ROBUST IMAGE REGISTRATION IN THE GRADIENT DOMAIN

Yeqing Li[1], *Chen Chen*[1], *Jinghao Zhou*[2] *and Junzhou Huang*[1]

[1]Department of Computer Science and Engineering, University of Texas at Arlington
[2]Department of Radiation Oncology, University of Maryland School of Medicine

ABSTRACT

In many real-world applications of image registration, the images have significantly different appearances due to the intensity variations. Many existing intensity based methods may fail to solve these challenging problems. In this paper, we propose a novel method based on the differential total variation (DTV) for image registration. It is inspired by the fact that the image gradients are much more stationary than the intensities, especially when there exist severe intensity distortions. Therefore, we prefer to register the images in the gradient domain, which intuitively leads to more accurate registration results. An efficient algorithm is presented to solve the DTV minimization problem. The proposed method is scalable and has no regularization parameter to be tuned, both of which are desired properties for image registration. We show the accuracy and efficiency of our method through extensive non-rigid registration experiments, on synthetic MR images and real retina and iris images.

Index Terms— image registration, sparse representation, differential total variation

1. INTRODUCTION

Image registration aims to find the geometrical transformation to align two or multiple images into the same coordinate system. The geometrical transformation to be estimated can be either rigid, affine, piecewise rigid or non-rigid. Non-rigid registration is the most challenging task. Based on the featured used in non-rigid registration, existing methods can be classified into feature-based registration and intensity-based registration [1]. In this paper, we are interested in two-image registration using their intensities.

In the past two decades, many non-rigid techniques have been proposed (e.g., [2, 3, 4, 5, 6]). Most of these works are based on minimizing an energy function containing a distance (or similarity) measure and a regularization term. The regularization encourages certain types of transformation related to different applications. The minimum distance should correspond to the correct spatial alignment. One of the most successful distance measures is based on the mutual information (MI) of images [7]. However, in many real-world applications, the intensity fields of two images may vary significantly. For example, slow-varing intensity bias fields often exist in brain magnetic resonance images; in temporal registration of retina images, the images may contain severe intensity artifacts [8]. As a result, many existing intensity-based distance measures are not robust to these intensity distortions. Although some methods are proposed for simultaneous registration and intensity correction (e.g., [9]), they often involve much higher computation complexity and suffer from multiple local minima. Recently, a sparsity-inducing similarity measure called residual complexity (RC) has been proposed [8], where the discrete cosine transform (DCT) is used to code the residual. It shows that RC is more robust and accurate than MI. For dealing with intensity distortions, RC is one of the state-of-the-art methods for dealing with intensity distortions.

In this paper, we propose a novel method for intensity-based image registration. We observe that the image gradients or edges are much more stationary than image pixels under spatially-varying intensity distortions. Motivated by this observation, we define a new similarity measure to match the edges of two images. Therefore, the edges of the residual image are encouraged to be sparse. Any misalignment will increase the sparseness. This leads to a differential total variation (DTV) minimization problem, whose sparsity-inducing performance has been guaranteed in theory [10]. Compared to RC [8], our model is entirely parameter free, which is an advantage. An efficient algorithm based on the backtracking gradient descend method is proposed to solve this problem. In each iteration, the computation complexity is linear to the number of pixels. Experimental results on several synthetic and real-world applications demonstrate that our method outperforms the state-of-the-arts in terms of robustness, accuracy and efficiency.

2. ROBUST IMAGE REGISTRATION WITH DIFFERENTIAL TOTAL VARIATION

Our method comes from the intuition that the locations of the image gradients (edges) should almost keep the same, even under spatially-varying intensity distortions. Therefore, we

Corresponding: Junzhou Huang. Email: jzhuang@uta.edu. This work was partially supported by NSF IIS-1423056, CMMI-1434401, CNS-1405985.

