

MRFSEG+GAMIXTURE - Software bundle documentation. Version 1.1

Jussi Tohka

July 25, 2007

1 INTRODUCTION

MRFSEG+GAMIXTURE is a collection of open-source software tools implementing a flexible voxel classification framework. The framework is based on a novel genetic algorithm based finite mixture model (GAMIXTURE) and a standard 3-D Markov random field (MRF) based on the iterative conditional modes (ICM) algorithm (MRFSEG). The key novelty in the MRFSEG program is the application of the configuration file to control MRF parameters. The software package is to be extended with a novel inhomogeneous MRF algorithm allowing even more flexible spatial and intensity information modeling.

At the moment the software supports Analyze 7.5 images. Future releases will add support for MINC files as well.

1.1 LICENSE

See the source code files. Please note that no representations about the suitability of this software for any purpose are made. It is provided "as is" without express or implied warranty.

1.2 PROGRAMS IN THE BUNDLE

The main programs are:

- **MRFSEG** for brain image segmentation using Markov random fields.
- **GAMIXTURE** for fitting FMMs to the image data using genetic algorithms.

The purpose is to 1) estimate the image parameters by gamixture, and 2) use these parameters to classify voxels in a 3-D image. There is a ready made shell script implementing this pipeline available (see Section 5 of this document). More detailed descriptions of the two main programs are provided in the sections 4.2 and 4.3.

The software package contains also smaller programs. The most important of these are:

- **erodemask**: Performs morphological erosion to a binary volume.
- **knnfilter**: Filters 3-D image with a KNN-filter
- **medfilt3D**: Performs 3-D median filtering of an input image
- **relabel_analyze**: Switches labels of an Analyze Image
- **showatlasspec**: Pretty prints the content of an atlas specification file, mainly for validation purposes. See the section about atlas specification files for further information.

2 INSTALLING THE SOFTWARE

2.1 SYSTEM REQUIREMENTS

The software has been developed using Linux operating system. It has been tested on Linux kernels 2.4 and 2.6. However, the software should work on Unix systems as well.

The only requirements are the C++ compiler and makefile utility. The software has been tested using g++ (versions 3.3 and 4.1).

2.2 INSTALLATION

Untar the tar file to the directory to which you wish to base the installation using `tar xvf mrfseg.tar`. Then go to the `src` sub-directory and type `make`. In the case the standard C libraries are located in an extraordinary directory, you may have to edit relevant parts of the `Makefile`.

3 USAGE OF THE MAIN PROGRAMS

3.1 GENERAL GUIDELINES

- The input images are expected to be in the Analyze 7.5 format.
- The filenames should be given with FULL PATHNAMES. They can be given with `hdr` or `img` extensions or in some cases without extensions. The latter choice is not reliable in all cases (particularly when path name contains the character `'.'`), but it is often the most convenient one.
- A command line help summary can be obtained by calling programs without arguments,.
- Saying "default" instead of the brainmask filename causes programs to assume that voxels with zero intensities belong to the background.
- The Analyze 7.5 specifications mandate that images represented with more than 8 bits per voxel are stored in signed units. Giving `-unsigned` as the second argument will make the programs assume that the data in image volumes is unsigned.
- There are probably other implicit limitations that I don't just now remember/know of.

3.2 GAMIXTURE

Genetic algorithm for mixture model optimization

The algorithm is described in:

[1] J. Tohka , E. Krestyannikov, I.D. Dinov , A. MacKenzie-Graham, D.W. Shattuck , U. Ruotsalainen, and A.W. Toga. Genetic algorithms for finite mixture model based voxel classification in neuroimaging. IEEE Transactions on Medical Imaging , 26(5):696 - 711, 2007.

[2] J. Tohka, E. Krestyannikov, I. Dinov, D. Shattuck, U. Ruotsalainen, A.W. Toga. Genetic algorithms for finite mixture model based tissue classification in brain MRI. In Proc. of European Medical and Biological Engineering Conference, IFMBE Proceedings vol. 11, pp. 4077 - 4082, Prague, Czech Republic, 2005.

