DOI: 10.1007/s11432-006-0484-1

Unsupervised segmentation of medical image based on difference of mutual information

LÜ Qingwen & CHEN Wufan

School of BME, Southern Medical University, Guangzhou 510515, China
Correspondence should be addressed to Chen Wufan (email: chenwf@fimmu.com)
Received January 8, 2005; accepted March 16, 2006

Abstract In the scope of medical image processing, segmentation is important and difficult. There are still two problems which trouble us in this field. One is how to determine the number of clusters in an image and the other is how to segment medical images containing lesions. A new segmentation method called DDC, based on difference of mutual information (dMI) and pixon, is proposed in this paper. Experiments demonstrate that dMI shows one kind of intrinsic relationship between the segmented image and the original one and so it can be used to well determine the number of clusters. Furthermore, multi-modality medical images with lesions can be automatically and successfully segmented by DDC method.

Keywords: medical image, unsupervised segmentation, mutual information, pixon.

Image segmentation is a very important research field in the scope of image processing. It has extensive application and involves almost all fields such as image understanding, pattern recognition and image encoding, etc. Furthermore, research of image segmentation is conducted on all types of images. In its application in medicine, it is achieved to segment the brain magnetic resonance (MR) image into different tissue classes, especially gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF)[1−6].

According to the general definition of segmentation, the segmented region should satisfy the conditions of uniform and continuity, of which the former means that all the pixels in the region are in similarity in terms of gray level, texture, color or other characteristics and the latter that there is path to connect any two points of the region. The segmentation calculation is tremendously complex and difficult to satisfy the above conditions to which researchers have been committing much effort. So far, most research achievements have targeted specific type of image or specific application, while general method and strategy are still a big challenge.

Among the traditional image segmentation approaches, clustering-based segmentation method is one of the most representative and typical methods. Fuzzy C-mean (FCM) belongs to this kind of algorithm[7,8]. The defect of the method is that it only takes care of the information of the gray level while ignoring the information of the space that brings about the inaccurate result for the images without significant difference in gray level or images with major overlapping in the gray level range. At the same time, it may be possible to divide a significant identical region with gray level difference into different classes (or clusters)[3].

Markov random field (MRF) well describes the relationship between the current pixels and their nearby pixels, due to which it has been gaining wide concerns and applications in the field of image processing[1−4,6,9]. In ref. [3] an improved GFCM algorithm is proposed, in which a prior in the form of Gibbs energy is used to overcome the defect of FCM mentioned above. Experiments show that the algorithm is effective to segment images with noise.

In ref. [4], a novel pixon-based adaptive scale method for image segmentation is introduced, of which the key idea is that a pixon-based image model is combined with an MRF model under a Bayesian framework. Experimental results demonstrate that this algorithm performs well and computational costs decrease dramatically. However, the definition of pixon is far away from the original idea of "fuzzy pixon"[10]. From this point, the modification in this method is as masterly as that one in ref. [11].

As ref. [5] saying, image segmentation is a challenging research work. There are still some issues troubling us in the field of medical image segmentation, concerning the automatic determination of the number of clusters (Nc) in an image and the segmentation of the medical image with lesions.

Herein, we introduce a new optimized segmentation measurement—difference of mutual information (dMI) and propose a new segmentation algorithm—DDC (rough Division, subtle Division and fuzzy Cross) method based on pixon, mutual information and generalized fuzzy theory. DDC method involves "cross calculation" with the subset by pixon segmentation and that one by dMI-based GFCM segmentation to realize the multi-resolution segmentation. We successfully apply this measurement dMI to the determination of Nc and DDC method to the segmentation of the medical images with lesions. The new measurement and the new algorithm proposed here meaningfully explore the two issues mentioned above.

