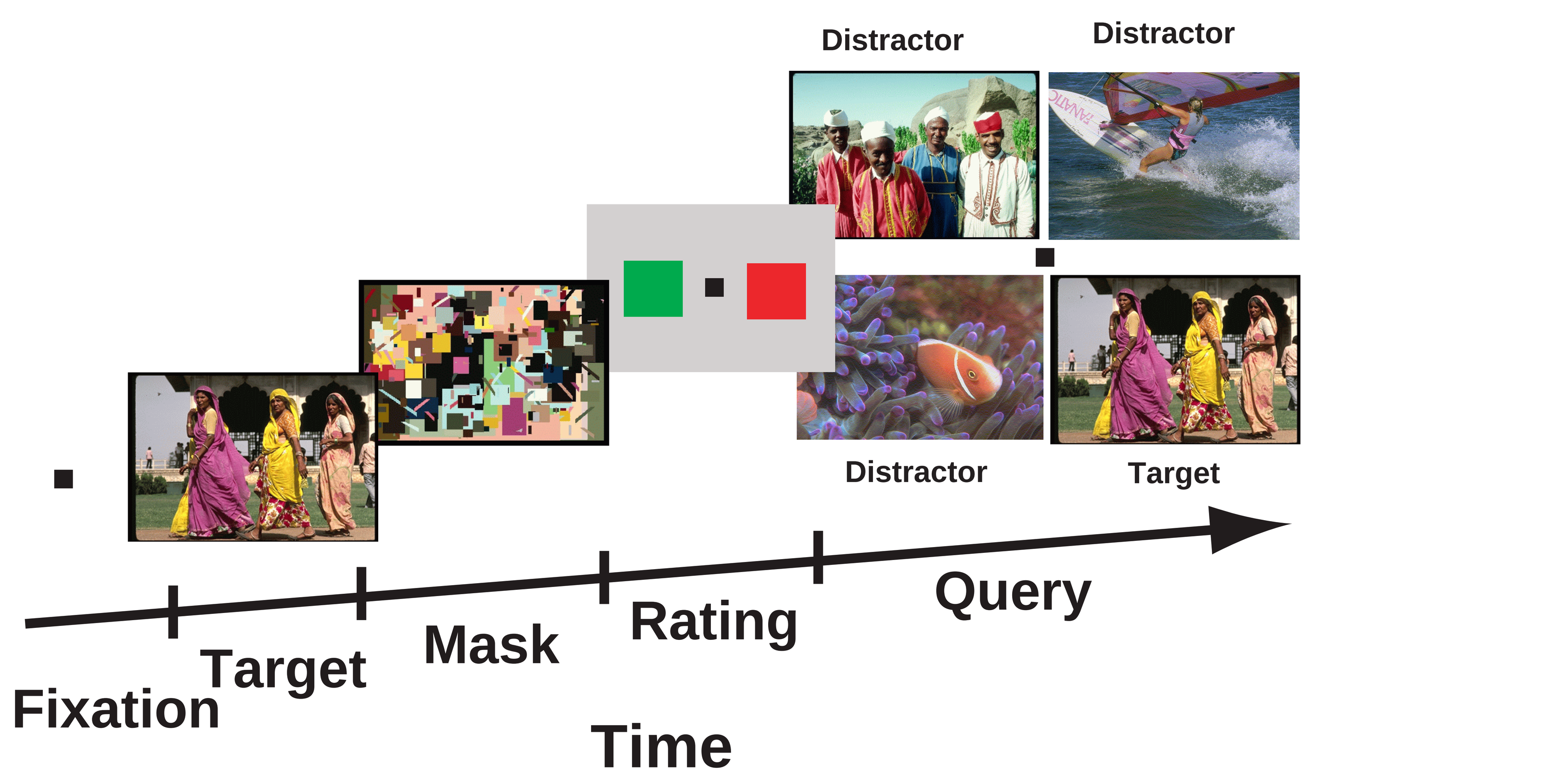
Single trial MEG recordings can predict the subjects ability to recognize a natural scene.