Usage: `gamixture [-unsigned] imagefile maskfile atlasfile fmmparamfile [parameters]`

Specifying `-unsigned` right after `gamixture` makes the program assume that data is unsigned. Unsigned integers are not supported by default by the Analyze 7.5 file format which is the reason behind this switch.

Parameters for the program:

- `imagefile` is the input image
- `maskfile` is the file specifying the brain mask. All voxels with the value greater than 0.5 (in the mask) are assumed to belong to the brain.
- `atlasfile` is the name of the atlas specification file (see the next section).
- `fmmparamfile` is the name of the file where the mixture parameters are to be written (see the next section).
- Optional parameters are :

-alpha:	parameter for blended crossover, defaults to 0.5
-size:	population size, defaults to 100
-terminationthr:	threshold for terminating the algorithm, defaults to 0.0005
-xoverrate:	crossover rate, defaults to 1
-maxgenerations:	maximum number of generations, defaults to 500
-sortpop	whether to use the permutation operator, defaults to 1
-parzenn	number of points for Parzen estimate (for approximate ML), defaults to 101
-parzensigma	window width parameter for the Parzen estimate, defaults to 1
-equalvar	whether component densities should have equal variances, defaults to 0
-restarts	the number of runs of the GA. The one with the highest likelihood score (or the lowest KL divergence) is selected as the result.

3.3 MRFSEG

Tissue classification based on Markov Random Fields and the ICM algorithm.

Reference: J. Besag. On the statistical analysis of dirty pictures, J R Stat Soc Series B, 48(3):259–302, 1986.

Usage: `mrfseg img brainmask atlas_def mixture_params labeling [pvelabelimg] [beta1] [beta2]`

- `img` is the name of the input image
- `brainmask` is the file containing the brain mask
- `atlas_def` is the name of the atlas specification file
- `mixture_params` is the file containing the FMM parameters (e.g. by GAMIXTURE). Note that the mixture parameter files must match the atlas specification file.
- `labelimg` is the filename for the tissue classified image (only pure tissue types)

- `pvelabelimg` is the name of the tissue classified image (also PVE classes).
- `beta1` is the parameter controlling the strength of the 1st order component of the MRF (defaults to 0.0)
- `beta2` is the parameter controlling the strength of the 2nd order component of the MRF (defaults to 0.1)

4 FILE FORMAT SPECIFICATIONS

File format definitions: The programs in the bundle read and write several types of files in addition to the image files. More precisely, files can contain information about the Parzen estimates, FMM parameters and atlas specifications. All these files are text files. No comments are allowed in the files.

4.1 ATLAS SPECIFICATION FILES

We begin with the atlas specification files:

These files are used for inputting the information about SVPA, MRF and FMM constraints to the programs. There are two types of such files, those that impose constraints to certain mixture components and those that don't.

The former type of file should start with the identifier "p" and these files are the ones that should be used for the most time. (If you don't want any constraints just give 0.0 for the lower limit and 1.0 for the upper limit.) Thereafter, the number of atlas regions should be given (this should be 0 if there is no atlas), and thereafter the number of labels should be given. Then, the atlas regions should be defined (name, filename, constraints related to that region). Then, the definitions of labels should follow (name, whether partial volume label or pure tissue label, components of the label (give 0 0 if pure label). Thereafter, the matrix specifying the 2 nd order MRF should be given.

Note that the background is counted as a label, but the images are constrained to contain no background voxels in the brain region. Therefore, the background label should not be defined. The background has always the label 0.