propose to register the images in the gradient domain. If the images are not aligned optimally, ghost will appear on the residual image, i.e., the edges of the residual image will be less sparse. With this rational, we seek the sparsest solution of the composite image. Let $\mathbf{I}$ be the reference image and $\mathbf{S}$ be the source image. We use the vector ℓ_1 norm the encourage the sparseness of the residual image gradients:

$$\min_{\mathcal{T}} E(\mathcal{T}) = ||\nabla \mathbf{I} - \nabla \mathbf{S}(\mathcal{T})||_1, \tag{1}$$

where $\mathbf{S}(\mathcal{T})$ denotes the source image warped by a transformation $\mathcal{T}$. Let $\mathbf{r}$ denotes the vectorized residual image $\mathbf{I} - \mathbf{S}(\mathcal{T})$. The registration can be written as:

$$\min_{\mathcal{T}} E(\mathcal{T}) = ||\mathbf{r}||_{TV}, \tag{2}$$

where the TV for an image with N pixels is defined as $\|\mathbf{x}\|_{TV} = \sum_i^N (|\nabla_1 \mathbf{x}_i)| + |\nabla_2 \mathbf{x}_i|)$, and ∇_1 and ∇_2 denote the forward finite difference operators on the first and second coordinates. We call the TV of the residual in (1) as differential total variation. The TV minimization has been successfully applied on a wide range of image reconstruction tasks (e.g. [11, 12, 13]), and the performance has been theoretically proved [10]. To our best knowledge, this is the first study to define the TV as a similarity measure in image registration. It is worthwhile to note that our method is substantially different from the TV regularization methods [14, 15], where the widely used sum-of-squared-difference (SSD) is their actual similarity measure. There is no regularization term in our model and no regularization parameter to be tuned.

3. ALGORITHM

The TV function is convex but not smooth. To efficiently minimize the non-differentiable TV function, we can have a tight approximation for the absolute value: $|x| = \sqrt{x^2 + \epsilon}$, where ϵ is a small constant (e.g. 10^{-10}). Now, this approximation enables us to obtain the gradient of the energy function $\nabla E(\mathcal{T})$ by the chain rule.

$$\nabla E(\mathcal{T}) = \sum_{i=1,2} \mathbf{J}^T \nabla_i^{-1} \frac{\nabla_i \mathbf{r}}{\sqrt{\nabla_i \mathbf{r} \circ \nabla_i \mathbf{r} + \epsilon}}, \tag{3}$$

where ∇_i^{-1} denotes the inverse operation of ∇_i; J denotes the image Jacobian matrix with respect to the transformation parameters; $\circ$ denotes the Hadamard product. We can find the computational complexity of Eq. (3) is linear to the image size. Gradient descent with backtracking is used to minimize the energy function (1), which is summarized in Algorithm 1. We set the initial step size $t_0 = 1$ and $\eta = 0.8$ for all experiments.

Algorithm 1 Gradient descent with backtracking

input: $\mathbf{I}, \mathbf{S}, t_0, \eta < 1, \mathcal{T}_0, k = 0$.
repeat
1) compute $\mathcal{T}_{k+1} = \mathcal{T}_k - t_k \nabla E(\mathcal{T}_k)$
2) if $E(\mathcal{T}_{k+1})/M_{k+1} > E(\mathcal{T}_k)/M_k$, set $t_k = \eta t_k$ and go back to (1)
3) $t_{k+1} = t_k$
4) $k = k + 1$
until Stop criterions

The function value is calculated on the overlapped area of two images. To avoid trivial solutions such as zooming in on a dark region, we use the normalized function value here (divided by the overlapped pixels M). When there is no overlapping, the function value will be infinity. We found this approach could effectively rule out the trivial solutions. The non-rigid transformation is modeled using the free form deformation (FFD) transformation with B-spline control points [16]. After the transformation is estimated, the image is warped by this transformation and we repeat this process. In addition, the above process is implemented in a coarse-to-fine hierarchical framework.

