Similarity Measures with Spatial Dependency for Brain Image Registration

Qolamreza R. Razlighi1 , Arno Klein1,

1 Molecular Imaging and Neuropathology Division, Columbia University/NYSPI,

New York, NY 10032, USA

[qr2108@columbia.edu](mailto:qr2108,%20ak2104%7d@columbia.edu), arno@binarybottle.com

**Abstract.** The quality of intensity-based image registration methods hinges on the use of an effective similarity measure. Information theoretic similarity measures, such as mutual information, are some of the most widely used and accurate measures for human brain image registration. However, classical mutual information does not take into account spatial information in the images being registered. The lack of spatial information in the computation of mutual information in the image registration has been discovered early and the attempted to overcome this issue has been started since its first introduction. Here we give a comprehensive review of attempts to incorporate spatial dependency into the computation of classical mutual information. We also provide a solution to the problem of translational misregistration in a recently introduced spatially dependent similarity measure and demonstrate that the resulting similarity measure does not have the issue associated with translational misregistration.

**Keywords:** Mutual Information, Image Registration, Spatial Information.

1 Introduction

One may consider image registration as the most fundamental concern in human brain image analysis, because it is used in an attempt to establish correspondences across brains, which is the basis for all subsequent analyses and interpretations. Registration methods are predominantly based on a global similarity measure (SM), since they do not require the time-consuming and subjective manual extraction of landmarks or other features. Information theoretic-based SMs are among the most accurate, approaching that of landmark-based gold standard registration methods.

An information theoretic measure was first introduced as an image registration similarity measure by Viola [1] and Maes [2] in 1995. Mutual information (*MI*) based on Shannon’s definition of entropy [3] has been demonstrated to be one of the most effective measures of image similarity. Unlike measures based on correlation of image intensities or differences of intensities, MI does not assume a linear relationship among the intensities in the images, which makes it well suited for multimodal image registration. Despite generally promising results of MI, it can result in misregistration [4-6]. Research into how to improve MI is ongoing and has included work on multiresolution methods, invariance with respect to overlap, and “higher-order” mutual information. Our focus in this study is on the latter.

The calculation of MI is typically based on a global joint histogram, expressing the joint intensity probabilities over the entire image. The underlying assumption is that there is no statistical relationship among neighboring pixels/voxels over the whole image domain. Contrast this with statistical metrics such as the correlation coefficient, which measures the relationship among neighboring pixels/voxels. This shortcoming was recognized soon after MI was introduced. For instance, Studholme *et al.* [4] attempted to incorporate spatial information into the computation of *MI* in 1996, a year after its introduction for image registration. Since then researchers have tried to find a method in which the inter-pixel/voxel dependency is taken into consideration rather than assume independence among pixels/voxels.

Conceptually, classical *MI* measures image similarity based on a single pixel/voxel correspondence. However, spatially dependent *MI* is based on the correspondence of multiple pixels/voxels. It is more likely that two random variables (representing pixel/voxel intensities) become similar/dissimilar due to noise, distortion or difference in modality than two sets of pixels/voxels in spatial order. Thus, spatially dependent *MI* should be more robust to image degradation and consequently more accurate in image registration. This advantage is the main reason behind all the recent attempts to incorporate spatial dependency into the computation of *MI*, but these attempts must overcome the curse of dimensionality when computing similarity across spatially dependent terms. Markovianity is an effective tool to confront dimensionality problems. Spatial mutual information (SMI) [5] is the only systematic method of computing spatially dependent MI (as defined by Shannon) while confronting dimensionality problems by applying the Markovianity constraint. The closest measure related to this work is the second-order *MI* introduced by Rueckert *et al.* [6]. Most of the spatially dependent SMs are either an *ad hoc* combination of *MI* with different image features to capture the image spatial information or a heuristic use of different definitions of entropy instead of the conventional Shannon entropy. The next section gives a comprehensive overview of existing approaches that incorporate spatial dependency into a similarity measure. Section 3 details the recently introduced similarity measure, spatial mutual information (SMI) [5], and discusses the advantages and primary drawback of this new SM. A new implementation is proposed to overcome its drawback related to translational misregistration and results are shown in section 4. Section 5 concludes the paper.

