Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

A survey of medical image registration – under review

Max A. Viergever^{a,*}, J.B. Antoine Maintz^b, Stefan Klein^c, Keelin Murphy^d, Marius Staring^e, Josien P.W. Pluim^{f,a}

^a Image Sciences Institute, University Medical Center Utrecht, Utrecht, The Netherlands

^b Faculty of Science, Utrecht University, Utrecht, The Netherlands

^c Biomedical Imaging Group Rotterdam, Departments of Medical Informatics and Radiology, Erasmus MC, Rotterdam, The Netherlands

^d INFANT Research Centre, University College Cork, Cork, Ireland

^e Division of Image Processing, Leiden University Medical Center, Leiden, The Netherlands

^f Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands

A R T I C L E I N F O

Article history: Received 4 April 2016 Revised 17 June 2016 Accepted 17 June 2016 Available online 21 June 2016

Keyword: Medical image registration

ABSTRACT

A retrospective view on the past two decades of the field of medical image registration is presented, guided by the article "A survey of medical image registration" (Maintz and Viergever, 1998). It shows that the classification of the field introduced in that article is still usable, although some modifications to do justice to advances in the field would be due. The main changes over the last twenty years are the shift from extrinsic to intrinsic registration, the primacy of intensity-based registration, the breakthrough of nonlinear registration, the progress of inter-subject registration, and the availability of generic image registration software packages. Two problems that were called urgent already 20 years ago, are even more urgent nowadays: Validation of registration methods, and translation of results of image registration has evolved, but still is in need of further development in various aspects.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

One of the early articles published in Medical Image Analysis was "A survey of medical image registration" by Maintz and Viergever (1998). The aim of the article was to present a comprehensive and structured record of approaches to registration of medical images. The article has been influential in the medical image analysis literature ever since, with > 3600 citations in Google Scholar and still 200 citations/year in the past few years.

This anniversary issue of the journal is a suitable occasion to review the contents of the article, in particular to take stock of what has changed over the last two decades in medical image registration. Is the classification proposed in the article still useful? Have observed trends continued, increased, or decreased? Are other striking observations still valid? Has the field changed in a way that was not foreseen then? And have the major problems identified at that time been addressed and solved?

These issues will be discussed in the following sections.

2. Is the classification proposed in 1998 still useful?

The article of Maintz and Viergever was not just a survey of image registration papers published until then, but in addition proposed a scheme to classify image registration methods in terms of nine distinctive characteristics. Slightly to our surprise, the classification setup is still quite functional, with as criteria (i) dimensionality (spatial or spatiotemporal 2D/2D, 2D/3D, 3D/3D), (ii) nature of the registration basis (extrinsic, intrinsic, non-image based), (iii) nature of the transformation (rigid, affine, projective, curved), (iv) domain of the transformation (global, local), (v) degree of interaction (interactive, semi-automatic, automatic), (vi) optimization procedure (parameters computed or searched for), (vii) modalities involved (mono-modality, multi-modality, modality to model, patient to modality), (viii) subjects involved (intra-subject, inter-subject, atlas), (ix) objects involved (e.g., brain, heart, breast).

The article typified extrinsic vs. intrinsic registration as the main dichotomy of the classification scheme. This is no longer valid. While extrinsic registration is not completely obsolete, it only features in a restricted number of applications. Furthermore, we would nowadays formulate some of the criteria slightly differently, and maybe add one or two as, e.g., pairwise (n=2 images) vs. groupwise (n > 2 images) registration or asymmetric vs. symmetric formulations. Also, subdivisions of some of the categories



Editorial



^{*} Corresponding author. *E-mail addresses*: max@isi.uu.nl (M.A. Viergever), j.b.a.maintz@uu.nl (J.B.A. Maintz), s.klein@erasmusmc.nl (S. Klein), keelinm@gmail.com (K. Murphy), m.staring@lumc.nl (M. Staring), j.pluim@tue.nl (J.P.W. Pluim).

would be due. For example, the category of optimization procedures could be divided into continuous and discrete methods, and for the category of curved transformations one could consider distinguishing small-deformation (or: elastic) and large-deformation (or: fluidic, based on integration of velocity fields) methods. And finally, the recent literature on curved registration comprises innovative proposals for transformation modelling, regularization, and optimization, which often appear intertwined. However, this does not preclude classification according to the original framework. It is still fairly straightforward to categorize these methods by the nature of the transformation and by the optimization procedure. So, overall, the classification scheme seems very usable a score of years after its conception. It could be readily updated to comprise all state-of-the art registration approaches, but this is beyond the scope of the present article.

