Intesity Normalization and Data Integrity Analysis with FS-FAST

`Comments or questions: analysis-bugs@nmr.mgh.harvard.edu\\`
`$Id: inorm.tex,v 1.1 2005/05/04 17:00:48 greve Exp $`

1 Introduction

This document describes how the FreeSurfer Functional Analysis STream (FS-FAST) performs global intensity normalization and analyzes the integrity of fMRI data. Intensity normalization is the rescaling of the data so that all runs/subjects have the same global within-brain average intensity. This reduces variance when data are combined across sessions/subjects. The motivation for this is that differences between sessions can largely be due to differences in fMRI signal intensity (eg, due to different shimming). The integrity of the data is evaluated by computing statistics of the temporal waveform averaged across all brain voxels as well as those of the temporal waveform averaged across all out-of-brain voxels. Both of these operations are accompished by the same set of programs, namely **fast_inorm.m inorm**, and **selxavg**.

2 Segmentation

The first step in both processes is to determine which voxels belong to the brain and which are outside of the brain. This is accomplished with two passes through the data. In the first pass, the global intensity average is computed across all voxels and frames from which a segmentation threshold is computed. By default, this threshold is set to 0.75 times the global average. In the second pass, the average intensity of each voxel across frames is computed to yield a mean intensity volume image. Voxels above the threshold are assumed to be inside the brain. An lower threshold is also computed as 0.25 times the global average; voxels under this threshold are assumed to be out-of-brain.

3 Intensity Normalization

Once the volume has been segmented into in-brain and out-of-brain voxels, the average of all the in-brain voxels is computed. A rescaling factor is then computed as:

$$\text{Rescale Factor} = \frac{\text{Rescale Target}}{\text{In-Brain Average}}, \tag{1}$$

where *Rescale Target* is the desired in-brain average. The actual value of the *Rescale Target* is irrelevant as long as it is not zero and is the same for all volumes that will be combinded in later stages of processing. The actual intensity normalization is accomplished by multiplying all intensities in the entire 4-D volume by the *Rescale Factor*. If the intensity normalzed volume were to be re-processed along the same lines, then the *In-Brain Average* would equal the *Rescale Target*, and the *Rescale Factor* would be 1. Note: the actual tasks of in-brain average computation and intensity rescaling are accomplished in seperate programs. **fast_inorm.m/inorm** perform the segmentation, compute the in-brain average, and evaluate the integretity of the data. **selxavg** performs the actual intensity rescaling.

4 Data Integrity Evaluation

The data integrity is evaluated by examining *scanner stability* and *image-formation fidelity*. *Scanner stability* is defined as the amount of change over time in the intensity of a voxel when there is no reason for the intensity of that voxel to change. Intensity changes may arise due to scanner noise, drift, or spiking. Of course, when a biological sample is scanned using fMRI, every voxel has the potential to have biologically-based signal changes in it, even those voxels far away from the sample. This leads to the second criterion, *image-formation fidelity*, which is a measure of how much of the actual signal "leaks" into spatial locations where there is no signal source. This leakage is a well known phenomenon for echo-planar acquisitions. Statistics for both of these factors are computed by **fast_inorm.m/inorm**.

To evaluate scanner stability, two average temporal waveforms are computed: one averaged across in-brain voxels, the other averaged across out-of-brain voxels. The waveforms are averaged across space in the hopes of averaging out signal changes due to the changes in the sample. Statistics such as the standard deviation, minimum, maximum, range, average absolute devation, average z-score, maximum z-score, and average linear drift are computed for both in-brain and out-of-brain waveforms. Spiking can be detected by examining the maximum z-score. Maximum z-scores over 3.5 should be investigated. Scanner drift can be evaluted by examining the linear drift. The standard devation can be used to evaluate scanner noise.

Two measures are computed to evalute image-formation fidelity. The first is the ratio of the in-brain average to out-of-brain average. The higher the ratio, the less the leakage. Ratios of 30 or greater are good. The second is the correlation coefficient between the in-brain waveform and the out-of-brain waveform. High coefficients indicate that the signal in the out-of-brain regions is indeed a leak from the in-brain region.

