Maastricht University

Reading Imagined Letter Shapes from the Mind's Eye Using Real-time 7 Tesla fMRI

Citation for published version (APA):

Goebel, R., van Hoof, R., Bhat, S., Luhrs, M., & Senden, M. (2022). Reading Imagined Letter Shapes from the Mind's Eye Using Real-time 7 Tesla fMRI. In *10th International Winter Conference On Brain-computer Interface (Bci2022)* IEEE. https://doi.org/10.1109/BCI53720.2022.9735031

Document status and date:
Published: 01/01/2022

DOI:
10.1109/BCI53720.2022.9735031

Document Version:
Publisher's PDF, also known as Version of record

Document license:
Taverne

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy
If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 12 May. 2024

Reading Imagined Letter Shapes from the Mind's Eye Using Real-time 7 Tesla fMRI

Rainer Goebel
Department of Cognitive Neuroscience
Maastricht University
Maastricht, Netherlands
r.goebel@maastrichtuniversity.nl

Rick van Hoof
Department of Cognitive Neuroscience
Maastricht University
Maastricht, Netherlands
rick.vanhoof@maastrichtuniversity.nl

Salil Bhat
Department of Cognitive Neuroscience
Maastricht University
Maastricht, Netherlands
salil.bhat@maastrichtuniversity.nl

Michael Lührs
Department of Cognitive Neuroscience
Maastricht University
Maastricht, Netherlands
michael.luhrs@maastrichtuniversity.nl

Mario Senden
Department of Cognitive Neuroscience
Maastricht University
Maastricht, Netherlands
mario.senden@maastrichtuniversity.nl

***Abstract*—We present a 7 Tesla fMRI proof-of-concept study of the first letter speller BCI that decodes imagined letter shapes from activity patterns in early visual cortical areas. New tools are developed to enable real-time population receptive field retinotopic mapping for encoding and decoding. Using two different letter shapes (H and T), classification performance of generated activity patterns during imagery reaches 80% accuracy in each individual. Using a denoising autoencoder, recognizable letter shapes could be reconstructed and displayed as feedback to participants in the scanner.**

Keywords—fMRI, letter speller, imagery, population receptive fields, denoising autoencoder

I. Introduction

Recently we developed an offline procedure to reconstruct vividly imagined letter shapes from associated activity patterns in early visual cortical areas measured with 7 Tesla functional magnetic resonance imaging (fMRI) [1]. Reconstruction of a stimulus from brain activity patterns required only a short (10 minutes) additional scan for estimating the population receptive fields (pRFs) of voxels in early visual cortex [2]. The reconstruction (decoding) process then inverts the established relationship between visual field and cortex such that activation of voxels can be projected back into the visual field. In an effort to extend this work to a more natural letter-speller BCI, we test the feasibility of a "Mind's Eye BCI" by efficiently implementing the retinotopic encoding and decoding tools in a real-time fMRI analysis software package, and test and refine the novel BCI system on healthy subjects during real-time fMRI scanning.

II. Methods

Fig. 1 provides an overview of the developed fMRI real-time BCI. The most important components of the system are described below.

A. Data Collection

FMRI measurements were acquired for three subjects (mean age 27 years, two males) in four pilot scans. Hence, scanning parameters may vary. The scans included one retinotopy run (random bar paradigm, [3]), imagery and perceptual runs. We recorded anatomical (voxel size = 0.7 × 0.7 × 0.7 mm 3 , TR = 5000 ms) and functional (voxel size = 0.8 × 0.8 × 0.8 mm3, TR = 2000 ms) images with a Siemens Magnetom 7T scanner and a 32-channel head-coil.

