Learning-based Neuroimage Registration

Leonid Teverovskiy and Yanxi Liu[1]

October 2004
CMU-CALD-04-108, CMU-RI-TR-04-59

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Abstract

Neuroimage registration has been a crucial area of research in medical image analysis for many years. Aligning brain images of different subjects in such a way that same anatomical structures correspond spatially is required in many different applications, including neuroimage classification, computer aided diagnosis, statistical quantification of human brains and neuroimage segmentation. We combine statistical learning, computer vision and medical image analysis to propose a multiresolution framework for learning-based neuroimage registration. Our approach has four distinct characteristics not present in other registration methods. First, instead of subjectively choosing which features to use for registration, we employ feature selection at different image scales to learn an appropriate subset of features for registering a specific pair of neuroimages. Second, we use interesting-voxel selection to identify image voxels that have the most distinct image feature vectors. These voxels are then used to estimate the deformation field for registration. Third, we iteratively improve our choice of features and interesting voxels during registration process. Fourth, we create and take advantage of a statistical model containing information on image feature distributions in each anatomical location.

[1]This work is supported in part by NIH grants AG05133 and DA015900-01

Keywords: feature vectors, feature selection, interesting voxels, deformable registration, image pyramid, thin plate splines, RANSAC.

Contents

1 Introduction

Neuroimage registration is an essential problem in medical image analysis. Transforming neuroimages so that their corresponding anatomical structures become aligned is essential for statistical quantification of human brain, computer aided diagnosis, neuroimage segmentation and study of normal aging.
Neuroimage registration, particularly cross-subject registration, presents a number of challenges.

1. Corresponding anatomical brain structures of different subjects may differ in shape and topology. While some structures have relatively simple form that does not change much from subject to subject, others, like sulci, have a complex shape that varies significantly between people. These variations generally increase with aging, and can be further intensified by the presence of neurological disorders.

2. In the cases where registration involves neuroimages of pathological brains, some lesions, like tumors, can be present in one brain but not in the other. In such cases the problem of aligning corresponding anatomical structures becomes ill-posed since not every structure in one brain has a corresponding one in the other brain.

In this work we propose a new learning-based method for neuroimage registration. Our method addresses the challenges by learning which features to use for each pair of images and which voxels of the images to use in order to drive the registration process.
This paper is organized as follows. In section 2, we outline strength and weaknesses of the existing methods for neuroimage registration. Then we present our approach in section 3. Experimental results follow in section 4; section 5 contains a brief discussion of the future work, and we conclude the paper in section 6.

2 Existing Approaches

Existing methods for neuroimage registration can be divided into three groups:

1. Deformation model driven registration [18, 20, 14]. Methods in this group maximize some similarity metric between two images as a function of transformation parameters. Most popular similarity metrics are mean sum of squared differences (MSSD) between the reference and registered input image and mutual information (MI). Advantages of these methods are that

   (a) they are fully automatic;

   (b) they do not require computation of features.

   However, these methods also have a number of disadvantages:

   (a) they are susceptible to converging to suboptimal solutions because currently used similarity measures, like MI and MSSD are not convex functions of transformation parameters [20, 21];

   (b) they require a parametrized deformation model to be chosen beforehand;

(c) their performance depends on the number of parameters in a deformation model [14]. The fewer parameters the deformation model has, the easier it is for a method to find a near optimal solution in the space of transformations defined by the model. However, spaces induced by deformation models with few parameters are often not rich enough to contain an adequate transformation for cross-subject registration. On the other hand, more descriptive deformation models with many degrees of freedom are computationally expensive and make it harder to escape local extrema because of the need to estimate a large number of parameters. [21, 20];

(d) their performance depends on the initial orientation of the images [20].

2. Landmark-based registration. Methods in this group compute registering transformation based on the user-specified correspondences between certain voxels(landmarks) in the reference image and the input image. Advantages of the methods in this group are that

   (a) deformation does not have to be parametrized in advance. The registering transformation is computed based on the given correspondences;

   (b) user-specified correspondences are accurate.

   Shortcomings of these methods include human intervention and time-consuming landmark specification.

