A NON-PARAMETRIC APPROACH TO AUTOMATIC CHANGE DETECTION IN MRI IMAGES OF THE BRAIN

Hae Jong Seo and Peyman Milanfar

Electrical Engineering Department
University of California at Santa Cruz
1156 High Street, Santa Cruz, CA, 95064

ABSTRACT

We present a novel approach to change detection between two brain MRI scans (reference and target.) The proposed method uses a single modality to find subtle changes; and does not require prior knowledge (learning) of the type of changes to be sought. The method is based on the computation of a local kernel from the reference image, which measures the likeness of a pixel to its surroundings. This kernel is then used as a feature and compared against analogous features from the target image. This comparison is made using cosine similarity. The overall algorithm yields a scalar dissimilarity map (DM), indicating the local statistical likelihood of dissimilarity between the reference and target images. DM values exceeding a threshold then identify meaningful and relevant changes. The proposed method is robust to various challenging conditions including unequal signal strength.

Index Terms— change detection, magnetic resonance imaging (MRI), local regression kernel

1. INTRODUCTION AND OVERVIEW

The automatic analysis of subtle change between images of the same subject over time is a very important component in a large number of applications in diverse disciplines. Areas where such analyses are deployed include computer-aided diagnosis (CAD), video surveillance, and remote sensing, to mention just a few. In particular, change detection in medical diagnosis may be applicable to a broad range of diseases including cancers, Multiple Sclerosis, Alzheimer's and more. In general, a change detection method consists of three stages: 1) geometric registration of images, 2) intensity adjustments, and 3) image comparison to identify changes. We refer the interested reader to [1] and references therein for a good summary.

The generic problem of interest addressed in this paper focuses on the third component and can be briefly described as follows: We are given a set of brain Magnetic Resonance Imaging (MRI) scans of the same subject acquired over time, and we are interested in identifying pixels which are "significantly" different between the two MRI scans. Even in the absence of registration errors, estimating diagnostically significant changes is still challenging due to such factors as signal nonuniformity or presence of noise. A variety of MRI artifacts also introduce a wide range of confounding factors, making standard change detection methods unreliable. In order to deal with these problems, multispectral MRI scans were employed for the purpose of lesion detection by many researchers. For example, there are at least five different MRI modalities including T1 weighed, inversion recovery (IR), proton-density-weighted (PD), T2-weighted, and fluid attenuation inversion recovery (FLAIR). For statistical change detection in multispectral MRI scans, Bosc et al. [2] used the Generalized Likelihood Ratio Test (GLRT) followed by nonlinear joint histogram normalization. However, their approach tends to fail when noise is non-stationary. Patriarche et al. [3] also used multispectral MRI scans to detect progression of brain tumors. Recently, Rousseau et al. [4] proposed an *a contrario* approach to detect Multiple Sclerosis in multispectral MRI scans. However, in the majority of clinical situations, only one type of anatomical MRI scan is collected, since the acquisition of multispectral MRI scans is more time consuming and costly. Longer scanning times are further not feasible in many patients due to the severity of their conditions.

The main objective of this paper is to present an automatic change detection method given two images from a single MRI modality (See Fig. 1 for a graphical overview.) Our proposed method is based on the calculation and use of what we call *local regression kernels* which are local features computed directly from the given pixels in both the reference image and the target images, as elaborated below. The key idea behind local regression kernels is to robustly obtain local geometric structures of images by analyzing the radiometric (pixel value) differences based on estimated gradients, and use this structure information to determine the shape and size of a canonical kernel. The motivation to use these local regression kernels is the earlier successful work on adaptive ker-
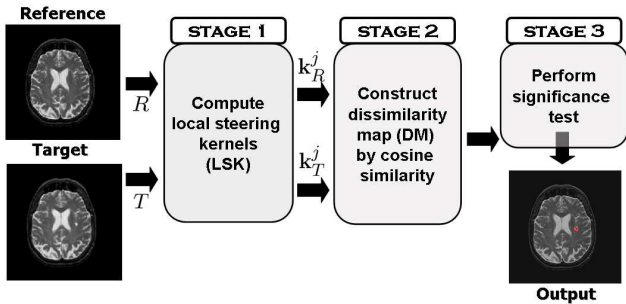
Fig. 1. System overview

nel regression for image denoising and interpolation [5]. Recently, Seo and Milanfar [6, 7] proposed to use local regression kernels as descriptors for generic 2-D object detection and demonstrated a high detection accuracy with challenging sets of real-world objects.