1 Mutual information and difference of mutual information

Mutual information (MI) is a basic concept from Shannon information theory, measuring the statistical dependence between two random variables or the amount of information that one variable contains about the other. MI has been widely used in image registration[12,13] before and the MI registration criterion presented by Maes[12] states that the MI of the image intensity values of corresponding voxel pairs is maximal if the images are geometrically aligned. Recently, image segmentation based on MI has been reported[14,15]. But the method using difference of MI (dMI) to determine the number of clusters in an image is seldom or never reported.

For the specific images A and B, the mutual information between them is

$$MI(A,B) = H(A) + H(B) - H(A,B),$$

where $H(A)$ is the entropy of image A and $H(B)$ is that of B, $H(A, B)$ is the joint entropy of them. If $p_A(a)$ and $p_B(b)$ respectively represent the probabilistic distribution of A and B, and $p_{AB}(a, b)$ is the joint probabilistic distribution of them, we have

$$MI(A,B) = \sum_{a,b} p_{AB}(a,b) \log \frac{p_{AB}(a,b)}{p_A(a) \cdot p_B(b)}.$$

It has been found in our study that the MI value between the original image I and its segmentation image S increases with the increasing Nc value in image S and converges at its maximum value $MI(I,I)$, i.e. MI between the original image and itself is maximal. At the same time, the difference of two adjacent MI value (dMI) decreases vibrantly or non-vibrantly (depends on the segmentation method used, we think), and converges at 0.

The definition of dMI is as the following:

$$dMI_n(I) = MI(I,S_n) - MI(I,S_{n-1}),$$

where I is original image, S_n is its segmentation with Nc equal to n, and S_{n-1} is its segmentation with Nc equal to $n-1$, apparently, $n > 1$. In order to compare dMI value of different images, it is normalized as the following:

$$dMI_n(I) = \frac{MI(I,S_n) - MI(I,S_{n-1})}{MI(I,I)}.$$

In Fig. 1 the images to be segmented are shown. The relationship between $MI(I,S_{Nc})$ and Nc and that between $dMI_{Nc}(I)$ and Nc are illustrated in Fig. 2, where MI, dMI have been normalized. Curves 1, 2 and 3 in Fig. 2 represent respectively the corresponding original images in Fig. 1(a)—(c); and FCM is used in segmentation for common. It can be found that with Nc increasing, MI gradually increases, while dMI decreases or waves down, showing that dMI gradually decreases in images (a) and (b) and waves down in image (c). The reason for its waving down is discussed in section 5.

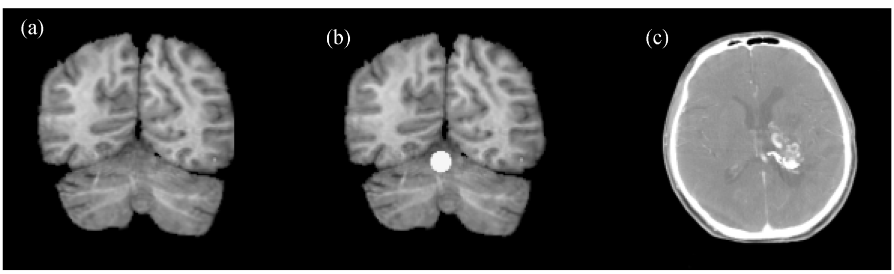

Fig. 1. Original images to be segmented. (a) Brain MR image; (b) brain MR image added with a simulated lesion; (c) brain CT image with a lesion.