3. Feature-vector based registration methods [27] . Methods in this group compute a feature vector for each voxel. Correspondences between voxels in the reference image and voxels in the input image are estimated based on the similarity of their feature vectors. Attractive characteristics of such methods are that

   (a) they are automatic;

   (b) they are less prone to converging to suboptimal transformations;

   (c) they do not require parametrized deformation model, but estimate the registering transformation based on the computed correspondences.

   However, such methods

   (a) use pre-selected set of features;

   (b) involve many hand-tuned parameters.

Our approach builds on feature-based registration methods. As does the method presented in [23], we utilize various features to construct an attribute vector describing a voxel in a neuroimage. Also, as in [27], we use gaussians to model distribution of features belonging to a given voxel. However, our method has a number of distinguishing traits:

1. We estimate parameters of gaussians based on affinely transformed copies of the reference image.

2. We utilize decision theoretic framework to find interesting voxels in the reference image. These interesting voxels are far from other voxels in the features space, i.e. they are different from the others. The same framework is then used to find matches for the interesting voxels among the voxels in the input image.

Table 1: Existing approaches

	Deformation model driven	Landmark based	Feature-vector based
Require human intervention	no	yes	no
Increasing degrees of freedom of the deformation model makes the method more prone to converge to local extrema	yes	no	no
Depend on the initial orientation of the reference and input images.	yes	no	no
Select driving voxels	no	yes	yes
Learn features	no	no	no

3. We do not select features manually in advance but learn which feature subset to use for every pair of neuroimages automatically. This gives our method the power to adapt to particularities of the specific pair of neuroimages.

3 The algorithm

3.1 Introduction

In this section we describe our algorithm for neuroimage registration. Characteristics of our algorithm are:

1. automatic;
2. multiscale;
3. adaptive: for every pair of images the algorithm learns which features and which voxels to use for registration.
4. independent of the initial orientation of the images;

As a preprocessing step for our algorithm we must estimate a statistical model that describes how feature vectors for every anatomical location are distributed. The model is estimated based on a reference image. The model does not have to be reestimated every time the algorithm is run as long as the registration is done to the image of the same subject. Since we can register any two images to each other by registering each of them to the third image, in most cases we can avoid relearning the model.
The algorithm contains the following major components (Figure 3):

1. Feature selection: this mechanism selects a subset of features to use for registration. It uses a feature pool - set of all features available to the algorithm.
2. Correspondence matching: this mechanism uses the model to find correspondences between a voxel in the reference image and a voxel in the input image.

