Medical Image Analysis (1998) volume 2, number 1, pp 1–37

A Survey of Medical Image Registration

J. B. Antoine Maintz* and Max A. Viergever

Image Sciences Institute, Utrecht University Hospital, Utrecht, the Netherlands

Abstract
The purpose of this chapter is to present a survey of recent publications concerning medical image registration techniques. These publications will be classified according to a model based on nine salient criteria, the main dichotomy of which is *extrinsic* versus *intrinsic* methods The statistics of the classification show definite trends in the evolving registration techniques, which will be discussed. At this moment, the bulk of interesting intrinsic methods is either based on segmented points or surfaces, or on techniques endeavoring to use the full information content of the images involved.

Keywords: registration, matching

Received May 25, 1997; revised October 10, 1997; accepted October 16, 1997

1. INTRODUCTION

Within the current clinical setting, medical imaging is a vital component of a large number of applications. Such applications occur throughout the clinical track of events; not only within clinical diagnostis settings, but prominently so in the area of planning, consummation, and evaluation of surgical and radiotherapeutical procedures. The imaging modalities employed can be divided into two global categories: *anatomical* and *functional*. Anatomical modalities, *i.e.*, depicting primarily morphology, include X-ray, CT (computed tomography[a]), MRI (magnetic resonance imaging[b]), US (ultrasound[c]), portal images, and (video) sequences obtained by various catheter "scopes", *e.g.*, by laparoscopy or laryngoscopy. Some prominent derivative techniques are so detached from the original modalities that they appear under a separate name, *e.g.*, MRA (magnetic resonance angiography), DSA (digital subtraction angiography, derived from X-ray), CTA (computed tomography angiography), and *Doppler* (derived from US, referring to the Doppler effect measured). Functional modalities, *i.e.*, depicting primarily information on the metabolism of the underlying anatomy, include (planar) scintigraphy, SPECT (single photon emission computed tomography[d]), PET (positron emission tomography[e]), which together make up the *nuclear medicine* imaging modalities, and fMRI (functional MRI). With a little imagination, spatially sparse techniques like, EEG (electro encephalography), and MEG (magneto encephalography) can also be named functional *imaging* techniques. Many more functional modalities can be named, but these are either little used, or still in the pre-clinical research stage, *e.g.*, pMRI (perfusion MRI), fCT (functional CT), EIT (electrical impedance tomography), and MRE (magnetic resonance elastography).

Since information gained from two images acquired in the clinical track of events is usually of a complementary nature, proper *integration* of useful data obtained from the separate images is often desired. A first step in this integration process is to bring the modalities involved into spatial alignment, a procedure referred to as *registration*. After registration, a *fusion* step is required for the integrated display of the data involved. Unfortunately, the terms *registration* and *fusion*, as well as *matching*, *integration*, *correlation*, and others, appear polysemously in literature, either referring to a single step or to the whole of the modality integration process. In this paper, only the definitions of registration and fusion as defined above will be used.

An eminent example of the use of registering different modalities can be found in the area of epilepsy surgery. Patients may undergo various MR, CT, and DSA studies

---

*Corresponding author
(e-mail: Twan.Maintz@cv.ruu.nl)

[a] Also formerly and popularly CAT, computed axial tomography.

[b] Also referred to as NMR, nuclear magnetic resonance, spin imaging, and various other names.

[c] Also echo(graphy).

[d] Also SPET, single photon emission tomography.

[e] SPECT and PET together are sometimes referred to as ECAT (emission computerized axial tomography).

for anatomical reference; ictal and interictal SPECT studies; MEG and extra and/or intra-cranial (subdural or depth) EEG, as well as ^{18}FDG and/or ^{11}C-Flumazenil PET studies. Registration of the images from practically any combination will benefit the surgeon. A second example concerns radiotherapy treatment, where both CT and MR can be employed. The former is needed to accurately compute the radiation dose, while the latter is usually better suited for delineation of tumor tissue.