Jochem W. Rieger[1], F. Plum, K. R. Gegenfurtner[1], C. Braun[2], H. Preißl[2] & H.H. Bülthoff[1]
rieger@kyb.tuebingen.mpg.de
[1]Max-Planck-Institut für biologische Kybernetik
Spemannstraße 38 • 72076 Tübingen
[2]MEG-Zentrum • Otfried-Müller-Str. 47 • 72076 Tübingen

1. Introduction

MEG recordings reflect ongoing brain activity in real time with reasonable spatial resolution. To improve the low signal to noise ratio recordings from single trials are commonly averaged over trials and even over subjects. Our goal was to predict from single trial MEG recordings whether a subject will recognize a briefly presented photography of a natural scene.

2. Experimental Methods

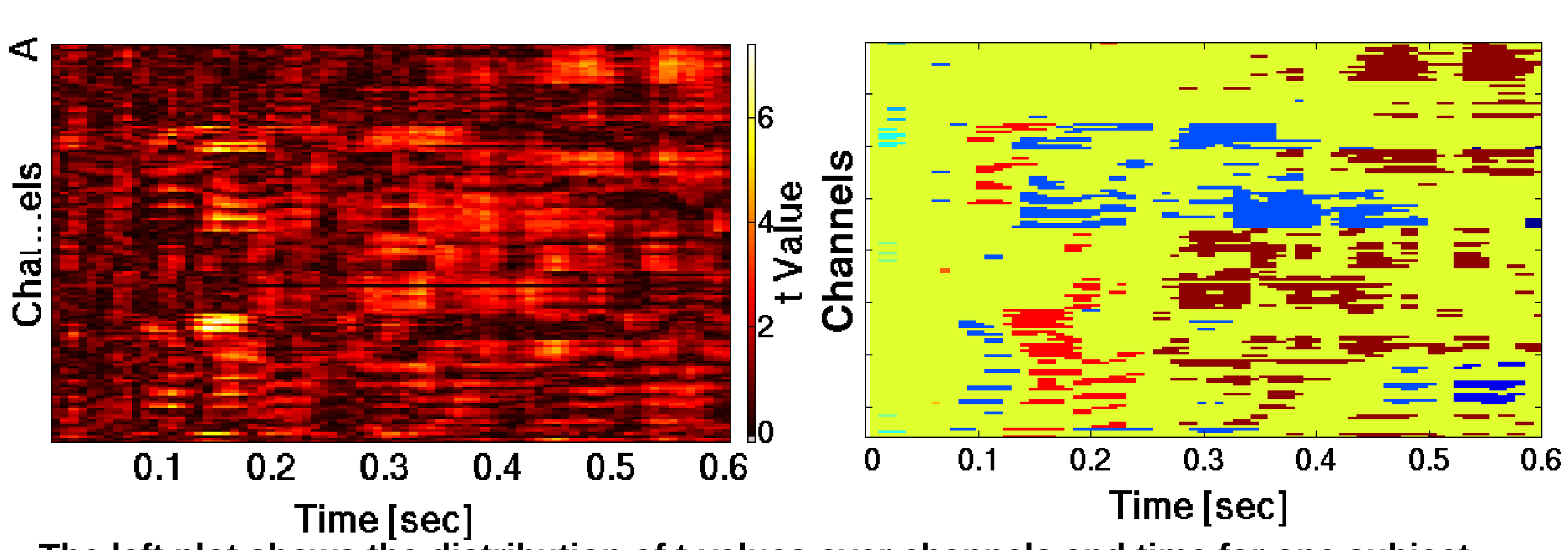


- After a fixation spot the target was presented for 37ms, followed immediately by a mask which was shown stayed for a randomly chosen duration between 1000 ms and 1400 ms.
- The mask was followed by a confidence rating screen. By moving the finger on the side of the green triangle subjects indicated a high confidence in their ability to recognize the target picture.
- The query phase immediately followed the confidence rating. The target was presented together with three distractor pictures (4 AFC). The subject had to indicate the target image by finger movement.
- Every picture was presented only once. The mask consisted of randomly sized and oriented rectangles and lines, mimicking the statistical properties of natural images. The colors for the mask were randomly chosen from all four pictures shown on that trial.
- Magnetic fields were recorded with a CTF 151-channel MEG-whole head system while target and mask were presented.
- Seven subjects participated in the experiment.