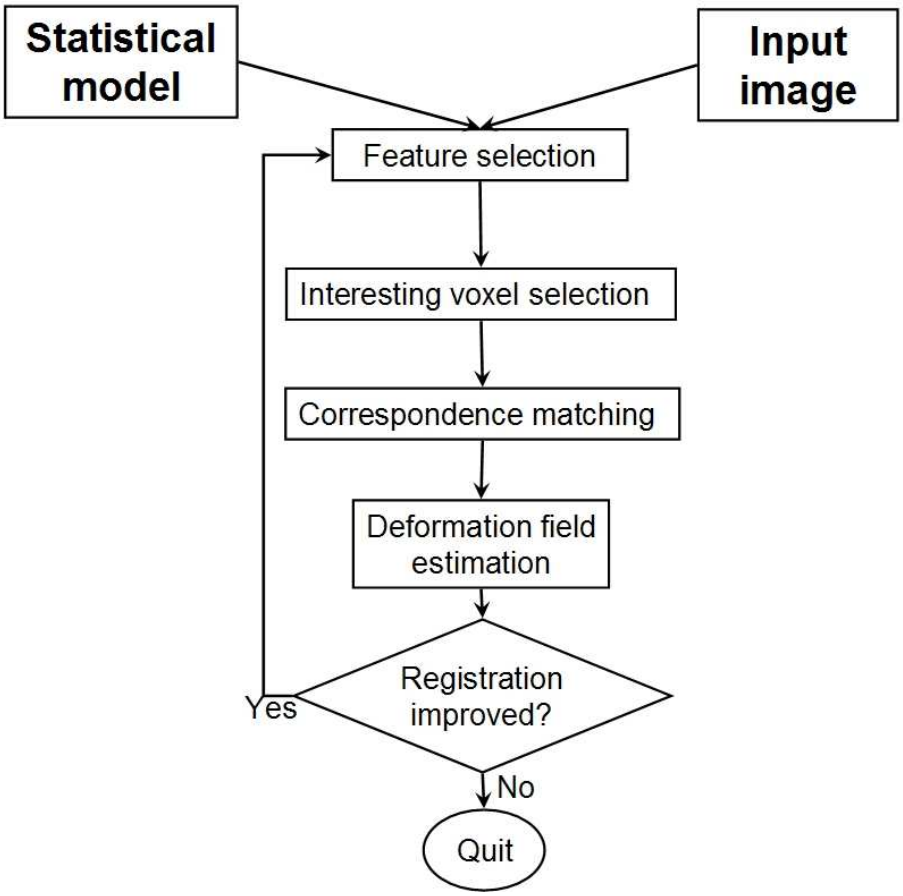

Figure 1: Block diagram of the proposed algorithm. Inputs to the algorithm are statistical model computed based on the reference image and an input image to be registered to the reference image. The algorithm starts by choosing a feature subset at random. Then we find interesting voxels of the reference image, i.e. voxels that can be matched correctly and with high confidence in this feature subspace to the corresponding voxels of the model. Under the same feature space we find corresponding voxels in the input image. Transformation between input and reference images is computed based on these correspondences and registration error is calculated. Then a new feature is added to the feature subset and the registration is repeated until the addition of any feature does not reduce registration error. Once the best feature subset is determined, we use it to find a potentially different set of interesting voxels, and repeat the above procedure again. The algorithm terminates if registration error is not decreasing any longer.

Table 2: Affine transformation parameters. A transformation is formed by composing rotation, skewing and scaling in the direction of X-axis. $18 \times 11 \times 9 = 1782$ affine transformation are constructed and applied to the reference image

	Minimum	Maximum	Step
Angle, ^{0}	0	340	20
Skew	0	0.5	0.05
Scale	0.8	1.2	0.05

3. Interesting voxel selection : it determines which voxels in the reference image are used as landmarks.

4. Deformation field estimation: this component fits a thin plate spline (TPS) transform [3] to the set of candidate correspondences.

5. Registration evaluation: this component computes the similarity measure between the reference image and the registered input image. We used mutual information and mean sum of square intensity differences as similarity measures.

3.2 Statistical model for image features

The statistical model for image features consists of two components. First component is a reference image, which prescribes what image coordinates each anatomical location should have in the registered image. The second component is a set of image features distributions. It contains information about how feature vectors are distributed for voxels in the reference image.
For a reference image we simply need to select an MR neuroimage of a healthy individual. Then image coordinates of, say, left tip of the corpus callosum in the input image after the registration should be the same as those of left tip of the corpus callosum in the reference image.
As for the second component, we have to estimate probability density function $f(X|v_i)$ which tells us the likelihood of observing feature vector X at the voxel at anatomical location v_i. Under the assumption that the components x_j of the feature vector X are independent we can factorize $f(X|v_i)$:

$$f(X|v_i) = \prod_{j=1}^{m} f(x_j|v_i), \tag{1}$$

where m is the dimensionality of X. Now we face an easier task of estimating $f(x_j|v_i)$ for every individual component of X. An ideal training set for this task would be a set of neuroimages of different subjects, where we know voxel by voxel correspondences between our selected reference image and every neuroimage in the set. However, such correspondences are sometimes semantically ambiguous [25] and are time consuming to produce by hand. Instead, we create our training set by applying 1782 different affine transformations to our reference image (see Table 2). We have an advantage of knowing exact voxel to voxel correspondences between the reference image and each of the transformed images. After we calculate feature vectors for every voxel in every image, we obtain a sample of 1782 feature vectors for each anatomical location v_i (see Figure 2).