3. Have observed trends continued, increased, or decreased?

Several trends in image registration approaches were formulated by Maintz and Viergever.

First, a shift from extrinsic to intrinsic registration was noted, even though clinically employed methods were generally extrinsic then. This trend has continued apace. In image registration research, extrinsic approaches are hardly found any more. In clinical applications where image registration is used, intrinsic methods are gaining ground, although in surgical and radiotherapeutical procedures, extrinsic matching remains in use.

Second, while surface-based methods were the most often used type of intrinsic registration at that time, it was observed that they had to give way to methods based on properties of individual voxels. This trend has certainly increased. Computational hurdles to applying voxel-based registration have rapidly diminished, so that it became feasible to take the full image contents into account in registration procedures rather than having to rely on segmentation of image objects that subsequently had to be aligned. It is noteworthy that point-based (often anatomical landmark-based) approaches still have their place in image registration, much more so than surface-based methods.

Third, it was mentioned that the need for creating public data bases of representative images and for assembling image registration validation protocols was emerging. These issues are still urgent, even though noticeable progress has been made on each of them. Several data sets with expert landmark annotations have become available in the last decade. Most of these concern manually delineated segmentations of structures, which are intended for evaluation of image segmentation methods but may also be used for evaluation of registration approaches. For example, public data sets of segmented MR brain images as IBSR (http://www.nitrc.org/ projects/ibsr) and LPBA40 (http://www.loni.usc.edu/atlases/Atlas_ Detail.php?atlas_id=12) have been used for this purpose in studies on evaluation of registration accuracy, see e.g. Klein et al. (2009). We would, however, like to draw the readers' attention to the study by Rohlfing (2012), which shows that the approach of evaluating registration algorithms on the basis of image similarity and tissue overlap measures has severe shortcomings and hence should be used with caution. Just a few data bases have been set up specifically for evaluation of registration methods, all concerning deformable thoracic image registration, and primarily aimed at registration of inspiration/expiration scans of the lungs. These annotated data sets are provided by: DirLab (http://www. dir-lab.com), POPI (http://www.creatis.insa-lyon.fr/rio/popi-model), and EMPIRE10 (http://empire10.isi.uu.nl). EMPIRE10 was launched as an evaluation challenge in conjunction with MICCAI 2010. Training data were made publicly available, and research groups could participate in the challenge by describing their approach and submitting its results, whereupon feedback was provided. The challenge is described in Murphy et al. (2011a). It is still open for submission, and currently lists 41 algorithm results from 28 first authors. Remarkably enough, it is the only challenge on image registration listed in the Grand Challenges repository (http: //grand-challenge.org), the more so since one of the earliest medical image analysis evaluation challenges, if not the first, dealt with image registration. It was the Retrospective Registration Evaluation Project (RREP), set up by J. Michael Fitzpatrick (West et al., 1997). It concerned an evaluation of algorithms for rigid registration of CT, MR and PET images of the human head, aimed at support of neurosurgical procedures. The gold standard was obtained by registration of markers screwed into patients' heads (as part of the clinical protocol). The challenge was continued as the Retrospective Image Registration Evaluation (RIRE) project, and is hosted by Kitware since 2007 (http://www.insight-journal.org/rire). It is still active, and currently counts > 400 submitting authors (!).

