At Land' directed by Maya Deren in 1944 (duration 14′40″). The fMRI images were acquired with a Sigma Excite 3 T MRI scanner (General Electric, Milwaukee, WI, USA). Functional images were obtained using a gradient echo-planar-imaging sequence with following parameters: TR 2.015 s, TE 32 ms, FA 75°, 34 oblique axial slices, slice thickness 4 mm, matrix 64 × 64, voxel size 3.4 × 3.4 × 4 mm, field of view (FOV) 22 cm. Total 489 EPI volumes were collected per subject. Structural images were scanned with 3-D T1 spoiled gradient imaging, matrix 256 × 256, TR 10 ms, TE 3 s, flip angle 15°, preparation time 300 ms, FOV 25.6 cm, slice thickness 1 mm, voxel size = $1 \times 1 \times 1$ mm^3 and number of excitations 1. The video feed obtained through the eye-tracking system iView X™ MRI-LR (long range) system (Sensomotoric Instruments GmbH, Germany) was used to evaluate subjects' alertness (eyes open or closed).

Data analysis

Data preprocessing

The fMRI data. The preprocessing in SPM8 (Wellcome Trust Centre for Neuroimaging; http://www.fil.ion.ucl.ac.uk) included realignment, co-registration, normalization into MNI space, and smoothing with 8 mm (full-width-at-half-maximum) Gaussian filter. We employed the ArtRepair toolbox (Mazaika et al., 2005) to remove artifacts from individual voxel-wise time-series. Finally high-pass temporal filter with 100 s (0.01Hz) cut-off was applied (SPM8's *spm_filter*). Out of all 489 volumes, the first 11 (dummy scans and opening credits) and the last 43 (fixation cross and another experiment) volumes were omitted, and only the remaining 435 volumes containing the movie were used in the data-analysis.

Annotations. All 36 continuous binary annotation time-series (regressors) were first sampled with 1 kHz and then convolved with a standard double-gamma hemodynamic response function (SPM8's *spm_hrf*) with 5 s lag. The convolved time-series were then down-sampled to match 0.5Hz sampling rate of the fMRI data and, similarly to fMRI data, high-pass filtered with 100 s cut-off. Preprocessing of the tactile annotation was otherwise similar, but the original data was obtained at the sampling rate of 5Hz.

ICA

The spatial ICA was done using GIFT toolbox (http://mialab.mrn.org/software/gift). In this method, individual data were temporally concatenated and processed via a dimension reduction procedure through two stages of principal component analysis (PCA) and the resulting data-set was then decomposed by ICA procedure using InfoMax algorithm implemented in GIFT. The chosen model order of 40 ICs had been found appropriate for our dataset in the previous study by Pamilo et al. (2012). The group-ICA decomposition was repeated at minimum 200 times with randomly initialized decomposition matrices using ICASSO (Himberg et al., 2004) method implemented in GIFT. Default GIFT parameters were used, except for the number of first PCA component count which was increased by 25%, and GICA3 algorithm was applied for back-reconstruction (for reasoning, see Appendix S1 in Pamilo et al., 2012). Subsequent analyses were carried on using centrotype estimates of stable components as produced by ICASSO. Population-level spatial maps of ICA components were determined with two-tailed t-test from subject-wise coefficient maps with $p < 0.05$ family-wise error (FWE) correction as implemented in SPM8. Throughout the remaining text, components are identified by the order given by GIFT (i.e., from IC1 to IC40). Although some components are typically artifacts and/or located outside the grey matter, all components underwent full analysis without pre-selection. This ensured that the analysis remained data-driven. Components were labeled by comparing their spatial overlap with various previously reported functional networks (Allen et al., 2011; Heine et al., 2012; Pamilo et al., 2012; Smith et al., 2009).

IC ranking based on ISC

To separate stimulus-related ICs that were inter-subjectively shared across subjects (extrinsic) from those that may relate to more idiosyncratic cognitive processes (intrinsic), the spatial and temporal ISC ranking methods were applied (Malinen and Hari, 2011). Temporal ISC (t-ISC) ranking is determined by computing the mean correlation between all subject-wise IC time-series, while spatial ISC (s-ISC) ranking relies on voxel-wise ISC values and requires one to define a proper ranking measure (i.e. the order parameter). We defined the order parameter for s-ISC with a formula $\sum_i ISC_i|t_i| / \sum_i |t_i|_i$, where ISC_i is an ISC value and t_i is a t-value for IC coefficients in voxel i. Due to the chosen divisor, this measure also takes into account the differences in components sizes (i.e., large component results in large divisor). Unlike the mask-based order parameter proposed in Ref. (Malinen and Hari, 2011), our order parameter does not require statistical binarization of IC and ISC maps and therefore allows a threshold-free sorting. Statistical significance of s-ISC and t-ISC values were compared against null distributions computed via 100,000 permutation iterations, resulting in positive-tail percentile thresholds (p-values). Each iteration included picking an IC randomly, resampling the data and computing s-ISC and t-ISC measures. For s-ISC, t-maps were spatially permuted voxel-wise inside the brain, and for t-ISC, subject-wise time-series were randomly circularly shifted. Those ICs that have a high ranking in both s-ISC and t-ISC are likely to represent 'extrinsic' activity.

IC visualization based on Isomap

To visualize the ICs and their relation to the different features of the film content, the annotation time-series and ICs' time-series were compared qualitatively using Isomap algorithm. Isomap is a non-linear dimension reduction technique that determines geodesic distance metrics between high-dimensional data-points (Tenenbaum et al., 2000). Unlike spatial ICA, the Isomap dimension reduction was applied on temporal dimension (435 timepoints), which was reduced into two. In order to reduce noise and number of data-points, IC time-series were first averaged over subjects, before all the time-series were standardized.

Full-model linear regression

To relate the brain regions (ICs; ROIs) with the annotated content features, the full-model ordinary least squares (OLS) linear regression analysis was applied. Using a set of 37 annotation time-series, the design matrix of 435 time-points and 37 regressors (features) was formed. All regressors were demeaned and, in order to set the scale, their maximum values were set to unity. Before the regression, the pair-wise Pearson correlations between all annotated time-series (total 666) were computed in order to reveal their mutual similarity. To quantify the severity of collinearity in OLS regression, the variance inflation factor (VIF) was computed for each regressor to identify those most affected by correlations (Hair, 2009). For VIF value 1 no correlations exists, while values over 5 or so indicate problems (see discussion in O'brien, 2007). If several regressors are correlated, the situation is known as *multicollinearity* (Andrade et al., 1999; Gantmacher, 1984). In such situations the OLS solution may become unstable, i.e., small changes in the data can cause large changes in the fit coefficients that are, thus, no longer reliable measures of the importance of regressors. Furthermore, if the number of relevant regressors for the measured data is actually smaller than the applied full set, the OLS fit tends to become overfitted (i.e., also models noise). In order to address these issues, in addition to the full-model OLS regression, we applied the elastic-net regularization and partial least squares regularization methods with the cross-validation scheme (Arlot and Celisse, 2010; Hastie et al., 2009).

Elastic-net regression

To address the issues of multicollinearity and overfitting, the elastic-net regularization with the cross-validation scheme was applied. By adding penalty function to the magnitude of OLS fit coefficients, the elastic-net regularization algorithm tends to leave out unnecessary features and is known to perform well with correlated regressors (Zou and Hastie, 2005). In short, when given the full design matrix and the data, the elastic-net algorithm returns the fit coefficients as a function of a real-valued regularization parameter λ. For $\lambda = 0$ no regularization is applied, while positive values lead to smaller models by setting some coefficient to zero.

We applied Glmnet, which offers a fast implementation of the elastic-net (Qian et al., 2013). For the main parameter α in [0,1], which sets the ratio between the *ridge* ($\alpha = 0.0$) and the *least absolute shrinkage and selection operator* (LASSO) regression ($\alpha = 1.0$), we used 0.80. Further, for the cross-validation we used 10-folds (i.e. 90% of data for training and 10% for testing) with 25 different randomized data-partitioning (Monte Carlo steps). This procedure generated a mean squared error (MSE) for each model order. The final model order was then determined within one standard error away from the global minimum of the MSE curve towards the smaller model order (known as a "one-standard error" rule, Hastie et al., 2009).

The elastic-net regression was applied to all $12 \times 40 = 480$ subject-wise IC time-series. After regularized feature selection, we computed OLS fit and Pearson correlation coefficient between IC and fitted time-series. For each IC, statistical significance of the mean temporal correlation coefficient over subjects was tested against bootstrapped correlation coefficients, resulting in approximated p-values. Bootstrapped correlation values were obtained by (1) picking a random IC, (2) randomly phase-mixing its time-series (same mixing operation for all subjects), (3) running Glmnet to find an optimal design matrix of a given size (determined previously for the unmixed time-series), and finally (4) computing the mean Pearson correlation coefficient over subjects between the IC and fitted time-series. This was repeated 20,000 times to build null-distributions, separately for each IC. Before computing the mean correlation, the Fisher transformation (*arctanh* function) was applied to all correlation coefficients. This approach is similar to that applied by Lahnakoski et al. (2012b) to create statistics for un-regularized OLS linear regression analysis, however, here due to regularization the design matrices varied across ICs and subjects, and hence separate distributions were needed for each IC. Note that by fixing the model size to 37 in step 3 (i.e., using the full design matrix), resulting statistics are reduced to those applied in (Lahnakoski et al., 2012b). Also, here we applied phase mixing instead of simpler shift-permutations where number of permutations is limited. In phase mixing, the phase of each frequency component is randomized in frequency domain and the randomized time-series is then transformed back to the time domain and therefore the number of available permutations is very large.

The same statistical procedure was applied in ROI-analysis (see below), except that the number of permutations was reduced to 7,000 per ROI and Monte Carlo steps to 15 in order to reduce computational cost. Those ICs and ROIs with a mean correlation coefficient high enough to surpass $p < 0.05$ FDR-corrected (Benjamini and Hochberg, 1995) against the null-distributions were declared statistically significant.

