# An efficient strategy based on an individualized selection of registration methods. Application to the coregistration of MR and SPECT images in neuro-oncology

Jean-Marc Tacchella, Elodie Roullot, Muriel Lefort, Mike-Ely Cohen, Rémy Guillevin, Grégorio Petrirena, Jean-Yves Delattre, Marie-Odile Habert, Nathanaëlle Yeni, Aurélie Kas, et al.

## HAL Id: hal-01084141
## https://hal.sorbonne-universite.fr/hal-01084141

# An Efficient Strategy Based on an Individualized Selection of Registration Methods. Application to the Coregistration of MR and SPECT Images in Neuro-Oncology.

Jean-Marc Tacchella[1,2], Elodie Roullot[2,3], Muriel Lefort[1], Mike-Ely Cohen[4], Rémy Guillevin[5], Grégorio Petrirena[6], Jean-Yves Delattre[6], Marie-Odile Habert[1,7], Nathanaëlle Yeni[1,7], Aurélie Kas[1,7] and Frédérique Frouin[1*]

[1]Sorbonne Universités, UPMC Univ Paris 06, CNRS, INSERM, Laboratoire d'Imagerie Biomédicale, F-75013, Paris, France

[2] ESME Sudria, Laboratoire PRIAM, Ivry-sur-Seine, France.

[3] Université Paris-Est, LISSI, Créteil, France.

[4] Université de Montréal, Centre de Recherche de l'Institut Universitaire de Gériatrie de Montréal, Montréal, Québec, Canada

[5] Université de Poitiers and CHU Poitiers, Service de Radiologie, Poitiers, France.

[6] AP-HP, CHU Pitié-Salpêtrière, Service de Neurologie, Paris, France.

[7] CHU Pitié-Salpêtrière, Service de Médecine Nucléaire, Paris, France.

* Corresponding author: frederique.frouin@inserm.fr

**Abstract**

An efficient registration strategy is described that aims to help solve delicate medical imaging registration problems. It consists in running several registration methods for each dataset and selecting the best one for each specific dataset, according to an evaluation criterion. Finally, the quality of the registration results, obtained with the best method, is visually scored by an expert as excellent, correct or poor. The strategy was applied to coregister Technetium-99m Sestamibi SPECT and MRI data in the framework of a follow-up protocol in patients with high grade gliomas receiving antiangiogenic therapy. To adapt the strategy to this clinical context, a robust semi-automatic evaluation criterion based on the physiological uptake of the Sestamibi tracer was defined. A panel of eighteen multimodal registration algorithms issued from BrainVisa, SPM or AIR software environments was systematically applied to the clinical database composed of sixty-two datasets. According to the expert visual validation, this new strategy provides 85% excellent registrations, 12% correct ones and only 3% poor ones. These results compare favorably to the ones obtained by the globally most efficient registration method over the whole database, for which only 61% of excellent registration results have been reported. Thus the registration strategy in its current implementation proves to be suitable for clinical application.

# 1. Introduction

Gliomas are the most frequent primary central nervous system tumors in adults. Conventional treatments for gliomas include surgery and radiotherapy, and possibly chemotherapy in case of tumor recurrences. Recent studies also address new emerging therapies such as antiangiogenic treatments that aim at reducing the neovascularization involved in the tumor growth. However, the prognosis of patients with high-grade gliomas remains poor and these new therapies induce strong side effects; early evaluation of the tumor response using non-invasive neuroimaging is therefore essential to establish the most suitable treatment strategy. Technetium-99m labeled Sesta-methoxyisobutyl-isonitrile (Sestamibi) is a radiopharmaceutical that accumulates in malignant gliomas(Soler *et al* 1998). The usefulness of Sestamibi brain Single-Photon Emission Computed Tomography (SPECT) in the follow-up of patients with high grade gliomas under chemotherapy was suggested (Bleichner-Perez *et al* 2007), (Prigent-Le Jeune *et al* 2004). In addition, it allows tumor recurrences to be distinguished from radionecrosis due to radiotherapy. A new investigation protocol was proposed to assess the potential value of Sestamibi SPECT imaging in addition to Gadolinium (Gd) enhanced MRI studies in patients with high grade gliomas and under antiangiogenic treatment. In order to combine anatomical and functional local information, a very accurate co-registration of both modalities is necessary. To avoid the labor-intensive task of manual registration and the associated intra- and inter-operator variability, automatic registration algorithms (Hellier *et al* 2003, Brown 1992, Zitová and Flusser 2003, Holden 2008, Maintz and Viergever 1998, Lester and Arridge 1999, Hill *et al* 2001, Jannin *et al* 2001, Crum 2004, Gholipour *et al* 2007, Oliveira and Tavares 2012) may be used. However, because of the specific physiological uptake of Sestamibi within intracranial structures, this step proved to be much more difficult than originally assumed. Indeed, despite the high number of registration algorithms available for multimodal brain data (either commercial or free software), no individual method manages to correctly register a high number of datasets.

A registration algorithm consists in iteratively determining the optimal settings of a transformation model in order to optimize a similarity criterion, allowing the superimposition of a source image onto a reference image. The optimal choice and parameterization of each component of a registration algorithm directly depends on the characteristics of the application: the matching type (inter- or intra-patient data), the body parts investigated, the imaging modalities implied. This upper level optimization is complex to determine. For multimodal registration problems, many iconic similarity criteria have been proposed, such as the Correlation Ratio (CR) (Roche *et al* 1998), the Normalized Cross Correlation (NCC) (Sarvaiya *et al* 2009), the Entropy Correlation Coefficient (ECC) (Maes *et*

*al* 1997), the Mutual Information (MI) (Collignon *et al* 1995, Viola and Iii 1997) , the Normalized Mutual Information (NMI) (Studholme *et al* 1998) and Woods criterion (W) (Woods *et al* 1993). Although less used, some feature-based similarity criteria such as the Chamfer distance (CD) (Mangin *et al* 1994) have proved to be useful.