4. Are other striking observations still valid?

The article (Maintz and Viergever, 1998) furthermore contains several interesting observations, not explicitly formulated as trends. These include:

- Registration is seldom used in diagnostic clinical practice, even though for some procedures the advantages of using registered images are obvious. This assertion still largely holds true for diagnostic medical specialties, including notably radiology. Rigid registration, which is generally present in commercial medical image analysis packages, may be used for some multi-modality protocols. For many diagnostic processes, however, nonlinear registration would be due, e.g. to detect changes in disease progression. While relatively fast methods for nonlinear registration have been developed in research settings, such methods have not reached the status of inclusion in commercial software that supports clinical diagnoses, for lack of genericity and robustness. The possibility to build fast and reliable image analysis pipelines using generic modules (preprocessing, registration, segmentation) may change this for the better at short notice, at the very least within another score of years.
- Intra-operative registration in surgical procedures and image registration for patient positioning in radiotherapy are used in the clinic with good results. This observation appears to have been a bit optimistic as concerns surgical procedures. At that time, neurosurgery pioneered with image registration methods for surgical guidance, true, but these methods have not found their way to routine clinical practice with the exception of rigid registration based on fiducial markers for neuronavigation. Registration is more widespread, however, in the presurgical stages of therapy selection and therapy planning. Especially in functional neurosurgery - and in the associated discipline of clinical neurophysiology - registration of images from quite a few different modalities is part of the clinical workflow. Image registration is furthermore on the rise in interventional radiology and cardiology, where 3D/2D registration aimed at integrating pre-interventional 3D information (CT, MRI, 3DRX) with 2D Xray intervention images for navigation purposes, is becoming available in clinical intervention software. And finally, for radiotherapy the picture is much more favourable. In fact, radiotherapy is probably the clinical specialty where image registration is used most prominently. Not only are fast and trusted rigid registration techniques at the disposal of the radiation oncologist for patient positioning in linear accelerators, image registration is also increasingly used in diagnosis and tumour staging, in treatment planning and guidance, and in response monitoring.

- Global rigid registration is currently the most frequently used registration approach in clinical procedures. This is, alas, still correct. While image registration research focuses almost exclusively on nonlinear methods, such methods have not nearly attained broad clinical acceptance.
- The level of accuracy needed for clinical purposes is generally not known, and cannot be readily quantified even by expert clinicians. This is an interesting point, which has only sparsely been addressed in the literature, but is nonetheless very topical as yet. Apart from the fact that the needed level of accuracy is difficult to assess, if at all, validation in a clinical setting has proven notoriously cumbersome, since anatomical and especially pathological variations are not readily included in a validation protocol. Accordingly, evaluation of – in particular nonlinear – registration methods has generally been restricted to controlled studies using, e.g., simulations or phantoms. Among the few exceptions of evaluation of nonlinear registration accuracy that do employ clinical data is the EMPIRE10 challenge mentioned earlier (Murphy et al., 2011a).
- There are as yet few publications on curved transformations and on inter-subject registration. This is no longer valid. There is a multitude of publications on nonlinear registration, witness the > 400 references on such methods in the survey article of Sotiras et al. (2013). The number of publications on intersubject registration is not that large, but nevertheless quite considerable.
- Registration and visualization are still quite separate topics. This was further specified: not many registration approaches use state-of-the-art visualization, not many visualization approaches use registered images as input. We have witnessed that these areas have learned from each other, and thereby strengthened each other. So, while the two topics have continued to be quite separate, the specification no longer holds true.
- Many mono-modal registration problems appear to have been solved satisfactorily. Another quite optimistic statement, which is possibly true for global rigid registration, but not for nonlinear registration nor for the related problem of motion correction. The amount of ongoing research in mono-modal registration is not nearly negligible. So, with regard to the above statement, there is no progress to be reported, unless progressive insight counts as such.

5. Has the field changed in a way that was not foreseen then?

While it is remarkable that quite a few of the observations made 20 years ago are still valid, the field of medical image registration has witnessed huge progress in this period. Rather than the then standard problem of just aligning two multi-modality images of the same patient, image registration has become multi-faceted, with issues as change detection, motion detection and correction, atlas-based segmentation, and groupwise registration. Four developments stand out:

- Registration research has focused largely on nonlinear registration (or 'curved' registration, as it was called in the original article); see (Sotiras et al., 2013) for a record of recent advances on this topic.
- Intensity-based ('voxel-based') registration has become the method of choice also in multi-modal applications. The increased use of mutual information as a similarity measure has played a prominent role in this process. See (Wells et al., 1996; Maes et al., 1997) for introductory articles on mutualinformation-based image registration, and (Pluim et al., 2003) for a survey.
- Inter-subject registration has gotten a larger share in registration research and applications. One important application is

atlas-based segmentation, where image registration is a key element to align a set of reference (atlas) images to a new target image, which is subsequently segmented by transferring labelled (often expert-annotated) structures of interest in the atlas images to it. Another widespread application is the comparison of – registered – images of a group of patients with those of a control group as an approach to finding imaging biomarkers.