2 Existing Approaches

In [4], the source image is segmented by thresholding and labeling connected components into four regions (air, fat, bone, and abdominal tissue). Then the mutual information of the intensity images and the labeled images is considered to be a new SM. However, the results reported do not show a significant improvement. So-called Jumarie entropy is used in [7] to define a new SM. The suggested expression for joint entropy in [7] resembles the normalized entropy of the absolute difference image. That is, for **D**=(**X**-**Y**), *H*(**D**)/*Dij* is defined as the joint entropy, where *Dij* denotes the intensity of the difference image **D** at the coordinate (*i*,*j*). However, the concept of incorporating spatial information is not taken into consideration.

*Second Order Mutual Information* (*SOMI*) involves the use of the co-occurrence matrix or Aura matrix to estimate the four-dimensional joint pdf of an image pair [6]. This measure is given by

|  |  |  |
| --- | --- | --- |
|  |  | (1) |

where denotes a finite discrete label set, *pX*(*x*,*x’*) the probability that *x* and *x’* are adjacent in image **X**, *pY*(*y*,*y’*) the probability that *y* and *y’* are adjacent in image **Y**, *pX,Y*(*x,x’,y,y’*) the joint probability that (*x*,*x’*) are adjacent in image **X** and (*y*,*y’*) are adjacent in image **Y**, and (*x*,*y*) denotes the corresponding pixels/voxels in the image pair. Unfortunately, the four-dimensional joint histogram for estimating *pX,Y*(*x,x’,y,y’*) becomes sparse in practice since there is insufficient data in a typical medical image to adequately fill all its histogram bins. In [6], Rueckert addressed this issue by reducing the number of the discrete label set to 16. However, such reduction has an adverse effect on the effectiveness of *SOMI* as a similarity measure. This drawback is thoroughly studied by Gao in [8] for the classical *MI*, and also in [9] for the *SOMI*.

*Gradient Mutual Information* (*GMI*) is a spatial similarity measure that is formed by combining *MI* and a gradient measure [10]. *GMI* is formulated as follows:

|  |  |  |
| --- | --- | --- |
|  |  | (2) |

|  |
| --- |
|  |

where *G*(*X*,*Y*) is the gradient part of the SM contributing to the spatial information, |*∇x*(*σ*)| denotes the magnitude of the gradient vector of image **X** at point *x* with the scale of *σ*, and |*∇y*(*σ*)| is the magnitude of the gradient vector of image **Y** at point *y* with the scale of *σ*, and *I*(*X*,*Y*) is the traditional mutual information. The registration outcome when using this similarity measure has shown some improvement for multimodal affine registration in comparison to classical MI [10]. Another attempt has been the use of gradient intensity images [11]. However, the mutual information of the gradient intensity of two images provides a measure of directional similarities regardless of their relative size and position and becomes maximal when images are rotationally aligned. This approach is proposed to accelerate the convergence of the optimization process rather than to capture image spatial mutual information. The maximum distance gradient magnitude for capturing image spatial information is another approach discussed in [12]. In this approach, MI is obtained from a 4D joint histogram of two images and corresponding maximum distance gradient magnitudes. However, the lack of available data samples to fill the histogram bins still remains an unsolved problem.

Shen *et al.* [13] have developed a similarity measure that determines image similarities based on an attribute vector for each pixel/voxel incorporating gray matter, white matter, and cerebral spinal fluid interfaces. The attribute vector is derived from the pixel/voxel's edge type and geometric moment invariants calculated from voxel intensities in a spherical neighborhood. This SM is specifically devised for inter-modal, inter-subject MR brain image registration requiring the segmentation of different tissue types.

Mutual information of regions was introduced by Russakoff *et al.* [14]. In this approach, a vector of intensity values is created for every pixel/voxel in the image. The components of these vectors are the respective intensity values of the neighboring pixels/voxels. These vectors form a matrix in which the rows are assumed to be normally distributed. Thus, the entropy is directly computed from the determinant of the covariance matrix. The entropy of the multivariate normal distribution is given by

|  |  |  |
| --- | --- | --- |
|  |  | (3) |

where **Z** is the set of multivariate normal random variables, |**Z|** is the determinant of the covariance matrix of **Z**, and *d* denotes the dimensionality of **Z**. It is worth pointing out that the rows of the composed matrix have the same histogram. In fact, they are the circular shift of each other with the difference being only on the border pixels. This makes the above composed matrix almost *circulant*, and consequently the rows carry almost the same information. On the other hand, the row-wise enumeration of the pixels of an image into one row actually destroys the 2D spatial information and consequently the information gets lost in the computation. Another similar attempt is reported in [15] by creating the same matrix with the intensity values of the image pixel/voxel and the average value of the neighboring pixels/voxels. Again by assuming a normal distribution for each row of the matrix, its entropy can be estimated by its covariance value. However, the effect of such a substandard assumption is still unknown.