4. RESULTS

4.1. Simulation

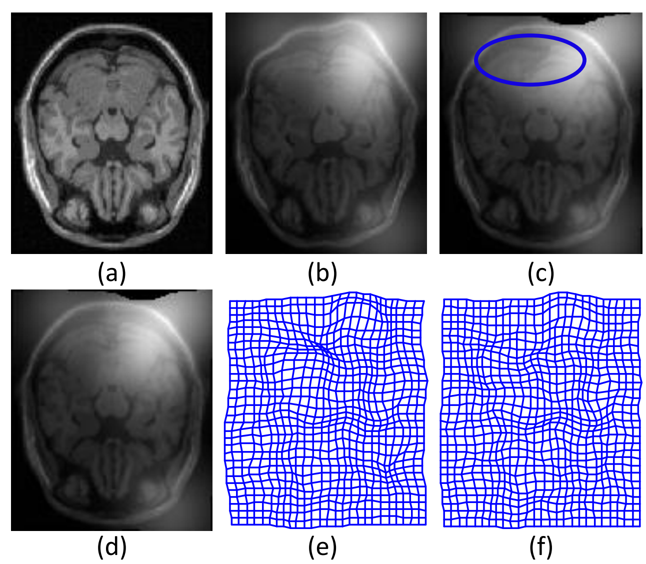

Fig. 1. Synthetic experiment with non-rigid transformation. (a) The reference image. (b) The source image with intensity distortion. (c) Registration result by RC. (d) Registration result by our method. (e) The transformation estimated by RC. (f) The transformation estimated by our method. Best viewed in $\times 2$ sized pdf file.

First, we conduct a simulation on a brain MRI image [17].

The source image is warped by a non-rigid transformation generated by Gaussian perturbations. We add a few Gaussian fields to simulate the intensity distortion and rescale the images to [0,1]. Fig. 1 shows the input images and the results by RC [8] and DTV. SSD is not compared in this case, as it always fails although different settings were tried. As we could see, both results are very close to the ground truth. A visible artifact can be observed in the image recovered by RC, which is highlighted by the blue circle. The transformation by DTV are more smooth than that by RC. Since the ground truth transformation is by Gaussian perturbations, the proposed method is obvious more accurate. Under this severe intensity distortion, it shows that our method is more robust than RC, for recovering the image details in particular.

For quantitative comparisons, we evaluate RC and the proposed method with random distortions and random transformations. The root-mean-square error (RMSE) is used as the metric for error evaluation of both image intensities and transformations. A few Gaussian intensity fields (from 1 to 6) are added on the brain image in Fig. 1. The reference image without intensity distortions is used as ground-truth.

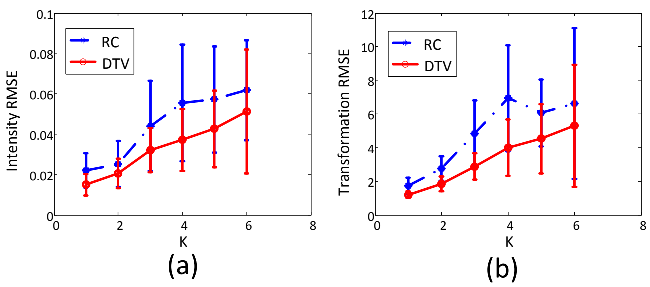

Fig. 2. Registration performance comparisons with random transformation perturbations and random intensity distortions. (a) Intensity RMSE on the brain image. (b) Transformation (non-rigid) RMSE on the brain image.

We run each setting 50 times and the results are plotted in Fig. 2. It can be observed that the proposed DTV is consistently better than RC, in terms of both intensity RMSE and transformation RMSE. The registration speed of our method is often faster than that of RC. The average speed for DTV is 6.5 seconds per registration on the brain image (216×180) while that of RC is 13.7 seconds per registration. This is because the DCT transform in RC costs $\mathcal{O}(N \log N)$ in each iteration [8], which is higher than the linear complexity in our method. All experiments are conducted on a desktop computer with Intel i7-3770 CPU with 12GB RAM.

4.2. Real-world medical image registration

We validate the performance of our method on real-world applications. The proposed method is compared with RC on two images from a iris video sequence [8] (shown in Fig. 3). The deformation between the source image and reference image is highly nonlinear. The intensity artifact in the source image makes this problem more challenging. The composition image without registration is shown in Fig. 3(c) using green and magenta colors. The vessels are blurry due to the misalignment. After registration, both RC and the proposed DTV provide accurate alignments on the vessels. However, the image registered by RC has been partially distorted due to the severe intensity variance.

