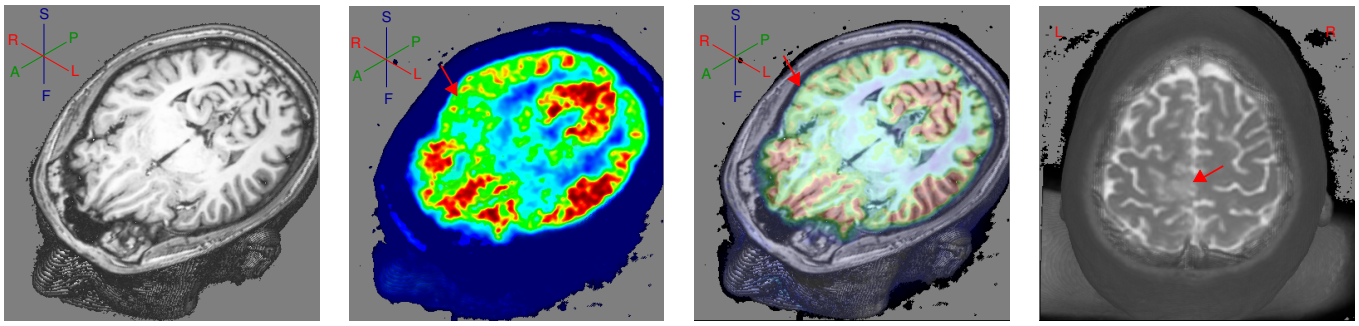
Pre-Alignment for Co-registration in Native Space

Wu, Shin-Ting, Augusto C. Valente, Lionis S. Watanabe
Dept. of Computer Engineering and Industrial Automation
University of Campinas
São Paulo, Brazil
www.dca.fee.unicamp.br/∼{ting,avalente,watanabe}

Clarissa L. Yasuda, Ana C. Coan, Fernando Cendes
Dept. of Neurology
University of Campinas
São Paulo
Email: yasuda.clarissa@gmail.com



(a) (b) (c) (d)

Fig. 1. Co-registration in native space: (a) non-lesional T1-weighted MR volume; (b) PET volume colored with rainbow palette: red and blue correspond, respectively, hyper and hypo-metablism; (c) Overlay of (a) and (b) provides accurate localization of hypometabolism on the anterior and middle part of the right temporal lobe; (d) Co-registration of T1- and T2-weighted MRI images make subtle signs of focal cortical dysplasia more visible.

***Abstract*—For nonlesional patients, the correct localization of the epileptogenic foci in native space remains a great challenge. Non-invasive functional PET images that provide information about cerebral activities may reveal the origin of seizure activity, but without precise anatomical detail. Co-registration of the functional images with MR images on the basis of maximization of mutual information (MMI) has shown to be very promising in improving presurgical evaluation. Nevertheless, a mutual information (MI) function is non-convex and the convergence of an algorithm to its optimum is guaranteed only if the initial estimate lies in its convex vicinity. We present in this paper a generally applicable method that pre-aligns the DICOM images such that their relative position becomes close to an optimum. The key to our solution is a robust user-guided interactive procedure to extract valid voxels, for both the centroid estimation and the registration. Aiming at comparative analysis, we introduce a numerical condition to quantify registration errors. The results are acceptable when we consider the intrinsic problems of the MMI-based registration algorithm we implemented.**

***Keywords*-Pre-alignment;co-registration; multimodal visualization.**

I. Introduction

The outcome of surgical treatment for refractory epilepsy depends on the precise localization of the epileptogenic focus in native space [1], [2]. It is important not only for the resection of the entire epileptogenic area, but also for the identification of its relationship with eloquent areas in the cortex, such as motor and language cortex [3]. Especially for non-lesional patients, the combination of both structural, i.e. high field MRI, and functional modalities, such as positron emission tomography (PET) and single-photon emission computed tomography (SPECT), is an essential part of pre-operative investigation [4]. The very blurred functional images provide information regarding the area with abnormal metabolism, that points to origin of seizures [1], while detailed anatomical images offer enough information for surgical planning, either surgical resection or electrodes implantation. The co-registration of different modalities in presurgical planning is technically challenging, as multimodal images may present large differences in intensity value distribution, in spatial resolution and orientation [5].

Image registration methods are divided into three categories: point-based methods, surface-based methods and intensity-based methods [6]. Point-based methods rely on a set of corresponding fiducial points that are clearly discernible. Their identification is usually based on interactive visual identification of anatomical landmarks. Surface-based methods consist of matching 3D boundary surface of anatomic structures. Image segmentation pre-processing is required for extracting the surface features. In principle, totally blind to spatial features, intensity-based methods are applicable directly on image raw data and require less anatomy-dependent pre-processing or user interactions, because they take into account only the similarity of voxel intensity. Looking for a data-independent solution, a variety of similarity measures has been proposed. The most successful similarity measures are information theoretic measures. In particular, the mutual information (MI)

measure, thoroughly presented in [7], [8], performs better in the cases that the volumes are partially overlapped as the measures are normalized with respect to the partiality.

The MMI-based registration is, indeed, an optimization problem that looks for a geometric transformation $\hat{\theta}$ which maximizes the mutual information $I(\boldsymbol{S}, \boldsymbol{F})$ of the voxel intensities in $\boldsymbol{S}$ and $\boldsymbol{F}$

$$\hat{\theta} = \arg\max_{\theta} I(\boldsymbol{S}(\mathbf{x}), \boldsymbol{F}(T_\theta \mathbf{x})), \tag{1}$$

where $\boldsymbol{F}$ is a floating image, $\boldsymbol{S}$ is a reference image, and T is a geometric transformation with parameters θ applied to a point $\mathbf{x}$.