Time [sec]

3. Data Preprocessing

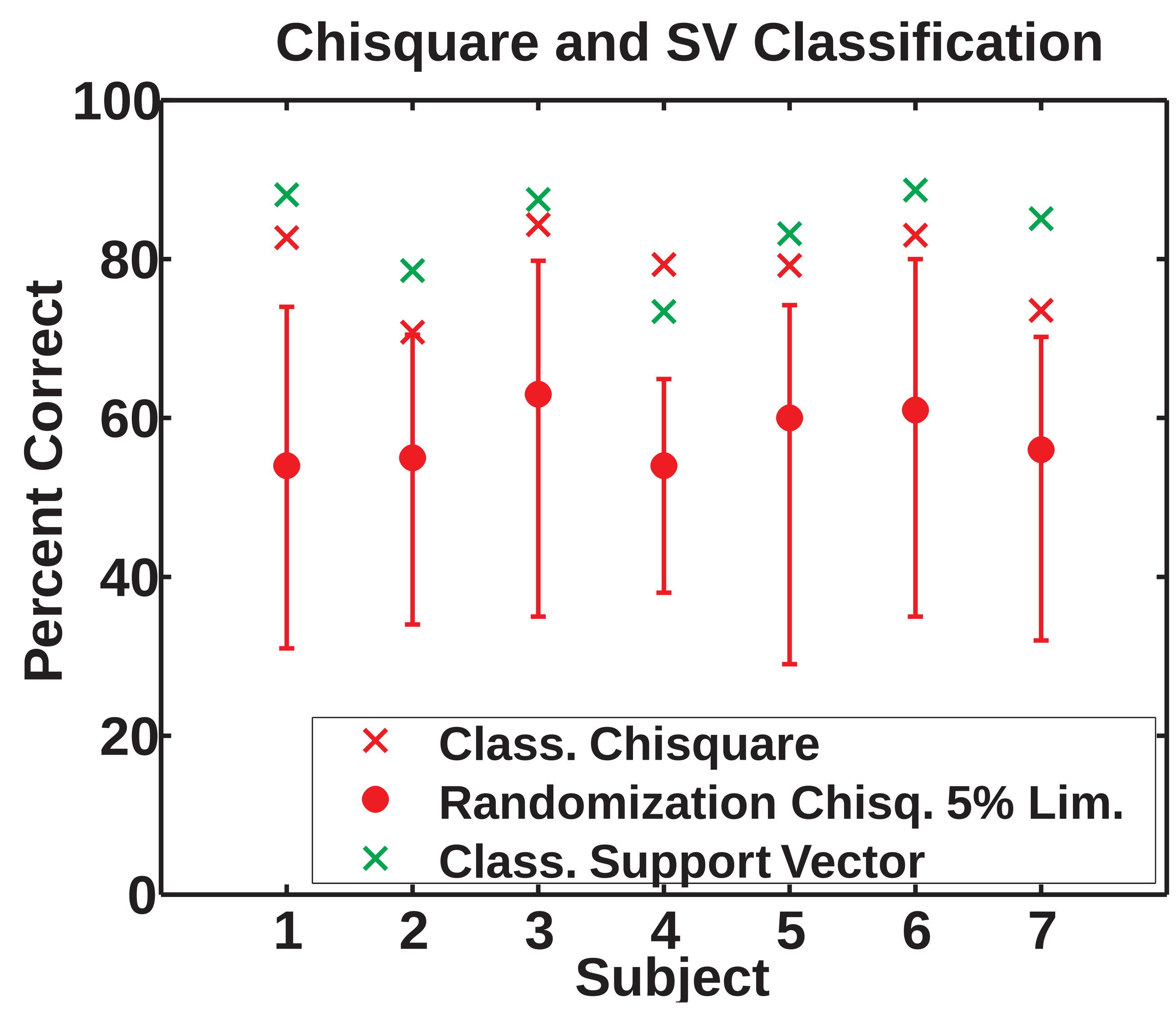
- We only included trials in which the confidence rating matched the actual response, in order to avoid lucky guesses and careless errors.
- Data recorded within 600 ms from target onset were included in the classification.
- Trials were grouped into correct and false trials. For each channel and each point in time t-values were calculated between these two groups.
- Data reduction was achieved by applying a region growing algorithm to cluster the data. Two parameters were specified: First only samples above a selected t-value were included. The second parameter was the minimum number of spatial and temporal neighbors that must fulfill the t-criterion. Contiguous samples that satisfy these two criteria are then combined into a cluster. The cluster means were then included into the further analysis



The left plot shows the distribution of t-values over channels and time for one subject. The right plot shows the distribution of samples that are combined into a cluster. Different colors indicate different clusters. Ten clusters were found by the algorithm. The parameters for the region growing algorithm were t ≥ 2 and neighbours 2.

