

REGISTRATION OF DYNAMIC RENAL MR IMAGES USING NEUROBIOLOGICAL MODEL OF SALIENCY

Dwarikanath Mahapatra and Ying Sun

Department of Electrical and Computer Engineering, National University of Singapore
4 Engineering Drive 3, Singapore 117576, Singapore

ABSTRACT

In this paper we propose the use of a neurobiology-based saliency measure to improve the performance of a quantitative-qualitative measure of mutual information for rigid registration of 4D renal perfusion MR images. Our registration method assigns greater importance to more salient voxels by applying a soft thresholding function to normalized saliency values. The resulting saliency map is a better representation of what is truly visually salient than an entropy-based saliency map. Our tests on real patient datasets show that incorporating this saliency measure produces better registration results than traditional entropy-based approaches.

Index Terms— Neurobiological model, saliency, mutual information, rigid registration, renal MR.

1. INTRODUCTION

In renal perfusion MRI, the abdomen is scanned at regular intervals following bolus injection of a contrast agent. Renal perfusion MR image sequences are often affected by the motion induced by patient breathing or movement, resulting in misalignment of successive image samples. For automated analysis of the acquired images, it is important that image samples at different time instants be optimally registered.

Recent works on registration of kidney images include a phase difference movement detection method [1], an edge based registration method [2], and a registration framework based on wavelet and Fourier transforms [3]. Mutual Information (MI) based methods can deal with large intensity variations arising due to flow of contrast agent, and have also proved to be successful in multimodality image registration. However, most MI-based registration methods give equal importance to all voxels. Pluim *et al.* in [4] used gradient information to complement MI for better results. Luan *et al.* used a quantitative-qualitative measure of MI [5] that weighs voxels according to their saliency values or importance. In [5] the saliency of a voxel is determined by considering the entropy of the intensity distributions over a neighborhood. This approach although producing good results has certain limitations. By considering the intensity distribution over a small neighborhood, an entropy-based saliency measure does not

take into account what is truly salient for the human eye and is subject to the effects of noise. In addition, there is the inherent problem of the choice of an appropriate scale.

To overcome these problems, we propose to calculate the saliency using a neurobiology-based approach. In the neurobiological model of attention [6], the degree of saliency is determined by measuring the contrast between different image scales, in terms of low level features such as intensity and edge orientation. As a closer imitation of the human visual system, the neurobiology-based saliency measure has the following three advantages: 1) multiple scale representation of an image overcomes the problem of choosing an appropriate scale for the neighborhood; 2) by subsequent subsampling of the image, information of larger neighborhoods is incorporated to give a more robust saliency map; and 3) by the use of lateral inhibition to suppress noise in individual feature maps, noisy regions in the saliency map are inhibited to produce a final map that has distinct salient regions.

The rest of the paper is organized as follows. In Section 2 we give a brief outline of how the saliency map is generated. Section 3 describes our registration algorithm. We discuss our results in Section 4 and conclude with Section 5.

2. NEUROBIOLOGY-BASED SALIENCY MAP

Saliency is a concept which states that there are regions in a scene that are more “attractive” than their neighbors and hence draw attention. Attention can be due to bottom-up cues or top-down influences. Here, we opt for the visual attention based model [6] that uses bottom-up cues such as local image contrast, presence of edges, color and texture.

A saliency map of an image is a representation of the salient regions of the image. Saliency at a given location is determined primarily by the contrast between this location and its surroundings. The image formed on the fovea of the eye is the central object on which a person is focussing his attention resulting in a clear and sharp image. Regions surrounding the central object have a less clearer representation on the retina. To simulate this biological mechanism, an image is represented as a Gaussian pyramid comprising of layers of subsampled and low-pass filtered images. The central representation of the image on the fovea is equivalent to the image

at higher spatial scales, and the surrounding regions are obtained from the lower spatial scales. The contrast is thus the difference between these scales.

Let $I(c)$ and $I(s)$ denote the intensity map at scale c and s , respectively. The contrast map $I(c, s)$ is defined as

$$I(c, s) = |I(c) \ominus I(s)|, \quad (1)$$

where the center is given by level $c \in \{1, 2, 3\}$ and the surround is given by level $s = c + \sigma$, $\sigma \in \{3, 4\}$ in the Gaussian pyramid. Different features (luminance, edge, etc.) are extracted from the images to construct feature maps which contribute to the final saliency map. Application of the process of lateral inhibition suppresses noise to a great extent resulting in a saliency map that has distinct salient regions. A detailed description of the method can be found in [6].