• The field has profited greatly from the creation of generic registration software packages. It would not be feasible to even try and mention all of them, so we confine ourselves to a few examples of currently frequently employed packages: ANTs (http://stnava.github.io/ANTs), NiftyReg (http://cmictig.cs.ucl.ac. uk/wiki/index.php/NiftyReg), elastix (http://elastix.isi.uu.nl), and registration modules of the ITK toolkit (http://www.itk.org). As an example of the popularity of these packages, we give some statistics on our own software toolbox elastix (Klein et al., 2010). The software has been downloaded > 30k times, with at present around 150 downloads/week. The mailing list has 325 subscribers, and in 2015 there were > 350 requests for assistance.

6. Have the major problems identified at that time been addressed and solved?

The major problems identified in (Maintz and Viergever, 1998) were not related to the methodology of registration, but rather to the situation after a registration has been obtained. Two major issues were identified and discussed from various viewpoints, viz. validation and clinical acceptance, whereby it should be noted that the second issue depends on the first. Quoting: "Many methods can still be considered barred from meaningful clinical application by the fact that they are as yet improperly validated".

Validation of registration methods and results was considered a major, but not insoluble problem. In fact, the sentence following the quote above stated that "the proper verification methods are known in most cases" and just needed "painstaking work of conducting the many experiments involved". This positive statement, by which the article was concluded, appears however somewhat in contradiction with two statements earlier in the article, viz. that "we cannot, with absolute certainty, quantify local registration errors" and that "the actual level of accuracy needed is still an unknown in many applications, and cannot be quantified accurately, even by the clinicians involved".

Currently, we still consider validation of – especially nonlinear – image registration methods and results non-trivial. We have made miles here in the past 20 years, but not as many as we had expected or even hoped. Evaluation of registration approaches in controlled studies using computer simulations or physical phantoms has limited value at best, see e.g. (Murphy et al., 2011a), and inclusion of the overwhelmingly rich variety in anatomy and pathology in a validation protocol poses a huge challenge. It is of paramount importance to turn our efforts not only to executing validation studies along well-established lines of action, but particularly also to reflecting upon the design of suitable validation protocols. Examples from our own group in this direction are (Murphy et al., 2011b; Muenzing et al., 2012).

As for clinical acceptance, the assertions in (Maintz and Viergever, 1998) again seem somewhat contradictory. On the one hand, "registration is rarely used in many clinical applications, even though such applications may benefit from registered images", whereas on the other hand "intra-operative registration and methods on patient positioning are in clinical use with apparent good results at a number of sites". The latter sentence shows that registration had earned its place in image-guided interventional research and – to a lesser extent – clinical practice, while the former

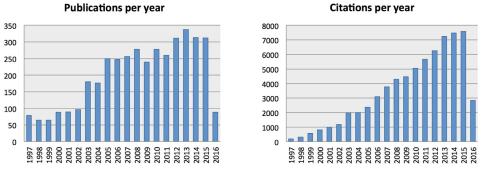


Fig. 1. Number of publications (left) and number of citations (right) for the search term "medical image registration", from Web of Science (Thomson Reuters), 21.06.16.

sentence largely refers to the almost complete absence of image registration in diagnostic procedures.

Over the past 20 years, image registration has significantly increased its role in a few medical specialties. Foremost amongst these is radiotherapy, where image registration has pervaded clinical practice in various ways, in diagnosis and tumour staging, in treatment planning and delivery, in dose accumulation, and in response monitoring. In neurosurgery and in interventional radiology and cardiology, navigation tools have been translated from a research setting to the clinic. Moreover, in functional neurosurgery and in clinical neurophysiology, multi-modality registration of images is part of presurgical patient management to optimize surgical planning. In other diagnostic and interventional disciplines, image registration still plays a modest role, although it is worthwhile mentioning that in nuclear medicine image fusion is omnipresent now thanks to the rapid developments in hybrid imaging scanners (PET-CT, SPECT-CT, PET-MRI), which provide integrated functional and anatomical images for diagnostic purposes.