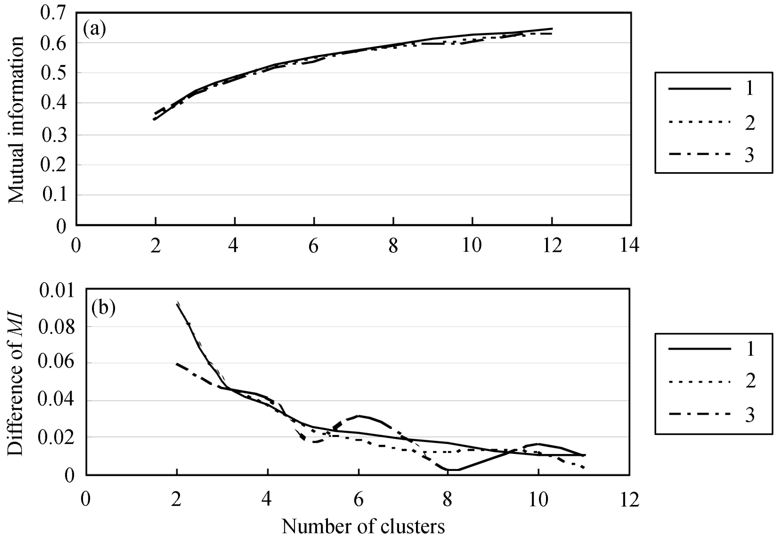

Fig. 2. (a) Relationship between *MI* and *Nc*; (b) relationship between *dMI* and *Nc*. Curves 1, 2 and 3 represent the images (a), (b) and (c) in Fig. 1, respectively.

2 GFCM segmentation

2.1 Traditional FCM algorithm

FCM is a method of clustering which allows each datum to belong to one of the clusters. This method (developed by Dunn in 1973[7] and improved by Bezdek in 1981[8]) is frequently used in pattern recognition and image segmentation. FCM is based on minimization of the following objective function:

$$J_{FCM} = \sum_{i,j} \sum_{k=1}^{Nc} \mu_k (i,j)^q \parallel \gamma(i,j) - v_k \parallel^2,$$

where q is any real number greater than 1, μ_k is the degree of membership of pixel (i,j) in the cluster k, $\gamma(i,j)$ is the gray value (color) of pixel (i,j), v_k is the center of the cluster k, ||*|| is a kind of distance between any measured data and the center, and Nc is the number of clusters.

2.2 GFCM algorithm

In ref. [3] the FCM algorithm is improved as the GFCM algorithm based on the MRF theory and its objective function is

$$J_{GFCM} = \sum_{i,j} \sum_{k=1}^{NC} \mu_k (i,j)^q (1 - p_k(i,j)) \parallel \gamma(i,j) - v_k \parallel^2,$$

where $p_k(i,j)$ is the prior probability of pixel (i,j) labeled as the k-th category under the function of the neighborhood system $N(i,j)$, and the others are the same as the above formula.

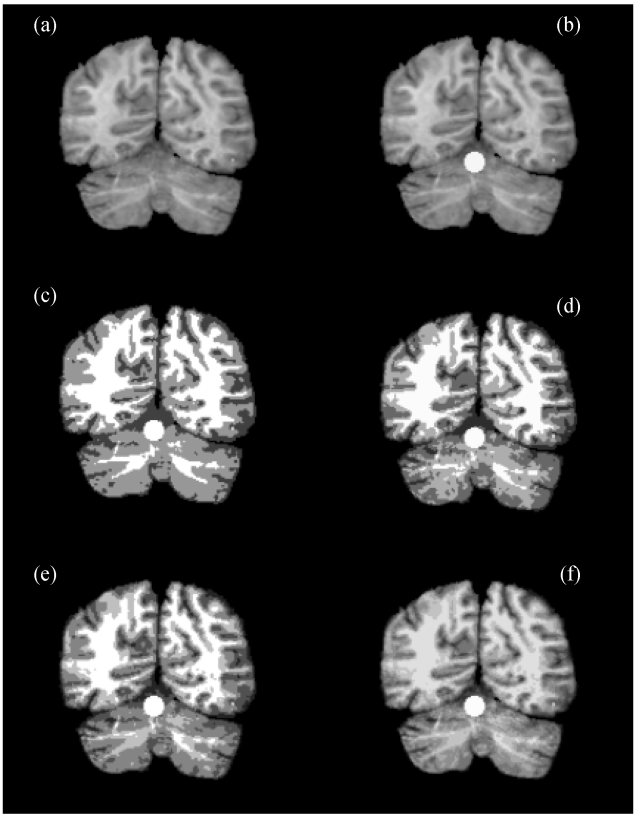

Fig. 3. Segmentation of image Fig. 1(b) by GFCM. (a) Original image, the same as Fig. 1(a); (b) simulated image, the same as Fig. 1(b); (c)—(e) segmentation results when the *Nc* are 4, 5, 6; (f) segmentation result when the *Nc* is 12.