Besides multimodality registration, important application areas exist in monomodality registration. Examples include treatment verification by comparison of pre- and post-intervention images, comparison of ictal and inter-ictal (during and between seizures) SPECT images, and growth monitoring, *e.g.*, using time series of MR scans on tumors, or X-ray time series on specific bones. Because of the high degree of similarity between these images, solving the registration is usually an order of magnitude easier than in the multimodality applications.

This paper aims to provide a survey of recent literature concerning medical image registration. Because of the sheer volume of available papers, the material presented is by necessity heavily condensed, and –except for a few interesting and “classic” cases– no papers written before 1993 are referred to. Concerning publications pre-dating 1993, we refer the reader to review papers such as van den Elsen, Pol & Viergever (1993) and Maurer, McCrory, & Fitzpatrick (1993). No complete review papers of a later date exist to our knowledge, except for the field of computer aided surgery (Lavallée, 1996). To narrow the field of available publications in such a way does not, however, impede us in reaching our primary goal, which is to paint a comprehensive picture of current medical image registration methods.

2. CLASSIFICATION OF REGISTRATION METHODS

The classification of registration methods used in this chapter is based on the criteria formulated by van den Elsen, Pol & and Viergever (1993). A version considerably augmented and detailed is presented. Nine basic criteria are used, each of which is again subdivided on one or two levels. The nine criteria and primary subdivisions are:

I. Dimensionality

II. Nature of registration basis

  a. Extrinsic
  b. Intrinsic
  c. Non-image based

III. Nature of transformation

  a. Rigid
  b. Affine
  c. Projective
  d. Curved

IV. Domain of transformation

V. Interaction

VI. Optimization procedure

VII. Modalities involved

  a. Monomodal
  b. Multimodal
  c. Modality to model
  d. Patient to modality

VIII. Subject

  a. Intrasubject
  b. Intersubject
  c. Atlas

IX. Object

A registration procedure can always be decomposed into three major pillars: the *problem statement*, the *registration paradigm*, and the *optimization procedure*. The problem statement and the choice of paradigm and optimization procedure together provide a unique classification according to the nine criteria mentioned. Although pillars and criteria are heavily intertwined and have many cross-influences, it can be said that the problem statement determines the classification according to criteria **VII**, **VIII**, and **IX**, and has a direct bearing on the criteria **I** and **III**. The paradigm influences the criteria **II**, **III**, **IV**, and **V** most directly, while the optimization procedure influences criterion **V** and controls **VI**. It is often helpful to remember the three pillars are independent, since many papers do not describe them as such, often presenting problem statement, paradigm, and optimization procedure in a compounded way.

In the following sections, we will discuss the separate criteria in more detail.

3. DIMENSIONALITY

I. Dimensionality

  a. Spatial dimensions only:

    1. 2D/2D

  2. 2D/3D
  3. 3D/3D

b. Time series (more than two images), with spatial dimensions:

  1. 2D/2D
  2. 2D/3D
  3. 3D/3D

3.1. Spatial registration methods

The main division here is whether all dimensions are spatial, or that time is an added dimension. In either case, the problem can be further categorized depending on the number of spatial dimensions involved. Most current papers focus on the *3D/3D* registration of two images (no time involved). *3D/3D* registration normally applies to the registration of two tomographic datasets, or the registration of a single tomographic image to any spatially defined information, *e.g.,* a vector obtained from EEG data. *2D/2D* registration may apply to separate slices from tomographic data, or intrinsically 2D images like portal images. Compared to *3D/3D* registration, *2D/2D* registration is less complex by an order of magnitude both where the number of parameters and the volume of the data are concerned, so obtaining a registration is in many cases easier and faster than in the *3D/3D* case. We reserve *2D/3D* registration for the direct alignment of spatial data to projective data, (*e.g.,* a pre-operative CT image to an intra-operative X-ray image), or the alignment of a single tomographic slice to spatial data. Some applications register multiple 2D projection images to a 3D image, but since a usual preprocessing step is to construct a 3D image from the 2D projection images, such applications are best categorized as *3D/3D* applications. Since most *2D/3D* applications concern intra-operative procedures within the operating theater, they are heavily time-constrained and consequently have a strong focus on speed issues connected to the computation of the paradigm and the optimization. The majority of applications outside the operating theater and radiotherapy setting allow for off-line registration, so speed issues need only be addressed as constrained by clinical routine.