The local steering kernel is modeled as

$$K(\mathbf{x}_i - \mathbf{x}_j; \mathbf{C}_i) = \frac{\sqrt{\det(\mathbf{C}_i)}}{h^2} \exp\left\{\frac{(\mathbf{x}_i - \mathbf{x}_j)'\mathbf{C}_i(\mathbf{x}_i - \mathbf{x}_j)}{-2h^2}\right\}, \quad (1)$$

where $i \in [1, \cdots, P^2]$, P^2 is the number of pixels in a local window, h is a global smoothing parameter, and the matrix $\mathbf{C}_i$ is a covariance matrix estimated from a collection of spatial gradient vectors within the local analysis window around a sampling position $\mathbf{x}_i = [x_1, x_2]'_i$. The covariance matrix $\mathbf{C}_i$ modifies the shape and size of the local kernel in a way which robustly encodes the local geometric structures present in the image.

In what follows, at a position $\mathbf{x}_j$, we will essentially be using (a normalized version of) the function $K(\mathbf{x}_i - \mathbf{x}_j; \mathbf{C}_i)$ as a function of $\mathbf{x}_i$ and $\mathbf{C}_i$ to represent an image's inherent local geometry (See Fig. 2 for example.) It is worth noting that normalization of this kernel function [6, 7] yields robustness to illumination, contrast, and color differences.

Very recently, Pecot et al. [8] introduced a change detection framework based on the so-called "patch-based Markov models" in image sequence analysis. It is worth mentioning that their method is to detect pixels with meaningful change for several frames by first constructing a difference image while our method directly computes local steering kernels from the reference image and the target image. The proposed method has an advantage over their method in that the calculation of LSK is stable even in the presence of uncertainty in the data [5] and is not sensitive to relatively large variations in illumination [6, 7]. To summarize the operation of the overall algorithm, given the reference image and the target image, we first calculate the LSK from both the reference image and the registered target image at all pixel locations. Comparison between LSKs computed from two images is carried out using the cosine similarity measure [9, 10]. This step produces a "dissimilarity map" showing the likelihood of dissimilarity between the reference and target images. The final output is given after a significance test. (See Fig. 1 for a graphical overview.)
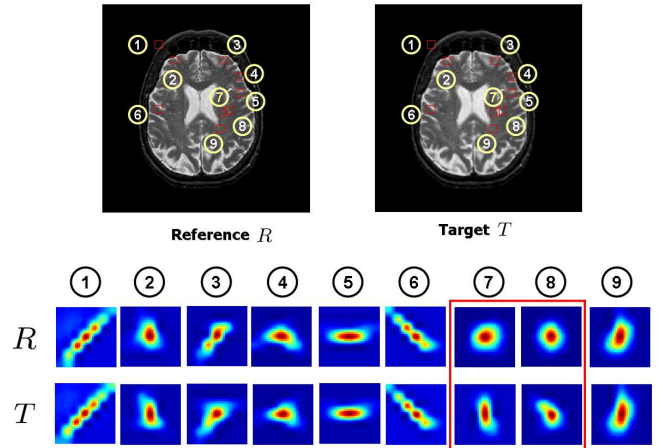

Fig. 2. Examples of LSK in various regions. Note that LSKs computed from various regions look alike except for regions 7 and 8 where small lesions exist in the target.

In the next section, we provide further details about the various steps outlined above. In Section 3, we demonstrate the performance of the system with some experimental results, and we conclude this paper in Section 4.

2. TECHNICAL DETAILS

Assume that we are given a target MRI scan T and that we have a reference MRI scan R. The first step in the proposed algorithm is to calculate the local steering kernel (LSK) measuring the relationship between a center pixel and its neighboring pixels, at each pixel from both R and T.