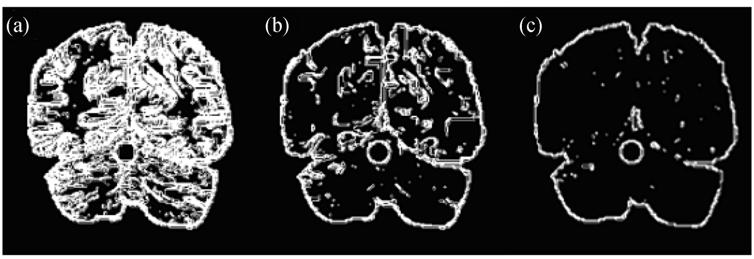

Fig. 4. Segmentation of image Fig. 1(c) by pixon. (a)—(c) *k* increases while the number of pixons decreases.

The segmentation of the image Fig. 1(b) using GFCM is shown in Fig. 3. Note that the simulated lesion cannot be accurately recognized until *Nc* is 12. For the general purposes, the experiment is also conducted using FCM algorithm and the same result is achieved. Fig. 3(d) is the segmentation result when *Nc* is 5 (right *Nc* value), with no way to correctly identify the simulated simple lesion and no way of correct segmentation (result in Fig. 8(b)) of CT (computer tomography) image with complex lesion (original in Fig. 1(c)).

3 Pixon segmentation

The "pixon" is a concept proposed by Pina and Puetter and firstly used for astronomic image restoration[10], of which the essence is that the spatial scale at each site of an image varies according to the information embedded in the image. Later, it is combined with MRF by Descombes[16] and Yang[4] to be applied in image description and image segmentation. Yang's new pixon definition scheme can be described as follows[4]:

$$IM = \bigcup_{i=1}^{n} P_i,$$

where IM is the pixon-based image model; n is the number of pixons; P_i is a pixon, which is made up of a set of connected pixels, a single pixel or even a sub-pixel. After the pixon-based model is defined, the image segmentation problem is transformed into a problem of labeling pixons. We call this process as pixon segmentation in this paper.

The process of pixon segmentation is documented in ref. [4] as the following procedures. 1) Obtaining the pseudo image. The pseudo image is a basic image to form the pixons and to obtain a segmented image. 2) Formulation of pixons. This process is based on the pseudo image. The pseudo image intensity is viewed as the temperature of the temperature field and transformation of the gradient as the diffusion coefficient. 3) Extraction of the pixons. Generally, the pixon segmentation is an over-segmentation, during which there is a constant k, larger value of k increasing the pixon size. The pixon segmentation of the image Fig. 1(b) is shown in Fig. 4, where white color represents the edges of the pixons.

4 DDC method

DDC (rough Division, subtle Division and fuzzy Cross) method proposed in this paper includes three steps, i.e. rough division step (D1 step), subtle division step (D2 step) and fuzzy cross step (C step). The flow chart of DDC method is shown in Fig. 5 and the steps of it are described as follows.

4.1 D1 step

The purpose of this step is to achieve rough segmentation of the original image with unknown Nc. During this step, GFCM algorithm is used and there emerges the same classical trouble, i.e. automatic determination of the number of clusters (Nc) in an image before segmentation, which has been extensively studied. Among them AIC (Akaike information criteria)[17−19] has gained wide application. However, the theoretical basis and the actual effects of AIC are not very satisfying. Therefore we propose a new measurement dMI to solve this problem. The process using dMI to determine Nc is as the following: for the image I, the number of clusters in the segmented images (written as n) gradually increases from 2, and $dMI_n(I)$ is calculated out, stop when $dMI_n(I)$ is less than ε, where ε is a specific threshold constant (here, 0.04). Then n is the Nc value to be determined.