Three processing stages are necessary before solving the optimization problem expressed in Eq. 1: the initial superposition of the reference and floating images, the calculation of probability density function, and the computation of mutual information. Because, in theory, the solution should be independent of the initial superposition, the existing proposals differ mainly in the two latter stages. They show that minor variations may yield remarkable differences either in the registration robustness or in the convergence rate to a locally optimal solution for T_θ [9], [10]. Nevertheless, we propose to study the impact of the first stage on numerical outcomes in this work.

Contributions: Instead of smoothing the MI-function with very sensitive application-dependent parameters, we are motivated from the known results to seek a simple and intuitive way for pre-aligning the multimodal volumes, represented in the Digital Imaging and Communications in Medicine (DICOM) format, such that their initial relative position is close to the correct solution. Based on a comparative analysis performed in this work, we propose to use the mask-based filtered volume centroids and the direction cosine information provided in raw image data for estimating the initial transformation of $\mathbf{x}$. The differential of our proposal for creating a mask is to involve an expert in the process. In the absence of gold standard, we also contribute in this work with a novel necessary condition, namely the direct-inverse matching condition, for quantifying the estimated registration errors to facilitate comparative analysis.

A. Related Work

Despite its success, MMI-based registration has still a number of well-known drawbacks. Its main disadvantage is that $I(\boldsymbol{S}, \boldsymbol{F})$ has typically many local maxima. A vast amount of research has been devoted to relieve some probable roots of the problem, namely probability distribution estimation and interpolation artifacts, the choice of optimization technique and the information-theoretic measure itself, in order to make MMI-based registration algorithms widely applicable without modality-specific pre-processing [9], [10]. In addition, a series of randomized experiments were performed to assess robustness and reliability of the MI criterion. For the computed tomography (CT) and MR images, it is reported in [10] that within a large range of -25 mm to 25 mm translation and -25° to 25° rotation around the correct position, there is only a single strong optimum, which coincides with the correct registration solution. In [7], it is reported 90% success rate of convergence to near the correct solution, starting from a randomized initial position using translational and rotational offsets for each axis uniformly selected in the range of [-25 mm,-25 mm] and [-10°,10°] around the ground truth.

We may assume as the rigid transformations around the origin

$$\begin{aligned} RT_\theta &= T_d \cdot R_{\phi_z} \cdot R_{\phi_x} \cdot R_{\phi_y} \\ &= \begin{bmatrix} 1 & 0 & 0 & d_x \\ 0 & 1 & 0 & d_y \\ 0 & 0 & 1 & d_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot R_{\phi_z} \cdot R_{\phi_x} \cdot R_{\phi_y}, \end{aligned}$$

where R_{ϕ_z}, R_{ϕ_x} and R_{ϕ_y} are rotations about x-, y- and z-axis, respectively. Since the rotations are relative to the reference volume's center, we should first translate the reference volume so that its center is at the origin by using the translation matrix M_{center}. Next, we apply RT_θ, and finally displace the results back to the original center by means of M^{-1}_{center}. The sequence of operations is

$$\begin{aligned} CT_\theta &= M^{-1}_{center} \cdot RT_\theta \cdot M_{center} \\ & \quad M^{-1}_{center} \cdot T_d \cdot R_{\phi_z} \cdot R_{\phi_x} \cdot R_{\phi_y} \cdot M_{center}. \end{aligned} \tag{2}$$

Maes *et al.* [8] briefly discuss pre-alignment by suggesting to initially position the reference and floating images such that their centers coincide and the corresponding scan axes of both images are aligned and have the same orientation. We wish to expand on this subject and, for this reason, we distinguish in our study two kinds of rotational pre-alignment:

1) M_{PAR} that aligns the scan axes, or the direction cosine information in DICOM image data, of the floating volumes with respect to the scan axes of the reference one [8]; and
2) M_{PAP} that aligns the scan axes of the floating volumes with respect to the axes of the patient-centered coordinate system.

Denoting the reference volume center as **rC=**$(rC_x, rC_y, rC_z, 1)$ and the floating volume center as **fC=**$(fC_x, fC_y, fC_z, 1)$, a trivial translational alignment is to displace the floating volume so that the centers of both volumes become the same. Once in the DICOM format the origin of a volume is at the upper left hand corner of the volume, the center fC should be rotated, when the floating volume is rotated with respect to the reference volume through $M_{PAR|PAP}$, i.e. M_{PAR} or M_{PAP}. For aligning two centers, we should actually displace the rotated center of the floating volume, $fC' = M_{PAR|PAP} fC$, toward the center of the reference volume.

However, the displacement to $\mathbf{rC}$ does not deliver a good solution yet. This is because that the patients may not be exactly centered with respect to the image volume. This motivates several works to replace geometric volume centers by geometric volume centroids $\mathbf{r\overline{C}}$=$(r\overline{C}_x, r\overline{C}_y, r\overline{C}_z, 1)$ and

$\mathbf{f\overline{C}}$=$(f\overline{C}_x, f\overline{C}_y, f\overline{C}_z, 1)$ [6], [11]. Since the centroid estimation is strongly dependent on the voxels of interest, some sort of filtering or segmentation is required to remove samples that are clinically irrelevant. Image processing techniques, such as threshold-based filtering, morphology, connected component feature extraction, may be employed [11]. In our analysis we consider two kinds of translational pre-alignment:

1) M_{PL} that makes the centroids of threshold-based filtered volumes coincide, and
2) M_{PLC} that leads the centroids of mask-based filtered volumes to be coincident.