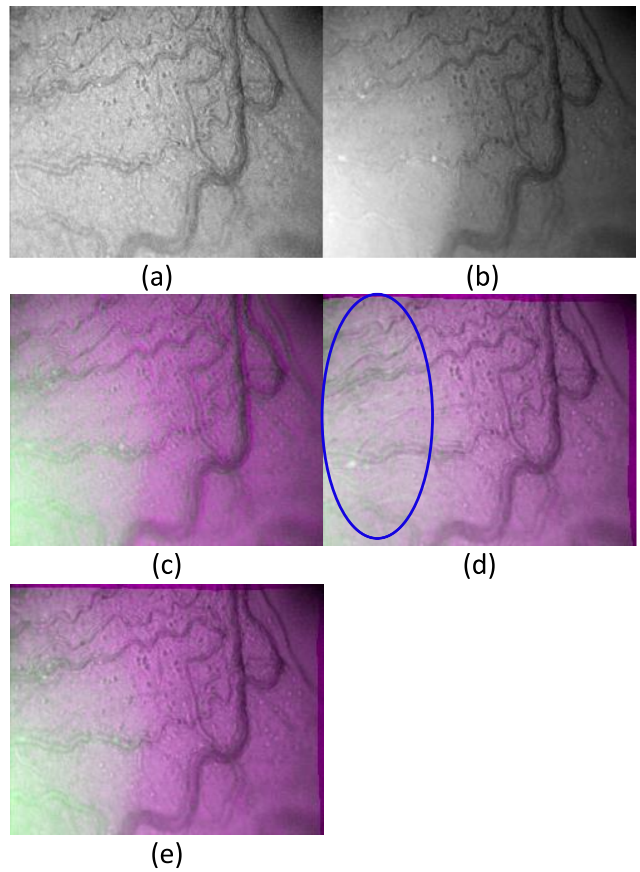

Fig. 3. Registration of two iris images [8]. (a) Reference image. (b) Source image. (c) The overlay before registration. (d) The overlay after registration by RC. (e) The overlay after registration by our method. Visible artifact is highlighted by the blue circle. Best viewed in $\times 2$ sized color pdf file.

Temporal and multimodal registration are performed on two retina images taken two years apart [18]. The reference image and source image are shown in Fig. 4 (a) and (b). These retina images are quite difficult to register by intensities, while most of existing methods are featured based [18, 19]. In order to avoid local minimum, we use affine transformation for preregistration and the result is shown in 4 (c). From the composition image in 4 (d), we could observe that there is still misalignment for the vessels at the bottom half. After non-rigid registration, the misalignment is eliminated.

There is a small error in the result by RC, while our result is very accurate from visual observation.

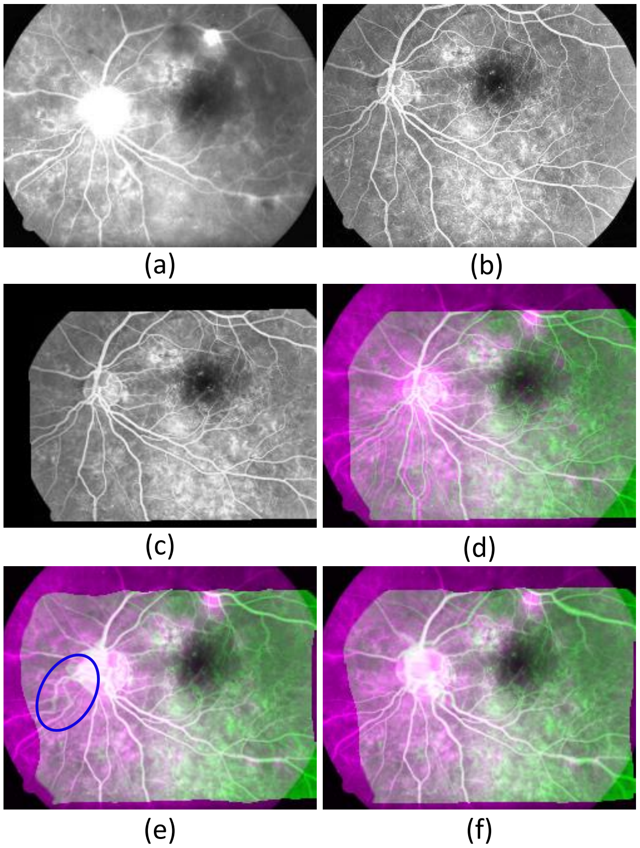

Fig. 4. Registration of two retina images [18]. (a) Reference image. (b) Source image. (c) The source image after affine preregistration. (d) The overlay before registration. (e) The overlay after registration by RC. (f) The overlay after registration by our method. Visual artifact is highlighted by the blue circle. Best viewed in $\times 2$ sized color pdf file.

