

# An Improved Mutual Information Similarity Measure for Registration of Multi-Modal Remote Sensing Images

Maha Shadaydeh and Tamas Sziranyi

Distributed Events Analysis Research Laboratory, Institute for Computer Science and Control,  
Hungarian Academy of Sciences. Budapest, Hungary.

## ABSTRACT

Registration of multi-modal remote sensing images is an essential and challenging task in different remote sensing applications such as image fusion and multi-temporal change detection. Mutual Information (MI) has shown to be successful similarity measure for multi-modal image registration applications, however it has some drawbacks. 1. MI surface is highly non-convex with many local maxima. 2. Spatial information is completely lost in the calculation of the joint intensity probability distribution. In this paper, we present an improved MI similarity measure based on a new concept in integrating other image features as well as spatial information in the estimation of the joint intensity histogram which is used as an estimate of the joint probability distribution. The proposed method is based on the idea that each pixel in the reference image is assigned a weight, then each bin in the joint histogram is calculated as the summations of the weights of the pixels corresponding to that bin. The weight given to each pixel in the reference image is an exponential function of the corresponding pixel values in a distance image and a normalized gradient image such that higher weights are given to points close to one or more selected key points as well as points with high normalized gradient values. The proposed method is in essence a kind of calculating similarity measure using irregular sampling where sampling frequency is higher in areas close to key-points or areas with higher gradients. We have compared the proposed method with the conventional MI and Normalized MI methods for registration of pairs of multi-temporal multi-modal remote sensing images. We observed that the proposed method produces considerably better registration function containing fewer erroneous maxima and leading to higher success rate.

**Keywords:** Multi-modal Image Registration, mutual information, joint intensity histogram.

## 1. INTRODUCTION

The increasing availability of satellite remote sensing image series has considerably increased research interest in multi-temporal change detection applications such as land cover use and environmental monitoring. Different change detection algorithms based on multi-modal multi-temporal images have been proposed so far (see [1] and references therein). However, the performance of change detection algorithms is affected by the accuracy of the registration of images taken at different times and possibly by different sensors. The aim of image registration is to find the geometric transformation which best aligns the position of features in the first image to the position of the corresponding features in the second Image. Key points based techniques attempt to align images based on the distance between selected key points<sup>2-3</sup>. Area-based techniques quantify the similarity between intensity values of images using similarity measures, such as mean square error and normalized cross-correlation. For images acquired from different sensors, statistical similarity measures based on minimizing the distance between the images' intensity probability distributions rather than intensity values are used. Mutual Information (MI), Normalized Mutual Information (NMI), Kullback-Leibler Divergence (KLD) and Cluster Reward Algorithm (CRA) are examples of such similarity measures.<sup>4</sup> Mutual information has shown to be successful similarity measure for multi-modal image registration applications in medical imaging<sup>5</sup> as well as remote sensing imagery.<sup>6</sup> MI measures the strength of the statistical relationship between two data-sets using their joint probability distribution. The maximum value of MI is expected to be achieved when the data-sets are geometrically aligned. In spite of its effectiveness, MI measure has some drawbacks. 1. MI surface is highly non-convex with many local maxima. 2. Spatial information is completely lost in the calculation of the global intensity joint distribution.

---

corresponding author: shadaydeh.maha@sztaki.hu

These MI-based registration problems have been approached so far either by combining the use of MI measure and other image features, or with appropriate joint probability distribution estimation methods. In [7] the authors proposed weighted mutual information method for medical image registration. The weights are given to the global MI and defined based on the ratio of the gradients of the two images for each possible geometric transformation. In [8] a weighted combination of local and global mutual information is optimized for medical image registration. In these two approaches however, weights are given to mutual information measures. When large estimation error exists in the calculation of mutual information these approaches may not improve performance. For remote sensing images in particular there are large regions of almost constant value (for example water surfaces); the mutual information of two local regions  $R_1$  and  $R_2$  of constant intensity values, say  $r_1$  and  $r_2$  respectively, is zero. Zero mutual information implies that  $R_1$  and  $R_2$  are different but they are actually not specially when  $r_1 = r_2$ . That is to say that mutual information contains large estimation error when the area of overlap in both images contain large regions of single intensity. This is due to the few nonzero bins in the joint histogram. This problem has been solved in [9] for change detection in remote sensing images using local similarity measure where we proposed to calculate local similarity measure mainly from pixels whose neighborhood variance has changed. For multi-source remote sensing imagery, the authors of [10] proposed general weighted MI method; In their method, a weight function for the joint probability distribution is defined using either Gaussian Mixture Model or kernel density estimation. In spite of its improved performance over the conventional MI, this method requires high computational load for the calculation of the weight function at each possible geometric transformation.