Partial least squares regression

To compare the application of elastic-net regularization with another similar regularization method, partial least squares (PLS) regression with the SIMPLS algorithm (de Jong, 1993) was applied. The PLS regression finds linear transformations that maximize the covariance between the design matrix and the response (Wold et al., 2001). Finding a proper model order, i.e., the number of components of the transformed design matrix to use is the key step. The optimal model order was selected following two empirical rules. Firstly, an optimal order was chosen close to the global minima of the MSE curves obtained via cross-validation. In order to compensate inaccuracies of MSE curves and to favor smaller model sizes in those cases where MSE curve had a long flat tail, a small deviation from the global minimum was allowed (up to 5% of the difference 'max(MSE)-min(MSE)'). For the cross-validation, 10-folds with 25 Monte Carlo steps were used. Secondly, the model order was bound between minimum 50% and 95% of the total variance explained by the full model. The second step was added as a fail-safe mechanism for situations when the cross-validation scheme, which contains a random element, would fail to produce a well-defined global MSE minimum (i.e., MSE curve had a complicated shape). The above two-step approach was set-up and verified manually with a few dozen MSE curves computed for IC (and later ROI) time-series. After determining the model order for each subject and time-series with the above procedure, statistical significance testing of OLS fitted time-series proceeded similarly as for the elastic-net method.

Full-brain ROI set

To compare the information gained by the application of the full and regularized regressions to ICs, the three regression methods were applied to a large set of full-brain cubic ROIs (edge length 6 mm). All voxels located outside the mask of grey-matter template and those with poor signal strength from one or more subjects were removed. For each subject, time-series were computed for each of the resulting 5346 ROIs with their volumes ranging between 120 and 216 mm^3 (i.e., 15–27 normalized voxels).

Results

Multicollinearity of annotations

We calculated variance inflation factor (VIF), which shows to what extent each individual annotation is affected by the multicollienarity (Fig. 1). It clearly distinguishes annotations that relate to cinematographic methods (e.g., camera, framing) with higher values above the red line (VIF = 5) and annotations that relate to character's bodily actions or aspects (e.g., feet, crawling body) below the red line. As most camera and motion related aspects could naturally co-exist in the same time-point (e.g. CAM_pan + FRAME_wide + CAM_high + MOV_all), in turn, many single-actor related aspects cannot co-exist, for instance, the main character cannot both run and stand simultaneously

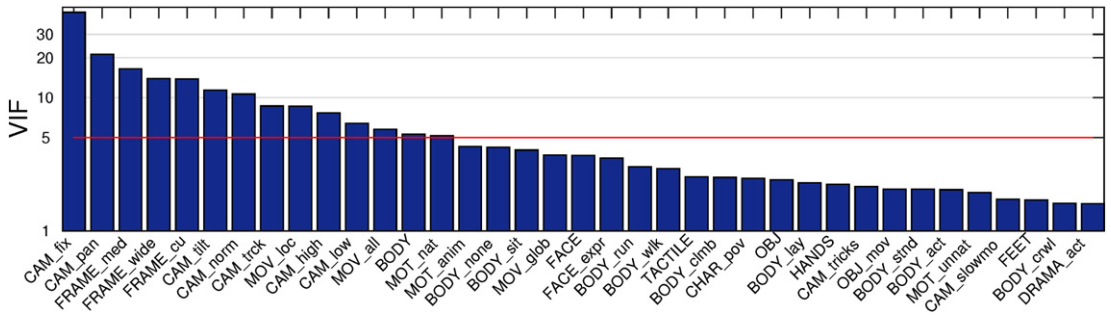

Fig. 1. Variance inflation factor (VIF) for 37 cinematic features used in the data-analysis. Red line shows VIF = 5 as a guide for an eye.

(BODY_run + BODY_stnd), the division is somewhat expected. The relatively high VIF values result from high correlations between regressors (see Supplementary Material section 1). In addition, pre-processing (i.e., convolution, downsampling and filtering) of regressor time-series was found to slightly modify and increase correlations against the original time-series (see Supplementary Material section 1).

Regional ranking of 40 ICs

We ranked the independent components to distinct functional networks by inspection of their spatial distribution maps compared to those reported in refs. (Allen et al., 2011; Heine et al., 2012; Pamilo et al., 2012; Smith et al., 2009). The 34 out of 40 independent components could be associated with one or multiple known functional networks as follows: visual networks [ICs 3, 8, 10, 13, 14, 19, 20, 33, 34, 35, 36, 37], auditory networks [ICs 4, 7], sensorimotor networks [ICs 4, 16, 18, 21, 24, 38, 39], default mode network [ICs 3, 17, 20, 22, 27, 31, 32], cerebellum [10, 19, 34, 36, 38], anterior cingulate/precuneus network [ICs 6, 31], salience network [ICs 6, 17, 18, 25, 30], striatum [IC 11], frontotemporal network [ICs 7, 25, 26, 32], frontoparietal network [ICs 21, 23, 24, 27, 28, 39, 40], and executive control network [ICs 16, 18, 26, 27, 29, 30]. Components [ICs 1, 2, 5, 9, 12, 15] were identified as artifacts based on their spatial distribution and temporal behavior. Since the actor's physical actions create apparent continuity for the otherwise non-narrative film, we were also able to identify a set of components [ICs 21, 18, 39, 24, 13, 23] ranked according to their amount of 'mirroring' coordinates (see Supplementary Material section 2).

Ranking ICs by t-ISC and s-ISC

Our results indicate that t-ISC and s-ISC methods are equally able to distinguish the most stimulus-driven components shared inter-subjectively by all subjects. The ranking of independent components based on t-ISC and s-ISC are depicted in Fig. 2. Total 21 ICs (out of 40) had a t-ISC above 0.060 indicating a significance of $p < 10^{-5}$ (uncorrected) obtained via permutation test. Similarly, s-ISC showed that total 17 ICs had order parameter above 0.062 obtained for spatially randomized t-maps for $p < 10^{-5}$ (uncorrected). The ranking can be considered consistent as the first 16 ICs are the same for both ISC-based rankings, although in slightly different order, and are assigned to visual networks [IC 3, 8, 10, 13, 14, 19, 20, 33, 34, 35, 36, 37], default mode networks [IC 3, 20, 32], cerebellum [IC 10, 19, 34, 36, 38], sensorimotor/frontoparietal network [IC 24, 38, 39] (as in Section 3.2; visualized in Supplementary Material section 3). Most of these components with high t-ISC and s-ISC can be associated with 'extrinsic' brain functions, in accordance with Malinen and Hari (2011). For the remaining ICs, both temporal correlation and spatial overlap between ISC maps were much weaker, and thus, these components could be associated with more 'intrinsic' brain functions, i.e., revealing individual, less stimulus-driven variations between our subjects, respectively. As all 40 ICs are included in the ranking, some of these ICs are artifacts.

Regularized linear regression analysis

In the following section, we describe our results of linking the cinematic features (annotation time-series) with the fMRI brain data (independent component time-series) by a means of un-regularized

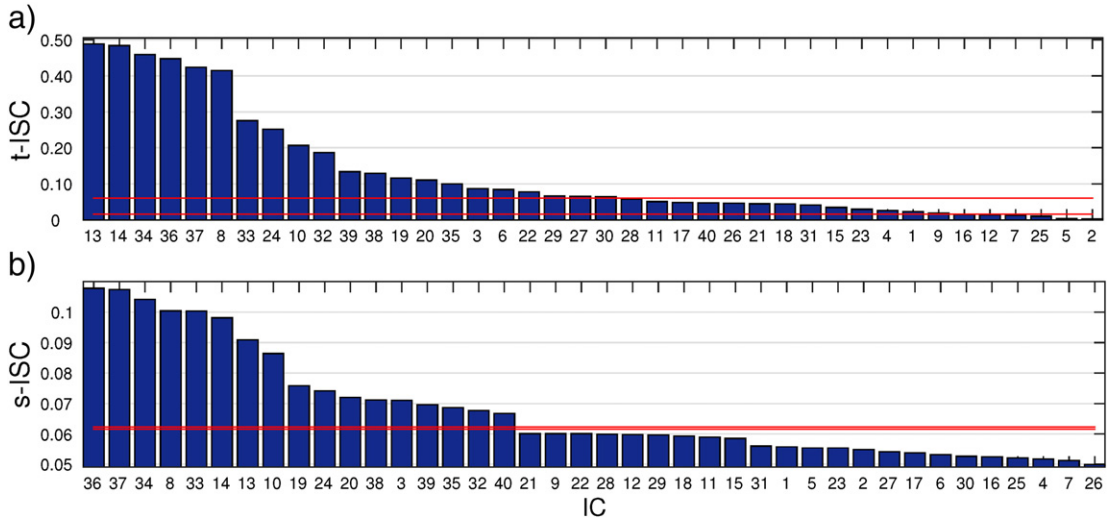

Fig. 2. ISC-based rankings of independent components. (a): Ranking ICs using temporal ISC (t-ISC). (b): Ranking ICs using the order parameter based on spatial ISC map (s-ISC). In both figures, red horizontal lines corresponds to (uncorrected) statistical significance levels $p = 10^{-5}$ (top) and $p = 0.05$ (bottom) obtained via permutation tests.