B. Technique overview

Composing the pre-alignment rigid transformation $M_{PL|PLC} \cdot M_{PAP|PAR}$ with CT_θ, we get a new expression for the searched rigid transformation:

$$\begin{aligned} T_\theta &= CT_\theta \cdot M_{PL|PLC} \cdot M_{PAP|PAR} \\ &= M^{-1}_{center} \cdot T_d \cdot R_{\phi_z} \cdot R_{\phi_x} \cdot R_{\phi_y} \cdot M_{center} \cdot \\ &\quad M_{PL|PLC} \cdot M_{PAP|PAR}, \end{aligned} \tag{3}$$

where the parameters of T_d, R_{ϕ_z}, R_{ϕ_y} and R_{ϕ_z} should be computed such that Eq. 1 is satisfied.

Once set the transformation model for T_θ and assumed $M_{PAP|PAR}$ derivable from the direction cosine triplets in the DICOM header, a question about implementation arises: how to rapidly estimate the centroids of a noisy image to build $M_{PL|PLC}$, once the noise may disguise valid samples?

We use the fact that the shape recognition is invariant for healthy individual and that GPU (graphics processing unit) resources allow rendering at interactive rate, and propose an environment in which an expert may interactively adjust the noise threshold as schematized in Figure 2. When the displayed volume reaches the expected shape, the depth maps of the visible surface are off-screen rendered and a volumetric binary mask is built for extracting a filtered volume. The procedure is detailed in Section III.

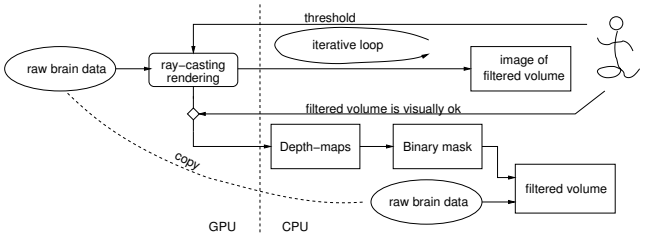


Fig. 2. Flowchart for a filtered volume.

II. Technical background

In this section, we present the classical formulation of rigid transformation applicable to medical image registrations.

A. Rotations

The geometric transformation M_{PAR} that aligns the orthonormal scan axes of a reference volume, $\{(r_{xx}, r_{xy}, r_{xz}, 0), (r_{yx}, r_{yy}, r_{yz}, 0), (r_{zx}, r_{zy}, r_{zz}, 0)\}$, and the orthonormal scan axes of a floating volume, $\{(f_{xx}, f_{xy}, f_{xz}, 0), (f_{yx}, f_{yy}, f_{yz}, 0), (f_{zx}, f_{zy}, f_{zz}, 0)\}$, is

$$M_{PAR} = \begin{bmatrix} r_{xx} & r_{yx} & r_{zx} & 0 \\ r_{xy} & r_{yy} & r_{zy} & 0 \\ r_{xz} & r_{yz} & r_{zz} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_{xx} & f_{xy} & f_{xz} & 0 \\ f_{yx} & f_{yy} & f_{yz} & 0 \\ f_{zx} & f_{zy} & f_{zz} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

To initially align the scan axes of the floating volume and the patient-centered coordinate system with M_{PAP}, the transpose of the matrix built by the scan axes of the floating volume may be used. This is because the scan axes are orthonormal and the patient coordinate system is built from orthogonal unit vectors.

B. Centroid

A centroid $\overline{\mathbf{C}}$ of a volume is obtained by averaging the position coordinates P_i of all n valid voxels, i.e. $\overline{C} = \frac{\sum_{i=1}^{n} P_i}{n}$.

III. New technique

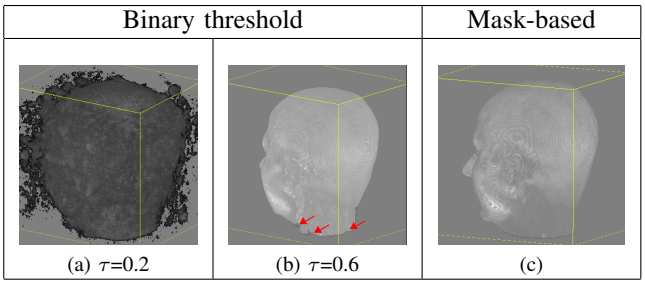
Our technique aims at improving the convergence of Eq. 1 by applying initially on each sample $\mathbf{x}$ the pre-alignment transformation $M_{PL|PLC} \cdot M_{PAP|PAR}$. To enhance the estimate of the volume centroids that are necessary for building $M_{PL|PLC}$, we propose a novel filtering technique to extract valid voxels.

It is worth remarking that previous survey reports draw attention to the importance of restraining the sample spaces being registered in practice [6], [10], [9]. Therefore, the extracted valid voxels also particularly suit to feed an MMI-based registration algorithm.

In order to avoid tampering with the sample spaces and to overcome the fact that the existing costly image processing techniques fall far short of covering a large variety of situations, we ask ourselves what is the best way to present intermediary results to an expert and how s/he can provide feedbacks in a simple and consistent way. Our solution is to adopt the paradigm presented in [12] focusing on the interface between an expert and the filtering algorithm. We also propose a necessary condition, namely the direct-inverse matching condition, for quantifying the estimated registration errors to facilitate comparative numerical analysis.

A. Filtering

From the user perspective, the simplest way to filter out the unwanted noises is to use the binary thresholding image filter. Only one parameter should be set. If the voxel value is above the specified threshold the output voxel value is assigned as valid. Otherwise the output voxels are discarded. In our experiments, we tried a series of normalized thresholds $\tau = \frac{specified\ threshold}{maximum\ intensity\ value}$ and empirically we get τ=0.2 for most resonance magnetic images acquired with our hospital's scanners.