5 Implementation

The program for performing the intensity normalization and data integrity computations is a Matlab function called **fast_inorm.m**. Another program, called **inorm** is a cshell script wrapper for **fast_inorm.m**. **inorm** checks to make sure that all the relevant files and directories exist before creating a temporary Matlab script which calls **fast_inorm.m**; **inorm** then calls Matlab and executes the script. All the computations are performed within Matlab. There is also a wrapper called **inorm-sess** which runs **inorm** within the sessions environment and so will automatically determine the command-line arguments.

6 Usage

When **inorm** is executed from the unix command-line, it prints out the following help message (note: **fast_inorm.m** take the same arguments):

```
USAGE: inorm [-options] -i instem
   instem   - prefix of input files
Options:
   -thresh  threshold : fraction of global mean to separate brain and air (.75)
   -TR TR
   -inplaneres mm  : pixel size
   -betplaneres mm : between-plane distance
   -seqname  name  : name of acquisition sequence
   -monly mfile       : don't run, just generate a matlab file
   -umask umask       : set unix file permission mask
   -version           : print version and exit
```

-i instem: this is the input stem of the functional input volume (bfile format is assumed). First slice and number of slices are autodetected. Required.

-TR TR: TR (time between frames). This is not used in the analysis but is stored in the output report for convenience. Not required.

-inplaneres mm: width of a pixel. This is not used in the analysis but is stored in the output report for convenience. Not required.

-betplaneres mm: distance between slices. Note this will not be the same as the slice thickness if a skip is used. This is not used in the analysis but is stored in the output report for convenience. Not required.

-seqname name: name of pulse sequence used to acquire the data. This is not used in the analysis but is stored in the output report for convenience. Not required.

-thresh threshold: fraction of mean global mean above which the mean of a voxel must attain in order to be considered "brain". Allowable range is 0 to 1. Default is .75.

-monly: only generate the matlab file which would accomplish the analysis but do not actually execute it. This is mainly good for debugging purposes. Not required.

7 Output

inorm inorm will create four files as output. First, a file called instem.meanval ("instem" is the value passed with the "-i" flag) in which the thresholded global mean value is stored. This file is used by **selxavg** to actually perform the intensity normalization. **inorm** also produces a file called instem.report in which the data integrity statistics stored (see below). Finally, it produces two files with average temporal waveforms: instem.twf-over and instem.twf-under. The TWF files will have a data matrix with the number of rows equal to the number of time points. The number of columns will be equal to the number of slices+3. The first column is the time point number. The second column is the global average time course, demeaned, detrended, and scaled so that the standard deviation is 1. The third column is simply the raw global time course. The remaining columns are the average time courses averaged within a slice. Note that the sum of the slice time courses will not equal to the global mean because each slice has a different number of voxels contributing. The difference between the "over" and "under" is that the "over" is derived from voxels that are over threshold (ie, tissue) whereas the "under" is derived from voxels that are under threshold (ie, air).

