

Brain Entropy Mapping Toolbox (BENtbx)

(May 01 2014)

<https://cfn.upenn.edu/~zewang/BENtbx.php>

Ze Wang

Depts. of Psychiatry and Radiology, Perelman School of Medicine,
University of Pennsylvania
zewang@mail.med.upenn.edu
redhatw@gmail.com

CONTENT

1. INTRODUCTION	3
1.1 THEORY AND BACKGROUND.....	3
1.2 SOFTWARE AND HARDWARE REQUIREMENT	6
1.3 IMAGE FORMAT	6
1.4 LICENSE	6
2. INSTALLING BENTBX.....	6
2.1 DOWNLOAD BENTBX	6
2.2 UNZIP AND INSTALL BENTBX	7
2.3 ADD BENTBX PATH TO MATLAB SEARCH PATH	7
2.4 SAMPLE DATA	8
2.4.1 <i>Data structure and parameters</i>	8
2.4.2 <i>Fast pass</i>	8
2.4.3 <i>Batch processing using the customized scripts</i>	8

(This manual is still in progress, please don't hesitate to drop me an email for any questions.)

1. Introduction

This is a brief tutorial¹ to the resting state fMRI based Brain Entropy Mapping toolbox (BENtbx). The main function for calculating entropy is written in C++. We provide compiled version for Windows, Linux, and Mac. The other scripts are for data preprocessing and are based on MATLAB (Mathworks Inc.) and SPM (Wellcome Department, UCL). The toolbox is free for academic users, and can be obtained from <https://cfn.upenn.edu/~zewang/> under the GPL license. The original GPL license and the file header should be included in any modified versions. Example datasets with customized settings are also available through the website. Both the toolbox and the sample data are not allowed for any commercial use without formal permission from the University of Pennsylvania. We (the author and University of Pennsylvania, the same for the following) are not and will not be responsible for any use which is made of this package. By providing exemplar data, no references or gold standards for the images, blood measures, and any kind of resultant outcomes are implied. We further disclaim any liability and accuracy of the outcomes arising from using this package. We are not responsible for any data interpretations. All the code and the data are provided **as they were**.

Please cite the toolbox and the related papers for any use of the toolkit.

1.1 Theory and background

Approximate Entropy calculation

Standard entropy calculation needs to know the probability density function, which requires a long series of data and is usually not easy to estimate. Many approximate ways have been invented for a practical entropy calculation, including the famous Approximate Entropy, approximate negentropy, Sample Entropy, Permutation Entropy, Wavelet Entropy, etc. Sample Entropy (SampEn) has been shown to be a reliable metric across various bio-medical studies and to be less sensitive to data length. In BENtbx, we provided Approximate Entropy and SampEn but we will focus on SampEn. SampEn is an extension of Approximate Entropy, which is estimated from a series of embedded vectors. For neuroimaging like fMRI, there are thousands of voxels, each has a time series. Consequently, calculating entropy for all voxels can be very exhaustive and time consuming without optimization. In BENtbx, we implemented an optimized version of SampEn (and Approximate Entropy) as stated below. It is straightforward from the definition of both entropy metrics.

A Fast SampEn calculation method.

Denote the rsfMRI data of one voxel by $\mathbf{x} = [x_1, x_2, \dots, x_N]$, where N is the number of acquisitions. SampEn starts with forming a series of vector sequences, each with m consecutive points extracted from \mathbf{x} : $\mathbf{u}_i = [x_i, x_{i+1}, \dots, x_{i+m-1}]$, where $i = 1$ to $N-m+1$, and m is the pre-defined dimension. For the purpose of description, we named these vectors as embedded vectors. Using a pre-specified distance threshold r , the number of embedded vectors \mathbf{u}_j ($j = 1$ to $N - m$, and $j \neq i$) whose distance from \mathbf{u}_i are less than r is recorded by $B_i^m(r)$. The same procedure is repeated for the dimension of $m+1$ to get $B_i^{m+1}(r)$. Then by averaging across all possible vectors, we have

$$B^m(r) = \frac{1}{(N-m)(N-m-1)} \sum_{i=1}^{N-m} B_i^m(r) \quad (\text{S1})$$

$$A^m(r) = \frac{1}{(N-m)(N-m-1)} \sum_{i=1}^{N-m} B_i^{m+1}(r) \quad (\text{S2})$$

And SampEn is calculated as:

$$\text{SampEn}(m, r, N, \mathbf{x}) = -\ln\left[\frac{A^m(r)}{B^m(r)}\right] \quad (\text{S3})$$

Note that the constant $\frac{1}{(N-m)(N-m-1)}$ in both Eqs. S1 and S2 will be canceled out in Eq. S3, therefore can be ignored during calculation.