3.2. Registration of time series

Time series of images are acquired for various reasons, such as monitoring of bone growth in children (long time interval), monitoring of tumor growth (medium interval), post-operative monitoring of healing (short interval), or observing the passing of an injected bolus trough a vessel tree (ultra-short interval). If two images need to be compared, registration will be necessary except in some instances of ultra-short time series, where the patient does not leave the scanner between the acquisition of two images. The same observations as for spatial-only registrations apply.

4. NATURE OF REGISTRATION BASIS

II. Nature of registration basis

- **a.** Extrinsic
  1. Invasive
     - **A.** Stereotactic frame
     - **B.** Fiducials (screw markers)
  2. Non-invasive
     - **A.** Mould, frame, dental adapter, *etc.*
     - **B.** Fiducials (skin markers)
- **b.** Intrinsic
  1. Landmark based
     - **A.** Anatomical
     - **B.** Geometrical
  2. Segmentation based
     - **A.** Rigid models (points, curves, surfaces)
     - **B.** Deformable models (snakes, nets)
  3. Voxel property based
     - **A.** Reduction to scalars/vectors (moments, principal axes)
     - **B.** Using full image content
- **c.** Non-image based (calibrated coordinate systems)

4.1. Extrinsic registration methods

Image based registration can be divided into *extrinsic*, *i.e.,* based on foreign objects introduced into the imaged space, and *intrinsic* methods, *i.e.,* based on the image information as generated by the patient.

Extrinsic methods rely on artificial objects attached to the patient, objects which are designed to be well visible and accurately detectable in all of the pertinent modalities. As such, the registration of the acquired images is comparatively easy, fast, can usually be automated, and, since the registration parameters can often be computed explicitly, has no need for complex optimization algorithms. The main drawbacks of extrinsic registration are the prospective character, *i.e.,* provisions must be made in the pre-acquisition phase, and the often invasive character of the marker objects. Non-invasive markers can be used, but as a rule are less accurate. A commonly used fiducial object is a *stereotactic frame* (Lunsford, 1988; Vandermeulen, 1991; Lemieux *et al.*, 1994b; Lemieux and Jagoe, 1994; Strother *et al.*, 1994; Hemler *et al.*, 1995c; Vandermeulen *et al.*, 1995; Peters *et al.*, 1996) screwed rigidly to the patient's outer skull table, a device which until recently provided the "gold standard" for registration accuracy. Such frames are used for localization and guidance purposes in neurosurgery. Since neurosurgery is one of the main application areas of registration, the use of a stereotactic frame in the registration task does not add an additional invasive strain to the patient. However, the mounting of a frame for the sole purpose of registration is not permissible. Sometimes other invasive objects are used, such as screw-mounted markers (Gall and Verhey, 1993; Leung Lam *et al.*, 1993; Maurer *et al.*, 1993; Li *et al.*, 1994b; Maurer *et al.*, 1994; Maurer *et al.*, 1995b; Maurer *et al.*, 1995a; Simon *et al.*, 1995b; Ellis *et al.*, 1996), but usually non-invasive marking devices are reverted to. Most popular amongst these are markers glued to the skin (Evans *et al.*, 1991; Maguire *et al.*, 1991; Malison *et al.*, 1993; Wang *et al.*, 1994b; Wahl *et al.*, 1993; Bucholz *et al.*, 1994; Li *et al.*, 1994b; Edwards *et al.*, 1995a; Edwards *et al.*, 1995b; Leslie *et al.*, 1995; Stapleton *et al.*, 1995; Wang *et al.*, 1995; Fuchs *et al.*, 1996), but larger devices that can be fitted snugly to the patient, like individualized foam moulds, head holder frames, and dental adapters have also been used, although they are little reported on in recent literature (Greitz *et al.*, 1980; Laitinen *et al.*, 1985; Schad *et al.*, 1987; Hawkes *et al.*, 1992; Evans *et al.*, 1989; Evans *et al.*, 1991).