To be more specific, the local steering kernel function $K(\mathbf{x}_i - \mathbf{x}_j)$ is calculated and normalized as follows

$$\overline{K}_I(\mathbf{x}_i - \mathbf{x}_j) = \frac{K_I(\mathbf{x}_i - \mathbf{x}_j)}{\sum_{i=1}^{P^2} K_I(\mathbf{x}_i - \mathbf{x}_j)}\Big|_{j=1,i=1}^{M,P^2}, \quad I \in \{R, T\}, \quad (2)$$

where M is the total number of pixels in I. Fig. 2 shows some examples of LSK in various regions of both the reference and the target. Note that LSKs computed from various regions in both reference and target look essentially identical except for regions 7 and 8 where small lesions exist.

At each pixel $\mathbf{x}_j$, with a preselected window size of $P \times P$, we arrive at an array of P^2 numbers by column-stacking (rasterizing) $\overline{K}_I(\mathbf{x}_i - \mathbf{x}_j)$ as $\mathbf{k}_I^j (I \in \{R, T\})$.

The next step in the algorithm is the measurement of a "distance" between the computed features, $\mathbf{k}_R^j$ and $\mathbf{k}_T^j$. Earlier works such as [9, 10] have shown that correlation based metrics perform better than the conventional Euclidean and Mahalanobis distances for classification and learning tasks. Motivated by the effectiveness of correlation-based similarity measures, we propose to use cosine similarity for change detection. Cosine similarity is defined as the inner product

between two normalized vectors as follows:

$$\rho(\mathbf{k}_R^j, \mathbf{k}_T^j) = \left< \frac{\mathbf{k}_R^j}{\|\mathbf{k}_T^j\|}, \frac{\mathbf{k}_R^j}{\|\mathbf{k}_T^j\|} \right> = \frac{\mathbf{k}_R^{j\,'} \mathbf{k}_T^j}{\|\mathbf{k}_R^j\| \|\mathbf{k}_T^j\|} = \cos \theta_j, \qquad (3)$$

where $\cos \theta_j \in [-1, 1]$. The cosine similarity measure therefore focuses only on the angle (phase) information while discarding the scale information. In general, there are two main types of correlation: Pearson's correlation coefficient which is the familiar standard correlation coefficient, and the cosine similarity (so-called non-Pearson-compliant). Cosine similarity coincides with the Pearson's correlation when each vector is centered to have zero-mean. However, Pearson correlation is less discriminating than the cosine similarity due to the fact that centered values are less informative than the original values, and the computation of centered values is sensitive to zero or small values in the vectors. Since the discriminative power is critical in the change detection framework, we focus on the cosine similarity.

When it comes to interpreting the value of "correlation", it is noted in [11] that $\rho_j^2 \in [0, 1]$ describes the proportion of variance in common between the two LSKs as opposed to ρ_j which indicates a linear relationship between two LSKs $\mathbf{k}_R^j, \mathbf{k}_T^j$. As for the final test statistic comprising the values in the dissimilarity map, we use the *proportion* of "residual" variance $(1 - \rho_j^2)$ to the shared variance ρ_j^2. More specifically, the test statistic at each point in the image is computed and dissimilarity map (DM) is generated at each point as follows:

$$\text{DM} : f(\rho_j) = \frac{1 - \rho_j^2}{\rho_j^2}. \qquad (4)$$

From a quantitative point of view, we note that $f(\rho_j)$ is essentially the inverse of the Lawley-Hotelling trace statistic [12], which is used as an efficient test statistic for detecting correlation between two data sets.

In order to detect salient and significant changes using the DM, we need a threshold τ. If we have a basic knowledge of the underlying distribution of $f(\rho_j)$, then we can make predictions about how this particular statistic will behave, and thus it is relatively easy to choose a threshold which will indicate whether the pair of features from the two images are sufficiently dissimilar. But, in practice, we do not have a very good way to model the distribution of $f(\rho_j)$. Therefore, instead of assuming a type of underlying distribution, we employ the idea of nonparametric testing. We compute an empirical PDF from the values of $f(\rho_j)$ across the image and we set τ so as to achieve, for instance, a 99 % confidence level in deciding whether a given value is in the extreme (right) tail of the distribution. This approach is based on the assumption that in the target image, most of pixels are not involved with significant change, and therefore, the few outliers will result in values which are in the tail of the distributions of $f(\rho_j)$.