To help users in their choices, a logical solution is to compare different registration algorithms in a specific context. This represents an active research field that requires procedures independent of the criteria used for registration (Murphy *et al* 2011a). A first approach is to test different algorithms on simulated data or phantom studies (Skerl *et al* 2006, 2008), for which a ground truth can be easily established. However, the necessary simplifications introduced by simulations or phantom objects make the generalization to clinical data sometimes hazardous. Some evaluation studies were also conducted on public clinical databases (Murphy *et al* 2011b, Klein *et al* 2010), (Sarkar *et al* 2005). To do the comparison, the segmentation of some structures of interest or the manual selection of some features are generally considered and conventional overlap criteria such as the Dice index (Dice 1945), the percentage of false positive or false negative voxels and some other evaluation criteria dedicated to specific applications (Murphy *et al* 2011a), are used. However, a fully automated approach has not yet been described for clinical data, and the visual assessment remains an unavoidable task. Furthermore, even the globally best method may fail to register some specific datasets, for which other methods may potentially succeed. Thus the concept of "global best method" is not always the most fruitful.


To take into account this potential complementarity of different registration algorithms, we recently proposed a new strategy that consists in selecting the best registration method among several for each dataset (Tacchella *et al* 2013). To achieve this selection a ranking of the different registration results is performed thanks to a dedicated quantitative evaluation criterion, using specific *a priori* knowledge. To deal with our specific clinical application, this criterion was related to the physiological uptake of Sestamibi. Compared to our previous work, a more efficient and more robust segmentation method was introduced to compute the criterion. Furthermore the panel of the registration methods was largely modified: 4 inefficient methods were removed, 13 additional methods were tested. With these modifications, the strategy applied to sixty-two datasets proves to be efficient and suitable for clinical application.

## 2. Methods

### 2.1.Database

The database was acquired in the framework of a follow-up protocol in patients presenting recurrent high grade gliomas and undergoing antiangiogenic treatment (Yeni *et al* 2014). All patients gave a written informed consent. The database consists in sixty-two datasets. The test protocol included two imaging exams performed within 3 days, providing a multimodal dataset (Figure 1) consisting in:

- one T1-weighted 3D MR volume of the head acquired after the injection of Gadolinium-DOTA with a 3.0 T MR scanner (SignaHDxt, GE), using TE = 3.568 ms and TR = 8.928 ms. The MR volumes contain 248 axial slices with a thickness of 0.69 mm and a pixel size of 0.48x0.48 mm².

- one Sestamibi-SPECT 3D volume of the head acquired 15 minutes or 3 hours after the injection of 740 ± 111MBq of Technetium-99m-Sestamibi with a three-headed gamma camera equipped with parallel high resolution collimators (Irix, Philips). Transmission data were not acquired for this protocol. SPECT images were post-filtered with a low-pass filter of order 4 with a cut-off frequency of 0.5 cm$^{-1}$, and then corrected for attenuation using Chang's method (Chang 1978) with $\mu = 0.12$ cm$^{-1}$. For the attenuation correction, an ellipsoid model of the head was considered and manually adjusted slice by slice by an expert when needed. The SPECT volumes are composed of 66 axial slices of isotropic voxels (2.3 mm in each direction).
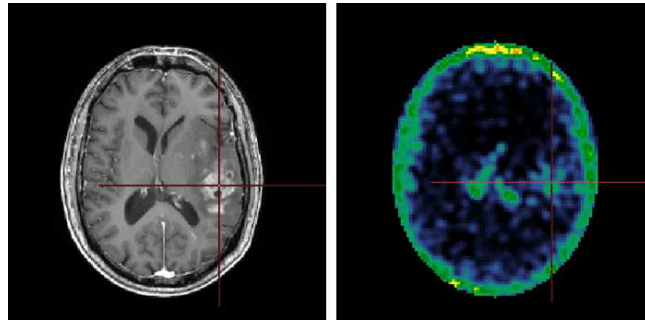


Figure 1 Dataset sample. Left: axial slice of the 3D T1-MR volume (gray color map) showing a contrast enhancement in the tumor. Right: axial slice of the SPECT volume (French colormap) showing high Sestamibi uptake in the tumor and in the choroid plexus.

### 2.2. Major issues of 3D multimodal registration

Co-registration of MRI and functional Sestamibi-SPECT data is a critical issue for the accuracy of the patient follow-up and for the comparative study of these two modalities. Manual 3D registration of multimodal images of the brain remains a very difficult and time-consuming task, even for experts, and is subject to inter- and intra-operator variability. Therefore the use of automatic methods is advised, even if the quality of the registration is always checked visually. For example, the Mutual

Information registration approach implemented in the SPM software was first used by clinicians, who reported a high number of registration failures using this method. Figure 2 shows an example of registration failure: on this specific slice, an unexpected high uptake of the Sestamibi can be observed in the left eye as well as a shift of the uptake at the level of the extraocular muscle. These kinds of failure motivated our new registration strategy.



