Rectified Registration Consistency for Image Registration Evaluation

Peng YE[†a)], *Nonmember*, Zhiyong ZHAO[†], *Member*, and Fang LIU[†], *Nonmember*

SUMMARY Registration consistency (RC) stands out as a widely-used automatic measure from existing image registration evaluation measures. However the original RC neglects the influence brought by the image intensity variation, leading to several problems. This letter proposes a rectified registration consistency, which takes both image intensity variation and geometrical transformation into consideration. Therefore the geometrical transformation is evaluated more by decreasing the influence of intensity variation. Experiments on real image pairs demonstrated the superiority of the proposed measure over the original RC.

key words: *registration consistency, rectified registration consistency, image registration evaluation, image processing*

1. Introduction

Image registration is a crucial step for many image analysis tasks like image fusion and change detection. Given the commercial and social impacts of lots of image registration applications, proper evaluation of image registration reliability and accuracy is of great importance. Though many image registration methods have been proposed, less attention has been paid to evaluating their performance. Lots of previous methods rely on visual inspection or root mean square error (RMSE) to assess the performance of image registration algorithms. However, visual inspection is not a satisfactory evaluation method, and RMSE needs operators to hand-choose control points, limiting its usage. Among existing image registration evaluation measures, registration consistency (RC) stands out as an automatic measure which does not need human involvement. It is widely used to evaluate remote sensing and medical-image registration methods [1]–[4]. However, though designed to reflect the geometrical difference between the original image and the transformed image, the original RC neglects the influence brought by image intensity variation, therefore leading to several problems. The two main problems are: 1. low distinguishability for images of narrow intensity variation range; 2. inability to indicate degree of wrong alignment across different images. The problems will be elaborated in Sect. 3.

After presenting the original RC and analyzing its deficiency, a rectified RC (rRC) is presented in this letter. The new measure rRC has better distinguishability with barely increased computation cost and could be used as a

qualitative indicator of displacement across different images.

2. Original Registration Consistency

RC measure is first proposed by Holden et al. [4] to evaluate the performance of intensity-based image registration methods in the absence of ground truth. Let $T_{I,R}$ denotes the transformation obtained by transforming the source image I over the reference image grid R. Similarly $T_{R,I}$ denotes the transformation obtained by transforming the reference image R over the input image grid I. For two dimensional images, RC [1] is formulated as Eq. (1):

$$RC = \left(\frac{1}{N}\right) \cdot \sum_{(x,y)\in(A_I \cap A_R)} \left\| (x,y) - T_{R,I} \circ T_{I,R}(x,y) \right\| \quad (1)$$

where the composition $T_{R,I} \circ T_{I,R}$ is a result of applying $T_{I,R}$ followed by $T_{R,I}$. $A_{I,R}$ is the overlap area of image I and R. A_I, A_R are their individual regions. N is the number of the pixels in the image that is being transformed. For better registration results, the difference between the original image and the transformed image is smaller, therefore generating smaller RC value. In an ideal registration case with ground truth parameters, value obtained from the RC equation should be equal to zero. As shown in Eq. (1), RC is automatic and easy to implement.

3. Deficiency of the Original RC

From Eq. (1) it could be seen that in the original RC there are two variables influence the RC value. One is the difference of the original image and the transformed image. This is brought by the error of geometrical transformations and is exactly the main designing aim of the original RC. The RC equation was supposed to reflect the extent of the transformation error. The other variable is the intensity variation of (x,y). For images of low intensity variation range, the value range of intensity difference is intrinsically smaller; therefore the RC is intrinsically low. For the same misalignment, images with wide intensity variation range would produce larger intensity difference than images with small intensity variation range, therefore yielding larger RC. Figure 1 shows three different images, their histogram, and their corresponding RC value (blue) and rRC (red) values. Each of the three images is geometrically transformed by a translation [n, n], n = 1, . . . , 25 to simulate a misaligned