Set the original image as I, and the segmented image (after the process in D1 step) as

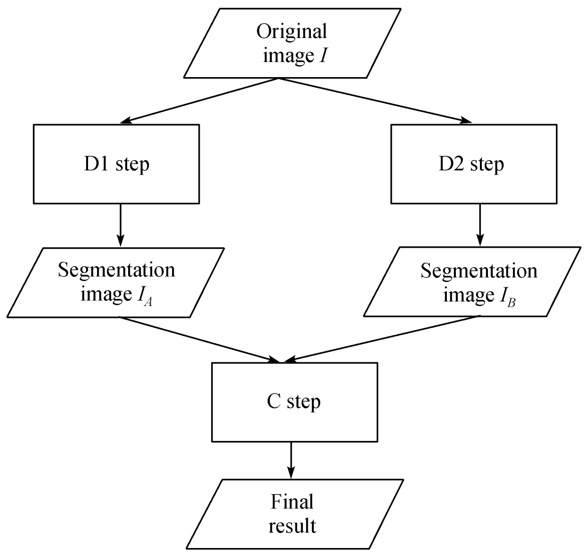

Fig. 5. Flow chart of DDC.

I_A, then we have

$$I_A = \bigcup_{i=1}^{n} A_i, \text{ and } A_i \cap A_j = \varnothing,\ i \neq j,\ i, j = 1, \cdots, n,$$

where n is Nc value, A_i is the set of pixels belonging to the i-th class.

4.2 D2 step

The purpose of this step is to acquire pixons. The algorithm for pixon segmentation is described in detail in ref. [4] and briefly introduced in section 3 of this paper. If the original image is I, and the segmented image (after the process in D2 step) is I_B, then we get

$$I_B = \bigcup_{i=1}^{m} B_i, \text{ and } B_i \cap B_j = \varnothing,\ i \neq j,\ i, j = 1, \cdots, m,$$

where m is the number of pixons and B_i is one specific pixon.

4.3 C step

The purpose of this step is to make "cross calculation" of I_A and I_B that acquired in the above two steps to achieve new segmentation.

Here we propose concepts of "first background" (BG1) and "second background" (BG2) to explain the "cross calculation". Generally, the image has both object and background and we call this background as BG1, such as the black parts in CT or MR image. In this paper, we take the segmented image acquired in D1 step as BG2.

In fact, the "cross calculation" in this step is to judge if the pixons acquired in D2 step can be merged into BG2 acquired in D1 step. The judgment is based on the generalized fuzzy operator $GF(s(C), \tilde{S}(A_j))$, which is based on generalized fuzzy theory[20], where C is the intersection of one specific pixon and one category of BG2; A_j is one category of

BG2; $s(C)$ is the value of one specific property of set C; $\tilde{S}(A_j)$ is the generalized fuzzy set based on the same property of set A_j. If the value of $GF(s(C),\tilde{S}(A_j))$ is less than 0, a new subclass is generated. All of the subclasses are united into a new class (category), in which the newer classes will be extracted until there is none for extraction.

The C step of DDC algorithm is described in Fig. 6 with details.

Furthermore, the property that extraction is based on can be changed each time and the mean of gray value, the local entropy or the texture features may be used. So DDC algorithm can be realized as a multi-resolution segmentation method.

```
C step:
  reapeat for all i, j (1≤i≤m, l≤j≤n)
      if ((C = Bi ∩ Aj)≠ ∅ )
          if (GF(s(C), S̃ (Aj))≥0)
              C merge into Aj
          else
              add a new subclass C and update IA
          end if
      end if
  end reapeat
  merge all new subclasses as a new class An+1
  reapeat
      extract a new class from An+1 and update IA ditto
    until no new class can be extracted
```

Fig. 6. Description of C step of DDC.