In this paper we present an improved MI similarity measure for image registration based on a new concept in integrating other image features as well as spatial information in the estimation of the joint intensity histogram which is used here as an estimate of the joint probability distribution. The proposed method is based on the idea that each pixel in the reference image is assigned a weight; each pixel contribute to the calculation of the joint histogram according to its weight only. That is, each bin in the joint histogram is calculated as the summations of the weights of the pixels corresponding to that bin and within the region of overlap. The weight given to each pixel in the reference image is chosen to be an exponential function of a distance image and normalized gradient image where higher weights are given to pixels with higher gradient value as well as to pixels closer to one or more of selected key-points. These key-points can be defined either arbitrarily by the user or automatically using for example Harris or Scale-Invariant Feature Transform. That is we calculate global similarity measure but from the pixels of set of local regions of interest, meanwhile we give higher weights to pixels with high gradient value which contain more relevant information about the images. In doing so we reduce the peaks in the joint histogram due to background or large homogenous areas. On the other hand, weights are calculated once for all geometric transformations, hence no considerable additional computational load is required.

In the following sections, after describing the image registration problem using MI and NMI similarity measures, we propose an improved MI and NMI similarity measures based on Weighted Joint Histogram (WJH), we refer to these methods hereafter as WJH-MI and WJH-NMI respectively. We then compare the performance of the proposed WJH-MI and WJH-NMI methods with the conventional MI and NMI methods for registration of multi-modal multi-temporal remote sensing image pairs.

## 2. IMAGE REGISTRATION USING MUTUAL INFORMATION

Suppose we have two images, a reference image  $X$  and floating image  $Y$ ; both images are represented by a matrix of discrete greyscale intensity values. For images represented by  $n$  bit/pixel, the greyscale intensity values are respectively  $\Omega_X = \{x_1, x_2, x_3, \dots, x_N\}$  and  $\Omega_Y = \{y_1, y_2, y_3, \dots, y_N\}$  where  $N = 2^n$ . Mutual Information  $I(X, Y)$  of the two images  $X$  and  $Y$  is defined as

$$I(X, Y) = H(X) + H(Y) - H(X, Y) \quad (1)$$

where  $H(X)$  is the entropy of the reference image,  $H(Y)$  is the entropy of the floating image and  $H(X, Y)$  is the joint entropy of the two images, all entropies are calculated over the region of overlap. The marginal and joint entropies are defined as

$$I(X) = \sum_x -P_X(x) \log P_X(x) \quad (2)$$

$$I(Y) = \sum_x -P_Y(y) \log P_Y(y) \quad (3)$$

$$I(X, Y) = \sum_{x,y} -P_{X,Y}(x, y) \log P_{X,Y}(x, y) \quad (4)$$

where the joint probability distribution  $P_{X,Y}(x, y) = Pr(X_s = x, Y_s = y)$ , for  $x \in \Omega_X, y \in \Omega_Y$  and  $(X_s, Y_s)$  is the corresponding pixel pair contained in the region of overlap; and marginal distributions are calculated by