7. Conclusion and discussion

We have loosely reviewed how the field of medical image registration has evolved over the past two decades, with the article "A survey of medical image registration" of Maintz and Viergever (1998) as guideline. We briefly summarize and discuss our major findings, as responses to the questions posed in the Introduction section.

The classification of the field introduced in the 1998 article appears still of use, subject to some modifications to take advances in the field into account. This conclusion is endorsed by the setup of a recent review article on 3D/2D registration methods for image-guided intervention (Markelj et al., 2012), where the classification scheme of (Maintz and Viergever, 1998) was adopted with only minor adaptations.

The major trends observed in the lead article were (i) the shift from extrinsic to intrinsic registration, which has at least continued; (ii) the shift from surface-based registration to intensitybased ("voxel-property-based") registration, which has certainly increased, with intensity-based techniques now forming the basis of the vast majority of registration approaches; and (iii) the emerging need of public data bases of representative, expert-annotated images and of validation protocols, which is still very much topical. Accordingly, the observed trends were quite indicative of some of the developments of the field.

Other striking observations comprised (iv) the rare use of registration in diagnostic clinical practice, which has not improved much, since it has proven difficult to devise registration methods that are robust against the many variations encountered in clinical practice of, e.g., scanner type, scanning protocol, and – pre-eminently – patient characteristics; (v) the clinical acceptance of registration in surgical procedures and radiotherapy, which has progressed especially in the latter discipline; (vi) the dominant position of rigid registration in clinical procedures, which has hardly decreased; (vii) the unfeasibility of quantifying the level of registration accuracy needed for clinical decisions, which has remained invariably true; (viii) the near absence of literature on curved transformations and on inter-subject registration, which has been more than made up for in the last two decades; (ix) the lack of interaction of the fields of registration and visualization, which has much improved and is hardly felt nowadays; and (x) the view that most mono-modality registration problems have been solved, which appears to have been too enthusiastic.

So, while progress in medical image registration has been considerable, several observations made twenty years ago continue to be valid for the present state of the art of the field, or incidentally have even proven optimistic.

The progress of the field can be assessed well by listing developments that were not foreseen in (Maintz and Viergever, 1998). The four most striking ones are the dominance of intensity-based approaches – taking into account the full contents of the images to be registered – over segmentation-based and landmark-based approaches, the upswing of nonlinear ('curved') transformations especially in registration research, the expansion of inter-subject registration e.g. to enable atlas-based segmentation, and the emergence of publicly available, generic and easy-to-use software packages for medical image registration.

This all suggests we are well underway with completing the development of the field of medical image registration. However, the two major problems mentioned in (Maintz and Viergever, 1998) – validation of registration methods and translation of these to the clinic – are major problems still, which have even been aggravated by the elaboration of registration methods. Consequently, medical image registration is far from done yet. The field is very much alive, witness Fig. 1, and has important issues to tackle still.

It will be interesting to see how the field will have advanced another 20 years from now. If we extrapolate the accomplishments of the past two decades into the next two, we may expect a wealth of methodological innovations comprising more intricate transformation models, similarity measures, and optimization techniques. It is unlikely that mutual information will be able to maintain its popularity, given the need for local measures of image similarity. Deep learning approaches to image registration could very well be the new game changer, although their use is severely challenged by one of the striking negative characteristics of the past score of years, the continuous lack of reference data sets for validation, and thus for learning. Both for validation of image registration methods and for translation of image registration approaches into daily clinical practice, the extrapolated developments yield a bleak outlook. We would like to see a radical change of this picture and therefore strongly advocate a shift of attention to the aspects of validation and clinical acceptance, possibly fueled by the desire to apply deep learning concepts, so that in 20 years (or preferably sooner) image registration is an integral part of the entire spectrum of routine clinical imaging.

We look forward to reviewing these advancements at the occasion of a next anniversary of the journal, but for now conclude by acknowledging the great progress we have seen on the topic of medical image registration in the past 20 years, and by commending Medical Image Analysis for the leading role it has played in realizing this. We congratulate the editors, the board and the staff of the journal on a job very well done. Cheers!