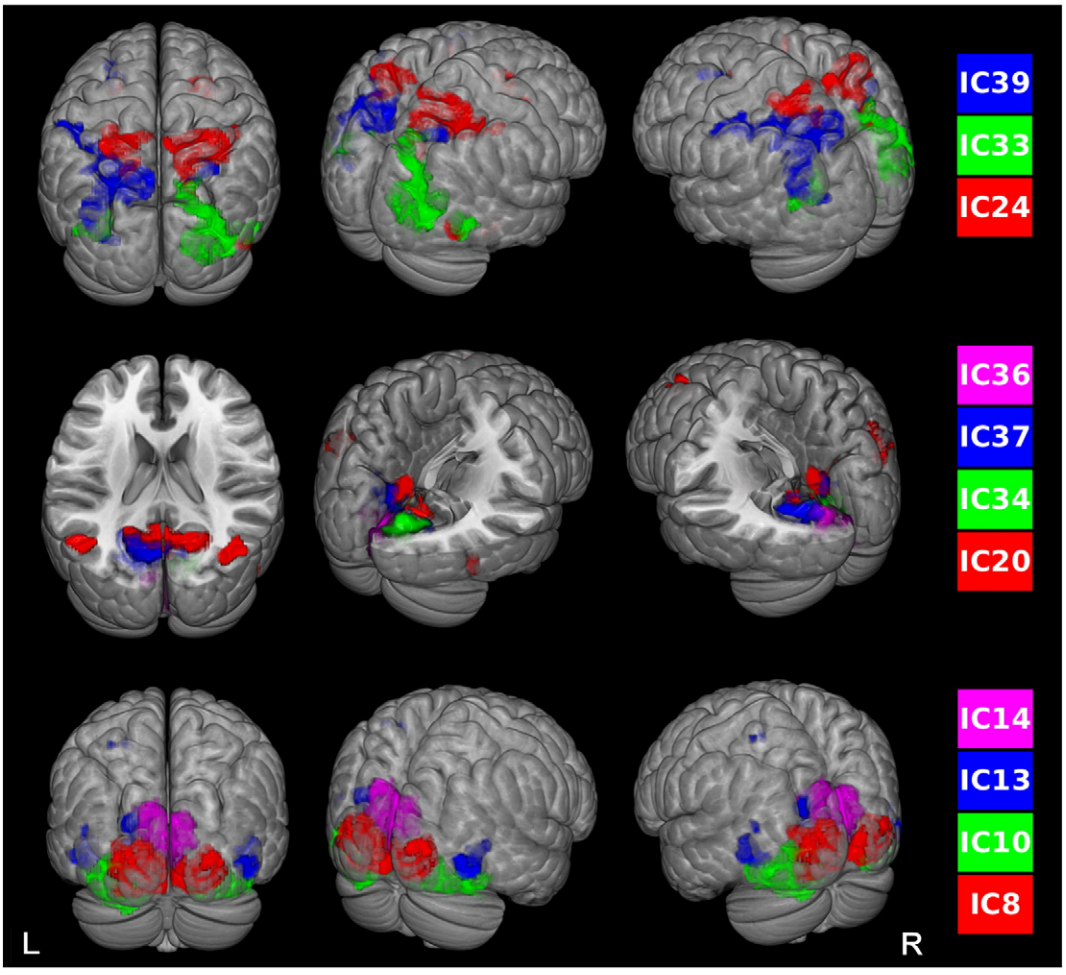

Fig. 3. The TOP-9 components identified by the elastic-net regularization method [IC 8, 10, 13, 14, 20, 24, 36, 33, 39], which also overlap with the top components identified by ISC-ranking. In addition, the figure shows two prominent ISC-ranking components that TOP-9 does not overlap [ICs 34, 37]. All IC maps are thresholded with $p < 0.05$ (FWE-corrected) and a minimum cluster size of 20 normalized voxels. Here, for the visualization purposes, the IC groups are created based on their local closeness, and do not necessarily represent functional connections.

full-model OLS linear regression and regularized PLS and elastic-net regression analysis.

Regression of independent components

The linear regression analysis was applied to the time-series of all 40 ICs using the 37 annotation regressors. For the full model, a high correlation with a statistical significance $p < 0.05$ FDR-corrected was found between annotations and IC time-series for seven ICs. Six are related mainly to visual network [ICs 8, 10, 13, 14, 33, 36] and one to sensorimotor/frontoparietal network [IC39]. Note that [ICs 10, 36] also overlap with cerebellum (as in Section Regional ranking of 40 ICs). In addition to these seven ICs, when we applied regularized PLS regression, also component assigned to sensorimotor/frontoparietal network [IC 24] was found significant. Finally, the elastic-net method turned out to be the most sensitive regularization method for this data as one more component related to visual/default mode network [IC 20] was identified, now summarizing the total number of nine statistically significant components (TOP-9). Fig. 3 shows the TOP-9 elastic-net ranking components together with two additional visual components [ICs 34, 37], which were not found significant in the regression analysis, but ranked amongst the five highest in the t-ISC and s-ISC rankings. The spatial locations for the major clusters of these ICs are summarized in Table 1 according to the Harvard-Oxford atlas (Jenkinson et al., 2012).

Mean correlation values and number of selected features (for PLS and elastic-net) are summarized in Table 2. As the mean number of chosen features for the elastic-net regression varied between 17 and 29, most of the regressors deemed relevant for the IC time-series. The problem of the full-model over-fitting becomes evident especially with IC20, which, on average, had only 17 relevant features out of 37.

The ordinary least squares (OLS) mean fit coefficients computed using the design-matrices chosen by the elastic-net method are shown in Fig. 4(a) for TOP-9 ICs. Despite the large number of selected features, only some of their corresponding coefficients were consistent over subjects. Only few regressors, e.g., 'FEET', 'BODY_wlk' and 'OBJ_mov', had both large and consistent coefficients. Apparent similarity between some coefficient vectors (e.g., IC33 vs. IC39 and IC13 vs. IC24) resulted from the fact that the time-series of those ICs were also highly correlated (data not shown), which was expected as these networks are also spatially close to each other (c.f., Fig. 3).

Fig. 4(b) depicts the two-dimensional Isomap projection of TOP-9 ICs and annotation time-series. Although distances between ICs are much smaller than those of annotations, IC20 and IC24 are somewhat different from others. While IC20 is apart from all other points, IC24 is closest to various body and action related annotations. Annotations were not clustered according to their assumed category prefix (e.g., BODY, MOV, and CAM), but instead have a more complicated relationship.

Mean IC time-series and their fitted annotation models are depicted in Fig. 5 for ICs 24, 13 and 20, illustrating the effect of the goodness of fit and temporal ISC value (t-ISC). Larger variability between subjects, corresponding to the smaller t-ISC, is noticeable for IC20 when compared against IC13 (c.f., Fig. 2a).

ICs can be also ranked based on their regression analysis p-values. Fig. 6 depicts such ranking obtained via the elastic-net method. The

Table 1
Largest clusters (up to three), peak coordinates and corresponding main cortical and subcortical regions according to Harvard-Oxford atlas (Jenkinson et al., 2012) for the selected 13 ICs (including TOP-9, presented in Section Regression of independent components). IC maps are thresholded at $p < 0.05$ FWE-corrected and a minimum cluster size of 20 normalized voxels. ICs are visualized in Fig. 3.

IC	Voxels	t-val	x	y	z	Label
8	2691	34.56	−24	−92	6	L occipital pole (34%), r occipital pole (26%)
	28	17.34	−6	−52	4	L precuneous cortex (54%) L cingulate gyrus, post. division (29%)
10	1597	23.73	−36	−62	−14	L lat. occipital cortex, inf. division (32%) L occipital fusiform gyrus (28%)
	974	29.52	40	−70	−14	R lat. occipital cortex, inf. division (37%) R occipital fusiform gyrus (30%)
13	362	18.32	48	−76	−4	R lat. occipital cortex, inf. division (99%)
	306	17.47	−46	−64	4	L lat. occipital cortex, inf. division (75%) L middle temporal gyrus, temporo-occipital part (8%)
	182	16.63	−22	−82	26	L lat. occipital cortex, sup. division (73%), L cuneal cortex (8%)
14	2017	26.5	2	−86	6	L cuneal cortex (22%), R cuneal cortex (13%)
20	1207	38.48	2	−64	16	L precuneous cortex (37%), R precuneous cortex (25%)
	635	22.42	−38	−72	38	L lat. Occipital Cortex, sup. Division (93%)
	609	24.25	38	−76	30	R lat. occipital cortex, sup. division (100%)
24	1494	25.89	18	−72	50	R sup. parietal lobule (50%) R lat. occipital cortex, sup. division (25%)
	577	25.81	−22	−54	62	L sup. parietal lobule (68%) L lat. occipital cortex, sup. division (26%)
	198	22.49	30	−8	58	R Precentral gyrus (49%), R sup. frontal gyrus (18%)
33	1755	39.49	28	−82	20	R lat. occipital cortex, sup. division (49%) R occipital pole (16%)
	350	24.18	22	−52	−16	R occipital fusiform gyrus (46%), r lingual gyrus (22%) R temporal occipital fusiform cortex (21%)
	288	23.63	−34	−78	16	L lat. occipital cortex, sup. division (80%) L lat. occipital cortex, inf. division (15%)
34	968	39.26	6	−72	4	R lingual gyrus (51%), r intracalcarine cortex (33%) L lingual gyrus (9%)
	26	16.26	8	−50	0	R lingual gyrus (85%)
36	1315	44.39	−6	−78	−2	L lingual gyrus (46%), l intracalcarine cortex (17%) R lingual gyrus (15%)
37	1778	29.56	−10	−68	4	L lingual gyrus (27%), l intracalcarine cortex (21%)
	87	13.86	22	−64	2	R lingual gyrus (71%), r intracalcarine cortex (11%)
39	2634	33.38	−28	−72	46	L lat. occipital cortex, sup. division (57%) L sup. parietal lobule (15%)
	158	24.08	22	−66	50	R lat. occipital cortex, sup. division (96%)
	85	16.79	−30	14	54	L middle frontal gyrus (68%), L sup. frontal gyrus (12%)

bottom diagram illustrates the relative order of components as obtained with the three regression methods. Apart from IC20, the TOP-9 remains the same through all the methods while major discrepancies between the ranking lists only appear for components with non-significant p-values and small correlation coefficients.