$$P_X(x) = \sum_y P_{X,Y}(x, y) \quad (5)$$

$$P_Y(y) = \sum_x P_{X,Y}(x, y). \quad (6)$$

Mutual information measures the distance between joint distribution  $P_{X,Y}(x, y)$  and the distribution in case of complete independence  $P_X(x) \cdot P_Y(y)$ . The two images  $X$  and  $Y$  shall be registered when mutual information measure  $I(X, Y)$  is maximal. MI has shown to be sensitive to the amount of overlap between the images. To overcome this problem the more robust Normalized Mutual Information<sup>11</sup> (NMI) measure was introduced as

$$I(X, Y) = \frac{H(X) + H(Y)}{H(X, Y)}. \quad (7)$$

The joint and marginal distributions can be estimated using joint and marginal histograms  $h(x, y)$ ,  $h(x)$ ,  $h(y)$  respectively. The joint histogram  $h(x, y)$  of the two images  $X$  and  $Y$  over their region of overlap can be estimated by counting the number of times the intensity pair  $(x, y)$  occurs in corresponding pixel pairs  $(X_s, Y_s)$ . Normalizing the joint histogram gives an estimate of the joint probability distribution  $P_{X,Y}(x, y)$ , i.e.

$$P_{X,Y}(x, y) = \frac{h(x, y)}{\sum_{x,y} h(x, y)} \quad (8)$$

Accordingly, estimation of the MI and NMI similarity measures requires only the calculation of the joint histogram. In the following section we propose new method for joint histogram estimation.

### 3. WEIGHTED JOINT HISTOGRAM BASED MUTUAL INFORMATION ESTIMATION METHOD (WJH-MI)

Estimation of the joint histogram  $h(x, y)$  of two images  $X$  and  $Y$  is usually done by calculating the number of times the intensity pair  $(x, y)$  occurs. This means that all pixels contribute similarly to histogram calculation. Here we propose to give a weight  $w(s)$  to each pixel  $s$  in the reference image; and each pixel contribute to the calculation of the joint histogram according to its weight only. That is, each bin in the joint intensity histogram is calculated as the summation of the weights of the pixels with intensity values corresponding to that bin. This is in essence similar to calculating similarity measure using irregular sampling.

The weight  $w(s)$  given to each pixel  $s$  in the reference image is chosen to be an exponential function of a distance image  $d(s)$  and normalized gradient image  $g(s)$ . That is

$$w(s) = \exp(-k d(s)/g(s)) \quad (9)$$

where  $k$  is a positive constant. The normalized gradient image  $g(s)$  is calculated using the Laplacian of Gaussian (LoG) filter of variance  $\sigma$ . To define the distance image, let  $\{c_m; m = 1, 2, \dots, M\}$  be a set of selected key-points defined as the centers of  $M$  regions of interest in the reference image  $X$ ; and let  $d_m(s)$  denotes the Euclidean distance of pixel  $s$  from selected key-point  $c_m$ . The distance image is then calculated by

$$d(s) = \begin{cases} \min(d_1(s), \dots, d_M(s)); & M \geq 1 \\ 1, & M = 0 \end{cases} \quad (10)$$

These  $M$  key-points can be defined either arbitrarily by the user or automatically using for example Harris or Scale-Invariant Feature Transform. For  $M = 0$ , the weight  $w(s)$  becomes a function of the gradient image only.

The constant  $k$  is used to control the slope of the exponential function and can be set to fixed value proportional to image resolution. For  $k = 0$ , the WJH-MI reduces to the conventional MI method.

Note that the use of this weight function allows us to calculate the joint histogram from pixels of different local regions of interest before calculating the MI similarity measure, hence reducing estimation error induced by the calculation of local MI from limited number of samples. That is to say that we calculate global MI based on irregular sampling where higher sampling frequency is used in selected local regions or areas with higher gradients. These local regions are not defined by specific neighborhood window size and shape but according to the distance image. The advantages of the proposed methods can be summarized in the following points:

- Reduction of the peaks in the joint histogram due to background or large homogenous areas by giving more weights to pixels with high gradient value which contain more relevant information for registration.
- Integration of spatial information through the distance image where points closer to the center of one of the local regions of interest are given higher weights in the calculation of joint histogram.
- The gradient and distance images depend on the reference image only and are calculated once for all geometric transformations of the floating image; hence no considerable additional computational load is required.
- Extension to integration of image features other than gradient image and/or different distance definition is straight forward.