In [16] the images’ joint histogram is computed along the corresponding points on random lines instead of the normal grid pixel/voxel pattern of the images. Since the random lines can be located in any orientation, the spatial information gets captured to some degree. However, the gain of this method depends strictly on the number of points on the line. Even with a minimal number (2 points), a 4D histogram is required to be computed which leads to the same dimensionality problem as discussed earlier.

There have also been a number of attempts to compute image spatial mutual information based on the multi-feature mutual information [17-20]. In these approaches, different features of an image are used to capture the image spatial information instead of dealing with neighboring pixels/voxels. These approaches transfer the dimensionality problem into various features of an image, yet they do not solve it. One solution is to estimate the image feature’s probability distribution as normal and obtain the joint entropy directly from the covariance matrix via equation (3). A reliable estimation of normal distribution using the covariance matrix requires much fewer samples, yet the error in estimating an image feature’s distribution by normal distribution is unknown. The feature extraction is also another issue when using these methods since this is done in an *ad hoc* manner and there is no systematic way of obtaining the most efficient features. For instance, image gradient is considered to be the other image feature in [17], Gaussian scale space expansion components are used as image features in [18], and in [20] a feature space called *Edgness* is used. Chappelow *et al.* [19] have developed an optimization algorithm to obtain the top five performing features of an image from its different gradient, first-order, or second-order statistical features. Despite the high computational complexity of these methods, the dimensionality problem still remains for computing 10-dimensional joint histograms.

A new similarity measure definition, α-MI, is introduced by Hero *et al.* [21, 22], which has been used on 2D data [23, 24]. As described in [25], a descent direction for Euclidean minimal spanning tree is derived, and some results for rigid registration are given. In [26], α-MI is applied to 3D tagged MR sequences of the heart. They have only used two features of intensity image from two different angles. In [27], α-MI is used for registration of MR images of the cervix; 15 different features of MR images (7 local structures at two scales of *σ* =1 and *σ* =2) including the image intensity itself have been selected. Then PCA is used to reduce the number of features because of the high computational complexity. The reported results indicate that there is not a significant improvement over classical MI despite the higher computational complexity. Furthermore, the algorithm requires various manual interventions to make the proposed α-MI a suitable SM in practice. In general, this prevents the use of α-MI as a SM for an automatic registration process.

The definition of Quantitative-Qualitative MI (QMI) is introduced in [28] and further developed in [29]. QMI is created by adding a utility coefficient into the formulation of the classical *MI* as follows:

|  |  |  |
| --- | --- | --- |
|  |  | (4) |

where *pX*(*x*), *pY*(*y*) are the intensity pdf of the image obtained from image histogram and *pX,Y*(*x,y*) is the joint histogram of the image **X** and **Y**. The coefficient *U*(*x*,*y*) is supposed to incorporate the spatial information of the image into this measure. In essence, *U*(*x*,*y*) is an *ad hoc* combination of the *Saliency* measure [30] and image gradient. Since an optimization process is required for every pixel/voxel of the image to compute this utility measure, its computational complexity is quite high. Even though the reported results reflect some improvements over classical MI, they may not be practically useful due to the considerable increase in computational complexity.



Fig. 1. Structure of neighboring pixels for a pair of QMRF of size 2.