The resulting component rankings given by t-ISC, s-ISC, and elastic-net regression (c.f., Figs. 2 and 6a) were somewhat different. Hence, we further computed the mutual mean ranking for these three methods. The first eleven [ICs 36, 13, 14, 33, 8, 34, 37, 10, 24, 39, 20] included all TOP-9 components.

Regression of full-brain regions of interest

The linear regression was repeated using the un-regularized full-model OLS regression and the regularized elastic-net and PLS regressions, for the 5346 ROIs inside the grey-matter. The regressions were conducted in similar manner as with the 40 ICs. As a result, total number of ROIs with a correlation coefficient high enough to pass $p < 0.05$ (FDR-corrected) for the elastic-net was 985, PLS 887, and the full-model OLS regression 751. For uncorrected $p < 0.001$ the number of ROIs was 570, 551, and 463, respectively. Hence, with regularization, the increase in the activation detection was over 20% when compared against un-regularized regression. All ROIs surpassing $p < 0.05$ (FDR-corrected) are shown in Fig. 7. Information related to elastic-net regression for 12 selected ROIs is listed in Table 3. With the elastic-net method, the mean number of chosen features was between 14 and 31 for all 985 statistically significant ROIs.

Table 2
List of ICs whose time-courses had highest correlations ($p < 0.05$ FDR-corrected) with the OLS fitted models constructed using the elastic-net (EN) and PLS methods, and a full-feature set. For ICs 20 and 24, non-significant values (three entries) are omitted. For PLS method the number of features is the number of chosen components. Standard deviation for the number of features over all subjects is given in parenthesis. Components marked with a single or double star remain significant at $p < 0.01$ or $p < 0.001$ FDR-corrected correspondingly. For spatial locations of ICs, see Fig. 3 and Table 1.

IC	Corr. (EN)	#Feat.	Corr. (PLS)	#Feat.	Corr. (full)
8	0.56	23.8 (5.2)	0.55	6.1 (3.4)	0.58
10	0.58*	23.5 (6.3)	0.57*	5.8 (2.2)	0.6
13	0.67**	26.9 (4.1)	0.67**	6.1 (1.7)	0.68**
14	0.62**	27.8 (5.4)	0.61**	6.9 (1.4)	0.63**
20	0.49	17.3 (8.3)	-	-	-
24	0.54	24.9 (9.5)	0.54	6.1 (2.2)	-
33	0.59*	28.6 (6.7)	0.59*	8.2 (2.5)	0.61*
36	0.62**	29.1 (5.3)	0.61**	7.0 (2.4)	0.63**
39	0.58	29.2 (5.0)	0.57*	6.9 (3.0)	0.59

Discussion

As shown by previous studies discussed earlier in this paper, any emotionally engaging movie that depicts motivated characters in interaction with other people in socially relevant situations, allows intersubjective correlation between different viewers (Hasson et al., 2004; Naci et al., 2014). In contrast, we proposed a novel data-driven computational approach in order to gain new insights into the viewers' cognitive processes during their engagement with relatively complex and ambiguous film, Maya Deren's 'At Land' that was intentionally designed to be non-narrative and not emotionally engaging. While even such an experimental art-film synchronizes viewers' brain activation to some extent, the underpinning neural dynamics may be rather different from those of story-driven films. This may partially explain the observation of Hasson et al. (2008a) that random surveillance footage would not allow as high intersubjective correlation as that of narrative film.

Our main results indicate that a large set of annotated movie features can successfully be regressed against ICs and ROI-based time-series using full and regularized (elastic-net and partial least squares) linear regression methods. We also demonstrated the usefulness of combining ISC and ICA methods for sorting ICs that most probably relate to the extrinsic, shared cognitive experience of viewing the film from those that relate to more idiosyncratic intrinsic cognitive processes (Malinen and Hari, 2011). In line with the chosen data-driven approach, non-parametric permutation testing was applied to validate statistical significance of the results.

The elastic-net regularized linear regression was proven the most sensitive method in (1) finding information from the data both with ICs and ROIs, (2) covering wider areas in the posterior brain areas, but also (3) revealing more clusters in the anterior brain areas, in comparison to un-regularized regression. As also PLS method provided more information, similar to elastic-net, it seems that adding regularization detects more time-dependent linkages between stimulus content and brain data than conventional un-regularized linear regression. To our best knowledge, this is the first detailed study of applying and

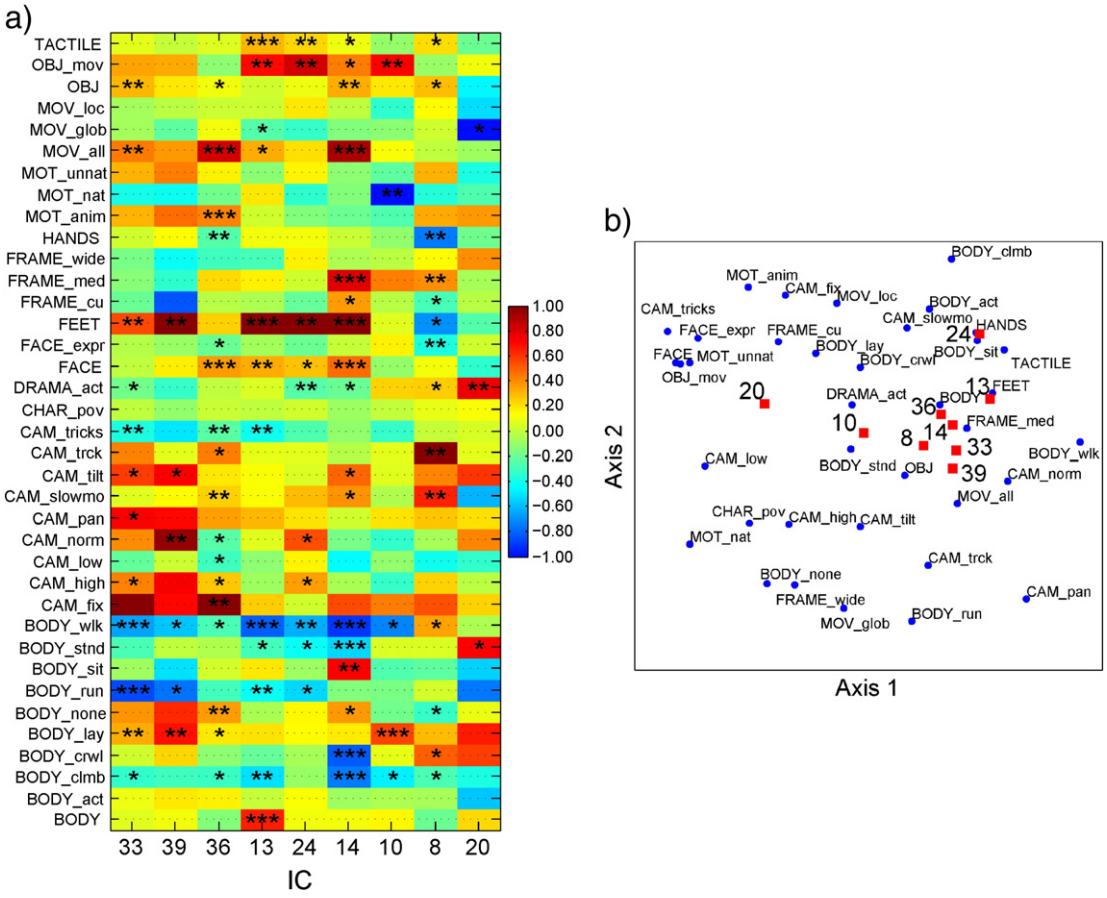

Fig. 4. (a): Mean linear OLS fit coefficients for ICs shown in Table 2 for the elastic-net method. The color indicates the magnitude and the sign of the mean fit coefficient, while stars indicate the consistence over subjects as measured via two-tail t-test. Number of stars from 1 to 3 indicates $p < 0.05$, 0.01 and 0.001 FDR-corrected. All coefficients are scaled column-wise between interval [-1, 1] using the maximum absolute value. Note that since each column (IC) has a different scaling coefficient, the row-wise values can be different even if the coloring remains identical. (b): Two-dimensional Isomap projection of all 37 pre-processed annotations (blue circles) and 9 ICs temporally averaged over subjects (red squares). The closer the points are spatially, the more similar their time-series are according to the two-dimensional Isomap projection.

comparing regularization in the analysis of the correlations between the fMRI data and a set of multicollinear cinematic annotations of movie content.

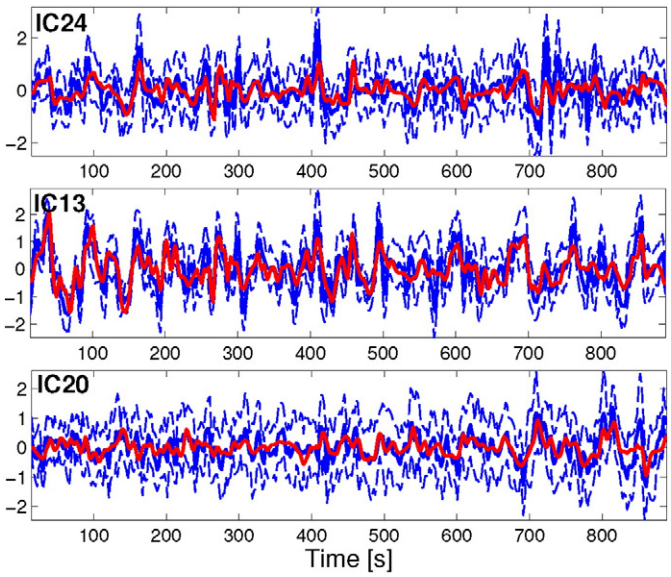

Fig. 5. Mean IC time-series (solid blue) and linear fit models (red) for ICs 24, 13 and 20. Dashed blue lines indicate one standard deviation limits over subjects. The mean correlation coefficients are 0.54 (IC24), 0.67 (IC13) and 0.49 (IC20).