#### 4. EXPERIMENTAL RESULTS AND DISCUSSION

In the first registration experiment we used two scanned images of Szada area in Hungary with the spatial resolution 1.5m/pixel. The images are co-registered and provided by the Hungarian Institute of Geodesy Cartography and Remote Sensing. As shown in Fig. 1, the 1984 image is a greyscale image while the 2007 image has been originally scanned in digital form (Nikon D3X, with AF-S Nikkor 50 mm 1.4G lens) in RGB color space; only the 'R' component is used in this experiment. The registration of these two images is challenging since the 1984 image has a degraded quality and several major changes appear between 1984 and 2007.

To evaluate the performance of the proposed algorithm on known transformation bases, the 1984 image is deliberately translated and rotated. The horizontal and vertical translations are resp.  $T_h = 63$  and  $T_v = 39$ , the rotation angle is  $\alpha = 4^\circ$ . The task now is to register the rotated and translated greyscale image of 1984 to the 'R' component of the 2007 image. The 1984/2007 image is considered as reference/floating image respectively. Three key-points were selected arbitrary in this experiment. The weight image is shown in Fig. 1 (c). The normalized gradient image was generated using LoG filter of variance  $\sigma = 1$  and filter length  $L = 11$ ; We searched the maximum value for WJH-MI and WJH-NMI, MI, and NMI similarity measures by iterating within the search space defined by horizontal translation range  $T_h = 20 : 1 : 70$  pixels, vertical translation  $T_v = 20 : 1 : 70$  pixels, and rotation angle  $\alpha = -7^\circ : 1 : -1^\circ$ . This search space is chosen in the neighborhood of the true transformation parameters to avoid possible local maxima. Transformation parameters  $\hat{T}_h$ ,  $\hat{T}_v$  and  $\hat{\alpha}$  are found as the position of the peak of the convergence surface. Profiles of the similarity measure surface on the horizontal and vertical planes are shown in Fig. 2. Table 1 shows the registration errors for WJH-MI, WJH-NMI, MI and NMI methods. Registration errors are defined as the difference between the true and estimated parameters,  $\Delta_h = T_h - \hat{T}_h$  (pixels),  $\Delta_v = T_v - \hat{T}_v$  (pixels), and  $\Delta_\alpha = \alpha - \hat{\alpha}$  (degree). These results show that the proposed WJH-MI and WJH-NMI methods result in similarity measure surfaces with more distinguished peaks and considerably less registration error.

In the second experiment, we used the data set shown in Fig. 3 (a) and (b) for Szada area in Hungary. The first image is  $912 \times 969$  pixels from year 2000. The second image is  $1100 \times 1000$  pixels from year 2007. Both images have 1.5m/pixel resolution. Some clear changes in the forest can be seen between these two dates. In this experiment, three local regions of interest were selected arbitrary but far from changed areas, weight map is shown in Fig. 3 (c). Since transformation parameters are unknown we adopted a coarse to fine strategy based on Gaussian pyramid starting from lower resolution. Results at certain resolution are then used to set approximate search space for the next higher resolution level. The search space at top resolution is defined by horizontal



Figure 1. Images used in Experiment 1. (a) Szada area, 2007 optical image; (b) Szada area, 1984 optical image rotated and translated; and (c) weight map.