(a) τ=0.2 (b) τ=0.6 (c)

Fig. 3. Filtering the volume depicted in Figure 7b.

The main problem we faced with the binary threshold filtering technique is when the noise has higher intensity values, as illustrates Figure 3a. However, if we increase τ to 0.6, valid samples (a part of the neck) are removed as shown in Figure 3b. This stems from the fact that, differently from the images acquired by CT scanners, the tissue responses to the magnetic and other radiation excitations are diffuse. Most brain components have large and overlapped response spectra, which makes their filtering a very hard task. As consequences, the joint probability distribution of the reference and the floating volumes may be critically altered and the estimation of centroids may deviate much from the correct ones.

Wu et al. present in [12] a way to remove only the samples before the visible surface of interest whose intensity value is above τ. In this way, all samples along the view direction and behind the surface are preserved. Figure 3c illustrates the visualization of the outcomes of the procedure applied on the same volume. Note that the noise removal has not affected the head's geometry.

(a) Top (b) Front (c) Back

Fig. 4. Filtering depth maps.

For estimating the volume centroids, the visible valid samples should be, however, numerically processable. Inspired by the idea of image masks, we apply the algorithm presented in [12] to generate, from the top, the front and the back views, three off-screen depth maps with the expert's specified threshold as illustrated in Figure 4. Each depth map separates the volume voxels in two sub-sets: the voxels that are on the interior side of the head and the voxels that are on the exterior side. Exploring this fact, we build with the intersection of the interior sides of three generated depth maps a 3D volume mask and apply it to extract brain region of interest. Figure 5 illustrates how our proposed algorithm works. The blue and the red arcs represent the top and the front depth maps, respectively, and the blank region corresponds to the region of valid samples that are further used in the centroid estimation and in the co-registration.

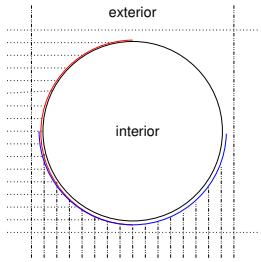


Fig. 5. Filtering with depth maps in 2D view.

B. Validation Test

It is a common practice to use the marker-based gold standard for evaluating the accuracy of registration methods [10]. However, it is not possible to implant fiducial markers in the acquisition sessions of the Boldrini Hospital and Hospital de Clinicas yet. To overcome this limitation, we devise other procedures to monitor the behavior of our proposed algorithm.

Maes *et al.* already observed in [8] that the choice of the floating image may affect the registration results. Nevertheless, in theory, if the accuracy of our algorithm is good enough, the role of the volumes to be registered should not impact much in the registration results if they have the similar voxel sizes. A necessary condition for a perfect registration of two volumes A and B is, therefore, that the estimated registration matrix $M_{A \rightarrow B}$, from A to B, is exactly equal to the inverse of the estimated matrix $M_{A \leftarrow B}$, from B to A.

Figure 6 depicts this condition. In the figure, volume A and volume B are drawn in red and green, respectively. The arcs in solid line represent direct transformations, while the arcs in dotted line denote inverse transformations. If the estimated transformations are correct, the distances $e1$ and $e2$ must be zero, because both direct (solid line) and inverse (dotted line) transformations should lead to the same solution. We will refer this necessary condition as direct-inverse matching condition. With this reasoning in mind, we interchange the roles of two unaligned volumes, A and B, in our experiments for checking the behavior of MMI-based registration algorithm combined with pre-alignment.

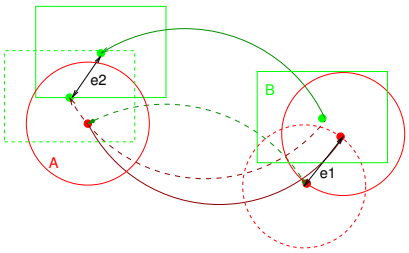


Fig. 6. Direct-inverse matching condition: $e1$ = $e2$ = 0.

IV. Implementation

As explained in Section III-A, the key for achieving our proposal is to get appropriate filtered volumes. Figure 2 schematizes our proposed flowchart to get such filtered volumes. Observe in the flowchart that we explore the GPU resources to render the off-screen depth maps and the on-screen filtered volumes. Two bottlenecks that may impair interactivity have already been identified [13]: the framebuffer bandwidth for writing depth values and the overhead of transferring large volume of data between CPU and GPU over PCI-e bus.

The modern GPUs are provided with up to 6GB video ram and texture mapping units (TMUs) that have not only texture caches but also up to 32-bit float processing capacity. Therefore, for reducing the first bottleneck, we shift the access of the depth buffer to the texture memory by storing the depth maps as textures in the GPU. And, for alleviating the second bottleneck, we transfer the raw volumes to the GPU as the first step in the filtering algorithm. In this way, only the normalized threshold (a floating-point value) should be sent to the GPU at each interaction in the iterative loop presented in Figure 2.

Another implementation strategy we adopted for improving the performance is to use 3 depth maps (top, front and back), instead of 6 depth maps (top, front, back, bottom, right and left), to construct the binary mask volume. Experimentally, we observe that the differences in the outcomes do not compensate the time and memory overhead.