Increasing sensitivity with data-driven regression

In order to relate the annotated content of the film 'At Land' with the activation in the observed brain regions (ICs, ROIs), we used 37 annotations as a set of regressors (for a list of annotations, see Appendix A). Multicollinearity and presence of possible non-relevant regressors entailed a serious risk of overfitting and corruption of fit coefficients when using OLS regression. Therefore, to address these issues, we employed both elastic-net and PLS regression schemes. The statistical significance of each regression was tested with a novel non-parametric permutation approach. The increased number of significant components indicated higher sensitivity of the introduced elastic-net method (9 vs. 8 for PLS and 7 for the full model). For the elastic-net method, mean number of chosen annotations ranged from 17 to 29.

The above findings were also supported by the full-brain ROI regression analysis. Significant activation was found in regions corresponding to previous TOP-9 components (Fig. 7) with additional smaller clusters found in the frontal pole (Table 3). The elastic-net regression approach was again the most sensitive of the three (i.e., the number of significant ROIs increased by one-fifth compared to the un-regularized regression). The difference between the PLS and elastic-net methods is likely related to the difficulty with PLS method in choosing the proper model order.

Compared to the standard OLS fit with all regressors included, our flexible data-driven regression increased sensitivity with multicollinear regressors. As a direct result of decreased model size, the regressors' fit coefficients are more reliable, which aids in their interpretation. We find the proposed elastic-net regularization approach especially useful in

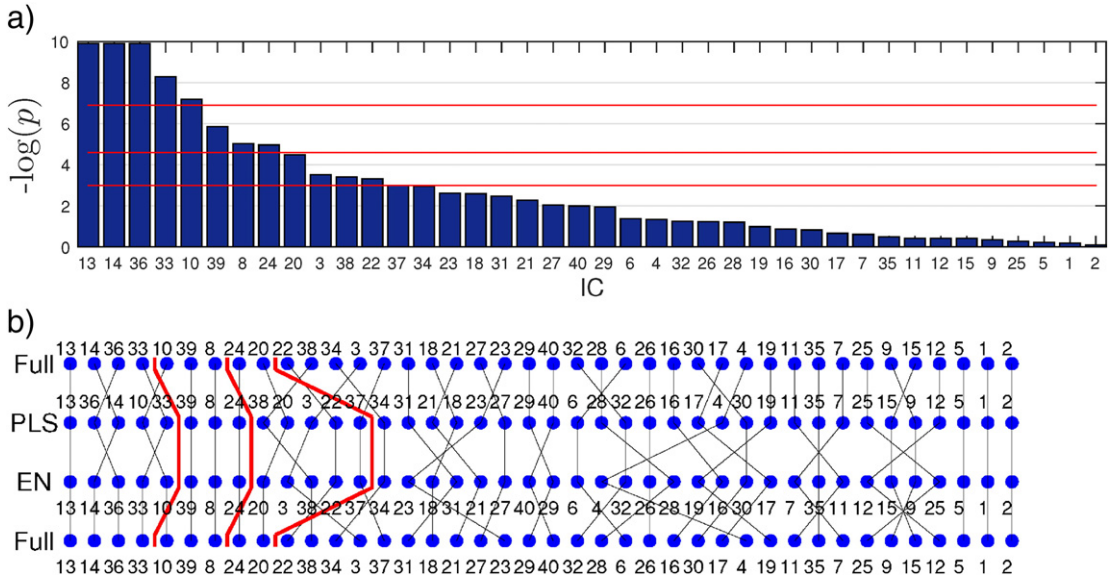

Fig. 6. Ranking ICs according to their linear OLS regression *p*-values. (a): Ranking according to the elastic-net method, red lines indicate uncorrected *p*-values 0.001 (top), 0.01 (middle), 0.05 (bottom). (b): Diagram of component order with all three methods; full model (Full, shown twice for easier comparison), elastic-net (EN) and partial least-squares (PLS). Black lines connect the same components and red lines indicate levels of (uncorrected) *p*-values 0.001 (left), 0.01 (middle), 0.05 (right).

fMRI studies utilizing naturalistic stimulus, e.g., movies, which include various simultaneously occurring and overlapping features.

Regularization has recently gained interest in fMRI community and has been applied in various context, such as anatomical alignment (Xu et al., 2012), classification and decoding (Kauppi et al., 2011; Li et al., 2009; Lorbert and Ramadge, 2013; Ng and Abugharbieh, 2011; Toiviainen et al., 2013) and functional connectivity (Ryali et al., 2012; Xu et al., 2013). When compared to standard general linear model approach, regularization techniques have been found to increase the sensitivity (Carroll et al., 2009; Li et al., 2010; McIntosh et al., 2004). In particular, regularized regression has been applied to manage thousands of annotations in order to construct cortical activation maps (Çukur et al., 2013; Huth et al., 2012; Mitchell et al., 2008; Naselaris et al., 2009; Nishimoto et al., 2011). In comparison, here we concentrated on smaller set of higher-level features. Our results demonstrated that regularization becomes useful already at the early stage of process (i.e., only tens of annotations). Naturally, if the number of annotations is counted in thousands, regularization becomes a practical necessity. The computational cost of applying regularization in fMRI analysis in voxel-wise fashion could be prohibitive, but combining regularization with ICA or ROI-analysis as done in our study makes it computationally inexpensive to apply.

Linking ICs to annotations

In accordance with previous free-viewing studies of movies (Bartels et al., 2008; Hasson et al., 2004; Jääskeläinen et al., 2008; Lahnakoski et al., 2012b) our top-ranked ICs were mainly located in the occipital and parietal regions. The distribution of the coefficients between ICs and the annotation features, shown in (Fig. 4a), revealed certain selectivity, however, no strict selectivity between ICs and annotation categories (e.g., BODY, CAM, FRAME, or MOT, see Appendix A) was found. This was also evident from the two-dimensional projection of annotations and TOP-9 components (Fig. 4b). In this article our specific focus has been on methods, and thus more detailed analysis of ICs and their linkage to movie content is left to another publication. In the following, a selection of components are only briefly discussed in relation to the annotations (Fig. 4a) and based on their corresponding cortical regions (Table 1).

Bartels et al. (2008) have previously pointed out the specificity of V5/MT+ and medial posterior parietal cortex (mPPC) to local motion and global motion, respectively, when subjects are freely viewing movie sequences. We were interested in seeing, if our ICs showed similar type of selectivity. A related set of TOP-9 components was comprised by ICs 13, 24 and 33 which all included the tentative V5/MT+ area (Huk et al., 2002): ICs 24 and 33 in the right hemisphere

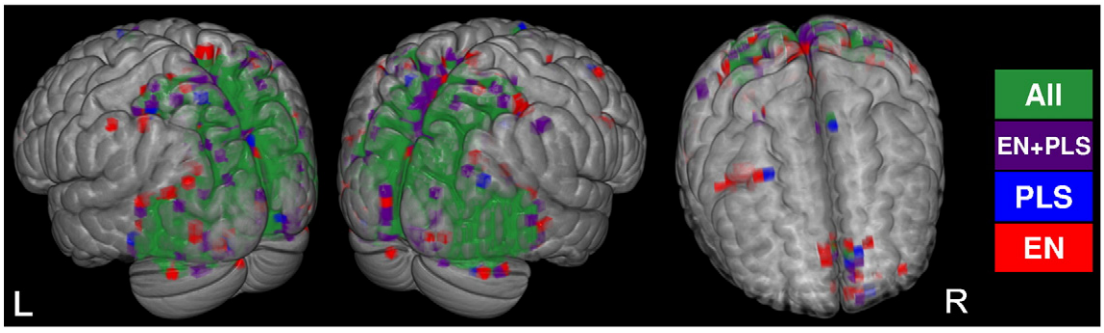

Fig. 7. A map of cubic ROIs whose time-courses significantly ($p < 0.05$, FDR-corrected) correlated with linear regression models created by elastic-net (red), PLS (blue), and full-model OLS regression (overlaps with the other two, no separate color area). The green shows overlap of all regression methods. The violet indicates overlap between PLS and elastic-net methods. Notably, while the regularization methods cover wider areas in the posterior parts of the brain, they also reveal more clusters in the anterior brain areas, in comparison to un-regularized OLS method.

Table 3
List of largest statistically significant ($p < 0.05$ FDR-corrected) ROI-clusters sorted by the cluster size (normalized voxels) with mean correlation and feature count for selected peak ROIs (standard deviation in parenthesis). Results are for the elastic-net regression analysis. Stars indicate statistical significance of $p < 0.001$ FDR-corrected.

Voxels	x	y	z	Corr.	#Feat.	Label
24418	−22	−52	66	0.54**	20.5 (5.7)	L sup. parietal lobule
	28	−50	62	0.57**	25.9 (7.6)	R sup. parietal lobule
	0	−82	4	0.60**	25.2 (5.5)	L intracalcarine cortex
	50	−66	2	0.64**	26.1 (5.1)	R lat. occipital cortex, inf. division
	−46	−74	4	0.61**	26.1 (5.9)	L Lateral occipital cortex, inf. division
274	4	62	20	0.51	20.7 (8.7)	R frontal pole
196	50	8	32	0.49	21.2 (11.2)	R precentral gyrus
162	0	−58	10	0.51	21.9 (9.4)	L precuneous cortex
147	−10	58	−16	0.52	18.1 (9.0)	L frontal pole
52	56	−42	50	0.52	22.8 (7.5)	R Supramarginal gyrus, post. division
49	0	12	68	0.48	23.8 (8.5)	L sup. frontal gyrus
44	32	−4	66	0.39	19.0 (10.2)	R precentral gyrus

and IC 13 bilaterally. IC13 in the tentative V5/MT+ showed sensitivity to feet, moving body, but also to annotated global motion. IC33 was sensitive to moments where all elements in the frame are moving, remarkably the camera. This IC33's pronounced selectivity to camera movements, e.g. panning, suggests involvement in the processing of optic flow, which is the property of MST subarea of MT+ (Dukelow et al., 2001; Huk et al., 2002). For the other two ICs, 13 and 24, mainly tactile, feet, face, and various body-related annotations were found particularly relevant suggesting their inclusion of MT area of MT+. Right-side lateralization of IC24 and 33 with the mute movie as stimuli is in line with previous fMRI studies on perception of biologic movement of hands, eyes, and face in the occipitotemporal regions (Pelphrey et al., 2005).