image (x-axis in Fig. 1 g, Fig. 1 h, Fig. 1 i). Figure 2 b shows Fig. 1 c after a misalignment of [13,13]. With the misaligned image and the original image, RC is computed as Eq. (1). Blue lines in Fig. 1 g, Fig. 1 h, and Fig. 1 i show the change of the RC value with respect to the translation error. Unlike Fig. 1 c, the RC of Fig. 1 a changes rather slowly, meaning RC has low distinguishability. Even for large translational displacement like [25,25] (x-direction, y-direction), the RC of Fig. 1 a is still as small as 0.06036 (normalized intensity), which is 0.2724 for Fig. 1 c. Table 1 shows RC value of these images with different displacements. It could be seen that RC fails completely to indicate the degree of wrong alignment across different images. The wide intensity variation image Fig. 1 c has the largest RC for the same displacement and much better distinguishability between
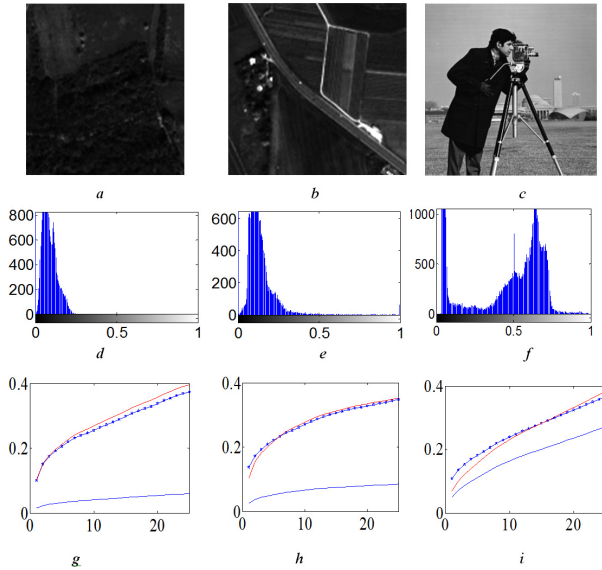
close displacements. Therefore, the original RC actually could only be used in narrow restricted situations. Although the intensity variation affects the RC value greatly, it is neglected in the original RC. To rectify this influence, there are two ways. One way is that histogram equalization is done for each image to obtain a wide intensity variation range before any RC value is computed. However, histogram equalization would change the original image data, which is not expected in many applications. Figure 2 a shows the histogram equalized image of Fig. 1 c. It could be seen that since the original image data has been changed, which would be inappropriate in many applications. The other way is to add a new item in Eq. (1) to decrease this negative impact.

4. Rectified Registration Consistency (rRC):

Compared with Eq. (1), a new item range is added in Eq. (2) to rectify the original RC. The new item *range* depicts the intensity variation range of majority pixels which is defined as in Fig. 2:

$$rRC = \frac{1}{range} \cdot \left(\frac{1}{N}\right) \cdot \sum_{(x,y)\in(A_I \cap A_R)} \left\| (x, y) - T_{R,I} \circ T_{I,R}(x, y) \right\|$$

(2)

To define the definite number of "majority pixels" in an image is a tricky problem. In this letter, to ensure the robustness of the new item *range*, several steps have been adopted: 1. "majority pixels" is defined to be certain number from 85% to 90% pixels of the whole image instead of a definite amount. The *range* is the largest intensity range of these majority pixels. 2. Intensities are sorted by their pixel numbers from the largest number to the smallest number. If the k-th intensity has N_k number of pixels, then $N_1 \geq N_k \geq N_{256}$. As in Fig. 3, the "majority pixels" starts from N_1 to N_k, covering pixels that have relatively larger numbers of quantity. The remaining pixels which are not included in the "majority pixels" have relatively smaller numbers and therefore

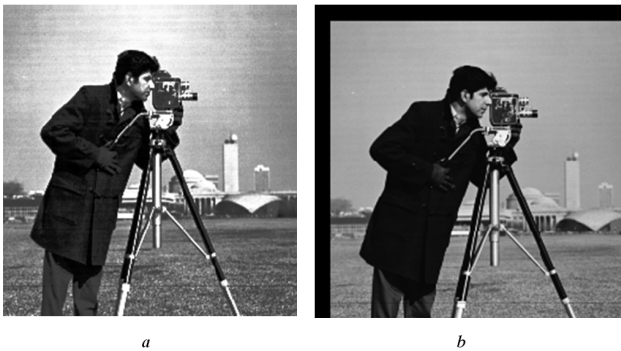
Fig. 1 Experimental results.

Fig. 2 (a) Histogram equalized image; (b) Misaligned image.

Table 1 RC of image a, b, c with different displacements.

Displacements	[1,1]	[3,3]	[13,13]
Image a	0.01565	0.02713	0.04518
Image b	0.02461	0.04331	0.07185
Image c	0.04934	0.08677	0.18800

The intensities in histogram are sorted by their pixel number from high to low. From the intensity with the highest number k=1 to the intensity with the lowest number k=256, denote the k-th intensity has N_k number of pixels, and initialize sum_pixels=0

    For k=1:256
        range_k: the max range between 1 to the k-th intensity;
        sum_pixels=sum(N_n; n=1,…,k);
        If sum_pixels<85%*all_pixels_of_the_whole_image
            range= range_k;
        elseif sum_pixels>90%*all_pixels_of_the_whole_image
            range= range_{k-1};
            break out;
        elseif range_k- range_{k-1}<5%* range_{k-1}
            range= range_k;
        else
            range= range_{k-1};
            break out;
    end
end

Fig. 3 Computation of the new item *range*.