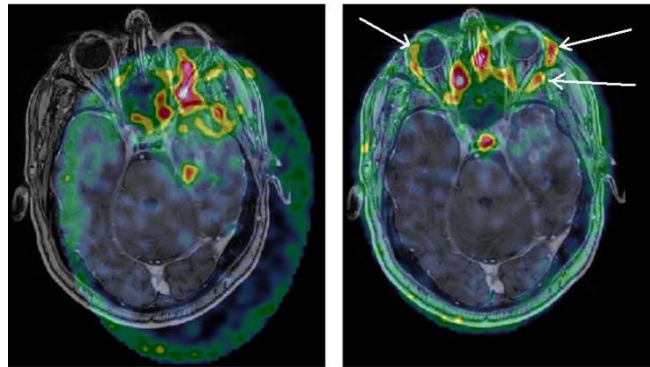Figure 2 Example of a poor SPECT / MRI registration using a mutual information based method. Left: before registration. Right: after registration. The arrows show an unexpected high uptake of the Sestamibi in the left eye and a shift of the uptake outside the extraocular muscle on the registered dataset.
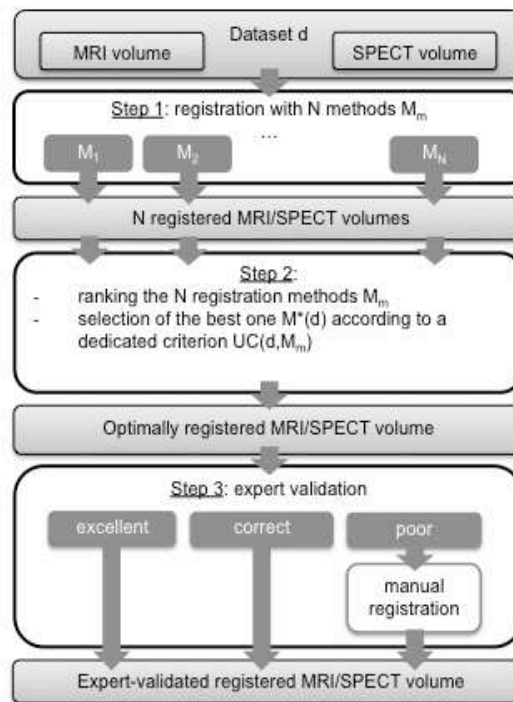
## 2.3. Individualized Registration Strategy



Figure 3 Flowchart of the registration strategy

Our strategy was based on the following observation: for a given dataset, a method based on Mutual Information could perform well, whereas for another dataset, the same method could fail, while a method based on Chamfer distance could provide satisfying registration results. To take into account these practical results in a general and coherent framework, we propose a registration strategy (Figure 3) that consists in running several complementary registration methods, evaluating these methods and selecting the best one for each individual dataset. Thus, the proposed method takes advantage of the possible complementarity of different registration algorithms. The process was fully integrated within the BrainVisa/Anatomist software environment (Rivière *et al* 2009).

### 2.3.1. Registration

This step consists in applying N registration methods $M_m$ ($1 \leq m \leq N$). In the present application, the registration is limited to intra-patient MRI/SPECT co-registration of brain. It is commonly accepted that rigid registration methods are well suited to intra-patient datasets on non-deformable regions (Lee *et al* 2007). Therefore, a rigid transformation model was selected, once the appropriate transformations for taking into account for different voxel sizes were taken into account. Standard registration methods used in neuroimaging and available as ready-to-use functions in the software environments BrainVisa 4.2.0 (Rivière *et al* 2009), SPM 8 (Ashburner 2012), and AIR 5.3.0 (Woods *et al* 1993) were all integrated within BrainVisa/Anatomist in order to be applied in a transparent way. The use of various parameter tunings led to a panel of eighteen fully automated registration methods, which are summarized in Table 1.

Registration algorithms implemented in SPM use conventional iconic criteria: Mutual Information (MI), Normalized Mutual Information (NMI), Entropy Correlation Coefficient (ECC), and Normalized Cross Correlation (NCC). The different algorithms using these criteria led to methods M9 to M16. The Mutual Information (MI) method is also available in BrainVisa with a slightly different implementation; it was used to derive the methods M1 and M2. The Correlation Ratio (CR) iconic criterion available in BrainVisa led to methods M3 and M4. Finally, the Chamfer Distance (CD) geometric criterion implemented in BrainVisa was tested in four different configurations leading to methods M5 to M8. As Chamfer Distance methods consist in minimizing the distance between corresponding contours in both volumes, the skull edges were chosen as features and were extracted from the MRI and the SPECT data. For the MR images, a thresholding (T) step was applied to remove the image background; the head was filled using smoothing morphological operations (MO), and a morphological gradient provided the skull edge. Regarding the SPECT images, the skull edge was extracted after application of Deriche filter (Deriche 1987) before (D-A) and after attenuation correction (D-AC). A threshold was also applied to SPECT images in order to

remove the low uptake areas (TKM-AC; M2 and M3), the threshold being automatically defined by a two-class k-means procedure. All MR images were subsampled at least with a factor of 8 (i.e. a factor of 2 in each direction) in order to reduce the computation time (SUB_MRI/8). For CD methods, in order to make the head edge extraction faster, the MRI was subsampled to the resolution of the SPECT (SUB_SPECT). Finally the two last methods M17 and M18 were derived from the Woods criterion (W)-based registration algorithm implemented in the AIR software.

| | | Similarity Measure | Preprocessing Steps | | Transformation direction |
| | | | MRI | SPECT | |
|---|---|---|---|---|---|
| Brainvisa4. | M1 | MI | SUB_MRI/8 | AC | I |
| 2.0 | M2 | MI | SUB_MRI/8 | TKM-AC | I |
| | M3 | CR | SUB_MRI/8 | TKM-AC | I |
| | M4 | CR | SUB_MRI/8 | AC | I |
| | M5 | CD | SUB_SPECT-T-MO | D-A | D |
| | M6 | CD | SUB_SPECT-T-MO | D-A | I |
| | M7 | CD | SUB_SPECT-T-MO | D-AC | D |
| | M8 | CD | SUB_SPECT-T-MO | D-AC | I |
| SPM8 | M9 | MI | SUB_MRI/8 | AC | D |
| | M10 | MI | SUB_MRI/8 | AC | I |
| | M11 | NMI | SUB_MRI/8 | AC | D |
| | M12 | NMI | SUB_MRI/8 | AC | I |
| | M13 | ECC | SUB_MRI/8 | AC | D |
| | M14 | ECC | SUB_MRI/8 | AC | I |
| | M15 | NCC | SUB_MRI/8 | AC | D |
| | M16 | NCC | SUB_MRI/8 | AC | I |
| AIR 3.5.0 | M17 | W | SUB_MRI/8 | AC | D |
| | M18 | W | SUB_MRI/8 | AC | I |

Table 1 Description of the eighteen registration methods

The transformation matrix for aligning an image volume A onto an image volume B can be estimated in two different ways: directly (D), by registering the volume A onto the volume B, or indirectly (I), by registering the volume B onto the volume A and then inversing the transformation matrix. In practice both ways can lead to quite different registration results; thus both ways were considered in the panel of methods. The registration of the SPECT onto the MRI is considered as direct (D).