3 Spatial Mutual Information

As stated earlier, a systematic attempt to incorporate the image pixel/voxel spatial dependency into the computation of image mutual information is done in [6]. However the dimensionality problem prevents making a significant improvement in computation of image spatial information. Markov processes were proven to be a useful tool in dealing with high-dimensional problems. Markov random fields (MRF), in particular are well-known in mathematical image analysis; for instance, MRF has established a strong record in image modeling in recent years [30,31]. The only attempt to compute image spatial information under the MRF constraint, which is a more relaxed constraint than the independency constraint, is due to [5]. In this approach a causal MRF model, called quadrilateral Markov random field (QMRF), is introduced and used to compute image spatial information under the general definition of Shannon entropy. The spatial entropy (SE) given by [5] is further simplified for homogenous anisotropic QMRF in [33] for neighboring structures of order 2 in the following equation:

(5)

where *m*x*n* is the image size, *H*(*X,Xu*) denotes the joint entropy of the pixel with its upper neighbor, *H*(*X,Xl*) the joint entropy of the pixel with its left neighbor, *H*(*Xu Xl*) the joint entropy of the left and upper neighbors, and *H*(*Xu ,Xr*) the joint entropy of the right and upper neighbors; see Fig. 1. Consequently, computation of spatial mutual information (*SMI*) is made possible from the spatial joint entropy [5]

(6)

where *mxn* is again the image size which is the same in both images, *H*(*X,Y*) is the joint entropy of the site *X* in image **X** with the corresponding site in image **Y**, *H*(*Y,Xu*) is the joint entropy of the site *Y* in image **Y** with the upper neighbor of its corresponding site in image **X**, *H*(*Yl,Xu*) is the joint entropy of the left neighbor of site *Y* in image **Y** with the upper neighbor of the corresponding site *X* in image **X**, *H*(*Yr,Xu*) is the joint entropy of the right neighbor of site *Y* in image **Y** with the upper neighbor of the corresponding site *X* in image **X**, *H*(*Xr,Y*) is the joint entropy of the site *Y* in image **Y** with the right neighbor of its corresponding site *X* in image **X**, and finally *H*(*Xl,Y*) is the joint entropy of the site *Y* in image **Y** with the left neighbor of its corresponding site *X* in image **X**.

One of the main concerns about equations (5) and (6) is that they do not contain all the possible joint entropy of the cliques in the first order MRF. For instance, it contains the clique but not . This is due to the type of scan used in the chain rule of the mathematical derivation of the in [5]. However, from an implementation standpoint the histograms of and are the same. Therefore their joint entropies should also be the same.

Next we discuss the drawback associated with SMI when it is used as a SM with spatial dependency, and present our solution.



(a) (b)

Fig. 2. The drawback of the original SMI as a similarity measure is demonstrated by its application in the case of translational misregistration.

4 New Implementation Method

The formulation of *SMI* in terms of two-dimensional joint entropies is made possible in equation (6) by a conditional independency assumption given in the following equation:

(7)

where *a*⊥*b*/*c* indicates that *a* and *b* are independent given *c*. This conditional independence assumption has an artifact in the case of translational misregistration which makes the SMI value negative for cases at or around pixel/voxel misalignments (see Fig. 2a). The theoretical reason behind this drawback is that when there is such a misalignment between **X** and **Y** images, the assumptions in equation (7) are not valid anymore. This problem is addressed in [5] by using the absolute value of the SMI. A better solution is to extend the definition of SMI to 3D brain MR images, as in [34].

(8)

where *SMIt* is the SM computed on the transverse slices, *SMIs* the SM computed on the sagittal slices and *SMIc* the SM computed on the coronal slices. The computed *SMIt*, *SMIs*, and *SMIc* are different due to the fact that SMI takes into account the in-plane spatial dependency of the image. For the case of translational misregistration, only two of the components of equation (8) become negative at or around the aforementioned misalignment points, which ensures the final product is always positive, as seen in Fig. 2b. Multiplying three SMIs obtained from three planes is more involved than the computation of one; this increases the optimization time significantly and often causes the optimization not to converge. Please note that this drawback only exists in the translational type of misregistration. Next we introduce a new method of computing SMI in which this drawback is removed.

If we assume that the source image is a translated version of the target image with a pixel/voxel misalignment, then the joint entropies or in equation (6) will drop to their minimum value of . This is the main reason that SMI becomes negative around these points. On the other hand, the conditional independency assumption in equation (7) implies that (data processing inequality in [35]; ). In fact, all the joint entropies in equation (6) should always be greater than . However, this condition becomes violated on or around pixel misregistration. To prevent such an artifact we have placed another condition on joint entropies in equation (6) to force them to always be greater than . Whenever the joint entropies of equation (6) become less than it means the registration process is on or about the critical point of pixel/voxel misalignment. By placing such an extra condition on the joint entropies of equation (6), the resulting SMI is guaranteed to always be positive. For convenience, we call this new method “Modified SMI” (MSMI). Fig. 3 shows the performance of the new method compared to the original SMI, and there is no visible artifact on or about pixel/voxel misalignment points.

