Neural Encoding and Decoding with Deep Learning for Dynamic Natural Vision

This document describes the main source code (Matlab or Python) related to Wen et al., 2017. The video-fMRI dataset used in this study is available on PURR at Purdue University (https://dx.doi.org/10.4231/R7SF2TCW).

fMRI data processing

The fMRI responses during watching training movies or testing movies are processed with Matlab. The code was developed in Red Hat Linux. It requires the workbench toolbox developed by Human Connectome Project (HCP) to read, write, and display fMRI data in CIFTI format. The workbench is available on https://www.humanconnectome.org/software/connectome-workbench.

```
movie_fmri_data_processing.m
```

Reproducibility during dynamic natural vision

The Matlab code maps the cortical locations activated by natural movie stimuli, by calculating the intra-subject reproducibility in voxel time series.

```
movie_fmri_reproducibility.m
```

movie data processing

The Matlab code processes the videos (Linux). The Python code extracts the feature maps from hierarchical layers in the AlexNet (Krizhevskey et al., 2012) by using Caffe (Jia et al., 2014).

```
video_processing.m

AlexNet_feature_extraction.py
```

CNN feature processing

The Matlab codes processes the CNN feature maps extracted from videos for bivariate analysis and building coding models.

```
AlexNet_feature_processing_bivariate.m

AlexNet_feature_processing_encoding.m
```

Retinotopic analysis

CNN feature maps are used to map the receptive fields of individual voxels by correlate the voxel time series to the unit time series in CNN.

```
retinotopic_analysis.m
```

Hierarchical mapping

The hierarchical feature maps in different CNN layers are mapped on the visual cortex by calculating the correlation between the voxel time series and the unit time series in CNN.

```
hierarchical_mapping_analysis.m
```

Neural encoding models

The code trains voxelwise encoding models by linearly mapping the CNN feature maps to individual voxel time series.

```
voxelwise_encoding.m
```

Visual reconstruction

The code trains decoding models to estimate the feature maps in the first CNN layer based on the fMRI responses.

```
visual_reconstruction.m
```

Semantic categorization

The code trains decoding models to estimate the representation in the CNN semantic space based on the fMRI responses. The representation can be further converted to category labels.

```
category_prediction.m
```

Sub-functions

```
amri_sig_corr.m    --- calculate the Pearson correlation coefficient between every
                       two time series.

amri_sig_detrend.m --- remove the n-order polynomial trend in a time series.

amri_sig_isvd.m    --- calculate the principal components by using SVD-updating
                       algorithm.

amri_sig_r2z.m     --- convert r to z by using Fisher's r-to-z transformation.

ciftiopen.m        --- read fMRI data in CIFTI format.

spm_hrf.m          --- define the canonical hemodynamic response function (HRF).

amri_sig_mlreg.m   --- perform multivariate linear regression.

rand_mlreg.m       --- randomly initialize the parameters in regression model.

retinotopic.m      --- quantify the eccentricity and the polar angle of a voxel's
                       receptive field.
```

Version

V 1.0 (original version) --- 2017/09/20

References

1. Wen, H., Shi, J., Zhang, Y., Lu, KH., Cao JY. & Liu, Z. (2017). Neural Encoding and Decoding with Deep Learning for Dynamic Natural Vision. Cerebral cortex. In press.
2. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).
3. Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., ... & Darrell, T. (2014, November). Caffe: Convolutional architecture for fast feature embedding. In Proceedings of the 22nd ACM international conference on Multimedia (pp. 675-678). ACM.