1 Normal aging dataset

198 subjects: 100 females, 98 males. Class labels: subjects in their fifties - 1, subjects in their sixties - 2, subjects in their seventies - 3. In addition, age of each subjects is also provided (see Data format section on how to obtain it). Leave-12-out cross-validation for males and leave-13-out cross-validation for males is used.

2 Data format

Each MAT file contains the following important variables (ignore the rest):

1. trainingSpread. N x 1001 matrix, where N is the number of subjects used for training. Each row corresponds to a subject, each column corresponds to a feature. The last column contains class labels.

2. testSpread. M x 1001 matrix, where M is the number of subjects used for testing. Each row corresponds to a subject, each column corresponds to a feature. The last column contains class labels (ground truth). **For leave-one-out experiments only first row of the testSpread matrix should be used.**

3. mergedFeatureIDs. 1 x 1000 vector containing feature type for each feature (there are 18 possible feature types). Values 1, 2, 3 denote x, y, z component of a deformation field; 4 - length of the deformation field vector (R); 5 - Jacobian (J); 6 - log of Jacobian (LJ). Values 7 through 12 denote absolute value of the symmetry score for x, y, z, R, J, LJ respectively. Values 13 through 18 denote signed symmetry scores for x, y, z, R, J, LJ, respectively.

4. mergedFeatureIDs. 1 x 1000 vector containing statistical moment for each feature. 3 - mean, 4 - standard deviation.

5. mergedScaleIDs. 1 x 1000 vector containing image scale for each feature. 0 - image subsampled by the factor of 8, 1 - by factor of 4; 2 - by the factor of 2; 3 - original image scale.

6. mergedBestAvrScores. Augmented variance ratio (AVR)[1] score of every feature.

2.1 Getting Ages

In order to get the ages for the subjects in the Aging dataset, one should use function getAgesForClass.m provided in the subfolder MATLAB_FILES inside MALES and FEMALES folders. Assuming that this function is called from the subfolder MATLAB_FILES, and one would like to get ages for a split number 3 of the data, the sample call looks like this:

$[test, training] = getAgesForClass(3,' agingPermutationsForAlzheimers.mat', ...$
$'agingSplitsForAlzheimers.mat',' allAges.mat');$

The function will return ages of the test subjects for the data split 3 in the variable *test*, and the ages of training subjects in the variable *training*.

NOTE: SPLIT_1 is NOT part of the cross-validation. It contains features ranked based on ALL subjects (the entire dataset is used for training). Otherise, each dataset contains 1000 highest ranked individual features (out of 55 million) for each cross-validation split of the data.

References

[1] Y. Liu, K. Schmidt, J. Cohn, and S. Mitra, "Facial Asymmetry Quantification for Expression Invariant Human Identification",Computer Vision and Image Understanding Journal, Vol. 91, No. 1/2, July, 2003, pp. 138 - 159.