Suppose \mathbf{u}_l and \mathbf{u}_k are two different embedded vectors built using the method stated above. Their distance calculation and the subsequent comparison with the tolerance threshold r are conducted twice in the original algorithm: one for $B_l^m(r)$ and the other for $B_k^m(r)$. This redundancy can be removed by changing the algorithm to be an iteration loop of distance calculation and threshold comparing for all 2-combination of the $N-m$ embedded vectors, so that the distance of \mathbf{u}_l and \mathbf{u}_k and the subsequent threshold comparison are performed only once but $B_l^m(r)$ and $B_k^m(r)$ are updated simultaneously. For fMRI data with N time points, the simplified routine will save $(N-m-1)*(N-m-2)/2$ times distance calculations and $(N-m-1)*(N-m-2)/2$ times distance comparisons. Suppose $N=200$, $m=3$, and there are

43366 intracranial voxels, 828724260 distance calculations and comparisons will be saved. Moreover, the first m elements of the dimension = $m+1$ embedded vector series are the same as the dimension = m vectors, so that distance calculation for a vector pair in the dimension of $m+1$ can share the mathematical operations for the first m elements during the distance calculation for the same vector pair in the dimension of m and it only needs to be updated with the last element of the vectors. This SampEn calculation algorithm was implemented in C++ computer language and was verified with Matlab code implementing the original algorithm.

Some guidance for choosing m and r . m is usually from 2-3 for fMRI; r can be from 0.1 to 0.4 or even 0.6. Small r will result in non-matching for both m and $m+1$ which will cause a problem for calculation SampEn (see Eq. S3).

BENtbx uses some batch processing scripts from ASLtbx:

- 1) Ze Wang, Geoffrey Aguirre, Hengyi Rao, JiongJiong Wang, Anna R. Childress, John A. Detre, *Empirical ASL data analysis using an ASL data processing toolbox: ASLtbx, Magnetic Resonance Imaging, 2008, 26(2):261-9.*

The main reference for BENtbx is:

- 2) Wang Z, Li Y, Childress AR, Detre JA (2014) *Brain Entropy Mapping Using fMRI. PLoS ONE 9(3): e89948. doi:10.1371/journal.pone.0089948.*

Some related abstracts are:

- 3) Wang Z (2012) Characterizing Resting Brain Information using Voxel-based Brain Information Mapping (BIM). 2012 Annual Meeting of the Organization for Human Brain Mapping. Beijing, China.
- 4) Wang Z (2012) Stable and Self-Organized Entropy in the Resting Brain. The Third Biennial Conference on Resting State Brain Connectivity. Magdeburg, Germany. pp. 208.
- 5) Ze Wang, Marcus Raichle, Anna Rose Childress, and John A Detre (2013) Mapping brain entropy using resting state fMRI. 2013 Annual Meeting of International Society of Magnetic Resonance in Medicine. Salt Lake City, USA. pp. 4861.
- 6) Ze Wang, Y Li, Z Singer, R Ehrman, A. V Hole, C. P O'Brien, Anna Rose Childress (2013) Human brain entropy mapping using thousands of subjects and its application in a drug addiction study. 2013 Annual Meeting of Society for Neuroscience. San Diego. pp. 7491.

1.2 Software and Hardware Requirement

As mentioned above, BENTbx contains a binary code for computing entropy and a bunch of Matlab scripts. If users just want to use the entropy calculation code, you can simply run the file name without any requirement for other software. The Matlab scripts need several functions from SPM, so both MATLAB (The MathWork, Inc., Natick, MA), version 5.3 or higher and SPM (5 and above) should be installed and the MATLAB search path should be updated to find SPM. Please refer to SPM's website to get more information. Depending on the dataset size, greater than 256 MB memory might be required. Otherwise, there are no additional specific hardware requirements for running BENTbx.