All these methods are fully automated and this first step does not require any manual interaction.

### 2.3.2. Individualized optimal method selection

The second step of the strategy consists in ranking the N methods to select the best one for the dataset d, according to an evaluation criterion UC, as first introduced in (Tacchella *et al* 2013). The method $M_m$ providing the higher UC value is considered the best one for this dataset and noted is $M^*(d)$:

$$M^*(d) = argmax_{M_m}\big(UC(d, M_m)\big) \tag{1}$$

The quantitative criterion UC is derived from the physiological and anatomical criteria that are used by experts for evaluating registered data and relies on structures that are outside the tumor area but are identifiable in both images: the pituitary gland and the extraocular muscles showing a high uptake of the Sestamibi on the SPECT images and the eyeballs presenting a low uptake of the Sestamibi on the SPECT images. The UC criterion is based on the intensity of the SPECT images within these specific structures segmented on the MR images.

The eyeballs are segmented automatically using a spherical Hough Transform within a box manually delineated around each eye, and the pituitary gland is coarsely segmented by defining a cylinder including the area, as described in (Tacchella *et al* 2013).

Furthermore, for the present study, the rectus extraocular muscles are segmented using a 3D iterative forest watershed algorithm (Felkel *et al* 2002) that requires for initialization the manual input of one seed for each muscle. This was proposed in order to avoid the labor-intensive task of manual segmentation of the extra ocular muscles that was previously presented (Tacchella *et al* 2013).
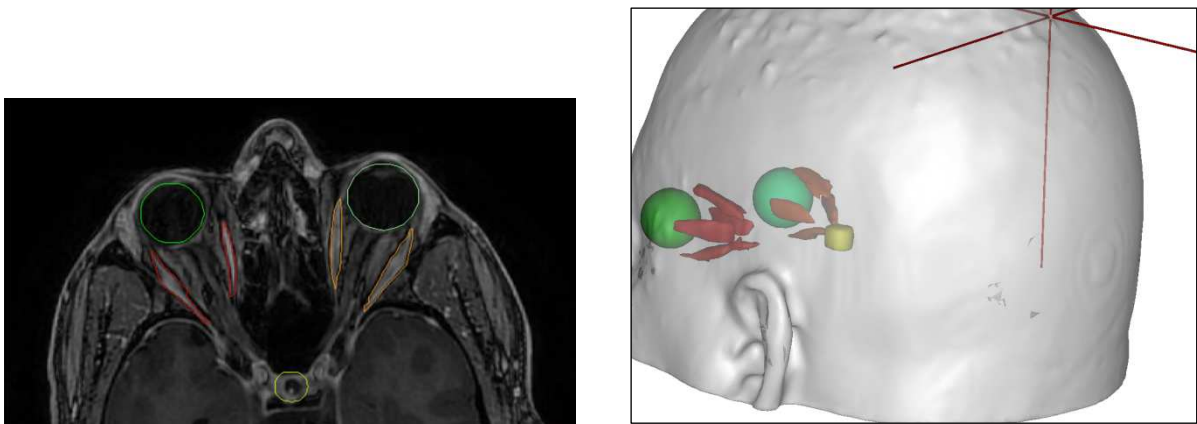


Figure 4 Segmentation of the eyeballs (green), the pituitary gland (yellow) and the extraocular muscles (red and orange) visualized on an MR axial slice (left) and on a 3D view (right).

The segmented areas (shown in Figure 4) are gathered into two regions $V_{hu}(d)$ (for high uptake volumes) and $V_{lu}(d)$ (for low uptake volumes). Let $I(d, M, p)$ denote the SPECT intensity of dataset d at voxel p after registration with method M, the expression of the criterion UC is given by Equation 2:

$$UC(d, M) = \frac{1}{Card(V_{hu}(d))} \sum_{p \in V_{hu}(d)} I(d, M, p) - \frac{1}{Card(V_{lu}(d))} \sum_{p \in V_{lu}(d)} I(d, M, p) \qquad (2)$$

For a given dataset, the computation of UC for the N registration methods allows us to compare the performances of the methods and to choose the best one according to Equation 1. This step was not fully automatic but could be achieved by a trained operator.

### 2.3.3. Expert validation

Although $M^*(d)$ is the best method for the dataset d, the registration provided by this method is not necessarily satisfying. In some cases, the registration can fail with all methods, even for the "best" one. Therefore, an expert was asked to provide a quality score for each registration among "excellent", "correct" and "poor".

## 3. Validation and results

The registration strategy was applied to the database. In order to evaluate and validate the complete process, it was necessary to first assess the robustness of UC toward the segmentation of anatomical structures, and then to estimate the improvement provided by the strategy.

### 3.1. Validation process

A preliminary study has shown the relevance of the UC criterion for ranking the different registration methods, using expert manual registration as a reference (Tacchella *et al* 2013). In that preliminary study, manual segmentation of extraoculor muscles was achieved by two operators $op_{m1}$ and $op_{m2}$. In the present study they were compared to the watershed segmentation that was provided by two operators $op_{w1}$ and $op_{w2}$. To compare the four segmentation results, the Dice index (Dice 1945), which measures the relative overlap of a region with respect to a reference region, was computed. Furthermore, to assess the influence of the segmentation on the ranking and on the choice of the best registration, the four following indicators were used:

I1: the number of times that the best method $M^*_{ref}(d)$ selected using the segmentation of one reference operator ($op_{ref}$) is different from the best one $M^*_{test}(d)$ using the segmentation of another operator ($op_{test}$). This indicator is called "$M^*$ inversions".

I2: the Target Registration Evaluation criterion (TRE) (Hill *et al* 2001, Aubert-Broche *et al* 2006) which measures a distance between two transformation matrices. It was used to compare the transformation matrices obtained by $M^*_{ref}(d)$ and $M^*_{test}(d)$. A TRE distance smaller than or equal to the spatial resolution of the SPECT image was considered to indicate equivalent registration methods.