3. EXPERIMENTAL RESULTS

In order to validate the proposed method quantitatively, we simulated lesions in normal brain MRI slices (sagittal, coronal, and axial views). These simulated lesions were generated using a 3-D region of interests (ROI) creation tool provided in MRIcro [13]. Exact sizes and locations of irregular shapes of simulated lesions were stored and treated as the ground truth. In order to cover a variety of lesions, we constructed a total of 168 (14 ROIs in different sizes $\times$ 3 different views $\times$ 4 intensity reduction of $0\%, 20\%, 40\%$, and 60%, respectively) target slices by following the procedure as in [14]. Besides, we further made the intensity range of targets (T) containing lesions different from that of the reference (R). We compute LSK of size 5×5 as descriptors from both R and T. As a consequence, each pixel in R and T yields a 25-dimensional local descriptor respectively[1]. By performing significance test on the resulting dissimilarity map with confidence level $\tau = 0.99$, we detected regions with anomalous and statistically significant changes. Fig. 3 illustrates three examples of the detected results at the simulated lesions with 20 % intensity reduction (i.e., degree of lesion transparency). As an overall measure of performance[2], we were able to achieve sensitivity= 0.877, specificity=0.998, and similarity index (SI)=0.879.

Rousseau et al. [4] evaluated their method on simulated lesion images and reported their SI value around 0.75. Shen et al. [14] tested their lesion detection method based on segmentation to lesions generated from MRIcro [13] and reported their SI values on the simulated target slices with $20\%, 40\%$, and 60% intensity reduction as 0.867, 0.879, and 0.724 respectively. Our method which obtained an average of 0.879 for SI performs comparably with the method in [14] and outperforms the method in [4] even though [4] used multispectral MR images[3].

4. CONCLUSION

In this paper, we have proposed a novel and relatively simple, but effective statistical change detection framework to detect meaningful changes between two MRI images. Our main contribution consists of a general nonparametric statistical framework based on local steering kernels, and calculation of test statistics derived from cosine similarity. The proposed

---

[1]It is worth mentioning that performance of our change detection system is not particularly sensitive to the choice of LSK size because the local covariance matrix $\mathbf{C}_i$ plays a role in automatically determining the shape and size of kernels.

[2]sensitivity = $\frac{A_{gt} \bigcap A_{dt}}{A_{gt}}$, specificity = $\frac{(I - A_{gt}) \bigcap (I - A_{dt})}{I - A_{gt}}$, SI = $2 \times \frac{A_{gt} \bigcap A_{dt}}{A_{gt} \bigcup A_{dt}}$, where A_{gt} represents the ground truth, which is regions with true lesions (i.e., simulated lesions in this paper). A_{dt} represents detected lesions. I refers to the whole image.

[3]As pointed out in [4], it is difficult to provide a fair comparison among automatic change detection algorithms due to the fact that there is no gold standard and codes of state-of-the art methods are not publicly unavailable.
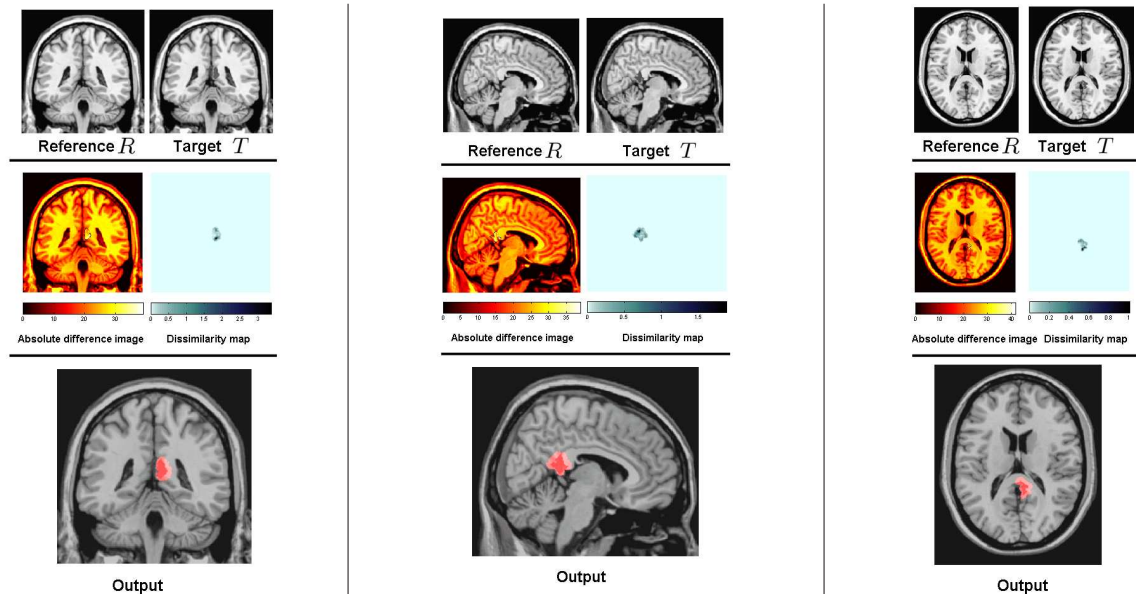
Fig. 3. Detected lesion on simulated images (left: coronal view, middle: sagittal view, right: axial view.) In all three cases, we used the same parameters $P = 5, h = 1.0, \tau = 0.99$. Note that absolute difference image can not identify lesions at all while the proposed method detected simulated lesions stably. Note that images are better illustrated in color.