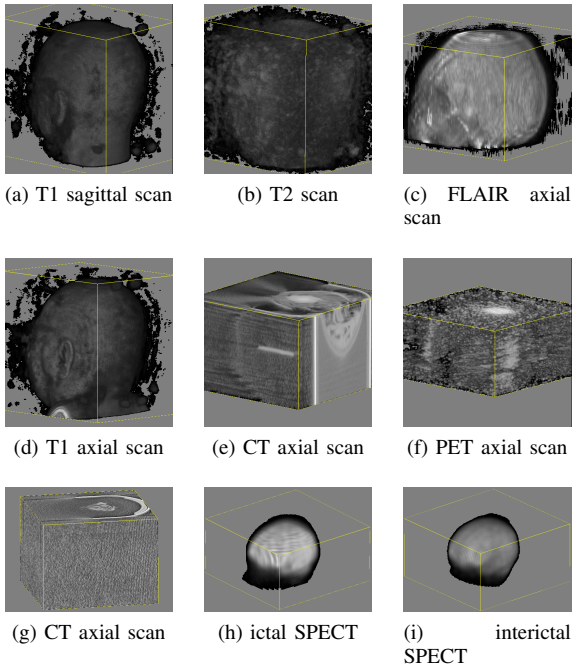
The last remark concerns the computation of T_θ from Eq. 3. In our implementation we consider that T_θ is composed of a sequence of 3D rotations followed by a translation, and apply Powell's multidimensional direction set method to minimize each parameter. For the sake of simplicity, we transform all the voxels of the floating image with $(M_{center} \cdot M_{PL|PLC} \cdot M_{PAP|PAR})$ before using Powell's method to get independently the 6 parameters in $(T_d \cdot R_{\phi_z} \cdot R_{\phi_x} \cdot R_{\phi_y})$. The initial values of these parameters are $\phi_{x0} = \phi_{y0} = \phi_{z0} = d_{x0} = d_{y0} = d_{z0} = 0$.

Our implementation platform is a notebook ASUS Intel®Core(TM)2 Duo T6600 2.2GHz with 2GB RAM and a NVIDIA GeForce GT 220M with 1GB VRAM.

V. Experiments

Nine LPS-oriented volumes of three patients have been chosen to illustrate the reliability and the robustness of our proposed procedure: (1) a set of T1-weighted (Figure 7a), T2-weighted (Figure 7b), and partial, oblique FLAIR (Figure 7c) MR volumes; (2) an off-centered T1-weighted MR volume (Figure 7d), the corresponding CT (Figure 7e) and PET (Figure 7f) volumes; and (3) a CT volume (Figure 7g), the corresponding ictal SPECT (Figure 7h) and interictal SPECT slices (Figure 7i). Note that the SPECT volumes have been pre-processed and most noise removed.

In Table I we may observe that those volumes differ greatly not only in the modality but also in the dimensions and the voxel sizes. Although we are interested in the co-registration of functional and structural images, we included in our experiments the co-registration of structural images to facilitate the accuracy assessment of the registration outcomes in a qualitative way and to draw a parallel between the visual quality and the numerical registration errors. It is worth remarking that all patients enrolled in the present study signed informed consent form approved by the Ethics Committee of University of Campinas.



(1) (a) T1 sagittal scan (b) T2 scan (c) FLAIR axial scan

(2) (d) T1 axial scan (e) CT axial scan (f) PET axial scan

(3) (g) CT axial scan (h) ictal SPECT (i) interictal SPECT

Fig. 7. Experiment volumes of three patients: (1), (2) and (3).

TABLE I
Volume Size (C.=column, R.=row and S.=slices) and acquisition resolution (voxel sizes) of the images in Figure 7

	(C., R., S.)	voxel sizes		(C., R., S.)	voxel sizes
(a)	(180,240,240)	(1,1,1)	(b)	(120,240,240)	(1.5,0.9,0.9)
(c)	(448,448,30)	(0.5,0.5,5)	(d)	(180,240,240)	(0.99,1,1)
(e)	(512,512,110)	(0.5,0.5,1.5)	(f)	(256,256,110)	(1.3,1.3,1.5)
(g)	(512,512,149)	(0.6,0.6,1.5)	(h)	(128,128,32)	(2.3,2.3,4.5)
(i)	(128,128,68)	(2.3,2.3,2.3)			

The structural MR images from Figure 7a to Figure 7d were acquired in the same scanner (Philips Achieva 3T). The other volumes were scanned in different hospitals – the Boldrini hospital and our university teaching hospital, Hospital de Clinicas. The CT and PET volumes were acquired in the Siemens s5vb20b multimodal imaging scanner and the SPECT images in the GE Millenium MG scanner. We remark that even if the volumes in Figures 7a–7c were obtained in the same session, in which the patient has not moved excessively between acquisitions, their orientations differ slightly.

A. Comparisons of Pre-alignment Alternatives

In this section we show the performance of possible pre-alignment alternatives presented in Section I-B. Because the pairs of volumes in Figures 7g–7h and Figures 7g–7i have

