Similarity Metrics for Medical Image Registration

Roger Tait
Nottingham Trent University, UK

Gerald Schaefer
Aston University, UK

INTRODUCTION

In general, image registration algorithms can be classified as either landmark or intensity-based. Landmark-based registration consists of four main stages:

- During the feature detection stage, distinguishing characteristics such as corners, edges, centres of gravity, and so forth are identified, either manually or automatically. This identification of landmarks is performed on both reference (fixed) and sensed (moving) images.
- The optimisation stage (Jenkinson & Smith, 2001) controls estimation of transform parameters that geometrically map landmarks between fixed and moving image.
- Upon selection of appropriate transform parameters, pixel values which are mapped into noninteger coordinates are interpolated in order to establish their value. This represents the image resampling stage (Grevera & Udupa, 1998).
- The feature matching stage is achieved through the use of a similarity metric in which a degree of closeness or accuracy of alignment between corresponding landmarks is calculated.

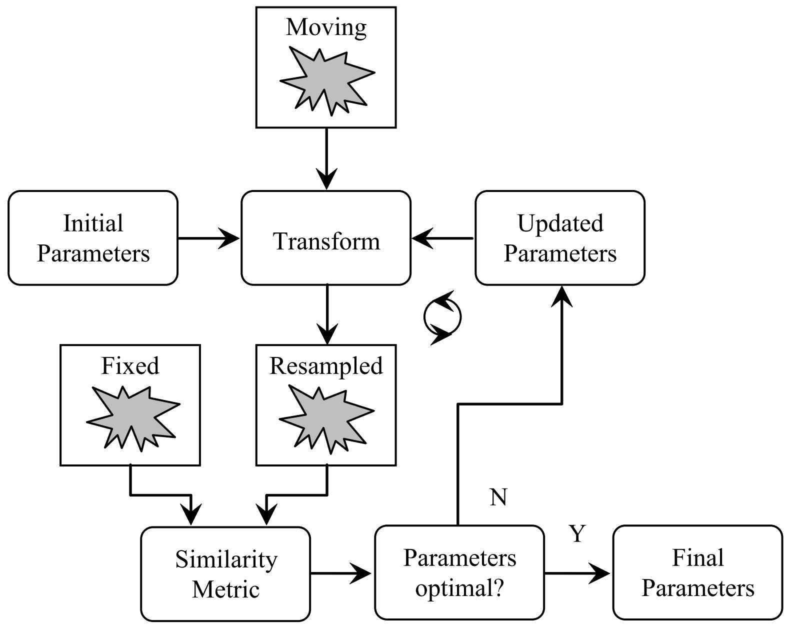
In intensity-based image registration methods, the feature detection stage is omitted. As a consequence, the transform optimisation and feature matching stages are performed using pixel intensities (or functions thereof) instead of landmarks. Intensity-based image registration algorithms comprise the following components:

- The spatial mapping of intensities throughout the alignment process is achieved with a transform component.
- An interpolation component is used to evaluate intensities at nondiscrete locations.
- The metric component calculates a measure of alignment accuracy.
- Optimisation of the similarity measure using a search space defined by transform parameters is achieved with an optimisation component.

The most important component of an image registration algorithm is the similarity metric used to determine when images are in accurate alignment (Penney, Weese, Little, Desmedt, Hill, & Hawkes, 1998). In Figure 1, the inputs to and output from a basic metric are illustrated. In general, a metric works by examining corresponding pixel values in both fixed and moving images and then formulating a measure of similarity based on the relationship between these intensities. The metric assumes that the relationship changes with variations in the spatial transformation used to map between images and a maximum similarity is achieved when the images are in close alignment (Brown, 1992).

Intensity equality which is high when pixels are similar is one such relationship employed as a similarity metric in single-modal registration where images are captured using the same sensor type. Total equality, however, is seldom reached due to noise and image acquisition inconsistencies. Additional robustness is therefore achieved by assessing the ratio of intensities and minimising the variance of such ratios. When images are acquired with different sensor types, as is typically the case in multimodal registration, an extension of the ratio method which maximises the weighted sum of variances can be employed. Alternatively, a relationship estimating the entropy of corresponding intensity pairs can be formulated where entropy, derived from information theory (Shannon, 1948), is the measure of the amount of information contained within a signal. Although many algorithms have been proposed, similarity calculation remains a complex task.

Figure 1. Flow diagram of similarity metric calculation



BACKGROUND

By comparing intensities, the metric component quantitatively measures how accurately fixed and moving images are aligned. The selection of a metric component is largely dependent on the type of registration problem to be solved (Roche, Malandain, Ayache, & Prima, 1999). For example, some metrics possess capture ranges that are well suited to images misaligned by a large transform. Other metrics, in contrast, are less computationally intensive but require initial transform parameters to be close to the optimum. During the alignment process, most metrics samples intensities over an entire image. Some metrics, however, employ a subset of samples drawn from the image. In both cases, similarity is calculated using intensities which fall within the boundary of the moving image. Intensity correlation has been used as a metric where a maximum similarity between fixed and moving images is searched for (Pratt, 1974). Using such an approach, high levels of alignment accuracy can be achieved by interpolating intensities before evaluating similarity. Such metrics are found predominantly in single-modal registration applications.

Although more difficult, the registration of images captured using different sensor types is commonplace in medical imaging applications. Viola (1995) suggests that for two images of differing modality, mutual information can be used as a measure of similarity where similarity is estimated using marginal and joint entropy based on probability distribution constructed using intensities from both fixed and moving images. As a consequence, mutual information can be described as the amount of information one image contains about another. Importantly, the ability to align multimodal images allows for the comparison of anatomical and functional data that can lead to a diagnosis which would be impossible to gain otherwise. For example, the evaluation of mutual information-based similarity metrics, used for the registration of brain scans, is presented by Holden et al. (2000).

SINGLE-MODAL REGISTRATION SIMILARITY METRICS

Intensity-based registration has been employed in the alignment of x-ray images and biomedical volume data (Russakoff, Rohlfing, & Maurer, 2003). In these applications, the cross-correlation of intensities can be used as a similarity metric. With such an approach, subpixel accuracy of alignment is achieved by interpolating intensities before similarity calculation. Traditionally, cross-correlation has been used to reg-

Related Content