5 Results and discussion

The DDC algorithm is applied to segmentation of medical images, including one image with simulated lesion (see Fig. 1(b)) and one image with a real lesion (see Fig. 1(c)).

In Fig. 7 the segmentation result of Fig. 1(b) using DDC algorithm is shown, among which (a) is the original image, (b) is the segmentation result using GFCM algorithm and its *Nc* value is 4, (c) is the result of pixon segmentation, (d) is the result of DDC segmentation and the final acquired *Nc* value is 5. From the result of the experiment, DDC algorithm correctly identified the brain tissues of GM, WM, CSF and the simulated lesion as well.

In Fig. 8 the DDC segmentation of a CT image with a lesion of AVM (artery and vein malformation) is shown, among which (a) is the original image, (b) is the result of GFCM segmentation and its *Nc* value is 5, (c) is the result of pixon segmentation, (d) is the result of DDC segmentation and the final acquired *Nc* value is 6. It is shown that the DDC algorithm basically reserves the result of GFCM segmentation and correctly recognizes the AVM lesion.

We also apply the DDC algorithm to segmentation of other medical images with different lesions under the modalities of CT, MR imaging. The results show that the DDC

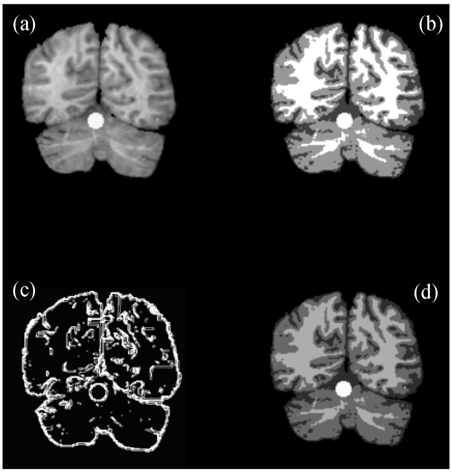

Fig. 7. Segmentation of image Fig. 1(b). (a) Original image; (b) GFCM; (c) pixon; (d)DDC.

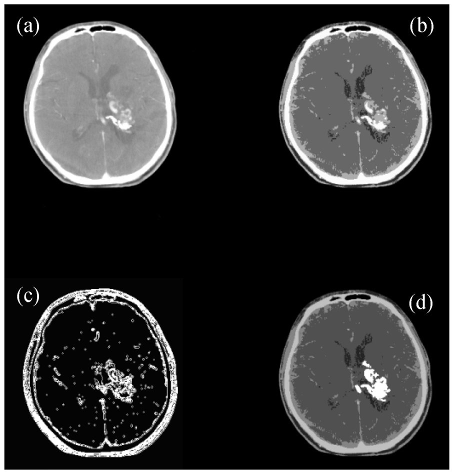

Fig. 8. Segmentation of image Fig. 1(c). (a) Original image; (b) GFCM; (c) pixon; (d) DDC.

algorithm can realize effective automatic segmentation of the medical images with lesions.

In this paper, the new proposed measurement dMI, which well describes the information increment of original image while the number of clusters in the segmentation image increases, actually reveals one kind of inner relationship between the segmentation image and the original one. It can be supposed that a proper balance between the number of clusters and the information scalar is gained when dMI is used in segmentation, so that the *Nc* value of the image can be well determined. Moreover, The DDC algorithm not only uses the gray level of the image, but also uses the spatial information, such as edge, texture, etc., to optimize the automatic segmentation of the complex medical image.

It should be pointed out that the C step consumes much less time than D1 step and D2 step in the DDC algorithm while D1 step and D2 step can perform in parallel. Therefore, the DDC algorithm can be approximately regarded as a parallel algorithm.