1.3 Image format

Both 3D and 4D NIfTI format and 3D Analyze images are acceptable for current version.

1.4 License

The University of Pennsylvania reserves all rights regarding BENTbx. BENTbx is free for academic users. You can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation, either version 3 of the License, or (at your option) any later version. Commercial or industrial users should contact the University and the author for the use of BENTbx.

BENTbx is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with BENTbx. If not, see <<http://www.gnu.org/licenses/>>.

2. Installing BENTbx

2.1 Download BENTbx

BENTbx can be downloaded from the following webpage: <http://cfn.upenn.edu/~zewang/software.html>. It's available as a RAR compressed file named "BENTbx.rar". The example dataset "BENrsmfmidata.rar" can be downloaded through the same webpage. Please spend a minute on typing the required information in the form. All recorded information will be kept confidential. We will only automatically pull out your emails and distribute messages when big updates are made. **Please don't fill non-words in the registration boxes.**

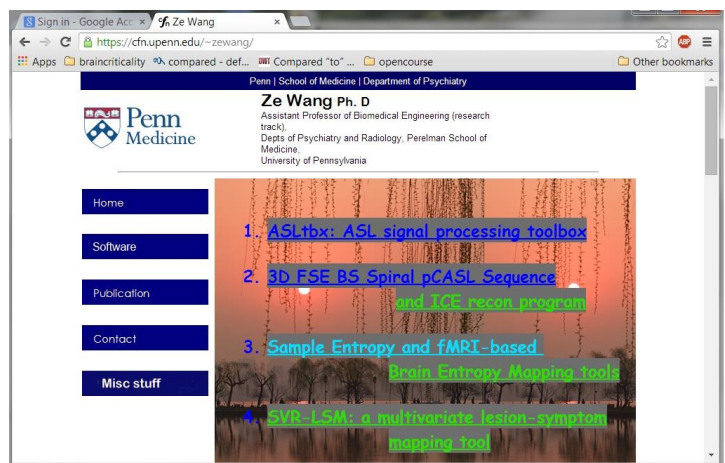


Figure 2.1. A snapshot of the download webpage.

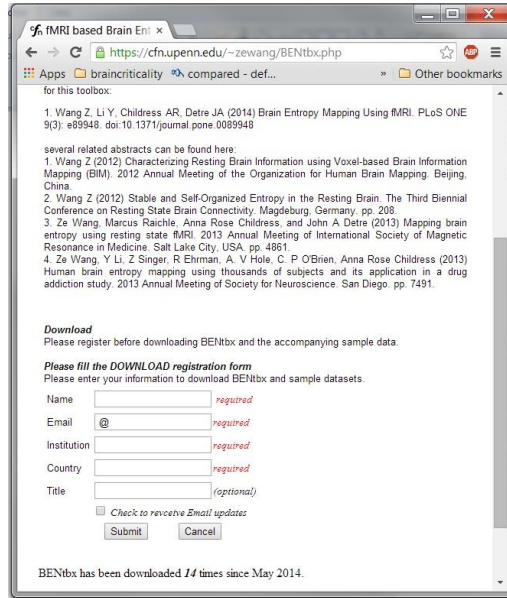


Figure 2.2. Screen snapshot of the registration page.

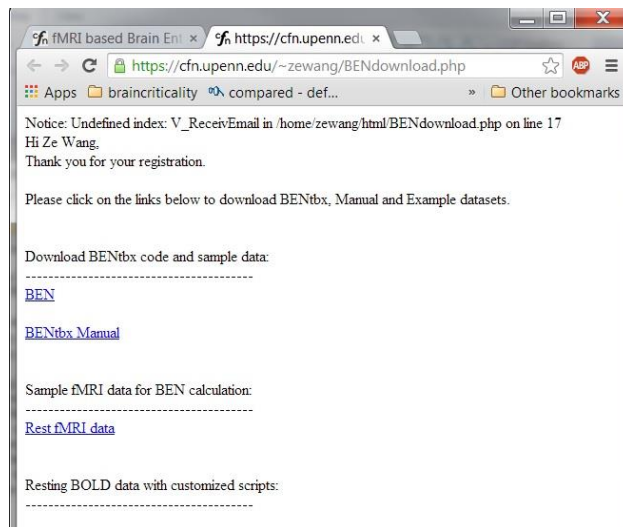


Figure 2.3. The webpage showing links to the scripts and sample data.