I3: the relative variation of UC ($\Delta$UC) computed using the segmentation of the reference operator ($op_{ref}$) between $M^*_{ref}(d)$ and $M^*_{test}(d)$:

$$\Delta UC = \frac{\left| UC(d, M^*_{test}(d)) - UC(d, M^*_{ref}(d)) \right|}{UC(d, M^*_{ref}(d))} \tag{3}$$

The variation was considered significant if the relative variation $\Delta$UC was greater than 2%.

I4: the quality of registration results evaluated by the expert, using the qualitative score assigned to the registration results corresponding to $M^*(d)$.

Indicators I2 and I3 were only computed when "$M^*$ inversions" occurred, and the expert evaluation was required if either TRE > 2.3 mm or $\Delta UC$ > 2 %. In such cases, $M^*_{ref}(d)$ and $M^*_{test}(d)$ are considered non-equivalent (NEq) if their visual score attributed by experts are different.

*3.2. Evaluation of the robustness of UC*

Using our preliminary study (Tacchella *et al* 2013), the segmentation and the subsequent ranking obtained by the manual segmentation provided by the first operator ($op_{m1}$) were chosen as the reference. The mean Dice indices (± standard deviation) computed for the eight rectus extraocular muscles between $op_{m1}$ and $op_{m2}$, $op_{w1}$, and $op_{w2}$ were respectively 0.64±0.15, 0.61±0.15, and 0.62±0.13. Furthermore, the mean Dice index between the segmentations provided by $op_{w1}$ and $op_{w2}$ was 0.89±0.13. Table 2 shows the results of the comparisons using the four indicators. The rankings between $op_{w1}$ and $op_{w2}$ or between $op_{m1}$ and $op_{m2}$ were always equivalent.

| | $op_{m1}$ / $op_{m2}$ | $op_{m1}$ / $op_{w1}$ | $op_{m1}$ / $op_{w2}$ | $op_{w1}$ / $op_{w2}$ |
|---|---|---|---|---|
| I1: $M^*$ inversions | 14 | 11 | 10 | 1 |
| I2: TRE > 2.3 mm | 5 | 3 | 3 | 0 |
| I3: $\Delta UC$ > 2% | 3 | 2 | 2 | 0 |
| I4: NEq | 0 | 0 | 0 | 0 |

Table 2: Influence of the segmentation of extraoculor muscles by different operators on the computation of UC and ranking

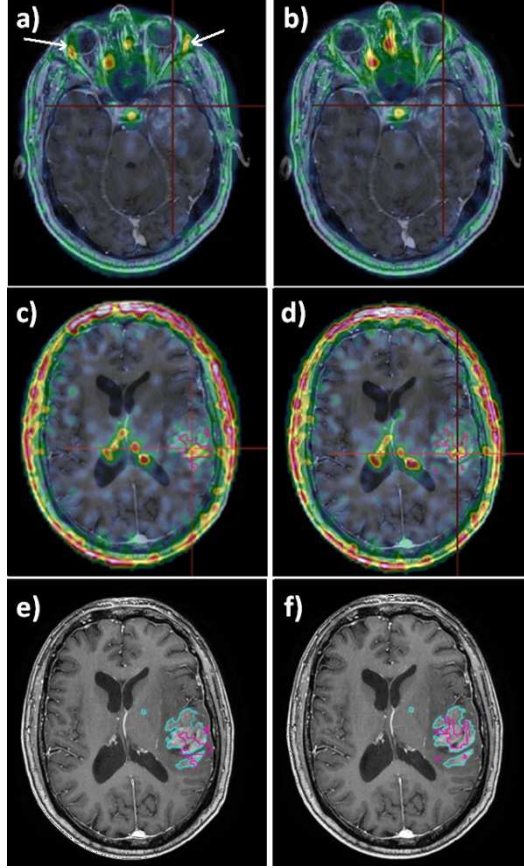### 3.3. Contribution of the individualized registration strategy



Figure 5 Example of a poor registration obtained by method M1 (5a, 5c, 5e) and an excellent registration obtained by M5 (5b, 5d, 5f). Fig 5a and 5b represent axial views showing eyes and extraocular muscles: an unexpected high uptake of the Sestamibi in the left eye and a shift of the uptake outside the extraocular muscles (arrows) are noticeable in case of poor registration (5a), conversely a high uptake of Sestamibi in the extraocular muscles and a low uptake in the eyeballs are shown in case of excellent registration (5b). Fig 5c and 5d represent axial views within the tumor showing the Sestamibi uptake (delineated in pink) and the Gd contrast-enhancement. The boundaries of the Sestamibi uptake (in pink) and of the Gd contrast-enhancement (in cyan) are reported on the MR images and show a different co-localization according to the registration methods (5e, 5f).

For each dataset d of the database, the best method $M^*(d)$ was selected after the ranking step, using the uptake criterion UC derived from the segmentation provided by the operator $op_{w1}$. Figure 5 illustrates for one sample dataset an excellent registration (Fig 5b) obtained by $M^*(d)$ ($M^*(d)$=M5 in this case) and a poor registration (Fig 5a) obtained by method M1. Figure 5e) and 5f) represent the superimposition of the boundaries of the Gd contrast enhancement in the tumor area manually segmented on the MRI and the boundaries of the uptake of Sestamibi corresponding to the tumor area segmented on the SPECT image. The Sestamibi uptake area was segmented by keeping all

voxels included in a manually defined region of interest with an intensity higher than 40% of the maximum intensity in this region (Yeni *et al* 2014).