3. METHODOLOGY

Given a 4D (time varying 3D volumes) renal perfusion MRI dataset of a patient, our objective is to align the volumes in a manner such that there is minimal registration error between corresponding voxels at different time instants. The major challenge in perfusion image registration is that the intensity variation of each individual voxel with time. Thus, it is important to devise a robust registration method that is able to deal with such intensity variations. Here we employ a quantitative-qualitative MI based method incorporating visual saliency as the utility of each voxel that works well for registration of multimodality brain images [5]. Our method is different in that instead of using an entropy-based approach to generate the saliency map, it uses the neurobiology-based approach as described briefly in Section 2, which leads to a more effective measure of the saliency. Another difference is that in our method the utility of each voxel is defined as a nonlinear function of the saliency measure, rather than using the saliency measure directly as the utility.

3.1. Quantitative-Qualitative Measure of MI

The mutual information of two images is a combination of their entropy values, both separately and jointly [7]. The MI of two images A and B is given by

$$MI(A, B) = \sum_{n=1}^N \sum_{m=1}^M p(A_n, B_m) \log \frac{p(A_n, B_m)}{p_n q_m}, \quad (2)$$

where $p(A_n, B_m)$ is the joint probability of image intensities A_n and B_m , while p_n and q_m are, respectively, the marginal probabilities of the image intensities A_n and B_m .

The traditional measure of MI only defines its quantitative aspect. However, the joint occurrence of events has different significance. While registering kidney images, the joint probability of voxel intensities in the region of the kidney hold far greater relevance than that of voxel intensities

in other regions. A mathematical representation of this qualitative importance factor is called the utility measure [8]. Let $u(A_n, B_m)$ represent the joint utility of the events A_n and B_m , the quantitative-qualitative measure of MI (QMI) is defined by

$$QMI(A, B) = \sum_{n=1}^N \sum_{m=1}^M u(A_n, B_m) p(A_n, B_m) \log \frac{p(A_n, B_m)}{p_n q_m}. \quad (3)$$

In our method, we use the neurobiological approach to calculate the saliency value of each voxel in the volume [6]. The resulting saliency values are then normalized such that the maximum saliency value is one at each time instant. Let $S(x)$ denote the normalized saliency value at a voxel x , its utility $U(x)$ is defined as a nonlinear function of $S(x)$:

$$U(x) = 0.5 \times \left[1 + \frac{2}{\pi} \times \arctan \left(\frac{S(x) - S_0}{\epsilon} \right) \right], \quad (4)$$

where S_0 and ϵ are the two free parameters. Here S_0 serves as a soft threshold, i.e., $U(x) = 0.5$ when $S(x) = S_0$; ϵ controls the width of the transitional region, i.e., the smaller ϵ is, the closer U resembles a step function. By adjusting the values of S_0 and ϵ , the utility defined in (4) allows us to place more importance on voxels with larger normalized saliency values, and at the same time, to ignore voxels with smaller normalized saliency values, resulting in improved registration accuracy. Experimentally we found that a value of 0.4 for S_0 and 2.0 for ϵ were robust for the renal MR images in our study.

Let $U_A(x)$ be the utility value of voxel x with intensity i in A and $U_B(y)$ the utility value of voxel y with intensity j in B , the joint utility value of intensity pair (i, j) is given by

$$u(i, j) = \sum U_A(x) \times U_B(y). \quad (5)$$

In order to compute the joint utility, the utility values at corresponding voxels are multiplied and summed over all voxels having a particular intensity pair.

3.2. Optimization

The optimal transformation T^* between the source image I_A and the target image I_B can be obtained by maximizing the QMI value between I_B and the transformed image I_{AT} under the transformation T as

$$T^* = \arg \max_T QMI(I_B, I_{AT}). \quad (6)$$

In this work we consider the kidney as a rigid body and restrict the transformation T to have 6 degrees of freedom accounting for both translation and rotation in 3D.

A multi-resolution framework, using four resolutions, was adopted to make the optimization algorithm run faster [9]. In the coarsest resolution an exhaustive search over the rotation parameters is performed and for each rotation parameter optimization for translation parameters is obtained. For the second coarsest resolution, local optimization is performed for

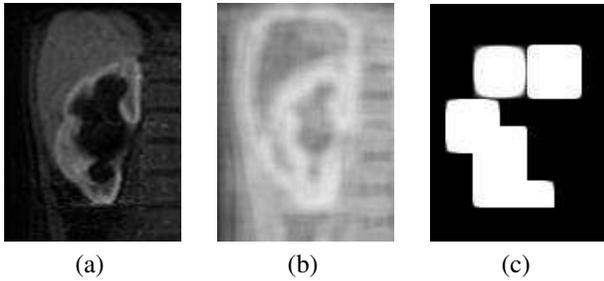


Fig. 1. (a) kidney in one typical frame; (b) saliency map generated by entropy-based method; and (c) by neurobiological model.

each of the candidate transformation parameters. For the subsequent finer resolutions local optimization is performed on the input candidate transformation parameters and the final registration result is obtained.