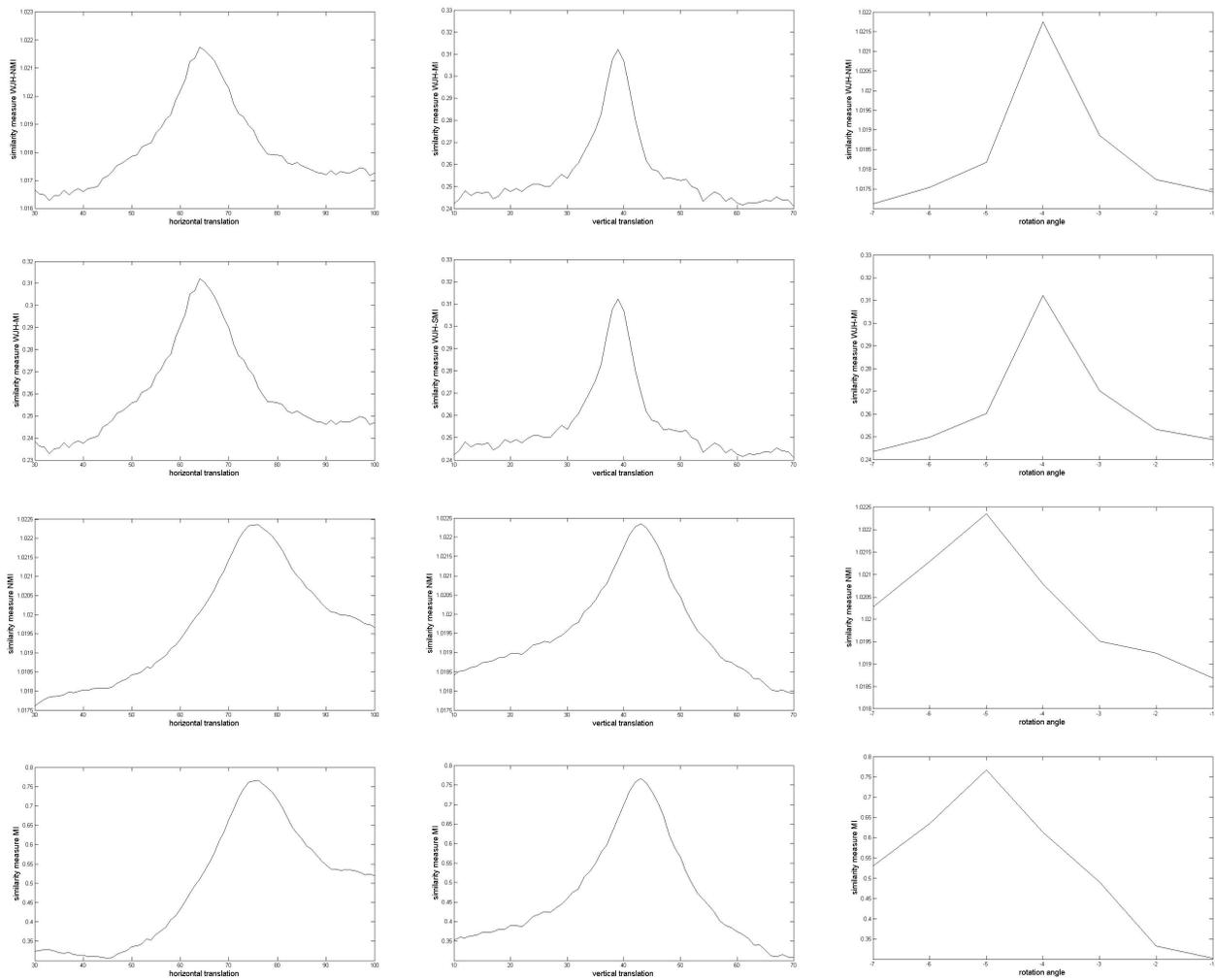


Figure 2. Similarity measures for Experiment 1. Rows 1-4: WJH-NMI, WJH-MI, NMI, MI. Columns 1-3: horizontal translation (pixels), vertical translation (pixels), rotation angle (degree)

Table 1. Registration errors for Experiment 1; horizontal translation error  $\Delta_h = T_h - \hat{T}_h$ , vertical translation error  $\Delta_v = T_v - \hat{T}_v$ , and rotation angle error  $\Delta_\alpha = \alpha - \hat{\alpha}$ .

parameter	MI	NMI	WJH-MI	WJH-NMI
$\Delta_h$ (pixels)	4	4	0	0
$\Delta_v$ (pixels)	13	13	1	1
$\Delta_\alpha$ (degree)	1	1	0	0