Figure 6 represents the frequency (in %) with which each method $M_m$ is considered as the best one ($M^*(d)$) and the fraction of excellent, correct, and poor registrations obtained when visually checking the results provided by $M^*(d)$.
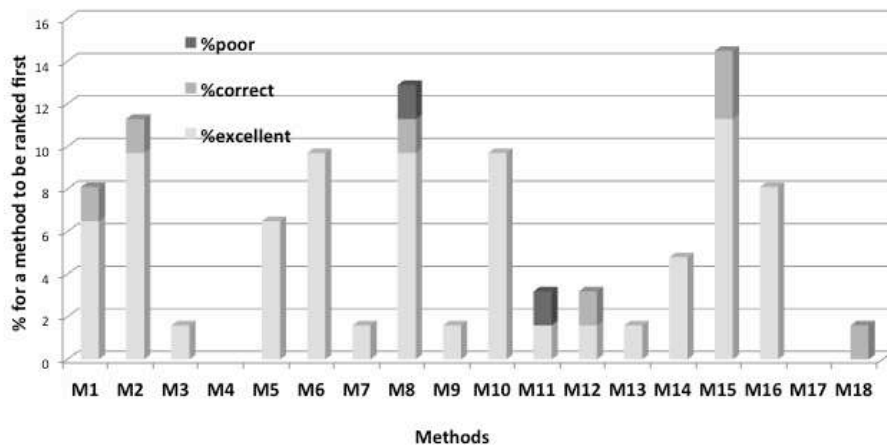


Figure 6 Frequency for each of the eighteen methods to be ranked first and associated quality of registration

Finally, the performance of our registration strategy was compared to the performance of the four methods providing the best results on the database (Figure 7). To identify these methods which, used individually, obtained the best results on a database, we computed the sum of ranks obtained by each method for all datasets and selected the methods that obtained the lowest sums (Murphy *et al* 2011b). This approach led us to identify the four best methods: M12, M14, M15 and M16. On the sixty-two datasets, methods M12, M14,M15 and M16 obtained respectively 42%, 50%, 61%, and 58% excellent registration results, 40%, 32%, 24%, and 29% correct registrations, 18%, 18%, 15%, and 13% poor registrations. On the same database, our strategy provided 85% excellent registrations, 12% correct registrations, and 3% poor ones.
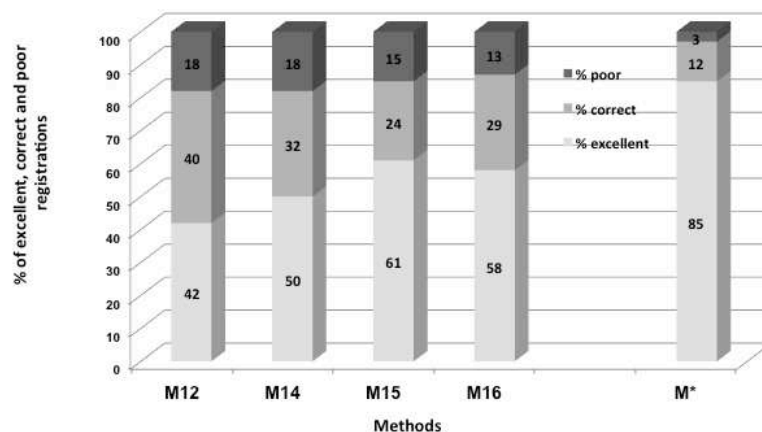


Figure 7 Quality of registration obtained by our strategy versus the four best method

## 4. Discussion

Our goal was to improve the registration results in a multimodal MRI/SPECT registration framework, where the specific distribution of the radiotracer (Sestamibi) used in SPECT made this task very challenging. A smart registration strategy based on the individualized selection, for each dataset, of the most suitable of eighteen registration methods, according to a dedicated uptake criterion was proposed.

### 4.1. Evaluation criterion UC

The evaluation criterion UC is one foundation of the strategy, since it allows us to rank the different registration methods and select the best one for each dataset. Its relevance was first demonstrated in (Tacchella *et al* 2013): the ranking obtained with this criterion was consistent with the registrations performed by an expert. Furthermore, its computation requires the segmentation of the eyeballs, the pituitary gland, and the extraocular muscles. The segmentation of the pituitary gland and the eyeballs is fast, easy and can be considered as operator-independent, since the interactive step is reduced to the delineation of a region of interest (box or cylinder) surrounding these structures. Extraocular muscles are more difficult to delineate. However we demonstrated that their delineation was essential for an appropriate selection of the best registration method. Indeed, the quality of registration obtained when using or not using extraocular muscles in UC computation differs in 30% of the cases, and worsened when not using them. To make the segmentation of extraocular muscles easier, the watershed algorithm with manual input of the seeds was proposed. Quantitative indicators were computed to evaluate the quality and the robustness of the criterion UC using watershed segmentation. The high Dice scores between the watershed segmentations performed by two operators indicate that the segmentation is reproducible. Moreover, the best method differs only in one case for which the methods were judged equivalent (Table 2): this confirms that the watershed segmentation used for the computation of UC is robust and reproducible. Comparing the rankings obtained using the manual segmentations performed by two operators shows that using manual segmentation for the ranking is less reproducible than using watershed segmentation (Table 2). However, the ranking remains relevant for all datasets. The lower Dice scores between the watershed segmentations and the manual segmentation of $op_{m1}$ suggest that the difference between watershed segmentations is smaller than the inter-operator variability. Furthermore, when comparing ranking, there are few cases of "$M^*$ inversions" and the quality of registration obtained was found similar in all cases. Thus, watershed segmentation for extraocular muscles leads to a robust, relevant and reproducible evaluation criterion UC and allows the processing time to be reduced (at least by a

factor of 4) when compared to the manual segmentation. Although this step is not fully automatic, it does not require an expert knowledge and can be achieved by a trained operator.