are less likely to be "majority pixels". 3. The third step to ensure a robust *range* is the auto-stopping criteria. After "majority pixels" excels 85% pixels of the whole image, the remaining pixels have small number. Although these pixels are small in amount, they might greatly influence the intensity range. For example, pixels which have intensity close to 1 in Fig. 1 e and Fig. 1 f are small in number but greatly influence the intensity range. They are not supposed to be deemed as "majority pixels". After 85% of all pixels, a 5% increase threshold of *range* is to ensure that the *range* increases smoothly. Therefore an abrupt change of the *range* by those pixels of small number is not valued. If the *range* increases smoothly, then it stops if "majority pixels" excels 90% pixels of the whole image. Other items have the same meaning as in Eq. (1). In our formulation of Eq. (2), image intensity is scaled to 0~1. By adding the new item, for smaller intensity range, *range* is smaller, leading to bigger $1/range$, therefore rRC is much bigger than the original RC; for wide intensity range like *range* = 1, rRC is the same as RC. Therefore images with narrow intensity variation range are punished more than images with good intensity diversity. The influence brought by the intensity variation is balanced by the new item. Since the *range* is a constant value in each rRC computation, the ascending characteristics of original RC is well-kept in rRC. If the intensity scale is from 0 to 256, then rRC is multiplied by 256 and range has a value range from 0 to 256.

5. Experimental Results

Red lines in Fig. 1 g, Fig. 1 h, and Fig. 1 i show the rRC of Fig. 1 a, Fig. 1 b, Fig. 1 c. It could be seen that rRC lines have larger slopes than RC lines, meaning rRC has better distinguishability between close misalignments. Blue dot line shows the RC values of histogram equalized images. It could be seen that their results are quite similar to the rRC. Since after histogram equalization, the influence brought by intensity variation range is eliminated, therefore the histogram equalization method produces the ideal results. However the histogram equalization method would change the original image data as shown in Fig. 2 a and cannot be accepted in many applications. On the other hand, rRC could achieve good result without changing the image data. Therefore the proposed method is more preferred. Table 2 shows the rRC version of Table 1. It could be seen that rRC has comparable values across different images and can be used as a qualitative indicator of displacement across different images. The time costs of RC and rRC (including the time costs of histogram construction and computations of *range* and rRC) for Fig. 1 a, Fig. 1 b, and Fig. 1 c are 0.0196s, 0.0119s, 0.0123s, and 0.0205s, 0.0163s, 0.0192s respectively. The computation costs of rRC are barely

Table 2 rRC of image a, b, c with different displacements.

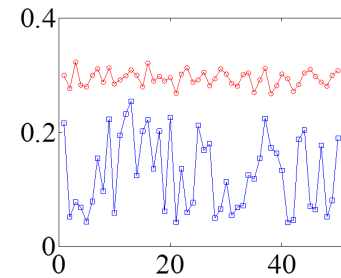| Displacements	[1,1]	[3,3]	[13,13]
Image a	0.10278	0.17814	0.29662
Image b	0.10331	0.18176	0.30154
Image c	0.06940	0.12206	0.26444

Fig. 4 RC and rRC of 50 image pairs.

increased compared with those of RC. Furthermore, due to the high computation power nowadays, it could be seen that the time costs of both methods are quite small and around 0.02s. Figure 4 shows the RC (blue line) and rRC (red line) value for 50 different image pairs under the translational displacement of [15, 15]. It could be seen that rRC is comparable between different images while RC is not.

6. Conclusion

This letter proposed an improved image registration evaluation measure. The new measure rectified the original registration consistency by decreasing the influence brought by image intensity variation, which is neglected in the original measure. The performance is improved with barely-increased computation cost, which is further proved with real image pair experiments. Though some thresholds are fixed in our implementation and proved to be effective in experiments, further work could be done to achieve full automatic threshold determination.

References

[1] S. Suri, Automatic Image to Image Registration for Multimodal Remote Sensing Images, Dissertation, 2010, Technische Universität München.
[2] H.M. Chen, M.K. Arora, and P.K. Varshney, "Mutual information-based image registration for remote sensing data," Int. J. Remote Sensing, vol.24, pp.3701–3706, Sept. 2003.
[3] H.M. Chen, P.K. Varshney, and M.K. Arora, "Performance of mutual information similarity measure for registration of multitemporal remote sensing images," IEEE Trans. Geosci. Remote Sens., vol.41, pp.2445–2454, Nov. 2003.
[4] M. Holden, D.L. Hill, E.R. Denton, J.M. Jarosz, T.C. Cox, T. Rohlfing, J. Goodey, and D.J. Hawkes, "Voxel similarity measures for 3-D serial MR brain image registration," IEEE Trans. Med. Imaging, vol.19, pp.94–102, Feb. 2000.