Since extrinsic methods by definition cannot include patient related image information, the nature of the registration transformation is often restricted to be rigid (translations and rotations only). Furthermore, if they are to be used with images of low (spatial) information content such as EEG or MEG, a calibrated video image or spatial measurements are often necessary to provide spatial information for basing the registration on. Because of the rigid-transformation constraint, and various practical considerations, use of extrinsic 3D/3D methods is largely limited to brain and orthopedic (Simon *et al.*, 1995b; Ellis *et al.*, 1996) imaging, although markers can often be used in projective (2D) imaging of any body area. Non-rigid transformations can in some cases be obtained using markers, *e.g.*, in studies of animal heart motion, where markers can be implanted into the cardiac wall.

4.2. Intrinsic registration methods

Intrinsic methods rely on patient generated image content only. Registration can be based on a limited set of identified salient points *(landmarks)*, on the alignment of segmented binary structures *(segmentation based)*, most commonly object surfaces, or directly onto measures computed from the image grey values *(voxel property based)*.

4.2.1. Landmark based registration methods

Landmarks can be *anatomical*, *i.e.*, salient and accurately locatable points of the morphology of the visible anatomy, usually identified interactively by the user (Evans *et al.*, 1989; Evans *et al.*, 1991; Hill *et al.*, 1991a; Hill *et al.*, 1991b; Maguire *et al.*, 1991; Zubal *et al.*, 1991; Henri *et al.*, 1992; Bijhold, 1993; Ding *et al.*, 1993; Fright and Linney, 1993; Gluhchev and Shalev, 1993; Hill *et al.*, 1993b; Morris *et al.*, 1993; Neelin *et al.*, 1993; Wahl *et al.*, 1993; Ge *et al.*, 1994; Harmon *et al.*, 1994; Moseley and Munro, 1994; Pietrzyk *et al.*, 1994; Strother *et al.*, 1994; Edwards *et al.*, 1995a; Edwards *et al.*, 1995b; Ge *et al.*, 1995; Hamadeh *et al.*, 1995b; Hamadeh *et al.*, 1995c; Leslie *et al.*, 1995; Meyer *et al.*, 1995; McParland and Kumaradas, 1995; Soltys *et al.*, 1995; Savi *et al.*, 1995; Stapleton *et al.*, 1995; Vandermeulen *et al.*, 1995; Zubal *et al.*, 1995; Christensen *et al.*, 1996; Evans *et al.*, 1996b; Evans *et al.*, 1996a; Erbe *et al.*, 1996; Fang *et al.*, 1996; Peters *et al.*, 1996; Rubinstein *et al.*, 1996), or *geometrical*, *i.e.*, points at the locus of the optimum of some geometric property, *e.g.*, local curvature extrema, corners, *etc*, generally localized in an automatic fashion (He *et al.*, 1991; Fontana *et al.*, 1993; Ault and Siegel, 1994; Eilertsen *et al.*, 1994; Thirion, 1994; Ault and Siegel, 1995; Uenohara and Kanade, 1995; Amit and Kong, 1996; Chua and Jarvis, 1996; Thirion, 1996a). Technically, the identification of landmark points is a segmentation procedure, but we reserve the classification *segmentation based* registration for methods relating to segmentation of structures of higher order, *i.e.*, curves, surfaces, and volumes. Landmark based registration is versatile in the sense that it –at least in theory– can be applied to any image, no matter what the object or subject is. Landmark based methods are mostly used to find rigid or affine transformations. If the sets of points are large enough, they can theoretically be used for more complex transformations. Anatomical landmarks are also often used in combination with an entirely different registration basis (Evans *et al.*, 1989; Evans *et al.*, 1991; Wahl *et al.*, 1993; Moseley and Munro, 1994; Hamadeh *et al.*, 1995c; McParland and Kumaradas, 1995; Zubal *et al.*, 1995; Christensen *et al.*, 1996; Evans *et al.*, 1996b): methods that rely on optimization of a parameter space that is not quasi-convex are prone to sometimes get stuck in local optima, possibly resulting in a large mismatch. By constraining the search space according to anatomical landmarks, such mismatches are unlikely to occur. Moreover, the search procedure can be sped up considerably. A drawback is that user interaction is usually required for the identification of the landmarks.