*4.2.Contribution of the individualized registration strategy*

As the performances of the registration methods vary from one dataset to another, it is relevant to select the most suitable method for each dataset. Indeed, Figure 5a represents a poor registration result obtained by method M1 and Figure 5b an excellent registration result obtained by method M5 which was the best one for this dataset. The features used to visually judge the registration are the following: an abnormal high uptake of Sestamibi is present inside the eyeballs and outside the extraocular muscles in Figure 5a, whereas the uptake of Sestamibi is high in the extraocular muscles and pituitary gland and low in the eyeballs in Figure 5b. The boundaries of the uptake of Sestamibi in the tumor differ according to the registration (Fig 5e and 5f). Thus, a small registration error may have a large impact on the local estimation of the tumor using both modalities. This result also underlines the necessity to choose an evaluation criterion for registration that is independent of the pathology. Figure 6 proves that many methods contribute to the strategy on the database; no method comes first in a majority of cases. However, this result can be tempered since there may be equivalent methods. Compared with our preliminary study (Tacchella *et al* 2013), additional methods issued from SPM environment clearly improved the results of our strategy, while the impact of AIR methods is less noticeable. Furthermore, to fully demonstrate the interest of our strategy, we compared M* to the four methods M12, M14, M15 and M16, which individually obtained the best results on the database. According to Figure 7, these methods provide at best 61% excellent registrations, and 13% poor ones. Thus, our strategy provides much more excellent registrations (85%) and much less poor registrations (3%). Its interest on the global quality of registration results on the database is thus clearly demonstrated.

*4.3.Future work*

Despite the good performance of the proposed strategy, two datasets were poorly registered. To deal with those cases, manual registration is a practical possibility, even if it is time-consuming and not an easy task. Some additional registration methods, including hybrid methods (Bullich *et al* 2009, Lee *et al* 2007) coupling iconic and feature-based information efficiently for other SPECT/MRI studies, could be tested. Of course, adding new methods to the strategy could improve the registration results on the database, as can be noticed when comparing the present results to those presented in (Tacchella *et al* 2013); however it increases the total computation time. This running time depends on the type

and the implementation of the registration methods. In the current implementation of the strategy, one MRI/SPECT dataset is registered in a few seconds by a feature-based method (M5-M8), whereas the registration with an iconic method takes about one minute. Thus the strategy requires about 15 minutes per dataset, in a sequential execution. However, parallel implementation could reduce that time. Furthermore, the combination of a minimal number of complementary methods in the strategy could also be defined. On the present database, it appears that the methods M4 and M17 are never ranked first (Figure 6). Thus, although they may have an interest for additional cases, they could be removed from the strategy. Of course, finding the ideal combination of methods is a complex problem. To achieve this goal, it is essential to go further into the definition of equivalent registration methods, using TRE and $\Delta$UC. Finally, this strategy could be applied to other registration problems. Of course, the methods to use and the quantitative evaluation criterion would have to be adapted according to the new context. For instance, studies in (Murphy *et al* 2011b, Klein *et al* 2010, Song *et al* 2010) although dealing with different data and different issues, could benefit from the results of the present study, using their own registration evaluation criteria. Furthermore, abdominal perfusion imaging studies using dynamic X-ray Computed Tomography in (Cohen *et al* 2013) is another potential application, where the evaluation criterion is based on the mean Dice index of bony structures (ribs and vertebrae). In this latter study, a non-rigid registration method has proven to perform better than a rigid one, but registration remains an open issue. Therefore, several complementary elastic approaches could be tested using the proposed strategy.

## 5. Conclusion

A new registration strategy was proposed that consists in using several registration methods (fully automated step) and selecting the best one for a given dataset according to a semi-automated quantitative criterion. This strategy was applied to a multimodal MRI/SPECT registration problem in patients with high grade gliomas undergoing antiangiogenic treatment. A context-specific evaluation criterion was defined that relies on the physiological uptake of the Sestamibi radiotracer, and not on the pathological one. The panel of registration methods included eighteen conventional iconic and feature-based algorithms available in the SPM, BrainVisa/Anatomist, and AIR free software environments. The practical efficiency of this approach was proven on sixty-two datasets: as a matter of fact the strategy provides far more registrations with excellent (85%) and correct (12%) quality than any other method used individually. Provided that adapted evaluation criteria and registration methods are chosen, this strategy could be applied to many registration studies, which remain open issues.

## 6. Acknowledgments

# 7. References

Ashburner J 2012 SPM: A history *NeuroImage* **62** 791–800

Aubert-Broche B, Grova C, Reilhac A, Evans A C and Collins D L 2006 Realistic simulated MRI and SPECT databases. Application to SPECT/MRI registration evaluation. *Med. Image Comput. Comput. Assist. Interv.* **9** 330–7

Bleichner-Perez S, Le Jeune F, Dubois F and Steinling M 2007 99mTc-MIBI brain SPECT as an indicator of the chemotherapy response of recurrent, primary brain tumors. *Nucl. Med. Commun.* **28** 888–94

Brown L G 1992 A survey of image registration techniques *ACM Computing Surveys* **24** 325–76

Bullich S, Ros D, Pavía J, Cot A, López N and Catafau A M 2009 Neurotransmission SPECT and MR registration combining mutual and gradient information. *Med. Phys.* **36** 4903–10

Chang L-T 1978 A Method for Attenuation Correction in Radionuclide Computed Tomography *IEEE Trans. Nucl. Sci.* **25** 638–43

Cohen M-E, Pellot-Barakat C, Tacchella J-M, Lefort M, De Cesare A, Lebenberg J, Souedet N, Lucidarme O, Delzescaux T and Frouin F 2013 Quantitative evaluation of rigid and elastic registrations for abdominal perfusion imaging with X-ray computed tomography *IRBM* **34** 283–6

Collignon A, Maes F, Delaere D, Vandermeulen D, Suetens P and Marchal G 1995 Automated multi-modality image registration based on information theory *Inf. Process. Med. Imag.* **3** 263 –274

Crum W R 2004 Non-rigid image registration: theory and practice *Br. J. Radiol.* **77** S140–S153

Deriche R 1987 Using Canny's criteria to derive a recursively implemented optimal edge detector *Int. J. Comput. Vis.* **187** 167–87

Dice L 1945 Measures of the Amount of Ecologic Association Between Species *Ecology* **26** 297 –302

Felkel P, Bruckwschwaiger M and Wegenkittl R 2002 Implementation and complexity of the watershed-from-markers algorithm computed as a minimal cost forest *Comput. Graph. Forum* **20** 26–35