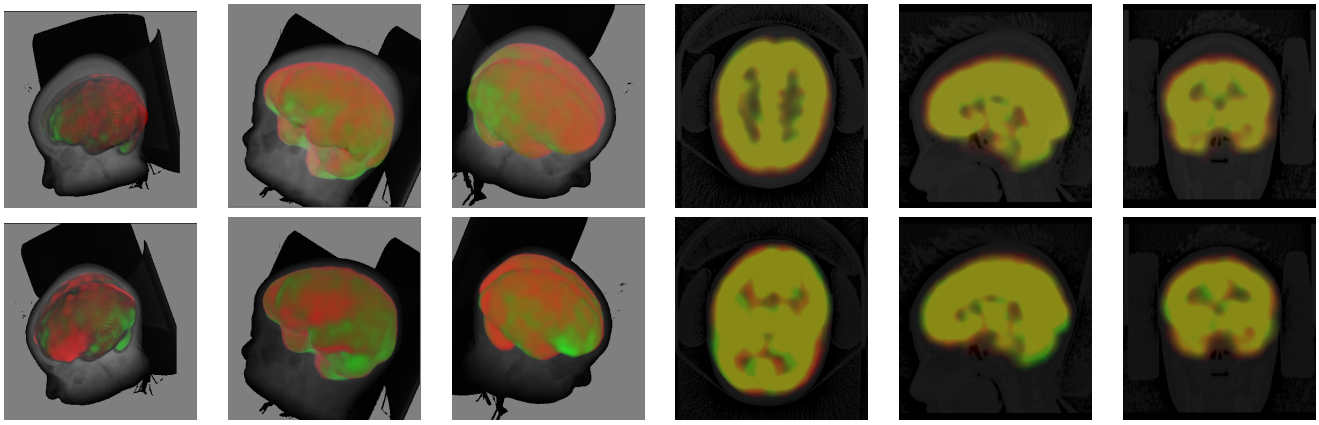


Fig. 8. Centroid estimation on the basis of binary thresholding volumes (M_{PL}): without angular pre-alignment and angular alignment with respect to the patient reference system M_{PAP} (first row), with respect to the scan axes of reference volume M_{PAR} (second row).

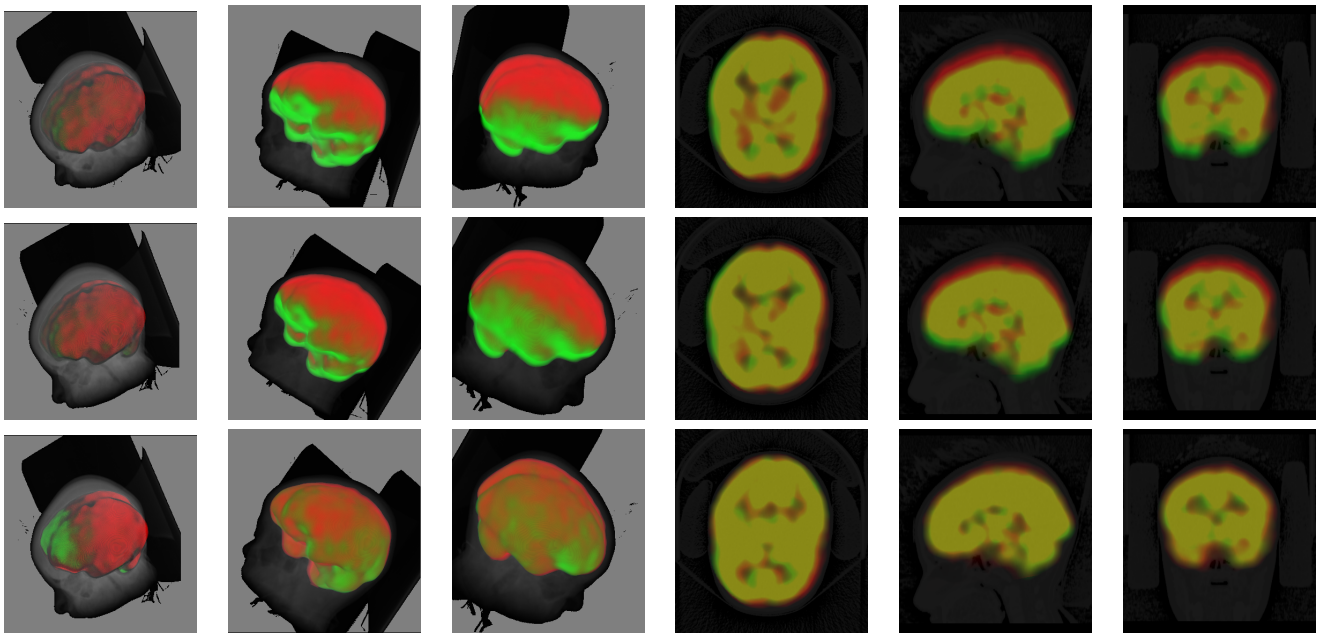
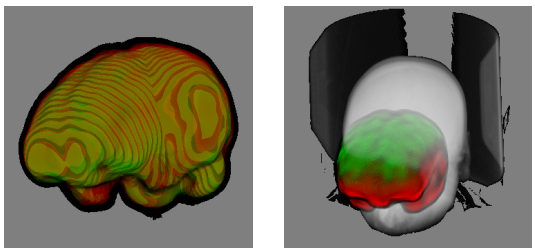


Fig. 9. Centroid estimation from binary mask filtering (M_{PLC}): without angular pre-alignment (first row), and angular alignment with respect to the patient reference system M_{PAP} (second row), with respect to the reference volume M_{PAR} (third row).



(a) Aligned SPECTs (b) raw CT, SPECTs data

Fig. 10. Ictal and interictal SPECTS and CT.

the largest discrepancy both in volume and in voxel sizes, they pose several challenging numerical issues. From them we may clearly perceive the difference in the registration results when we change the pre-alignment transformation matrices. Hence, they have been selected in our comparison experiments even though they are not clinically useful. The relative location and the orientation of three volumes before the registration are shown in Figure 10b. They are colored, respectively, in grayscale, red and green.

In our assessment tests, we use the same MMI-based registration algorithm described in [14]. The procedure we adopted is as follows. First, we co-register pairwise the volumes in Figures 7g–7h and in Figures 7g–7i. Then, we apply the obtained matrices to transform the coordinates of the SPECT volumes from its original reference to the reference of the CT volume. Finally, we compare visually the re-placed SPECT volumes with the co-registered SPECT volumes illustrated in Figure 10a.