Notably, although IC34 and IC37 hold high positions in both ISC-based rankings, they fall outside statistical significance in the regression analysis. This was expected as these ICs are located at early visual cortical regions and our regressors only contain cinematic features related to higher perceptual aspects (e.g., body parts, movements, or actions) while lacking low-level visual ones (e.g., salience, or contrast).

Intriguing enough, none of highest ICs in our ranking list cover, for instance, posterior superior temporal sulcus (pSTS), which has been related to socially valid action detection (Saxe et al., 2004), and to social contexts observed in story-driven narrative films (Lahnakoski et al., 2012a). An attractive explanation is that the ambiguous content of 'At Land' simply failed to elicit similar kind of synchronized pSTS activation as in films that elicit strong emotions. After all, according to the director herself the film was deliberately designed to *not* elicit emotions (Deren, 2005).

Limitations of the work

We found significant correlation between the annotations of the cinematic features and fMRI data with most of the activated areas in the occipital and parietal lobes. However, against our expectations, only smaller clusters in the anterior parts of the brain were detected. Although our annotations included – in addition to a set of bodily features – conceptually defined cinematic features (e.g., dramatic points), we failed to link these features to anterior cognitive processes. We tend to explain this by limitations of the current analysis. Perhaps, a simple linear relationship between BOLD signal and the time-locked annotation model lacks flexibility to count for multiplicity of dynamically varying conceptually defined content elements (in comparison to observed bodily features). In addition, to simplify the analysis and interpretation of fit coefficients, a standard HRF with 5 s lag was assumed for all subjects and brain regions. Yet, such standard assumptions by nature introduce oversimplifications to the analysis process. One solution to enhance the regression analysis would be to reduce the noise with additional pre-processing steps as suggested by Power et al. (2014).

Several established methods deal with multicollinearity in regression analysis, e.g., LASSO, support and sparse relevance vector machines, as well as random forest models (Acharjee et al., 2013; Dormann et al., 2013; Hastie et al., 2009; Tibshirani, 2011). A typical method is to orthogonalize the design matrix into a reduced number of regressors (e.g., via PCA with Varimax rotation) before regression (Alluri et al., 2012). Often a small number of such orthogonal regressors can explain most of the variance while the rest can be omitted. When applied to our 37-feature model, this approach, however, deemed impractical since 24 PCA regressors were required to cover 95% of the variance and these regressors were too complicated to allow a meaningful interpretation.

To estimate an optimal model order for elastic-net and PLS, a standard randomized cross-validation approach was applied as implemented in Glmnet. However, this does not take into account the autocorrelation that exists in BOLD-based signals; hence the training and testing datasets are not fully independent. This may lead to non-optimal model order. However, (1) since the similar autocorrelation exists for all ICs and regressors, (2) large number of random partitioning are used, and (3) we are only interested in the model size (not MSE values themselves), this issue was not deemed crucial.

One of many challenges in naturalistic neuroscience relate to the management of complexity of the collected data. The computational demand significantly increases when applying regularization, in comparison to the full-model linear regression. The increase was observed in our study even with the relatively small number of regressors. Consequently, the dimension reduction in ICA or ROI-based analysis is especially useful, when these methods are applied to large amounts of annotation data.

Despite the increase in computational demand, in our view, the more refined, time- and context-dependent non-binary annotations should be used to model the stimulus content. In the current study, the tactile annotation data created in real-time by a group of viewers during watching the film 'At land' represented such annotation data. In the future studies, the regression methods should be tested with content annotations that model the effect of memory decay and anticipation at each given moment of narrative nowness (Kauttonen et al., 2014).

Conclusions

Explorative data-driven analysis methods, such as independent component analysis (ICA), have been shown useful for studying the functional MRI data collected while subjects are freely viewing movies. In this study, we applied a data-driven method also for the annotated features of an experimental film 'At Land' in order to analyze their relation to the fMRI data. Adaptive regression with elastic-net method was found more sensitive than a full or partial least-squares regression. The computational cost of applying regularization in fMRI analysis in voxel-wise fashion could be prohibitive, but combining regularization with ICA or ROI-analysis as done in our study makes it computationally inexpensive to apply. We find the proposed elastic-net regularization approach especially useful in fMRI studies utilizing naturalistic stimulus, e.g., movies, which include various simultaneously occurring and overlapping features. Dealing with two complex systems, such as the human brain and a naturalistic movie stimulus, no single analysis method currently available is enough to give the full picture. The set of methods applied here provided multiple windows to how the brain operates when engaged in movie viewing.

Acknowledgments

We thank the film editor Jelena Rosic for annotating 'At Land' and Marita Kattelus for assistance in the fMRI scanning. We also thank Elina Pihko and Eero Smeds for providing us with the tactile annotation

data, and Juha Lahnakoski and Miika Koskinen for valuable comments. The fMRI data was measured at the AMI Centre Aalto NeuroImaging and computational resources were provided by Science-IT, Aalto University School of Science. The research was funded by aivoAALTO and Aalto Starting Grant for NeuroCine at the Aalto University.

Appendix A. List of annotations

Detailed descriptions of all 37 annotations used in the study. The corresponding abbreviations are given in parenthesis if applicable, as well as the number of annotated events in each category. The film has a total of 147 shots, i.e., uninterrupted series of frames separated with cuts.

1. BODY (49): The category describes visibility of the whole body, parts of the body, or several bodies in the frame.
2. FEET (49): The category describes visibility of feet in the frame.
3. HANDS (23): The category describes visibility of hands in the frame.
4. FACE (78): The category describes visibility of all faces in the frame.
5. BODY SITUATEDNESS (BODY_sit; 62): The category describes appearance of the body, or parts of the body interacting with objects or surroundings.
6. FACE EXPRESSIONS (FACE_expr; 68): The category describes basic emotions and recognizable facial expressions.
7. CAMERA ANGLE NORMAL (CAM_norm; 29): The camera angle category describes a normal eyelevel point in space from which the camera observes an event.
8. CAMERA ANGLE LOW (CAM_low; 19): The camera angle category describes a low point in space from which the camera observes an event happening above it.
9. CAMERA ANGLE HIGH (CAM_high; 30): The camera angle category describes a high point in space from which the camera observes an event happening below it.
10. CAMERA TRICKS (CAM_tricks; 8): The category describes any unnatural appearance, motion, lighting, as well as camera and editing tricks, such as motion backward, or stop motion.
11. CHARACTER'S POINT-OF-VIEW (CHAR_pov; 27): The category describes images linked to the main character's point-of-view, i.e., what the character sees.
12. BODY FRAMED IN WIDE SHOT (FRAME_wide; 25): The category describes moments where the full body is visible in the frame.
13. BODY FRAMED IN MEDIUM SHOT (FRAME_med; 48): The category describes moments where the upper body is visible in the frame.
14. BODY FRAMED IN CLOSE-UP SHOT (FRAME_cu; 68): The category describes moments where only face or other body details are visible in the frame.
15. BODY ACTION (BODY_act; 9): The category describes significant bodily actions between characters, objects, and space (e.g. starts climbing, crawls through the whole, or grabs a chess piece).
16. DRAMA ACTION (DRAMA_act; 17): The category describes significant time points for the dramatic story development (e.g. introducing a new character, new space, or surprising action).
17. BODY LYING (BODY_lay; 13): The category describes someone laying or leaning.
18. BODY WALKING (BODY_wlk; 18): The category describes someone walking.
19. BODY STANDING (BODY_stnd; 10): The category describes someone standing.
20. BODY RUNNING (BODY_run; 11): The category describes someone running.
21. BODY CRAWLING (BODY_crwl; 3): The category describes someone crawling.
22. BODY CLIMBING (BODY_clmb; 5): The category describes someone climbing.
23. OBJECTS (OBJ; 29): The category describes the temporal duration of object appearances.
24. OBJECT MOVING (OBJ_mov; 7): The category describes objects in movement annotated over the temporal duration of their appearance.
25. CAMERA FIXED (CAM_fix; 40): The category lists images, in which the camera is mounted on a tripod and doesn't move from a single position or frame size.
26. CAMERA TILT (CAM_tilt; 11): The category lists images, in which the camera is turned vertically up and down from a single position (as opposed to moving the whole camera up and down).
27. CAMERA TRACKING (CAM_trck; 19): The category lists images, in which the camera is mounted on a cart travelling on tracks.
28. CAMERA PAN (CAM_pan; 34): The category lists images, in which the camera is turned horizontally left and right from a single position.
29. CAMERA SLOW-MOTION (CAM_slowmo; 6): The category lists images, in which slow motion effect is applied.
30. MOTION ANIMATED (MOT_anim; 6): The category lists images, in which natural motion is changed with camera or editing techniques (e.g. the sea waves rolling backwards).
31. MOTION NATURAL (MOT_nat; 9): The category lists images, in which the motion is natural, including living beings, or objects.
32. MOTION UNNATURAL (MOT_unnat; 6): The category lists images, in which natural motion appears in unnatural context (e.g. reflection on the face).
33. MOVEMENT LOCAL (MOV_loc; 39): The category lists movement, in which camera is still and object moves in relation to the frame.
34. MOVEMENT GLOBAL (MOV_glob; 12): The category lists movement, in which camera moves but the object remains still in relation to the frame (camera follows the object or entity).
35. MOVEMENT ALL (MOV_all; 26): The category lists kinesthetic movement, in which camera moves and the object moves in relation to the frame.
36. NO BODY (BODY_none; 23): The category describes moments without people.
37. TACTILE motion: This category describes the strength of tactile experience averaged over separate annotations of 16 subjects. It gives information of how strong or weak the character's tactile experiences on the film are rated by several viewers. The rating data were collected after the neuroimaging experiment using an in-house built tool that allows subjects continuously annotate while viewing the movie on computer screen. In tactile annotation task, after viewing the film without any task, the participants were instructed as follows (in Finnish): “You will now view the previously seen black-and-white short film ‘At land’ by Maya Deren again. The film may occasionally elicit strong impressions of tactile sensing in the viewer. During viewing the film, we ask you to evaluate, how strongly you experience the film particularly by tacit sensing. The evaluation happens by moving the mouse up and down along the vertical line marked at the right side of the screen. When doing this, the arrow visible on the line will move accordingly. When your experience of tacit sensing is weak, move the arrow downwards, and when the experience is strong, move the arrow upwards”.