7.1 Report File

The data integrity statistics are stored in the report file. A sample is shown below:

```
# FS-FAST Intensity Normalization Report
# date:    08-Nov-2000
# Input Volume       005/f
# nrows              64
# ncols              64
# nslices            29
# ntrs               98
```

```
# inplaneres         3.125000
# betplaneres        6.000000
# TR                 2.500000
# seqname            unknown
#
# GlobalMean         158.518886
# Relative Threshold Over 0.750000
# Absolute Threshold Over 118.889164
# Relative Threshold Under 0.250000
# Absolute Threshold Under 39.629721
#
# Over-Threshold Stats
# OV NVox        35950
# OV PctVox       30.27
# OV Mean        469.195391
# OV StdDev      2.916084
# OV AvgAbsDev   2.686290
# OV Min         464.114743
# OV Max         475.170904
# OV Range       11.056161
# OV SNR         160.899149
# OV ZAvg        0.921198
# OV ZMax        2.049157
# OV ZMax Index  87
# OV Drift       0.098789
#
# Under-Threshold Stats
# UN NVox        69442
# UN PctVox       58.46
# UN Mean        14.265598
# UN StdDev      0.100661
# UN AvgAbsDev   0.194493
# UN Min         14.132254
# UN Max         14.537888
# UN Range       0.405633
# UN SNR         141.719883
# UN ZAvg        1.932167
# UN ZMax        2.705030
# UN ZMax Index  82
# UN Drift       0.002535
#
# Over/Under Stats f
# OU Mean          32.889991
# OU Cor           0.780557
# OU eCorStd       0.063468
# OU tCor          12.298495
# OU tSigCor       0.000000
# OU log10tSigCor  20.708152
#
# PctUnaccounted  11.27
#
# StackFix-Based Stats
# SF Mean      469.195391
# SF StdDev    2.686290
# SF SNR       174.662983
```

```
# SF Min       462.964506
# SF Max       475.850515
# SF NOut      2/2842
#
##Slc NOver  Mean    SFStd  SFSNR    Min     Max SFOut  ZMax  Trend
  0    205  203.29  1.1106 183.04  200.22  206.60  0     2.39  0.0233
  1    398  336.86  2.2067 152.65  331.51  343.99  0     2.60  0.0663
  2    678  379.17  2.0673 183.42  374.51  384.96  0     2.30  0.0716
  3    971  443.10  2.0949 211.51  437.53  449.03  0     2.29  0.0802
  4   1195  481.63  2.3164 207.92  476.05  487.42  0     2.07  0.0915
  5   1337  502.21  2.6271 191.16  495.59  508.18  0     2.11  0.1040
  6   1442  520.27  2.8812 180.57  513.63  526.13  0     1.95  0.1150
  7   1524  525.48  2.9463 178.35  518.67  531.46  0     1.95  0.1181
  8   1550  537.27  2.8425 189.01  530.52  542.77  0     1.99  0.1153
  9   1600  541.31  2.8771 188.14  534.48  546.91  0     2.00  0.1151
 10   1624  536.37  3.2358 165.76  526.73  542.96  0     2.51  0.1266
 11   1583  522.91  3.2742 159.71  514.61  529.21  0     2.12  0.1292
 12   1494  507.34  3.0448 166.62  499.70  513.76  0     2.12  0.1206
 13   1479  494.12  2.5936 190.52  487.57  500.01  0     2.11  0.1055
 14   1536  478.99  2.1257 225.33  472.95  484.24  0     2.35  0.0847
 15   1614  462.80  2.0096 230.29  457.97  467.58  0     1.99  0.0797
 16   1616  461.72  2.0094 229.78  456.28  466.49  0     2.17  0.0820
 17   1618  470.95  2.4207 194.55  464.96  476.33  0     2.08  0.0970
 18   1599  474.25  2.6244 180.71  468.02  479.77  0     1.92  0.1077
 19   1563  462.46  2.5541 181.06  455.83  468.80  0     2.12  0.1049
 20   1486  448.62  2.8020 160.11  442.76  456.99  0     2.38  0.1109
 21   1440  427.49  2.6306 162.51  423.11  436.79  1     2.66  0.0977
 22   1304  419.60  2.9046 144.46  415.22  429.65  0     2.54  0.0938
 23   1210  413.24  2.9748 138.91  408.42  423.42  0     2.50  0.0943
 24   1072  415.72  3.0885 134.60  411.97  425.90  0     2.54  0.0635
 25    948  421.36  4.2847  98.34  414.60  434.03  0     2.48  0.0555
 26    844  411.40  2.1453 191.77  407.93  418.86  0     2.65  0.0701
 27    673  363.95  2.5765 141.25  354.73  369.26  1     2.76  0.0960
 28    347  234.39  3.3303  70.38  223.16  240.49  0     2.72  0.0561
```