Figures 8 and 9 present the outcomes of six proposed pre-alignment combinations. The first column presents the

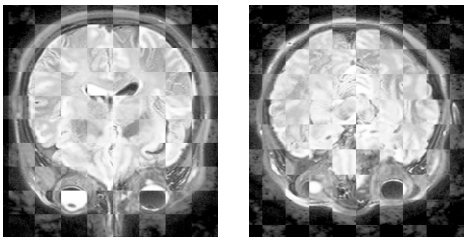
output of each pre-alignment procedure. This serves as input to the MMI-based registration algorithm. The second and the third columns show, respectively, the result of MMI-based registration from the right-side and the left-side views. Finally, the fourth, fifth and sixth columns present 200^{th} axial slice, 256^{th} sagittal slice and 75^{th} coronal slice of the registered volume.

In Figure 8 binary threshold filtering was applied to the volumes before the centroid of each volume is estimated. For the CT volume, the threshold τ is 0.2, while for the SPECT volumes, we chose 0.01. Observe that the first row of Figure 8 shows the outcomes corresponding to the results without angular pre-alignment and the results when we apply angular pre-alignment with respect to the patient reference system. This is because that the floating and the reference volumes have their orientation coincident with the patient reference system. In Figure 9 binary mask filtering is applied before the centroid of each volume is estimated. The noise threshold values for generating depth maps were 0.85 and 0.01 for CT and SPECT volumes, respectively. Although the combination of centroid estimation on the basis of filtered volumes and angular pre-alignment with respect to the reference volume (sixth alternative) does not yield the best pre-alignment, it delivers the best registration from our 2D visual inspection: not only the colors of SPECT volumes are well blended but also their placements with respect to the skull are almost aligned. From 3D visualization, we can perceive that there is almost no region with pure red or pure green. Therefore, we select the sixth procedure for pre-alignment and constrain our further discussions on the analysis of this pre-alignment alternative in terms of reliability and robustness.

B. Reliability

We consider a registration algorithm reliable if it has non-significant registration errors. If we simply apply our proposed registration algorithm on pairs of identical volumes, we expect that the rotation angles and the displacements are exactly zero. Nevertheless, due to numerical computation, we may have numerical rounding and propagation errors. Applying the procedure on the volumes depicted in Figure 7, the largest angle differences with the reference solution occur in an SPECT dataset for rotation around the x-axis (0.40 degrees) and the largest displacement differences are translations in the z-direction (-0.60mm). In sequence, the applied τ is: 0.4, 0.51, 0.4, 0.45, 0.85, 0.73, 0.78, 0.01, and 0.01.

For qualitative assessment, we investigate visually 2D spatial alignment by looking at continuity of tissue boundaries in the two interleaved structural images in a checkerboard pattern. The alignment is good if there is no discontinuity of contours and the tissues merge smoothly across the checkerboard borders, as depicts Figure 11. And, for quantitative evaluation, we assess the direct-inverse matching condition: we consider as registration errors the differences between the transformed position vector $\mathbf{x}$ by the estimated direct geometric transformation parameters, $M_{A\rightarrow B}$, and the position obtained by the inverse of the estimated transformation matrix $M_{B\rightarrow A}$. Since the errors are dependent on the distance of the points from the center of rotation [8], we compute two kinds of errors: centroid registration error (CRE), i.e. the distance of the transformed centroids, and box registration error (BRE), i.e. the maximum distance of the eight transformed scanned volume's corners. These corners are the vertices of the yellow box surrounding the brain's surface in Figures 3 and 7.

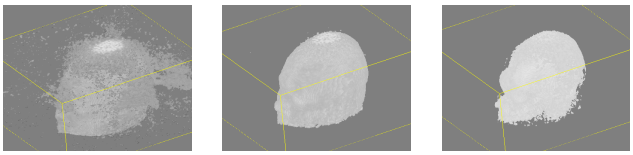


(a) Floating volume: FLAIR (b) Floating volume: T2-weighted

Fig. 11. Interchange of the roles of FLAIR and T2 volumes in registration.

C. Robustness

One may argue that our proposed interactive filtering procedure is prone to subjective interpretation which may affect registration results. Therefore, we also evaluate how noticeable the deviations of τ affect the results. For simplicity, we consider that the filtering threshold τ may vary between ± 0.10, and calculate the CRE and BRE errors of the points transformed both by the matrix obtained with the τ given in Section V-B and by the matrix estimated with its deviations in the worst situations. We consider as the worst situations when the filter of one volume is added to 10% and the filter of the other volume is subtracted from 10%. Figure 12 illustrates how these variations may affect a sample space.



(a) τ=0.63 (b) τ=0.73 (c) τ=0.83

Fig. 12. The influence of τ in the sample spaces.

VI. Results and Discussion

On the basis of our experiments, we may draw some concluding remarks with regard to reliability, robustness and performance.

A. Reliability

As expected from MMI-based registration techniques [9], [10], the volumes with larger voxel size, namely the FLAIR, the PET and the SPECT images, have larger registration errors when they are co-registered with volumes with smaller voxel size. Pre-alignment has improved the convergence to numerical solutions that seem close to the correct result.

Without resampling the images isotropically and considering the numerical errors in the registration of an image with itself, the registration of the volumes of the patient (1) seems fair under the direct-inverse matching condition. The errors $e1$ and $e2$ increase consistently in accordance with the ratio of voxel sizes. For the patient (2), we observe discrepancies between $e1$ and $e2$. It follows from the fact that the physical dimensions of the scanned volumes differ greatly. In the registration of images of the patient (3) involving SPECT modality, we observe that the CRE and BRE errors and discrepancies between $e1$ and $e2$ are much greater although their dimensions are of similar size. We believe that these results are due to the fact that the PV interpolation schema is insufficient to guarantee accuracy in the cases that the reference voxel size and the floating voxel size differ greatly from one another.