References

- Klein, A., Andersson, J., Ardekani, B.A., Ashburner, J., Avants, B., Chiang, M.C., Christensen, G.E., Collins, D.L., Gee, J., Hellier, P., Song, J.H., Jenkinson, M., Lepage, C., Rueckert, D., Thompson, P., Vercauteren, T., Woods, R.P., Mann, J.J., Parsey, R.V., 2009. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. NeuroImage 46, 786–802.
- Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P.W., 2010. Elastix A toolbox for intensity-based medical image registration. IEEE Trans. Med. Imaging 29, 196–205.
- Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P., 1997. Multimodality image registration by maximization of mutual information. IEEE Trans. Med. Imaging 16, 187–198.
- Maintz, J.B.A., Viergever, M.A., 1998. A survey of medical image registration. Med. Image Anal 2, 1–36.
- Markelj, P., Tomaževič, D., Likar, B., Pernuš, F., 2012. A review of 3D/2D registration methods for image-guided interventions. Med. Image Anal. 16, 642–661.
- Muenzing, S.E.A., Van Ginneken, B., Murphy, K., Pluim, J.P.W., 2012. Supervised quality assessment of medical image registration: Application to intra-patient CT lung registration. Med. Image Anal. 16, 1521–1531.

- Murphy, K., Van Ginneken, B., Klein, S., Staring, M., De Hoop, B.J., Viergever, M.A., Pluim, J.P.W., 2011. Semi-automatic construction of reference standards for evaluation of image registration. Med. Image Anal. 15, 71–84.
- Murphy, K., Van Ginneken, B., Reinhardt, J.M., Kabus, S., Ding, K., Deng, X., Cao, K., Du, K., Christensen, G.E., Garcia, V., Vercauteren, T., Ayache, N., Commowick, O., Malandain, G., Glocker, B., Paragios, N., Navab, N., Gorbunova, V., Sporring, J., De Bruijne, M., Han, X., Heinrich, M.P., Schnabel, J.A., Jenkinson, M., Lorenz, C., Modat, M., McClelland, J.R., Ourselin, S., Muenzing, S.E.A., Viergever, M.A., De Nigris, D., Collins, D.L., Arbel, T., Peroni, M., Li, R., Sharp, G.C., Schmidt-Richberg, A., Ehrhardt, J., Werner, R., Smeets, D., Loeckx, D., Song, G., Tustison, N., Avants, B., Gee, J.C., Staring, M., Klein, S., Stoel, B.C., Urschler, M., Werlberger, M., Vandemeulebroucke, J., Rit, S., Sarrut, D., Pluim, J.P.W., 2011. Evaluation of registration methods on thoracic CT: the EMPIRE10 challenge. IEEE Trans. Med. Imag, 30, 1901–1920.
- Pluim, J.P.W., Maintz, J.B.A., Viergever, M.A., 2003. Mutual-information-based registration of medical images: a survey. IEEE Trans. Med. Imag. 22, 986–1004.
- Rohlfing, T., 2012. Image similarity and tissue overlaps as surrogates for image registration accuracy: widely used but unreliable. IEEE Trans. Med. Imag. 31, 153–163.
- Sotiras, A., Davatzikos, C., Paragios, N., 2013. Deformable image registration: A survey. IEEE Trans. Med. Imag. 32, 1153–1190.
- Wells, W.M.III., Viola, P., Atsumi, H., Nakajima, S., Kikinis, R., 1996. Multi-modal volume registration by maximization of mutual information. Med. Image Anal 1, 35–51.
- West, J., Fitzpatrick, J.M., Wang, M.Y., Dawant, B.M., Maurer Jr., C.R., Kessler, R.M., Maciunas, R.J., Barillot, C., Lemoine, D., Collignon, A., Maes, F., Suetens, P., Vandermeulen, D., Van den Elsen, P.A., Napel, S., Sumanaweera, T.S., Harkness, B., Hemler, P.F., Hawkes, D.J., Hill, D.L.G., Studholme, C., Maintz, J.B.A., Viergever, M.A., Malandain, G., Pennec, X., Noz, M.E., Maguire Jr., G.Q., Pollack, M., Pelizzari, C.A., Robb, R.A., Hanson, D., Woods, R.P., 1997. Comparison and evaluation of retrospective intermodality brain image registration techniques. J. Comput. Assist. Tomogr. 21, 554–566.