In landmark based registration, the set of identified points is sparse compared to the original image content, which makes for relatively fast optimization procedures. Such algorithms optimize measures such as the average distance (L_2 norm) between each landmark and its closest counterpart (the *Procrustean* metric), or iterated minimal landmark distances. For the optimization of the latter measure the *Iterative closest point* (ICP) algorithm (Besl and McKay, 1992) and derived methods are popular. Its popularity can be accredited to its versatility –it can be used for point sets, and implicitly and explicitly defined curves, surfaces and volumes–, computational speed, and ease of implementation. The Procrustean optimum can sometimes be computed, using *e.g.*, Arun's method (1987), but is more commonly searched for using general optimization techniques. Such techniques are referred to in section 7. Yet other methods perform landmark registration by testing a number of likely transformation hypotheses, which can, *e.g.*, be formulated by aligning three randomly picked points from each point set involved. Common optimization methods here are quasi-exhaustive searches, graph matching and dynamic programming approaches.

4.2.2. Segmentation based registration methods

Segmentation based registration methods can be *rigid model based* (Chen *et al.*, 1987; Levin *et al.*, 1988; Guéziec and Ayache, 1992; Jiang *et al.*, 1992b; Ayache *et al.*, 1993; Collignon *et al.*, 1993a; Fritsch, 1993; Gee *et al.*, 1993; Gee *et al.*, 1994; Gee *et al.*, 1995a; Gee *et al.*, 1995b; Gee and Haynor, 1996; Gilhuijs and van Herk, 1993; Hill *et al.*, 1993a; Kittler *et al.*, 1993; Miller *et al.*, 1993; Rusinek *et al.*, 1993; Tsui *et al.*, 1993; Turkington *et al.*, 1993; Zhao *et al.*, 1993; Collignon *et al.*, 1994; Ettinger *et al.*, 1994b; Ettinger *et al.*, 1994a; Feldmar and Ayache, 1994; Fritsch *et al.*, 1994b; Fritsch *et al.*, 1994a; Grimson *et al.*, 1994a; Grimson *et al.*, 1994b; Grimson *et al.*, 1994c; Hemler *et al.*, 1994a; Hemler *et al.*, 1994b; Huang and Cohen, 1994; Hata *et al.*, 1994; Henderson *et al.*, 1994; van Herk and Kooy, 1994; Kanatani, 1994; Krattenthaler *et al.*, 1994; Kooy *et al.*, 1994; Lavallée *et al.*, 1994; Liu *et al.*, 1994; Maurer *et al.*, 1994; Mendonça *et al.*, 1994; Péria *et al.*, 1994; Philips, 1994; Petti *et al.*, 1994; Simon *et al.*, 1994; Serra and Berthod, 1994; Szelisky and Lavallée, 1994; Szeliski and Lavallée, 1994; Scott *et al.*, 1994; Strother *et al.*, 1994; Staib and Xianzhang, 1994; Taneja *et al.*, 1994; Wang *et al.*, 1994a; Zuk *et al.*, 1994; Ardekani *et al.*, 1995; Andersson *et al.*, 1995; Andersson, 1995; Betting and Feldmar, 1995; Betting *et al.*, 1995; Burel *et al.*, 1995; Christmas *et al.*, 1995; Feldmar *et al.*, 1995; Grimson *et al.*, 1995; Henri *et al.*, 1995; Hemler *et al.*, 1995c; Hemler *et al.*, 1995b; Hemler *et al.*, 1995a; Hamadeh *et al.*, 1995b; Hamadeh *et al.*, 1995c; Hamadeh *et al.*, 1995a; Kruggel and Bartenstein, 1995; Lavallée and Szeliski, 1995; Leszczynski *et al.*, 1995; Maurer *et al.*, 1995a; Pellot *et al.