5. CONCLUSION

In this paper, we have proposed a novel method based on DTV minimization for intensity based image registration. It is motivated by the stationarity of the image gradients under smooth intensity distortions. Intuitively, DTV is robust to a wide range of registration applications with intensity artifacts/outliers. To solve the DTV minimization problem, an efficient algorithm is presented based on the modified gradient descent method. In each iteration, the computation complexity of this algorithm is $\mathcal{O}(N)$, which is lower than $\mathcal{O}(N \log N)$ in RC [8], where N is the image size. Experiments on synthetic and real-world images demonstrate that our method are more robust, efficient and accurate than RC.

6. REFERENCES

[1] Aristeidis Sotiras, Christos Davatzikos, and Nikos Paragios, "Deformable medical image registration: A survey," *IEEE Transactions on Medical Imaging*, vol. 32, no. 7, pp. 1153–1190, 2013.

[2] Jan Modersitzki, *Numerical methods for image registration*, OUP Oxford, 2003.

[3] JB Maintz and Max A Viergever, "A survey of medical image registration," *Medical image analysis*, vol. 2, no. 1, pp. 1–36, 1998.

[4] Chen Chen, Yeqing Li, Wei Liu, and Junzhou Huang, "SIRF: Simultaneous image registration and fusion in a unified framework," *arXiv preprint arXiv:1411.5065*, 2014.

[5] Yigang Peng, Arvind Ganesh, John Wright, Wenli Xu, and Yi Ma, "RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2233–2246, 2012.

[6] Junzhou Huang, Xiaolei Huang, and Dimitris Metaxas, "Simultaneous image transformation and sparse representation recovery," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2008, pp. 1–8.

[7] Paul Viola and William M Wells III, "Alignment by maximization of mutual information," *International Journal of Computer Vision*, vol. 24, no. 2, pp. 137–154, 1997.

[8] Andriy Myronenko and Xubo Song, "Intensity-based image registration by minimizing residual complexity," *IEEE Transactions on Medical Imaging*, vol. 29, no. 11, pp. 1882–1891, 2010.

[9] Jan Modersitzki and Stefan Wirtz, "Combining homogenization and registration," in *Biomedical Image Registration*, pp. 257–263. 2006.

[10] Deanna Needell and Rachel Ward, "Stable image reconstruction using total variation minimization," *SIAM Journal on Imaging Sciences*, vol. 6, no. 2, pp. 1035–1058, 2013.

[11] Chen Chen, Yeqing Li, and Junzhou Huang, "Calibrationless parallel mri with joint total variation regularization," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 106–114. 2013.

[12] Chen Chen, Yeqing Li, Leon Axel, and Junzhou Huang, "Real time dynamic mri with dynamic total variation," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 138–145. 2014.

[13] Junzhou Huang, Shaoting Zhang, and Dimitris Metaxas, "Efficient mr image reconstruction for compressed mr imaging," *Medical Image Analysis*, vol. 15, no. 5, pp. 670–679, 2011.

[14] Noppadol Chumchob, "Vectorial total variation-based regularization for variational image registration," *IEEE Transactions on Image Processing*, vol. 22, no. 11, 2013.

[15] Claudia Frohn-Schauf, Stefan Henn, and Kristian Witsch, "Multigrid based total variation image registration," *Computing and Visualization in Science*, vol. 11, no. 2, pp. 101–113, 2008.

[16] Daniel Rueckert, Luke I Sonoda, Carmel Hayes, Derek LG Hill, Martin O Leach, and David J Hawkes, "Nonrigid registration using free-form deformations: application to breast MR images," *IEEE Transactions on Medical Imaging*, vol. 18, no. 8, pp. 712–721, 1999.

[17] Chris A Cocosco, Vasken Kollokian, Remi K-S Kwan, G Bruce Pike, and Alan C Evans, "Brainweb: Online interface to a 3d MRI simulated brain database," *NeuroImage*, 1997.

[18] Frédéric Zana and Jean-Claude Klein, "A multimodal registration algorithm of eye fundus images using vessels detection and hough transform," *IEEE Transactions on Medical Imaging*, vol. 18, no. 5, pp. 419–428, 1999.

[19] Yuanjie Zheng, Ebenezer Daniel, Allan A Hunter, Rui Xiao, Jianbin Gao, Hongsheng Li, Maureen G Maguire, David H Brainard, and James C Gee, "Landmark matching based retinal image alignment by enforcing sparsity in correspondence matrix," *Medical image analysis*, vol. 18, no. 6, pp. 903–913, 2014.