In addition, dMI value waves down with *Nc* value of the segmented image increasing, but not smoothly decreases on several occasions. We think that possibly results from the use of FCM algorithm: (1) FCM is over-dependent on the choice of primary clustering center; (2) FCM often converges at the local minimum but not global minimum.

Acknowlegments The authors are grateful to the anonymous reviewers for their significant work. The authors also thank G. Yan, Y. Yang, B. Li, L. Zhou, Q. Feng, Y. Feng, Z. Liu and R. Wu of Southern Medical Univ., Dr. Y. Wu of UCLA and J. Kim of MIT for their help. This work was supported by the National Basic Research Program of China (Grant No. 2003CB716101), and the National Natural Science Foundation of China (Grant No. 30130180).

References

1 Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans Med Imag, 2001, 20(1): 45－57

2 Ruan S, Moretti B, Fadili J, et al. Fuzzy Markovian segmentation in application of magnetic resonance images.

Comp Vision Image Under, 2002, 85: 54－69

3 Feng Y, Chen W. Brain MR image segmentation using fuzzy clustering with spatial constraints based on Markov random field theory. In: Proc of MIAR 2004 (LNCS 3150, Beijing). Berlin: Springer-Verlag, 2004. 188－195

4 Yang F, Jiang T. Pixon-based image segmentation with Markov random fields. IEEE Trans Image Process, 2003, 12(12): 1552－1559

5 Wang A, Shen L. Study surveys on image segmentation. Meas Control Tech (in Chinese), 2000, 19(5): 1－16

6 Lin Y, Chen W, Yang F, et al. New approach for image segmentation based on generalized fuzzy Gibbs random fields. Chin J Comput (in Chinese), 2003, 26(11): 1465－1469

7 Dunn J C. A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. J Cybern, 1973, 3: 32－57

8 Bezdek J C. Pattern Recognition with Fuzzy Objective Function Algorithms. New York: Plenum Press, 1981

9 Geman S, Geman D. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. IEEE Trans Pattern Anal Machine Intell, 1984, 6 (6): 721－741

10 Pina R K, Puetter R C. Bayesian image reconstruction: The pixon and optimal image modeling. Publ Astron Soc Pac, 1993, 105: 630－637

11 Yang Y, Huang B, Wang Q, et al. Meaningful region segmentation of an image combined with intensity distribution and spatial information field. Acta Electron Sinica (in Chinese), 2003, 31(2): 252－254

12 Maes F, Collignon A. Multimodality image registration by maximization of mutual information. IEEE Trans Med Imag, 1997, 16(2): 187－198

13 Studholme C, Hil1 D L G, Hawkes D J. Incorporating connected region labeling into automated image registration using mutual information. In: Proc of MMBIA'96 (San Francisco). Piscataway: IEEE Press, 1996. 3－31

14 Rigau J, Feixas M, Sbert M, et al. Medical image segmentation based on mutual information maximization. In Proc of MICCAI 2004 (Saint-Malo). Berlin: Springer-Verlag, 2004. 135－142

15 Kim J, Fisher III J W, Yezzi A, et al. A nonparametric statistical method for image segmentation using information theory and curve evolution. IEEE Trans Image Process, 2005, 14(10): 1486－1502

16 Descombes X, Kruggel F. A Markov pixon information approach for low-level image description. IEEE Trans Pattern Anal Machine Intell, 1999, 21(6): 482－493

17 Akaike H. A new look at the statistical model identification. IEEE Trans Auto Contr, 1974, 19(6): 716－723

18 Cheng B, Wang Y, Zheng N, et al. MRF model and FRAME model-based unsupervised image segmentation. Sci China Ser F-Inf Sci, 2004, 47(6): 697－705

19 Li S Z. Markov Random Field Modeling in Image Analysis. 2nd ed. Berlin: Springer-Verlag, 2000

20 Chen W, Lu X, Chen J, et al. A new algorithm of edge detection for color image: Generalized fuzzy operator. Sci China Ser A-Math, 1995, 38(10): 1272－1280