*, 1995; Pallotta *et al.*, 1995; Pajdla and van Gool, 1995; Pennec and Thirion, 1995; Ryan *et al.*, 1995; Rizzo *et al.*, 1995; Simon *et al.*, 1995b; Simon *et al.*, 1995a; Serra and Berthod, 1995; Scott *et al.*, 1995; Sull and Ahuja, 1995; Troccaz *et al.*, 1995; Turkington *et al.*, 1995; Vassal *et al.*, 1995; Vandermeulen *et al.*, 1995; Xiao and Jackson, 1995; Zubal *et al.*, 1995; Declerc *et al.*, 1996; Evans *et al.*, 1996b; Ettinger *et al.*, 1996; Feldmar and Ayache, 1996; Grimson *et al.*, 1996; Gilhuijs *et al.*, 1996; Ge *et al.*, 1996; Goris *et al.*, 1996; Hemler *et al.*, 1996; Jain *et al.*, 1996; Lavallée *et al.*, 1996b; Lavallée *et al.*, 1996a; Qian *et al.*, 1996; Szeliski and Lavallée, 1996; Wang *et al.*, 1996c), where anatomically the same structures (mostly surfaces) are extracted from both images to be registered, and used as sole input for the alignment procedure. They can also be *deformable model based* (Bajcsy *et al.*, 1983; Guéziec, 1993; Taubin, 1993; Davatzikos and Prince, 1994; MacDonald *et al.*, 1994; Sandor and Leahy, 1994; Tom *et al.*, 1994; Bronielsen, 1995; Bainville *et al.*, 1995; Mangin *et al.*, 1995; Sandor and Leahy, 1995; Thirion, 1995; Cuisenaire *et al.*, 1996; Davatzikos *et al.*, 1996; Davatzikos, 1996; McInerney and Terzopoulos, 1996; Thirion, 1996b), where an extracted structure (also mostly surfaces, and curves) from one image is elastically deformed to fit the second image. The *rigid model based* approaches are probably the most popular methods currently in clinical use. Their popularity relative to other approaches is probably for a large part due to the success of the "head-hat" method as introduced by Pelizzari and co-workers (Chen *et al.*, 1987; Levin *et al.*, 1988; Pelizzari *et al.*, 1989; Chen and Pelizzari, 1989), which relies on the segmentation of the skin surface from CT, MR and PET images of the head. Since the segmentation task is fairly easy to perform, and the computational complexity relatively low, the method has remained popular, and many follow-up papers aimed at automating the segmentation step, improving the optimization performance, or otherwise extending the method have been published. Another popularity cause is the fast *Chamfer matching* technique for alignment of binary structures by means of a distance transform, introduced by Borgefors (1988). A drawback of segmentation based methods is that the registration accuracy is limited to the accuracy of the segmentation step. In theory, segmentation based registration is applicable to images of many areas of the body, yet in practice the application areas have largely been limited to neuroimaging and orthopedic imaging. The methods are commonly automated but for the segmentation step, which is performed semi-automatically most of the times.

With *deformable models* however, the optimization criterion is different: it is always locally defined and computed, and the deformation is constrained by elastic model-

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | **sales@docketalarm.com** | **1-866-77-FASTCASE**

fastcase®
Smarter legal research.