framework is general enough as to be extendable to 3-D for tumor detection in serial MRI scans using analogous 3-D local steering kernels [15]. Due to its robustness to noise and other systemic perturbations, we also expect the present framework to be quite effective in other imaging modalities such as CT, PET, etc.

5. REFERENCES

[1] R.J.Radke, S. Andra, O. Al-Lofahi, and B. Roysam, "Image change detection algorithms: A systematic survey," *IEEE Transactions on Image Processing*, vol. 14, pp. 294–307, March 2005.

[2] M. Bosc, F. Heitz, J. Armspach, I. Namer, D. Gounot, and L. Pumbach, "Automatic change detection in multimodal serial MRI: applicaton to Multiple Sclerosis lesion evolution," *NeuroImage*, vol. 20, pp. 643–656, July 2003.

[3] J. W. Patriarche and B. J. Erickson, "Part 1. automated change detection and characterization in serial MR studies of brain tumor patients," *Journal of Digital Imaging*, vol. 20, pp. 203–222, September 2007.

[4] F. Rousseau, S. Faisan, F. Heitz, J. Armspach, Y.Chevalier, F.Blanc, J. Seze, and L. Rumbach, "An a contario approach for change detection in 3D multimodal images: Application to Multiple Sclerosis in MRI," *IEEE Engineering in Medicine and Biology Society (EMBS)*, pp. 2069–2072, August 2007.

[5] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Transactions on Image Processing*, vol. 16, no. 2, pp. 349–366, February 2007.

[6] H. J. Seo and P. Milanfar, "Using local regression kernels for statistical object detection," *Proceedings of IEEE International Conference on Image Processing (ICIP)*, pp. 2380–2383, October 2008.

[7] ——, "Training-free, generic object detection using locally adaptive regression kernels," *Submitted to IEEE Transactions on Pattern Analysis and Machine Intelligence*, October 2008.

[8] T. Pecot and C. Kervrann, "Patch-based markov models for change detection in image sequence analysis," *The International Workshop on Local and Non-Local Approximation in Image Processing*, August 2008.

[9] Y. Ma, S. Lao, E. Takikawa, and M. Kawade, "Discriminant analysis in correlation similarity measure space," *International Conference on Machine Learning*, vol. 227, pp. 577–584, 2007.

[10] Y. Fu, S. Yan, and T. S. Huang, "Correlation metric for generalized feature extraction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 12, pp. 2229–2235, 2008.

[11] R. J. Rummel, *Applied Factor Analysis*. Evanston, Ill.: Northwestern University Press, 1970.

[12] M. Tatsuoka, *Multivariate Analysis*. Macmillan, 1988.

[13] [Online]. Available: http://www.sph.sc.edu/comd/rorden/ mricro.html

[14] S. Shen, A. Szameitat, and A. Sterr, "Detection of infarct lesions from single MRI modality using inconsistency between voxel intensity and spatial location-a 3D automatic approach," *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, pp. 532–540, 2008.

[15] H. Takeda, P. Milanfar, M. Protter, and M. Elad, "Super-resolution without explicit subpixel motion estimation," *Submitted to IEEE Transactions on Image Processing*, 2008.