An example follows: (see explanation below)

```
jtohka@[10]% more icbm_image_spec_prob.txt
p 0 7
0.0 0.1
0.0 1.0
0.0 1.0
0.0 0.05
0.0 0.3
0.0 0.3
csf 1 0 0
gm 1 0 0
wm 1 0 0
csfbg 0 1 0
csfgm 0 1 2
gmwm 0 2 3
0 0 1 1 0 1 1
0 -1 1 1 0 0 1
1 1 -1 1 1 0 0
1 1 1 -1 1 1 0
0 0 1 1 -1 1 1
1 0 0 1 1 -1 1
1 1 0 0 1 1 -1
```

End of the example

The first line

```
p 0 7
```

means that the FMMs are constrained, and there are 7 labels (the background is automatically included and has always label zero.) The 0 after p is reserved for later amendments.

The lines

```
0.0 0.1
```

```

0.2 1.0
0.2 1.0
0.0 0.05
0.0 0.3
0.0 0.3

```

mean that there must be more than 0% and less than 10% of the label 1, more than 20% and less than 100% of the label 2 etc. in the volume The lower limit (i.e. the first number) should always be smaller than the upper limit.

The lines

```

csf 1 0 0
gm 1 0 0
wm 1 0 0
csfbg 0 1 0
csfgm 0 1 2
gmwm 0 2 3

```

mean that the label number 1 is called csf and it is a pure tissue label;
the label number 2 is called gm and it is a pure tissue label;
the label number 3 is called wm and it is a pure tissue label;
the label number 4 is called csfbg and it is a mixed tissue label consisting of csf (the label number 1) and the background (the label number zero);
the label number 5 is called csfgm and it is a mixed tissue label consisting of csf (the label number 1) and the gm (the label number 2);
the label number 6 is called gmwm and it is a mixed tissue label consisting of gm (the label number 2) and wm (the label number 3);

Important note: it is a good idea to keep the pure labels as first labels (i.e. 1, ... k) and thereafter number the pve labels. Also the order of the pure labels should be according to the order of the (expected) means within these labels in the image. For example, in T1-weighted MRI CSF should be label 1, GM should be label number 2, and WM should be label number 3.

Finally, the lines

```

0 0 1 1 0 1 1
0 -1 1 1 0 0 1

```



```

1 1 -1 1 1 0 0
1 1 1 -1 1 1 0
0 0 1 1 -1 1 1
1 0 0 1 1 -1 1
1 1 0 0 1 1 -1

```

specify the second order MRF based on pairwise interactions.

4.2 MIXTURE FILES

Mixture files are plain text files listing the mixture parameter values. The mixture file always requires an atlas file to accompany it and the mixture file must match the atlas specification. Mixture file contains a single line consisting of FMM parameter values. When there are k types of pure tissue, and $kpve$ types of mixed tissues, the file looks like:

$$\mu_1 \quad \sigma_1^2 \quad p_1 \quad \mu_2 \quad \sigma_2^2 \quad p_2 \cdots \mu_k \quad \sigma_k^2 \quad p_k \quad p_{k+1} \cdots p_{k+kpve}$$

5 USEFUL SCRIPTS

At the time of writing there are few helpful scripts:

- `mrfseg_classifier` combines the GAMIXTURE with MRFSEG to tissue classify an MR image
- `mrfseg_classifier_s2` does the same but smooths intensity histogram (Parzen estimates) more. Good for lower resolution images.
- `eae_classifier` implements the combination of programs for mouse brain image tissue classification. This requires also the program `relabel_analyze`.

6 READY MADE ATLAS SPECIFICATION FILES

There are a few ready made atlas specification files available in the `atlaspecs` sub-directory:

- `mouseatlas_prob.txt` : An atlas specification file for the tissue classification in mouse images (T2 weighted).
- `icbm_image_spec_prob.txt` : An atlas specification file for the tissue classification in high resolution human T1 weighted brain MRI.
- `image_lowres_spec_prob.txt`: An atlas specification file for the tissue classification in low resolution human T1 weighted brain MRI.
- `image_lowres_nopve_spec_prob.txt`: An atlas specification file for low resolution human T1 weighted brain MRI. PVE is modeled during the GAMIXTURE but PVE voxels are discouraged in the voxel classification step. Useful if the purpose is to only get three-class segmentation and PV information is required.