4. RESULTS

We tested our registration algorithm on 5 real patient datasets. The images were acquired on a GE 1.5T Signa scanner following bolus injection of Gd-DPTA contrast agent. The image matrix was 256×256 pixels and the number of volumes acquired varied from 24-44, with 10-18 slices in each volume.

4.1. Saliency Map

Fig. 1(a)-(c) show respectively the cropped image corresponding to the right kidney (of a patient) in one slice, the saliency map generated by the entropy-based approach in [5], and our neurobiology-based approach. We used only *luminance* and *edge* information for determining the saliency map. The saliency map shown in Fig. 1 (b) gives a lot of importance (in terms of saliency values) to less important regions surrounding the kidney. With the entropy-based method, regions with uniform intensity distributions have lower entropy, which implies that the kidney having uniform intensity (especially during the wash-in of contrast agent) will have an importance that is not much different from its surrounding areas. In contrast, the saliency map in Fig. 1 (c) overcomes the drawbacks of the entropy-based approach, showing distinct salient regions corresponding to the kidney. Although Fig. 1 (c) gives the impression of a mask, replacing the *mathematically* calculated saliency map with a *subjective* hand-drawn mask may not lead to good registration results. Moreover, such masks would be required for frames at all sampling instants making it a very tedious approach.

4.2. Qualitative Evaluation

We first evaluate the registration results qualitatively by visual examination. Fig. 2 (a) shows the kidney at the instant just prior to the wash-in of the contrast agent and Fig. 2 (b) shows the kidney at the next sampling instant when the wash-in of the contrast agent has already occurred. As shown, there

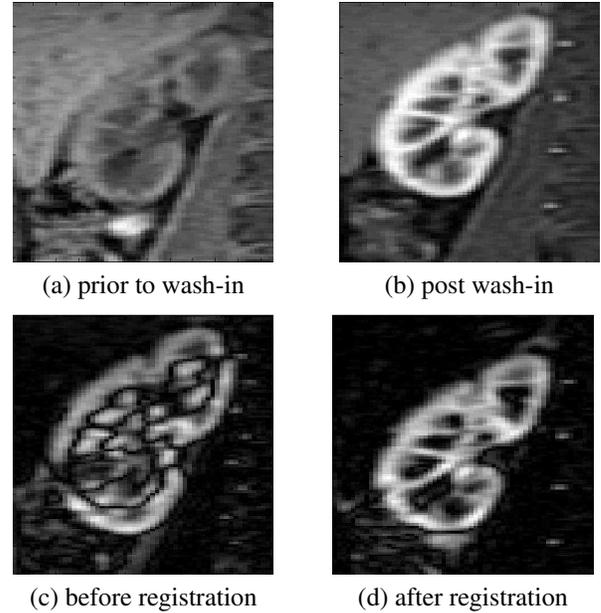


Fig. 2. Difference images before and after registration.

has been a slight displacement of the kidney in the upward direction due to patient breathing. Consequently, the difference image before registration in Fig. 2 (c) shows the intensity changes due to both patient motion and the wash-in of the contrast agent. Our objective is to obtain a difference image that displays the intensity changes of the kidney purely due to wash-in of the contrast agent. The difference image after applying our registration algorithm is shown in Fig. 2 (d), which verifies that the intensity difference is largely between voxels corresponding to the same physical location of the kidney. We have evaluated all the 5 datasets by observing the difference images as well as the registered datasets, and have obtained satisfactory registration results. Next, we attempt to evaluate the registration results quantitatively.

4.3. Quantitative Evaluation

In the absence of the ground truth for the clinical datasets, we tested the performance of our algorithm by first carefully selecting a dataset without any noticeable motion during the wash-in of the contrast agent; and then simulating the patient breathing motion by displacing the kidney volumes with a set of *random but known* translations and rotations. The translation parameters were randomly chosen from a uniform distribution between $(-10, 10)$ voxels for head to feet (HF) and left to right (LR) direction and between $(-3, 3)$ voxels in the anterior to posterior (AP) direction. Similarly, the rotation parameters were chosen between $\pm 10^\circ$ along all directions.