Gholipour A, Kehtarnavaz N, Briggs R, Devous M and Gopinath K 2007 Brain functional localization: a survey of image registration techniques. *IEEE Trans. Med. Imag.* **26** 427–51

Hellier P, Barillot C, Corouge I, Gibaud B, Le Goualher G, Collins D L, Evans A, Malandain G, Ayache N, Christensen G E and Johnson H J 2003 Retrospective evaluation of intersubject brain registration. *IEEE Trans. Med. Imag.* **22** 1120–30

Hill D L, Batchelor P G, Holden M and Hawkes D J 2001 Medical image registration. *Phys. Med. Biol.* **46** R1–45

Holden M 2008 A review of geometric transformations for nonrigid body registration. *IEEE Trans. Med. Imag.* **27** 111–28

Jannin P, Grova C and Gibaud B 2001 Fusion de données en imagerie médicale: revue méthodologique basée sur le contexte clinique *ITBM-RBM* **22** 196–215

Klein A, Ghosh S S, Avants B, Yeo B T T, Fischl B, Ardekani B, Gee J C, Mann J J and Parsey R V 2010 Evaluation of volume-based and surface-based brain image registration methods. *NeuroImage* **51** 214–20

Lee J-D, Huang C-H, Weng Y-H, Lin K-J and Chen C-T 2007 An automatic MRI/SPECT registration algorithm using image intensity and anatomical feature as matching characters: application on the evaluation of Parkinson's disease. *Nucl. Med. Biol.* **34** 447–57

Lester H and Arridge S 1999 A survey of hierarchical non-linear medical image registration *Pattern Recognit.* **32** 129–49

Maes F, Collignon A, Vandermeulen D, Marchal G and Suetens P 1997 Multimodality image registration by maximization of mutual information. *IEEE Trans. Med. Imag.* **16** 187–98

Maintz J B A and Viergever M A 1998 A survey of medical image registration *Med. Imag. Anal.* **2** 1–36

Mangin J, Frouin V, Bloch I, Bendriem B and Lopez-Krahe J 1994 Fast nonsupervised 3D registration of PET and MR images of the brain *J. Cereb. Blood Flow Metab.* **14** 749–62

Murphy K, Van Ginneken B, Klein S, Staring M, De Hoop B J, Viergever M a and Pluim J P W 2011a Semi-automatic construction of reference standards for evaluation of image registration. *Med. Imag. 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 and Pluim J P W 2011b Evaluation of registration methods on thoracic CT: the EMPIRE10 challenge. *IEEE Trans. Med. Imag.* **30** 1901–20

Oliveira F P M and Tavares J M R S 2012 Medical image registration: a review. *Comput. Methods Biomech. Biomed. Engin.* **17** 1–21

Prigent-Le Jeune F, Dubois F, Perez S, Blond S and Steinling M 2004 Technetium-99m sestamibi brain SPECT in the follow-up of glioma for evaluation of response to chemotherapy: first results. *Eur. J. Nucl. Med. Mol. Imag.* **31** 714–9

Rivière D, Geffroy D, Denghien I, Souedet N and Cointepas Y 2009 BrainVISA: an extensible software environment for sharing multimodal neuroimaging data and processing tools *NeuroImage* **47** S163

Roche A, Malandain G, Pennec X and Ayache N 1998 The correlation ratio as a new similarity measure for multimodal image registration *Med. Image Comput. Comput. Assist. Interv.* **1496** 1115–24

Sarkar A, Santiago R J, Smith R and Kassaee A 2005 Comparison of manual vs. automated multimodality (CT-MRI) image registration for brain tumors. *Medical dosimetry : official journal of the American Association of Medical Dosimetrists* **30** 20–4

Sarvaiya J N, Patnaik S and Bombaywala S 2009 Image Registration by Template Matching Using Normalized Cross-Correlation *2009 International Conference on Advances in Computing, Control, and Telecommunication Technologies* (IEEE) pp 819–22

Skerl D, Likar B and Pernus F 2008 A protocol for evaluation of similarity measures for non-rigid registration. *Medical image analysis* **12** 42–54

Skerl D, Likar B and Pernus F 2006 A protocol for evaluation of similarity measures for rigid registration. *IEEE transactions on medical imaging* **25** 42–54

Soler C, Beauchesne P, Maatougui K, Schmitt T, Barral F G, Michel D, Dubois F and Brunon J 1998 Technetium-99m sestamibi brain single-photon emission tomography for detection of recurrent gliomas after radiation therapy *Eur. J. Nucl. Med.* **25** 1649–57

Song J H, Christensen G E, Hawley J A, Wei Y and Kuhl J G 2010 Evaluating Image Registration Using NIREP *Bio. Imag. Reg.* **6204** 140–50

Studholme C, Hawkes D J and Hill D L 1998 Normalized entropy measure for multimodality image alignment *Medical Imaging '98* ed K M Hanson (International Society for Optics and Photonics) pp 132–43

Tacchella J-M, Roullot E, Lefort M, Cohen M-E, Guillevin R, Petrirena G, Delattre J-Y, Habert M-O, Yeni N, Kas A and Frouin F 2013 A new strategy to improve coregistration of spect and MR images in patients with high grade glioma. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2013** 4002–5

Viola P and Iii W M W 1997 Alignment by Maximization of Mutual Information *Int. J. Comput. Vis.* **24** 137–54

Woods R P, Mazziotta J C and Cherry S R 1993 MRI-PET registration with automated algorithm. *J. Comput. Assist. Tomogr.* **17** 536–46

Yeni N, Petrirena G, Tacchella J-M, Guillevin R, Martini J-B, Alain G, Habert M-O, Delattre J-Y and Kas A 2014 Early assessment of bevacizumab response using 99mTc-sestamibi single-photon emission tomography in recurrent high grade glioma. *J. Nucl. Med. , submitted*

Zitová B and Flusser J 2003 Image registration methods: a survey *Image Vis. Comput.* **21** 977–1000