Appendix B. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2015.01.063.

References

Abdollahi, R.O., Jastorff, J., Orban, G.A., 2013. Common and segregated processing of observed actions in human SPL. Cereb. Cortex 23, 2734–2753. http://dx.doi.org/10.1093/cercor/bhs264.

Acharjee, A., Finkers, R., Visser, R.G., Maliepaard, C., 2013. Comparison of regularized regression methods for ~Omics data. Metabolomics 3, 126. http://dx.doi.org/10.4172/2153-0769.1000126.

Allen, E.A., Erhardt, E.B., Damaraju, E., Gruner, W., Segall, J.M., Silva, R.F., Havlicek, M., Rachakonda, S., Fries, J., Kalyanam, R., Michael, A.M., Caprihan, A., Turner, J.A., Eichele, T., Adelsheim, S., Bryan, A.D., Bustillo, J., Clark, V.P., Feldstein Ewing, S.W., Filbey, F., Ford, C.C., Hutchison, K., Jung, R.E., Kiehl, K.A., Kodituwakku, P., Komesu, Y.M., Mayer, A.R., Pearlson, G.D., Phillips, J.P., Sadek, J.R., Stevens, M., Teuscher, U., Thoma, R.J., Calhoun, V.D., 2011. A baseline for the multivariate comparison of resting-state networks. Front. Syst. Neurosci. 5, 2. http://dx.doi.org/10.3389/fnsys.2011.00002.

Alluri, V., Toiviainen, P., Jääskeläinen, I.P., Glerean, E., Sams, M., Brattico, E., 2012. Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. Neuroimage 59, 3677–3689. http://dx.doi.org/10.1016/j.neuroimage.2011.11.019.

Andrade, A., Paradis, A.L., Rouquette, S., Poline, J.B., 1999. Ambiguous results in functional neuroimaging data analysis due to covariate correlation. Neuroimage 10, 483–486. http://dx.doi.org/10.1006/nimg.1999.0479.

Arlot, S., Celisse, A., 2010. A survey of cross-validation procedures for model selection. Stat. Surv. 4, 40–79. http://dx.doi.org/10.1214/09-SS054.

Bartels, A., Zeki, S., 2004a. Functional brain mapping during free viewing of natural scenes. Hum. Brain Mapp. 21, 75–85. http://dx.doi.org/10.1002/hbm.10153.

Bartels, A., Zeki, S., 2004b. The chronoarchitecture of the human brain–natural viewing conditions reveal a time-based anatomy of the brain. Neuroimage 22, 419–433. http://dx.doi.org/10.1016/j.neuroimage.2004.01.007.

Bartels, A., Zeki, S., Logothetis, N.K., 2008. Natural vision reveals regional specialization to local motion and to contrast-invariant, global flow in the human brain. Cereb. Cortex 18, 705–717. http://dx.doi.org/10.1093/cercor/bhm107.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B 57, 289–300.

Brugman, H., Russel, A., 2004. Annotating multimedia/multi-modal resources with ELAN. 4th International Conference on Language Resources and Evaluation. Max Planck Institute for Psycholinguistics. The Language Archive, pp. 2065–2068.

Calhoun, V.D., Adali, T., 2012. Multisubject independent component analysis of fMRI: a decade of intrinsic networks, default mode, and neurodiagnostic discovery. IEEE Rev. Biomed. Eng. 5, 60–73. http://dx.doi.org/10.1109/RBME.2012.2211076.

Calhoun, V.D., Adali, T., Pearlson, G.D., Pekar, J.J., 2001. A method for making group inferences from functional MRI data using independent component analysis. Hum. Brain Mapp. 14, 140–151. http://dx.doi.org/10.1002/hbm.1048.

Carroll, M.K., Cecchi, G.A., Rish, I., Garg, R., Rao, A.R., 2009. Prediction and interpretation of distributed neural activity with sparse models. Neuroimage 44, 112–122. http://dx.doi.org/10.1016/j.neuroimage.2008.08.020.

Caspers, S., Zilles, K., Laird, A.R., Eickhoff, S.B., 2010. ALE meta-analysis of action observation and imitation in the human brain. Neuroimage 50, 1148–1167. http://dx.doi.org/10.1016/j.neuroimage.2009.12.112.

Chen, H., Yao, D., 2004. Discussion on the choice of separated components in fMRI data analysis by spatial independent component analysis. Magn. Reson. Imaging 22, 827–833. http://dx.doi.org/10.1016/j.mri.2003.12.003.

Çukur, T., Nishimoto, S., Huth, A.G., Gallant, J.L., 2013. Attention during natural vision warps semantic representation across the human brain. Nat. Neurosci. 16, 763–770. http://dx.doi.org/10.1038/nn.3381.

De Jong, S., 1993. SIMPLS: An alternative approach to partial least squares regression. Chemom. Intell. Lab. Syst. 18, 251–263. http://dx.doi.org/10.1016/0169-7439(93)85002-X.

Deren, M., 2005. Essential Deren: Collected writings on film. Documentext.

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., McClean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D., Lautenbach, S., 2013. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. Ecography 36, 27–46. http://dx.doi.org/10.1111/j.1600-0587.2012.07348.x.

Dukelow, S.P., DeSouza, J.F., Culham, J.C., van den Berg, A.V., Menon, R.S., Vilis, T., 2001. Distinguishing subregions of the human MT+ complex using visual fields and pursuit eye movements. J. Neurophysiol. 86, 1991–2000.

Friston, K., Penny, W., Ashburner, J., Kiebel, S., Nichols, T., 2006. Statistical parametric mapping: The analysis of functional brain images. 1st ed.

Gantmacher, F.R., 1984. The theory of matrices. 3rd ed. Chelsea Pub Co.

Hair, J., 2009. Multivariate data analysis. Faculty Publications.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., Malach, R., 2004. Intersubject synchronization of cortical activity during natural vision. Science 303, 1634–1640. http://dx.doi.org/10.1126/science.1089506.

Hasson, U., Landesman, O., Knappmeyer, B., Vallines, I., Rubin, N., Heeger, D.J., 2008a. Neurocinematics: The neuroscience of film. Projections 2, 1–26. http://dx.doi.org/10.3167/proj.2008.020102.

Hasson, U., Yang, E., Vallines, I., Heeger, D.J., Rubin, N., 2008b. A hierarchy of temporal receptive windows in human cortex. J. Neurosci. 28, 2539–2550. http://dx.doi.org/10.1523/JNEUROSCI.5487-07.2008.

Hastie, T., Tibshirani, R., Friedman, J., 2009. The elements of statistical learning. 2nd ed. Springer.

Heine, L., Soddu, A., Gómez, F., Vanhaudenhuyse, A., Tshibanda, L., Thonnard, M., Charland-Verville, V., Kirsch, M., Laureys, S., Demertzi, A., 2012. Resting state networks and consciousness: alterations of multiple resting state network connectivity in physiological, pharmacological, and pathological consciousness States. Front. Psychol. 3, 295. http://dx.doi.org/10.3389/fpsyg.2012.00295.

Himberg, J., Hyvärinen, A., Esposito, F., 2004. Validating the independent components of neuroimaging time series via clustering and visualization. Neuroimage 22, 1214–1222. http://dx.doi.org/10.1016/j.neuroimage.2004.03.027.

Huk, A.C., Dougherty, R.F., Heeger, D.J., 2002. Retinotopy and functional subdivision of human areas MT and MST. J. Neurosci. 22, 7195–7205 (doi:20026661).

Huth, A.G., Nishimoto, S., Vu, A.T., Gallant, J.L., 2012. A continuous semantic space describes the representation of thousands of object and action categories across the human brain. Neuron 76, 1210–1224. http://dx.doi.org/10.1016/j.neuron.2012.10.014.

Jääskeläinen, I.P., Koskentalo, K., Balk, M.H., Autti, T., Kauramäki, J., Pomren, C., Sams, M., 2008. Inter-subject synchronization of prefrontal cortex hemodynamic activity during natural viewing. Open Neuroimaging J. 2, 14–19. http://dx.doi.org/10.2174/1874440000802010014.

Jastorff, J., Orban, G.A., 2009. Human functional magnetic resonance imaging reveals separation and integration of shape and motion cues in biological motion processing. J. Neurosci. 29, 7315–7329. http://dx.doi.org/10.1523/JNEUROSCI.4870-08.2009.

Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W., Smith, S.M., 2012. FSL. Neuroimage 62, 782–790. http://dx.doi.org/10.1016/j.neuroimage.2011.09.015.

Kanwisher, N., Yovel, G., 2006. The fusiform face area: a cortical region specialized for the perception of faces. Philos. Trans. R. Soc. Lond. B Biol. Sci. 361, 2109–2128. http://dx.doi.org/10.1098/rstb.2006.1934.

Kauppi, J., Huttunen, H., Korkala, H., Iiro, J., Sams, M., Tohka, J., 2011. Artificial neural networks and machine learning. In: Honkela, T., Duch, W., Girolami, M., Kaski, S. (Eds.), ICANN 2011, lecture notes in computer science. Springer, Berlin, Heidelberg, pp. 189–196 http://dx.doi.org/10.1007/978-3-642-21738-8.