These samples enable us to perform estimation of the distribution $f(x_j|v_i)$ of feature vector components x_j at each anatomical location v_i. In our experiments we chose to use parametric

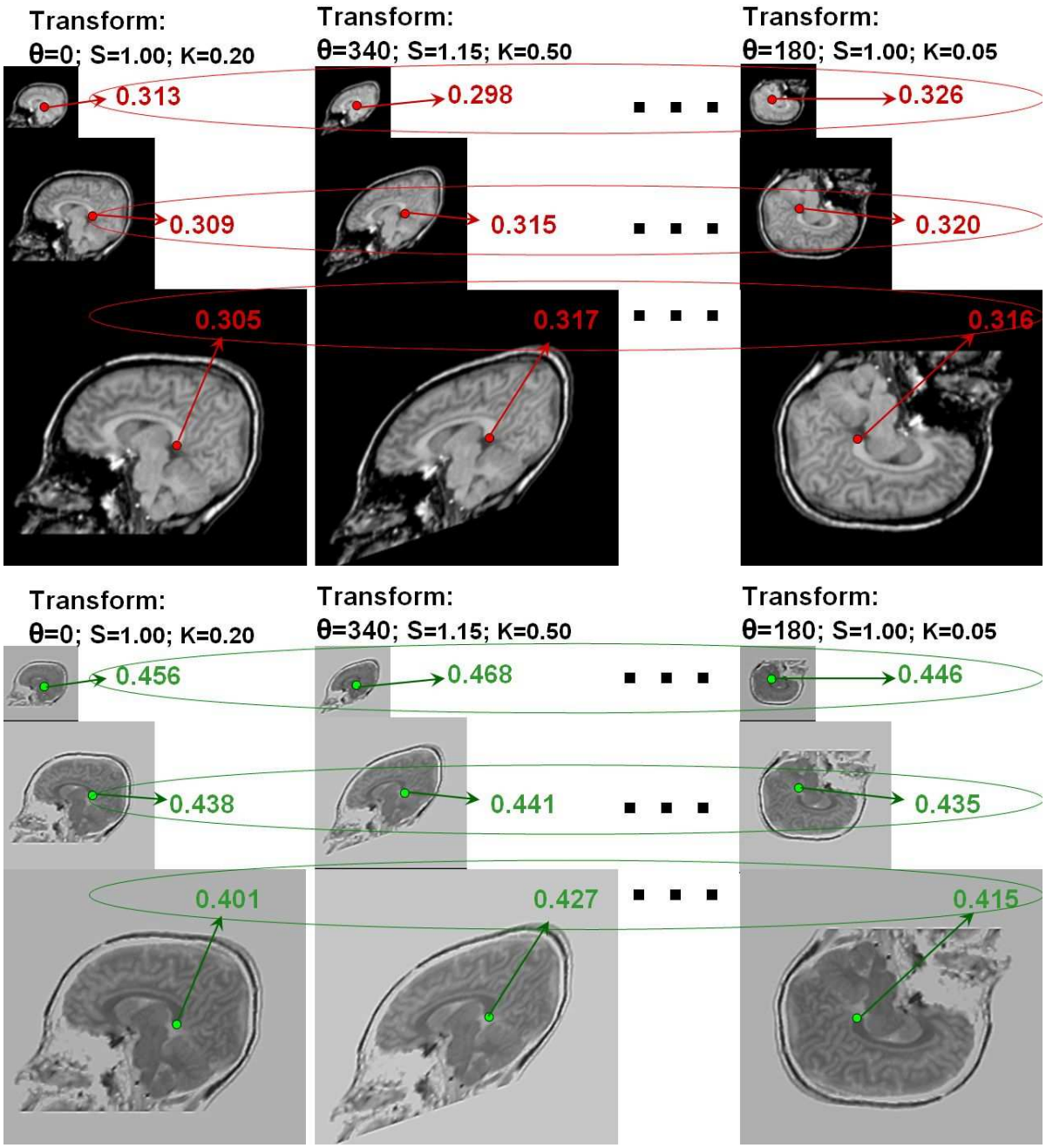

Figure 2: For each combination of angle θ, scaling S and skewing K shown in the Table 2 we compute an affine transform and apply it to the reference image. Features are computed for every transformed copy at three image scales. Values of *intensity_2_mean* (top) and *gabor_0_3* (bottom) features for a particular voxel are shown