5 Conclusion



Fig. 3. Comparison of MSMI and SMI as similarity measures in the case of translational mis-registration.

We have described the importance of incorporating spatial information into the computation of image mutual information and have given a comprehensive overview of existing attempts at computing mutual information with spatial dependency. We have also shown that the recently defined spatial mutual information (SMI, for a homogeneous anisotropic QMRF of order 2) has an artifact in the case of translational misregistration. Finally we introduced a new method of computing SMI, MSMI, that corrects the artifact. For future work, we intend to evaluate the new MSMI for the case of brain image registration.

References

1. Viola, P.A.: Alignment by maximization of mutual information. Ph.D. Thesis, MIT, Cambridge, MA (1995)

2. Maes, F.,Colligon, F., Vandermeulen, D., Marchal, G., Suetens, P.: Multimodality image registration by maximization of mutual information. IEEE Trans. Med. Imag., vol. 16, no. 2, pp. 187–198 (1997)

3. Shannon, C.E.: A mathematical theory of communication. Bell System Technical Journal, vol. 27, pp. 379-423, 623-656, (1948)

4. Studholme, C., Hill, D.L.G., Hawkes, D.J.: Incorporating connected region labelling into automated image registration using mutual information. Mathematical Methods in Biomedical Image Analysis, A.A. Amini, F. L. Bookstein, and D. C.Wilson, Eds. Los Alamitos, CA: IEEE Computer Soc. Press, pp. 23-31 (1996)

5. Razlighi, Q.R., Kehtarnavaz, N, Nosratinia, A.: Computation of image spatial entropy using quadrilateral Markov random field. IEEE Transactions on Image Processing, vol. 18, no. 12 (2009)

6. Rueckert, D., Clarkson, M.J., Hill, D.L.G., Hawkes, D.J.: Non-rigid registration using higher-order mutual information. in Proc. SPIE Conf. Medical Imaging: Image Processing, K. M. Hanson, Ed., vol. 3979, pp. 438–447 (2000)

7. Rodríguez-Carranza, C.E., Loew, M.H.: A weighted and deterministic entropy measure for image registration using mutual information. Proc. SPIE Conf. Medical Imaging: Image Processing, K. M. Hanson, Ed., vol. 3338, pp. 155-166 (1998)

8. Guo, Y., Lu, C.: Multi-modality image registration using mutual information based on gradient vector flow. Proc. 18th ICPR, Hong Kong, pp. 697–700 (2006)

9. Beijing, C., JunLi, L., Gang C.: Study of Medical Image Registration Based on Second-order Mutual Information. Proceedings of IEEE International Conference on Multimedia & Expo(ICBBE07), WuHan, pp. 7-9, China (2007)

10.Pluim, J.P.W., Maintz, J.B.A., Viergever, M.A.: Image registration by maximization of combined mutual information and gradient information. IEEE Trans. Med. Imag., vol. 19, no. 8, pp. 809–814, (2000)

11.Shams, R., Sadeghi, P., Kennedy, R.A.: Gradient Intensity: A New Mutual Information-Based Registration Method. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 1-8 (2007)

12.Gan, R., Chung, A.C.: Multi-Dimensional Mutual Information Based Robust Image Registration Using Maximum Distance Gradient-Magnitude. ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, vol. 3565, pp. 210-221 (2005)

13.Shen, D., Davatzikos, A.C.: HAMMER: Hierarchical Attribute Matching Mechanism for Elastic Registration. IEEE Transactions on Medical Imaging, vol.12, no.11 (2002)

14.Russakoff, D.B., Tomasi, C., Rohlfing, T., Maurer, C.R.: Image Similarity Using Mutual Information of Regions. ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, vol. 3023, pp. 596-607 (2004)

15.Yang, C., Jiang, T., Wang, J., Zheng, L.: A Neighborhood Incorporated Method in Image Registration. G.-Z. Yang et al. (Eds.): MIAR 2006, LNCS 4091, pp. 244 – 251, Springer-Verlag Berlin Heidelberg (2006)