2.2 Unzip and install BENTbx

Copy the rar file into a directory, for example, "Z:\BEN". Decompress the rar file and move all the .m files to a directory, for example, "Z:\BEN\BENTbx".

2.3 Add BENTbx path to Matlab search path

Suppose the toolkit is installed in Z:\BEN\BENTbx, in Matab command window, type:

`addpath Z:\BEN\BENTbx`

to add the BENTbx path to Matlab search path.

2.4 Sample data

One sample dataset was provided to show how to use BENTbx.

2.4.1 Data structure and parameters

The sample data can be found from the same webpage as that for BENTbx scripts. First, please unzip the rar file and put them into a new folder such as “Z:\BEN”. Please also copy the entire BENTbx folder over. That is to say, under the folder Z:\BEN, there will be two subfolders: one is “BENTbx”, the other is “sub06880” (the folder for holding the sample data here). Within the data subfolder, there are two directories: “anat_anlz” holding the structural image and “func_anlz” holding the resting state fMRI images.

2.4.2. Fast pass

batch_calc_BEN_manual.m provides a more friendly way to directly calculate entropy without much scripting.

2.4.3 Batch processing using the customized scripts

Most of the preprocessing steps have been included in ASLtbx, and users can find more descriptions about them in ASLtbx manual. Below is a copy of the content of batch_run.m included in BENTbx, and the processing pipeline includes:

1. Reset orientations for both T1 and EPI images;
2. T1 image segmentation;
3. Slice timing correction for EPI images;
4. Motion correction;
5. Generating a skullstripped mean EPI image to be used for coregistration;
6. Calculating temporal SNR and temporal variance map;
7. Registering T1 and the grey matter, white matter maps into the EPI image space;
8. Temporal filtering using CompCor;
9. Smoothing;
10. BEN calculation;
11. Registering EPI and BEN maps into T1;
12. Spatially normalization (registration to MNI space) for T1 and BEN maps;

13. Smoothing BEN maps.

sublist.m lists all the subjects to be processed. You can disable a subject by putting a “#” before the subject ID.

par.m is for setting the environment variables such as data path, output name, ... etc.

batch_reset_orientation; % resetting image origin to the center of the image volume. This step is required for late registration.

batch_nsegment; % segmenting T1 images. The c1, c2, c3 images (grey matter, white matter, and CSF segments) are required in batch_coreg_t12EPI, batch_filtering

batch_slicetiming; % correcting acquisition time difference between slices
% You should change line 5 in batch_slicetiming.m to the correct number of slices for your data.

batch_realign; % motion correction. This code will be updated later to consider the global signal effects.

% batch_create_msk; % creating brain mask. You need fsl installed. This file may not run in Windows. Then you can use the following substitute.

batch_create_msk2; % Creating a skull-stripped EPI using hard-thresholding.

batch_tstd; % calculating tsnr and temporal variance image

batch_coreg_t12EPI; % registering T1 to EPI image. This is to resample the c2 and c3 images into the BOLD image space to be used as ROIs for extracting white matter and CSF signal

% this file also creates a skull stripped T1 image: t1_noskull.nii to be used for registration

batch_filtering; % temporal nuisance filtering. A bandpass filter is used.

batch_smooth; % smoothing the filtered images

batch_calc_BEN; % calculating BEN maps

batch_coreg; % coregistering EPI images and BEN map to T1

batch_norm_spm8; % normalization using spm8 nsegmentation

batch_smooth_BENmaps; % smoothing BEN maps

batch_calc_BEN is the matlab scripts for calling the compiled c++ entropy calculation program. You can change the two parameters therein: the window length m and the cutoff r.

batch_calc_BENmultifreqband is for calculating SampEn with different scale but in a slightly different way. Rather than doing a Haar wavelete transform, an FFT is performed and half of the frequency band is taken for a down-scale entropy calculation.

2.4.4. Output

“batch_calc_BEN.m” will output BEN maps at different time scale. The default setting is from 0 to 1. You can change it from 0 to a bigger number in line 76 therein. Scale=0 means the finest level (the original

data); scale=1 means downsampling by 2; increasing scale by 1 means a further downsampling by 2. All BEN maps were multiplied by a constant 1000 (line 112).