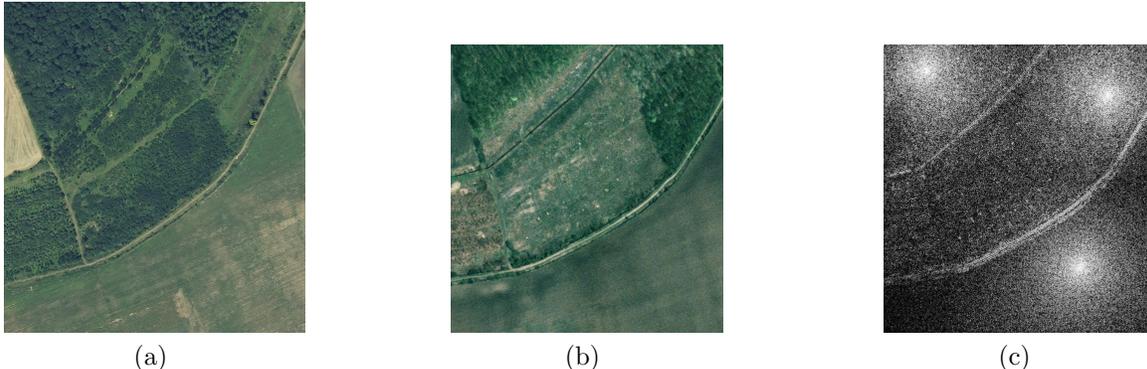


Figure 3. Images of Szada area, Hungary used in Experiment 2. (a) 2007 image ; (b) 2000 image; (c) weight map.

translation range  $T_h = 100 : 1 : 230$  pixels, vertical translation  $T_v = 20 : 1 : 150$  pixels, and rotation angle  $-7^\circ : 1 : -4^\circ$ . Registered images are shown in Fig. 4 (a) and (b). It can be clearly seen that the WJH-NMI has considerably better matching accuracy. Contour plots of the convergence surface are shown in Fig. 5 for top resolution level. These plots show that the WJH-NMI method has more distinguished single peak surface while the conventional NMI has several local maxima.

## 5. CONCLUSION

In this paper we have proposed an improved MI similarity measure based on the use of weighted joint histogram. The weights are given to each pixel in the reference image and defined according to the values of the normalized gradient of this pixel and to its Euclidean distance from the centers of selected regions of interest. The proposed method is in essence a kind of calculating global similarity measure using irregular sampling where sampling frequency is higher in areas close to key-points or areas with higher gradients.

We have compared the proposed method with the conventional MI and Normalized MI methods for registration of pairs of multi-modal remote sensing images taken at two different dates. We observed that the proposed method produces better similarity measure surfaces with more distinguished peaks and less registration error. Possibility of multi-scale analysis based on different values of the LoG filter parameters still needs further investigation.

## ACKNOWLEDGMENTS

This work has been supported by ESA-PECS DUSIREF project.

## REFERENCES

- [1] C. Benedek, M. Shadaydeh, Z. Kato, T. Sziranyi, and J. Zerubia, “Multilayer markov random field models for change detection in optical remote sensing images,” *ISPRS Journal of Photogrammetry and Remote Sensing* **107**, pp. 22–37, 2015.
- [2] S. Dawn, V. Saxena, and B. Sharma, “Remote sensing image registration techniques: A survey,” *Image and Signal Processing, Lecture Notes in Computer Science* **6134**, pp. 103–112, 2010.