We applied three registration algorithms to the simulated 4D dataset, namely, the method in [4], quantitative-qualitative MI as in [5], and our method. The search ranges for the translation and rotation parameters were set to be

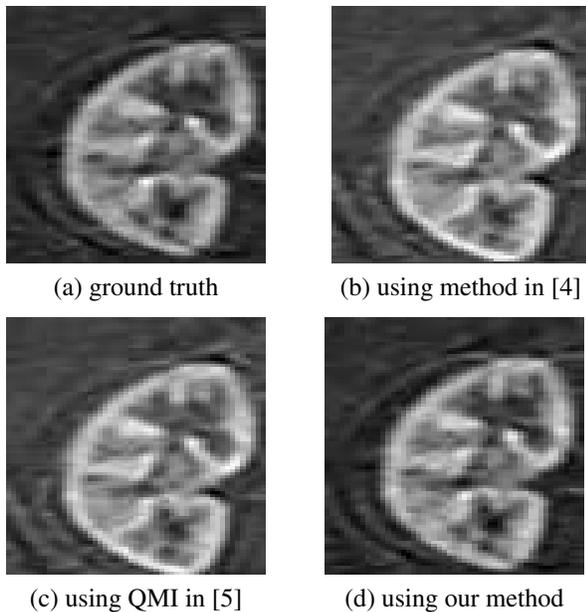


Fig. 3. Comparison of difference images obtained using three registration methods.

± 50 voxels in the HF and LR directions, ± 5 voxels in the AP direction, and $\pm 50^\circ$ for all rotations. The mean value of translation errors in voxels were (0.46, 0.42, 0.16) using the method in [4], (0.44, 0.43, 0.15) using the QMI in [5], and (0.3, 0.29, 0.1) using our method. The mean rotation errors were (0.0, 0.0, 0.57), (0.0, 0.0, 0.55), and (0.0, 0.0, 0.43) in degrees, respectively. Using respectively the intensity and edge information to quantify voxel importance, the methods in [5] and [4] have similar errors, larger than those of our method. This suggests that the intensity and edge information have nearly equal individual contributions in determining the importance of a region but a combination of these two leads to greatly improved results. Fig. 3 compares the difference images obtained using the three methods for one particular slice. It can be seen that the difference image obtained using our method, as shown in Fig. 3(d), is the most consistent with the ground truth difference image shown in Fig. 3(a). This is also reflected in the mean square error between the difference images shown in Fig. 3(a)-(c) and the ground truth; our method has the smallest mean square error. The above results indicate that our method is the most accurate among the three methods for this registration task.

5. CONCLUSION

In this paper, we proposed the use of a neurobiology-based visual attention model to determine voxel saliency for the purpose of image registration. Our approach overcomes the problem of scale selection. The resulting saliency map is less noisy and has distinctly salient regions, thanks to a lateral inhibition process and the combination of luminance

and edge information. By incorporating a soft thresholding function our registration algorithm is able to give more importance to voxels with higher saliency and hence leads to improved registration accuracy. We compared the performance of three registration methods, including our method, a traditional MI based method, and a qualitative-quantitative MI based method, on 4D renal perfusion MRI datasets. The experimental results show that our method outperforms the other two methods in terms of registration accuracy.

6. ACKNOWLEDGEMENT

We would like to thank Dr. Borys Shuter, Senior Lecturer in the Yong Loo Lin School of Medicine, National University of Singapore and Senior Imaging Physicist in the Department of Diagnostic Imaging, National University Hospital, for providing the datasets and insightful comments.

7. REFERENCES

- [1] E.L.W. Giele *et al.*, "Movement correction of the kidney in dynamic MRI scans using FFT phase difference movement detection," *Journal of Magnetic Resonance Imaging*, vol. 14, pp. 741–749, 2001.
- [2] Y. Sun, M.-P. Jolly, and J.M.F. Moura, "Integrated registration of dynamic renal perfusion MR images," in *ICIP*, 2004, vol. 3, pp. 1923–1926.
- [3] T. Song, V.S. Lee, H. Rusinek, M. Kaur, and A.F. Laine, "Automatic 4-D registration in dynamic MR renography based on over-complete dyadic wavelet and fourier transforms," in *MICCAI*, 2005, vol. 2, pp. 205–213.
- [4] J.P.W. Pluim, J.B.A. Maintz, and M.A. Vierger, "Image registration by maximization of combined mutual information and gradient information," *IEEE Transactions on Medical Imaging*, vol. 14, no. 8, pp. 809–814, 2000.
- [5] H. Luan *et al.*, "Multimodality image registration by maximization of quantitative-qualitative measure of information," *Pattern Recognition*, vol. 41, no. 1, pp. 285–298, 2008.
- [6] L. Itti and C. Koch, "A saliency-based search mechanism for overt and covert shifts of visual attention," *Vision Research*, vol. 40, pp. 1489–1506, 2000.
- [7] P. Viola and W.M. Wells III, "Alignment by maximization of mutual information," *International journal of Computer Vision*, vol. 24, no. 2, pp. 137–154, 1997.
- [8] M. Belis and S. Guiasu, "A quantitative-qualitative measure of information in cybernetic systems," *IEEE Trans. Inform. Theory*, vol. 14, pp. 593–594, 1968.
- [9] M. Jenkinson and S. Smith, "A global optimization method for robust affine registration of brain images," *Medical Image Analysis*, vol. 5, pp. 143–156, 2001.