Kauttonen, J., Kaipainen, M., Tikka, P., 2014. Model of narrative nowness for neurocinematic experiments. 5th Workshop on Computational Models of Narrative, pp. 77–87 http://dx.doi.org/10.4230/OASIcs.CMN.2014.77.

Lahnakoski, J.M., Glerean, E., Salmi, J., Jääskeläinen, I.P., Sams, M., Hari, R., Nummenmaa, L., 2012a. Naturalistic FMRI mapping reveals superior temporal sulcus as the hub for the distributed brain network for social perception. Front. Hum. Neurosci. 6, 233. http://dx.doi.org/10.3389/fnhum.2012.00233.

Lahnakoski, J.M., Salmi, J., Jääskeläinen, I.P., Lampinen, J., Glerean, E., Tikka, P., Sams, M., 2012b. Stimulus-related independent component and voxel-wise analysis of human brain activity during free viewing of a feature film. PLoS One 7, e35215. http://dx.doi.org/10.1371/journal.pone.0035215.

Lankinen, K., Saari, J., Hari, R., Koskinen, M., 2014. Intersubject consistency of cortical MEG signals during movie viewing. Neuroimage 92C, 217–224. http://dx.doi.org/10.1016/j.neuroimage.2014.02.004.

Large, M.-E., Cavina-Pratesi, C., Vilis, T., Culham, J.C., 2008. The neural correlates of change detection in the face perception network. Neuropsychologia 46, 2169–2176. http://dx.doi.org/10.1016/j.neuropsychologia.2008.02.027.

Li, Y., Namburi, P., Yu, Z., Guan, C., Feng, J., Gu, Z., 2009. Voxel selection in FMRI data analysis based on sparse representation. IEEE Trans. Biomed. Eng. 56, 2439–2451. http://dx.doi.org/10.1109/TBME.2009.2025866.

Li, X., Coyle, D., Maguire, L., McGinnity, T.M., Watson, D.R., Benali, H., 2010. A least angle regression method for fMRI activation detection in phase-encoded experimental designs. Neuroimage 52, 1390–1400. http://dx.doi.org/10.1016/j.neuroimage.2010.05.017.

Lorbert, A., Ramadge, P.J., 2013. The Pairwise Elastic Net support vector machine for automatic fMRI feature selection. 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 1036–1040 http://dx.doi.org/10.1109/ICASSP.2013.6637807.

Malinen, S., Hari, R., 2011. Data-based functional template for sorting independent components of fMRI activity. Neurosci. Res. 71, 369–376. http://dx.doi.org/10.1016/j.neures.2011.08.014.

Malinen, S., Hlushchuk, Y., Hari, R., 2007. Towards natural stimulation in fMRI-issues of data analysis. Neuroimage 35, 131–139. http://dx.doi.org/10.1016/j.neuroimage.2006.11.015.

Mazaika, P., Whitfield, S., Jeffrey, C., 2005. Detection and repair of transient artifacts in fMRI data. Human brain mapping conference.

McIntosh, A.R., Chau, W.K., Protzner, A.B., 2004. Spatiotemporal analysis of event-related fMRI data using partial least squares. Neuroimage 23, 764–775. http://dx.doi.org/10.1016/j.neuroimage.2004.05.018.

Mckeown, M.J., Makeig, S., Brown, G.G., Jung, T.-P., Kindermann, S.S., Kindermann, R.S., Bell, A.J., Sejnowski, T.J., 1998. Analysis of fMRI data by blind separation into independent spatial components. Hum. Brain Mapp. 6, 160–188.

Mitchell, T.M., Shinkareva, S.V., Carlson, A., Chang, K.-M., Malave, V.L., Mason, R.A., Just, M.A., 2008. Predicting human brain activity associated with the meanings of nouns. Science 320, 1191–1195. http://dx.doi.org/10.1126/science.1152876.

Naci, L., Cusack, R., Anello, M., Owen, A.M., 2014. A common neural code for similar conscious experiences in different individuals. Proc. Natl. Acad. Sci. http://dx.doi.org/10.1073/pnas.1407007111.

Naselaris, T., Prenger, R.J., Kay, K.N., Oliver, M., Gallant, J.L., 2009. Bayesian reconstruction of natural images from human brain activity. Neuron 63, 902–915. http://dx.doi.org/10.1016/j.neuron.2009.09.006.

Ng, B., Abugharbieh, R., 2011. Generalized sparse regularization with application to fMRI brain decoding. Inf. Process. Med. Imaging 22, 612–623.

Nishimoto, S., Vu, A.T., Naselaris, T., Benjamini, Y., Yu, B., Gallant, J.L., 2011. Reconstructing visual experiences from brain activity evoked by natural movies. Curr. Biol. 21, 1641–1646. http://dx.doi.org/10.1016/j.cub.2011.08.031.

O'brien, R.M., 2007. A caution regarding rules of thumb for variance inflation factors. Qual. Quant. 41, 673–690. http://dx.doi.org/10.1007/s11135-006-9018-6.

Pajula, J., Kauppi, J.-P., Tohka, J., 2012. Inter-subject correlation in fMRI: method validation against stimulus-model based analysis. PLoS One 7, e41196. http://dx.doi.org/10.1371/journal.pone.0041196.

Pamilo, S., Malinen, S., Hlushchuk, Y., Seppä, M., Tikka, P., Hari, R., 2012. Functional subdivision of group-ICA results of fMRI data collected during cinema viewing. PLoS One 7, e42000. http://dx.doi.org/10.1371/journal.pone.0042000.

Peelen, M.V., Downing, P.E., 2005. Within-subject reproducibility of category-specific visual activation with functional MRI. Hum. Brain Mapp. 25, 402–408. http://dx.doi.org/10.1002/hbm.20116.

Pelphrey, K.A., Morris, J.P., Michelich, C.R., Allison, T., McCarthy, G., 2005. Functional anatomy of biological motion perception in posterior temporal cortex: An fMRI study of eye, mouth and hand movements. Cereb. Cortex 15, 1866–1876. http://dx.doi.org/10.1093/cercor/bhi064.

Power, J.D., Mitra, A., Laumann, T.O., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2014. Methods to detect, characterize, and remove motion artifact in resting state fMRI. Neuroimage 84, 320–341. http://dx.doi.org/10.1016/j.neuroimage.2013.08.048.

Qian, J., Hastie, T., Friedman, J., Tibshirani, R., Simon, N., 2013. GLMNET for Matlab. http://web.stanford.edu/~hastie/glmnet_matlab.

Rizzolatti, G., Cattaneo, L., Fabbri-Destro, M., Rozzi, S., 2014. Cortical mechanisms underlying the organization of goal-directed actions and mirror neuron-based action understanding. Physiol. Rev. 94, 655–706. http://dx.doi.org/10.1152/physrev.00009.2013.

Ryali, S., Chen, T., Supekar, K., Menon, V., 2012. Estimation of functional connectivity in fMRI data using stability selection-based sparse partial correlation with elastic net penalty. Neuroimage 59, 3852–3861. http://dx.doi.org/10.1016/j.neuroimage.2011.11.054.

Saxe, R., Xiao, D.-K., Kovacs, G., Perrett, D.I., Kanwisher, N., 2004. A region of right posterior superior temporal sulcus responds to observed intentional actions. Neuropsychologia 42, 1435–1446. http://dx.doi.org/10.1016/j.neuropsychologia.2004.04.015.

Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., Filippini, N., Watkins, K.E., Toro, R., Laird, A.R., Beckmann, C.F., 2009. Correspondence of the brain's functional architecture during activation and rest. Proc. Natl. Acad. Sci. U. S. A. 106, 13040–13045. http://dx.doi.org/10.1073/pnas.0905267106.

Tenenbaum, J.B., de Silva, V., Langford, J.C., 2000. A global geometric framework for nonlinear dimensionality reduction. Science 290, 2319–2323. http://dx.doi.org/10.1126/science.290.5500.2319.

Tibshirani, R., 2011. Regression shrinkage and selection via the lasso: a retrospective. J. R. Stat. Soc. Ser. B Stat. Methodol. 73, 273–282. http://dx.doi.org/10.1111/j.1467-9868.2011.00771.x.

Toiviainen, P., Alluri, V., Brattico, E., Wallentin, M., Vuust, P., 2013. Capturing the musical brain with Lasso: Dynamic decoding of musical features from fMRI data. Neuroimage 88C, 170–180. http://dx.doi.org/10.1016/j.neuroimage.2013.11.017.

Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. Chemom. Intell. Lab. Syst. 58, 109–130. http://dx.doi.org/10.1016/S0169-7439(01)00155-1.

Xu, H., Lorbert, A., Ramadge, P.J., Guntupalli, J.S., Haxby, J.V., 2012. Regularized hyperalignment of multi-set fMRI data. IEEE Statistical Signal Processing Workshop, pp. 229–232.

Xu, P., Xu, H., Ramadge, P.J., 2013. Detecting stimulus driven changes in functional brain connectivity. 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, pp. 3507–3511 http://dx.doi.org/10.1109/ICASSP.2013.6638310.

Zacks, J.M., Braver, T.S., Sheridan, M.A., Donaldson, D.I., Snyder, A.Z., Ollinger, J.M., Buckner, R.L., Raichle, M.E., 2001. Human brain activity time-locked to perceptual event boundaries. Nat. Neurosci. 4, 651–655. http://dx.doi.org/10.1038/88486.

Zacks, J.M., Speer, N.K., Swallow, K.M., Maley, C.J., 2010. The Brain's cutting-room floor: Segmentation of narrative cinema. Front. Hum. Neurosci. 4. http://dx.doi.org/10.3389/fnhum.2010.00168.

Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. J. R. Stat. Soc. Ser. B 67, 301–320. http://dx.doi.org/10.1111/j.1467-9868.2005.00503.x.