B. Robustness

Surprisingly, we perceive that when we change the threshold τ the errors may be up to 1mm. Further investigation leads us to a plausible justification which is illustrated in Figure 12. The figure shows the variation of the sample spaces of the volume represented in Figure 7f. Note that their shapes vary noticeably in the range of pre-specified deviations. Certainly, this affects the outcomes. In practice, neuro-experts hardly make such large variations in their subjective decisions.

The most relevant result of this experiment is that the error pattern seems the same for all tested cases. In addition, our results support the conjecture presented in [8]: when an image possessing intensity values with higher spatial correlation is chosen as reference image, the registration is more robust.

C. Performance

The most costly stage is the computation of the joint histogram necessary for determining the mutual information, even a parallel approach based on reduction techniques is used [14]. Comparatively, the time for pre-alignment is negligible. Among the presented experiment tests, the registration time varies from 26s (the pair Figure 7h←Figure 7h) to 20 min (the pair Figure 7f←Figure 7e). We observe that the time performance depends critically on the volume data size of the floating volume. The higher the volume size, the greater the time spent. For example, when we change the role of the pair Figure 7f←Figure 7e to Figure 7e←Figure 7f, the time is reduced to 3.8 min.

D. Limitations

We have exhaustively tested our proposed procedure in all combinations of modalities available. Visually the outcomes are plausible, but its reliability must be validated with more objective means, such as intracranial eletroencephalogram, results from pathology, and finally marker-based gold standard. In addition, due to the hardware requirement, the implemented procedure cannot be installed in any clinician's personal computer yet.

VII. Conclusion

Aiming to improve the accuracy of the epileptogenic focus localization in nonlesional patients, we investigate in this work an alternative to enhance the behavior of the MMI-based registration algorithm in native space. Differently from the previous approaches focusing essentially on smoothing the estimate MI-function, we analyze the impact of pre-alignment on the registration of non-invasive functional and structural images. Figure 1 illustrates the application of our proposed procedure in clinical examinations.

Acknowledgment

We aknowledge Dr. Bárbara Juarez Amorim for all the functional images presented in this paper. We would like to thank São Paulo Research Foundation (FAPESP), National Council for Scientific and Technological Development (CNPq) and Coordination for the Improvement of Higher Education Personnel (CAPES) for the grants without which it would not be possible to turn our ideas into a working prototype.

References

[1] D. C. Alessio and L. Louis, "Localisation of epileptic foci using novel imaging modalities," *Current Opinion in Neurology*, vol. 26, no. 4, pp. 368–373, 2013.

[2] J. S. Duncan, "Imaging in the surgical treatment of epilepsy." *Nature reviews. Neurology*, vol. 6, no. 10, pp. 537–550, Oct. 2010. [Online]. Available: http://dx.doi.org/10.1038/nrneurol.2010.131

[3] R. Rodionov, C. Vollmar, M. Nowell, A. Miserocchi, T. Wehner, C. Micallef, G. Zombori, S. Ourselin, B. Diehl, A. W. McEvoy, and J. S. Duncan, "Feasibility of multimodal 3d neuroimaging to guide implantation of intracranial eeg electrodes," *Epilepsy research*, vol. 107, no. 1, pp. 91–100, November 2013.

[4] J. Zhang, W. Liu, H. Chen, H. Xia, Z. Zhou, S. Mei, Q. Liu, and Y. Li, "Multimodal neuroimaging in presurgical evaluation of drug-resistant epilepsy," 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2213158213001472?np=y

[5] E. Haber and J. Modersitzki, "Intensity gradient based registration and fusion of multi-modal images," in *Methods of Information in Medicine, Schattauer Verlag*, 2006, pp. 726–733.

[6] *Handbook of Medical Imaging: Medical Image Processing and Analysis*. SPIE Press, 2000, vol. 2, ch. Image Registration.

[7] W. M. Wells, III, P. Viola, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *Medical Image Analysis*, vol. 1, no. 1, pp. 35–51, 1996.

[8] F. Maes, A. Collignon, D. V, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE transactions on Medical Imaging*, vol. 16, pp. 187–198, 1997.

[9] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE Transcations on Medical Imaging*, pp. 986–1004, 2003.

[10] F. Maes, D. Vandermeulen, and P. Suetens, "Medical image registration using mutual information," *Proceedings of the IEEE*, vol. 91, no. 10, pp. 1699–1722, 2003.

[11] G. Bartoo and W. Hanson, "Multi-modality image registration using centroid mapping," in *Engineering in Medicine and Biology Society, 1989. Images of the Twenty-First Century., Proceedings of the Annual International Conference of the IEEE Engineering in*, 1989, pp. 550–551 vol.2.

[12] S.-T. Wu, J. E. Y. Vidalón, and L. Watanabe, "Snapping a cursor on volume data," in *SIBGRAPI*, 2011, pp. 109–116.

[13] C. Cebenoyan, "Graphics pipeline performance," in *GPU Gems*, R. Fernando, Ed. Addison-Wesley, 2004, pp. 473–486. [Online]. Available: http://http.developer.nvidia.com/GPUGems/gpugems_ch28.html

[14] A. C. Valente and S. T. Wu, "Registration and fusion with mutual information for information-preserved multimodal visualization," in *Proceedings of Workshop of Theses and Dissertations – Sibgrapi 2012*, 2012, pp. 1–6, accessed in November 2013.