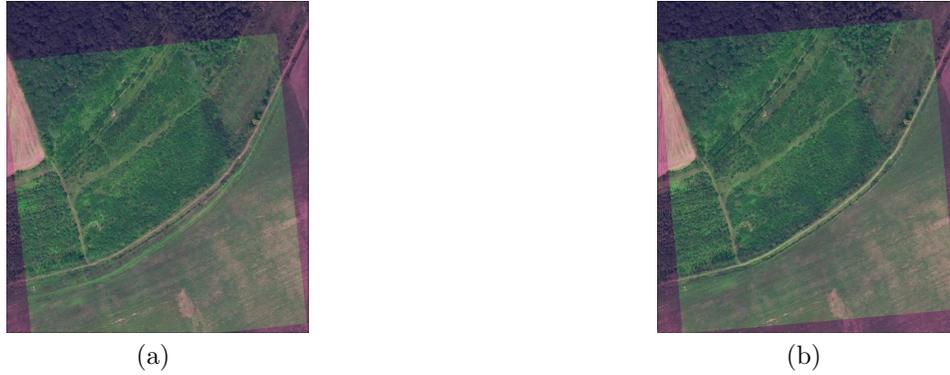


Figure 4. Registration results of Experiment 2. (a) registered image using NMI; (b) registered image using WJH-NMI.

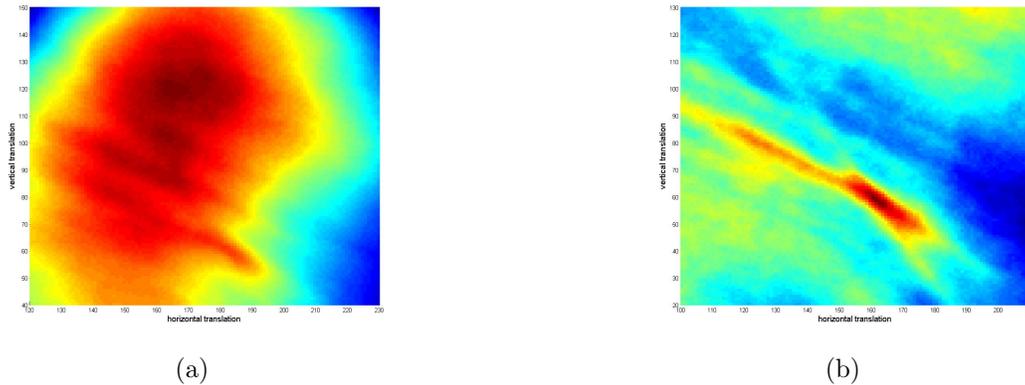


Figure 5. Contour plot of similarity measure convergence surface in Experiment 2. (a) using NMI; (b) using WJH-NMI.

- [3] B. Zitova and J. Flusser, “Image registration methods: A survey,” *Image and Vision Computing* **21**, pp. 977–1000, 2003.
- [4] J. Inglada and A. Giros, “On the possibility of automatic multisensor image registration,” *IEEE Trans. Geosci. Remote Sens.* **42**, pp. 2104–2120, 2004.
- [5] J. P. Pluim, J. B. Maintz, and M. A. Viergever, “Mutual-information-based registration of medical images,” *IEEE Transactions on Medical Imaging* **22**, pp. 986–1004, 2003.
- [6] H. M. Chen, P. K. Varshney, and M. K. Arora, “Performance of mutual information similarity measure for registration of multi-temporal remote sensing images,” *IEEE Trans. Geosci. Remote Sens.* , pp. 986–1004, 2003.
- [7] J. P. Pluim, J. B. Maintz, and M. A. Viergever, “Image registration by maximization of combined mutual information and gradient information,” *IEEE Transactions on Medical Imaging* **19 (8)**, 2000.
- [8] K. Wilkie and E. Vrscay, “Mutual information-based methods to improve local region-of-interest image registration,” *Lecture Notes in Computer Science* **3656**, pp. 63–72, 2005.
- [9] M. Shadaydeh and T. Sziranyi, “An improved local similarity measure estimation for change detection in remote sensing images,” *2014 IEEE International Conference on Aerospace Electronics and Remote Sensing Technology* , pp. 234–238, 2014.
- [10] E. Parmehr, C. Zhang, and C. Fraser, “Automatic registration of multi-source data using mutual information,” *Proceedings of ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* , pp. 303–308, Sept. 2012.
- [11] C. Studholme, D. L. G. Hill, and D. J. Hawkes, “An overlap invariant entropy measure of 3d medical image alignment,” *Pattern Recognition* **32 (1)**, pp. 71–86, 1999.