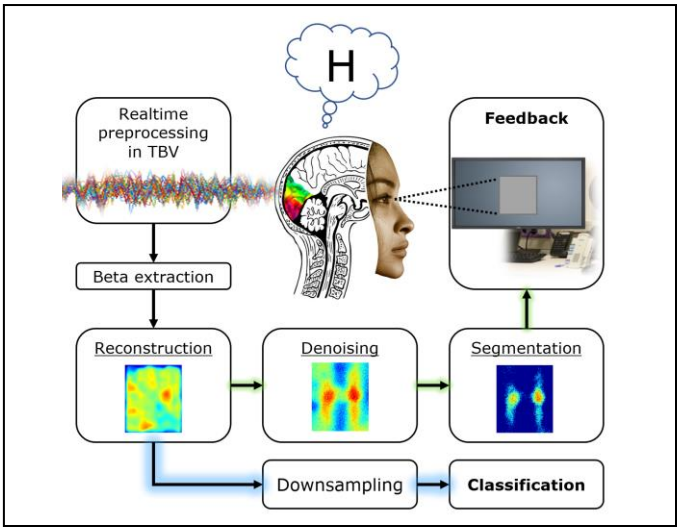

Fig. 1. Overview of "Mind's Eye BCI". Participants were cued (auditory) to create vivid mental images of the letters 'T' and 'H' (trial=10s). After a short resting period (10s), cortical activity related to visual mental imagery was reconstructed into 150x150 pixel images and presented as feedback on the scanner display.

B. Real-time pRF-mapping using Gradient Descent on Hashed-Gaussian Tiles

Conventional pRF mapping [2] requires too much computational time and resources for our real-time application. Therefore, a new pRF mapping approach has been developed [4,5] based on tile coding and hashing [6]. Here, we exhaustively partitioned the input space (mapping stimulus) into overlapping regions (tiling). 2D Gaussians were used as tiles as they tend to produce smooth receptive fields. Each tiling was then pseudo-randomly collapsed into a smaller set of tiles such that individual tiles consist of non-contiguous regions (hashing). The stimulus response was modeled as a linear combination of the overlap of tiles and the presented stimulus. Gradient descent was performed on the weights of the tiles. By computing the dot product between tiles and their corresponding weights, receptive fields can be obtained.

C. Denoising Autoencoder

The auto-encoder was trained on an independent dataset of six participants [1]. The input contained reconstructions of individual trials of letters 'H' and 'T' from perception data.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

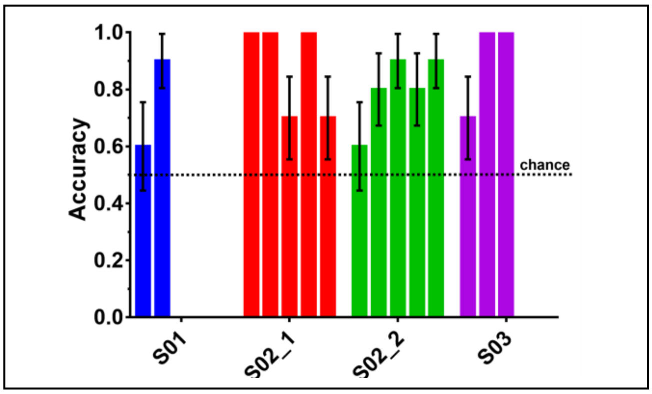

Fig. 2. Classification accuracies. Each bar represents ConvNet classification accuracy for a single run (10 trials). Mean classification accuracy of imagery trials was 80%. SE was taken from the binomial distribution.

To these reconstructions, we added reconstructions of noise observed in the same runs. The variance of added noise was a factor of 8 larger than the variance of signal to mimic and exaggerate the lower SNR for imagery as compared to perception and helps to force the autoencoder to restore noisy data. The model attempts to reproduce output based on the input and contains a single hidden layer with the number of units amounting to 1% of pixels in the reconstructed image. Hidden units had a rectified linear activation function while output units activated linearly. The learning rate was 10^{-6}, momentum was 0.9, batches had a size of 100, and loss was measured by the sum of squared distances. Training lasted 2500 iterations.

D. *Real-Time Decoding*

Each functional volume was fed to Turbo-BrainVoyager for incremental linear detrending of the time-series and 3D motion correction. Beta values were estimated by convolving trial predictors with a standard hemodynamic response function. Next, each trial was reconstructed into estimated images of the visual field by projecting voxel activations back through their receptive fields. The reconstructions were subsequently fed into two streams:

• **feedback**: reconstructions of mental imagery were denoised by the autoencoder, segmented to distinguish the letter from the background and finally presented to the participant in the scanner.

• **classification**: reconstructions of imagery trials were classified by a convolutional neural network (ConvNet).