16.Bardera, A., Feixas, M., Boada, M., Sbert, M.: High-Dimensional Normalized Mutual Information for Image Registration Using Random Lines. J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): LNCS 4057, pp. 264–271, Springer-Verlag Berlin Heidelberg (2006)

17.Tomazevic, D., Linkar, B., Pernus, F.: Multifeature mutual information. Proc. SPIE Med. Imag., J. M. Fitzpatrick and M. Sonka, Eds., vol. 5370, pp. 143-154 (2004)

18.Holden, M., Griffin, L.D., Saeed, N., Hill, D.L.G.: Multi-channel Mutual Information Using Scale Space. C. Barillot, D.R. Haynor, and P. Hellier (Eds.): MICCAI 2004, LNCS 3216, pp. 797–804, Springer-Verlag Berlin Heidelberg (2004)

19.Chappelow, J., Madabhushi, A., Rosen, M., Tomaszeweski, J., Feldman, M.: A Combined Feature Ensemble based Mutual Information Scheme for robust Inter-Modal. Inter-Protocol Image Registration. Proceedings of 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, page(s): 644-647 (2007)

20.Butz, T., Thiran, J.P.: Affine Registration with Feature Space Mutual Information. W. Niessen and M. Viergever (Eds.): MICCAI 2001, LNCS 2208, pp. 549–556, Springer-Verlag Berlin Heidelberg (2001)

21.Hero, A., Ma, B., Michel, O., Gorman, J.: Applications of entropic spanning graphs. IEEE Signal Processing Magazine, vol. 19, no. 5, pp. 85-95 (2002)

22.Neemuchwala, H., Hero, A.: Entropic Graphs for Registration. ser. Signal Processing and Communications. New York: CRC Press, ch. 6, pp. 185-235 (2005)

23.Neemuchwala, H., Hero, A., Carson, P.: Image matching using alpha-entropy measures and entropic graphs. Signal Processing, vol. 85, no. 2, pp. 277-296 (2005)

24.Oubel, E., Frangi, A.F., Hero, A.: Complex wavelets for registration of tagged MRI sequences. Proceedings of IEEE International Symposium on Biomedical Imaging, pp. 622-625 (2006)

25.Sabuncu, M.R., Ramadge, P.J.: Spatial Information in Entropy-Based Image Registration. ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, vol. 2717, pp. 132-141 (2003)

26.Oubel, E., Craene, M., Gazzola, M., Hero, A., Frangi, A.: Multiview registration of cardiac tagging MRI sequences. Proceedings of IEEE International Symposium on Biomedical Imaging (ISBI), Arlington, VA, USA, pp. 388-391 (2007)

27.Staring, M, van der Heide, A., Klein, S, Viergever, A., Pluim, P.W.: Registration of Cervical MRI Using Multifeature Mutual Information. IEEE Transactions on Medical Imaging, Vol 28, no. 9, pp. 1412-21 (2009)

28.Luan, H., Qi, F., Shen, D.: Multi-Modal Image Registration by Quantitative- Qualitative Measure of Mutual Information (Q-MI). ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, vol. 3765, pp. 378-387 (2005)

29.Luan, H., Qi, F., Xue, Z., Chen, L., Shen, D.: Multimodality image registration by maximization of quantitative-qualitative measure of mutual information. Pattern Recognition, vol. 41, Issue 1, pp. 285-298 (2008)

30.Kadir, T., Brady, M.: Saliency, scale and image descriptions. Journal International Journal of Computer Vision, vol. 45, no. 2, pp. 83-105 (2001)

31.Stan Z. L.: Markov Random Field Modeling in Image Analysis (3rd Edition). Series: Advances in Pattern Recognition. Springer (2009)

32.Winkler, G.: Image Analysis, Random Fields and Markov Chain Monte Carlo Methods. A Mathematical Introduction, 2nd edition. Springer Verlag, New York (2003)

33.Razlighi, Q.R, Rahman, M., Kehtarnavaz, N.: Fast computation methods for estimation of image spatial entropy. Journal of Real-Time Image Processing, Springer Berlin / Heidelberg (2010)

34.Razlighi, Q.R, Kehtarnavaz, N., Yousefi, S., Klein, A.: New Method in Evaluation of Similarity Measures for Brain Image Registration. Neuroimage (2011)

35.Cover, T.M., Thomas, J.A.: Elements of Information Theory. Wiley-Interscience, 2nd edition (1991)