3. Chi-Square Classification

- The means of the clusters from single trials were classified in a leave-one-out cross-validation, e.g. the models for the classification were calculated from all but the current trial.
- Mean and variance of the cluster means over trials were estimated. For the cluster data of the given trial two z-values were calculated from these parameters, one for the correct and one for the false trials.
- The trial was assigned to the class in which the sum of the squared z-values was lowest (χ^2).
- On average 79 percent of the trials were correctly classified.
- Confidence limits for the classification performance were determined in a randomization test. Classification performance obtained with the measured datasets was significantly above chance for all subjects (t 2 and neighbours 2).

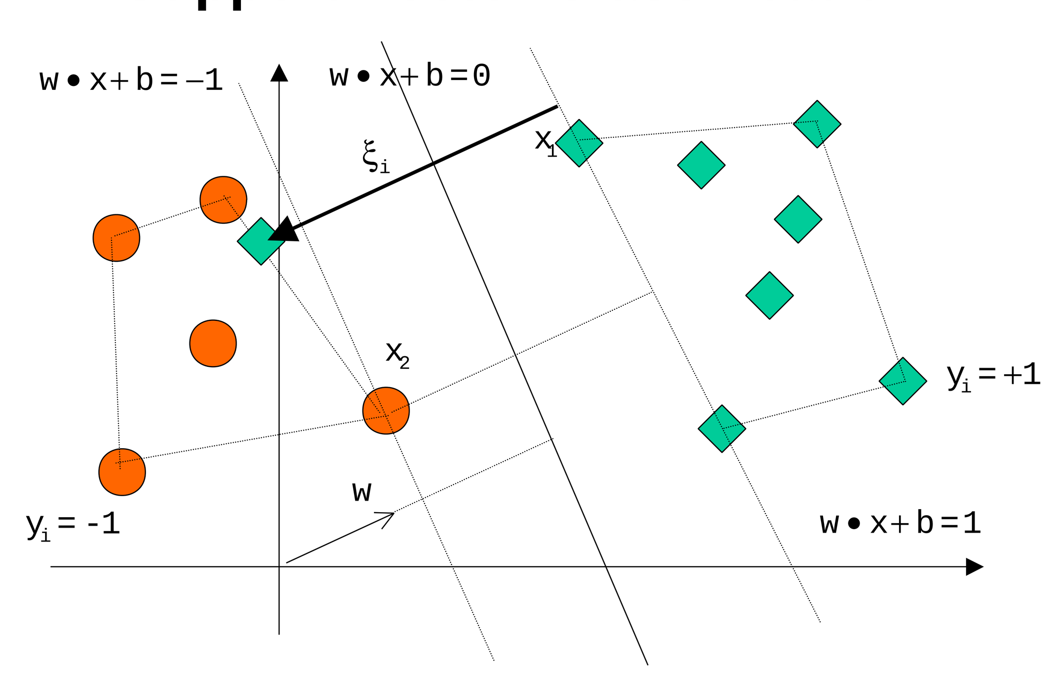


4. Support Vector Classification

- As in the Chi-square classification the means of the clusters from single trial were classified in a leave-one-out cross-validation by Support Vector Classification (SVC).
- SVC determines an optimal separating hyperplane with largest margin to the classes.
- In cases where the datasets are not clearly separable this leads to the following optimization problem: Minimize the distance between the class boundary and corresponding misplaced samples while maximizing the margin between the class boundaries. The trade-off between this terms has to be determined by experiment.

(Cortes and Vapnik 1995)

Support Vector Classification



- The classification results were obtained by using different projections into feature spaces. This is the space where the SVC searches the separating hyperplane. We used the following transformations (kernel functions): a linear kernel, an inhomogenous cubic polynomial, and a radial-basis-function (RBF) kernel.
- Except for subject 4 the best classification results are better than these from Chi-square classification. On average 83% of the single trials were correctly classified.

4. Conclusions

➨ It is possible to predict with reasonable performance (83% correct classifications) from single trial MEG activation measured during the first 600ms after target presentation whether a subject will subsequently recognize a natural scene.

➨ The classification results of the linear chi-square classification are statistically significant above chance in a randomization test.

➨ The nonlinear transformation applied in SVC improves the classification performance in most cases.