The classifier was trained on reconstructions of imagery trials (down sampled to 30x30) of letters 'H' and 'T' from 5 of the 6 subjects in [1]. Data of the sixth subject was used for validation. The ConvNet consisted of two blocks of convolution batch normalization, rectification and max pooling, followed by another convolutional layer, batch normalization, rectification and a fully connected layer with a softmax activation. Loss was measured by cross-entropy and was minimized using the Adam optimizer [7] with an initial learning rate of 0.001.

III. Results

Fig. 2 shows classification accuracy across three subjects, who were trained to imagine the letters H and T. Each bar indicates a single run in which there were 10 trials.

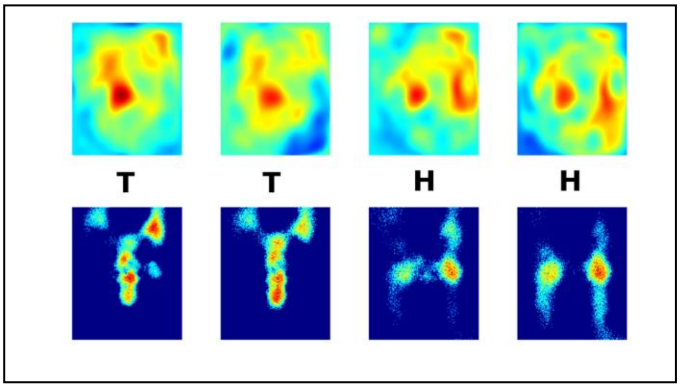

Fig. 3. Example visual field pattern reconstructions of single trials during an imagery run. In the top row the correct label and the resulting reconstructions are shown without denoising. In the bottom row the same images are shown after they were 'denoised' by the autoencoder and had their background removed.

While participants seem to struggle in their first attempt, they achieve an overall accuracy of around 80%. Fig. 3 depicts examples of letter shape reconstructions. The top row shows raw imagery reconstructions, the middle row shows the correct letter shapes and the bottom row shows the denoised versions. As you can see the raw reconstructions are simply unrecognizable. However, in most cases the neural network brings the image closer to a recognizable shape.

IV. Conclusons

Utilizing the preserved retinotopic organization in visual imagery, we set out to implement the first non-invasive BCI that can directly decode contents of the mind's eye. Our work is an extension of our previous offline letter reconstruction to a real-time decoding BCI. Preliminary results indicate that real-time reconstruction of imagined letters is possible (80% accuracy), even when using a generalizable convnet classifier which does not require training examples of the individual being scanned. In addition, we produced recognizable (denoised) reconstructions of imagined letter shapes on a single trial basis. To achieve a natural letter-speller communication BCI, further optimization of the real-time processing tools is needed and the paradigm needs to be extended to more letter shapes.

References

[1] M. Senden, T. C. Emmerling, R. Van Hoof, M. A. Frost, R. Goebel, "Reconstructing imagined letters from early visual cortex reveals tight topographic correspondence between visual mental imagery and perception," Brain Structure and Function, 224(3), 2019, pp. 1167-1183.

[2] S. O. Dumoulin, B. A. Wandell, "Population receptive field estimates in human visual cortex," Neuroimage 39(2), 2008, pp. 647-660.

[3] M. Senden, J. Reithler, S. Gijsen, R. Goebel, „Evaluating population receptive field estimation frameworks in terms of robustness and reproducibility," PloS one, 9(12), 2014, e114054.1338.

[4] S. Bhat, M. Lührs, R. Goebel, M. Senden, "Extremely fast pRF mapping for real-time applications," Neuroimage. 2021, 245:118671.

[5] mario.senden, & salil-maastricht. (2022). ccnmaastricht/CNI_toolbox: M220P121 (m2.2.0p1.2.1). Zenodo. https://doi.org/10.5281/zenodo.5867065

[6] R. S. Sutton, A. G. Barto, "On-policy Prediction with Approximation," in Reinforcement learning: An introduction. The MIT Press, 2018, pp. 217–222.

[7] D. P. Kingma, J. Ba., "Adam: A method for stochastic optimization," arXiv, 2014, 1412.6980.