Computers & Geosciences 88 (2016) 41–53

Contents lists available at ScienceDirect

Computers & Geosciences

journal homepage: www.elsevier.com/locate/cageo

Case study

A contour-line color layer separation algorithm based on fuzzy clustering and region growing

Tiange Liu [a], Qiguang Miao [a,*], Pengfei Xu [f], Yubing Tong [d], Jianfeng Song [a], Ge Xia [c], Yun Yang [b,e], Xiaojie Zhai [a]

[a] School of Computer Science and Technology, Xidian University, Xi'an, Shaanxi 710071, China
[b] Xi'an Research Institute of Surveying and Mapping, Xi'an, Shaanxi 710054, China
[c] Lafayette College, Easton, PA 18042, USA
[d] Radiology Department, University of Pennsylvania, Philadelphia, PA 19104, USA
[e] State Key Laboratory of Geo-Information Engineering, Xi'an, Shaanxi 710054, China
[f] School of Information Science and Technology, Northwest University, Xi'an, Shaanxi 710127, China

ARTICLE INFO

Article history:
Received 11 August 2014
Received in revised form
23 December 2015
Accepted 25 December 2015
Available online 28 December 2015

Keywords:
Contour-line
Color layer separation
Fuzzy clustering
Single-prototype region growing

ABSTRACT

The color layers of contour-lines separated from scanned topographic map are the basis of contour-line extraction, but it is difficult to separate them well due to the color aliasing and mixed color problems. This paper will focus us on contour-line color layer separation and presents a novel approach for it based on fuzzy clustering and *Single-prototype Region Growing for Contour-line Layer* (SRGCL). The purpose of this paper is to provide a solution for processing scanned topographic maps on which contour-lines are abundant and densely distributed, for example, in the condition similar to hilly areas and mountainous regions, the contour-lines always occupy the largest proportion in linear features and the contour-line separation is the most difficult task. The proposed approach includes steps as follows. First step, line features are extracted from the map to reduce the interference from area features in fuzzy clustering. Second step, fuzzy clustering algorithm is employed to obtain membership matrix of pixels in the line map. Third step, based on the membership matrix, we obtain the most-similar prototype and the second-similar prototype of each pixel as the indicators of the pixel in SRGCL. The spatial relationship and the fuzzy similarity of color features are used in SRGCL to overcome the inaccurate classification of ambiguous pixels. The procedure focusing on single contour-line layer will improve the accuracy of contour-line segmentation result of SRGCL relative to general segmentation methods. We verified the algorithm on several USGS historical maps, the experimental results show that our algorithm produces contour-line color layers with good continuity and few noises, which verifies the improvement in contour-line color layer separation of our algorithm relative to two general segmentation methods.

1. Introduction

Historical topographic maps contain rich cartographic information, such as locations of buildings, roads, contour-lines and hydrography (Chiang et al., 2013). Essentially, these geographic elements consisting of color point, linear, and area features are used to represent topographic and geographic information about the parts of the Earth (Chen et al., 2006). Because the topographic maps are invaluable carriers of information about the landscape in the past over large areas (Chiang et al., 2014), lots of research efforts had been made in extracting geographic elements from maps (Chen et al., 2006; Khotanzad and Zink, 2003; Chiang et al., 2009; Gamba and Mecocci, 1999; Leyk et al., 2006; Cao and Tan, 2002). Among all graphical elements, contour-line is the most important one to characterize three-dimensional terrain on two-dimensional map sheets. Without contour-lines, a topographic map degenerates into a planimetric map providing no three-dimensional data about the terrain (Chen et al., 2006; Khotanzad and Zink, 2003; San et al., 2004; Salvatore and Guitton, 2004). An example is shown in Fig. 1, where contour-lines are usually drawn in a very dense way. The close space between lines and their widely distribution make contour-line extraction the most time-consuming process. Meanwhile, color aliasing and mixed color will emerge from the frequent overlapping of contour-lines and background or other area geographic features, which makes the contour-line extraction process even worse. However, the accurate topographic height information is extremely valuable for terrain analysis and change detection and plays the key role in constructing three-

* Corresponding author.
E-mail address: qgmiao@mail.xidian.edu.cn (Q. Miao).

http://dx.doi.org/10.1016/j.cageo.2015.12.017

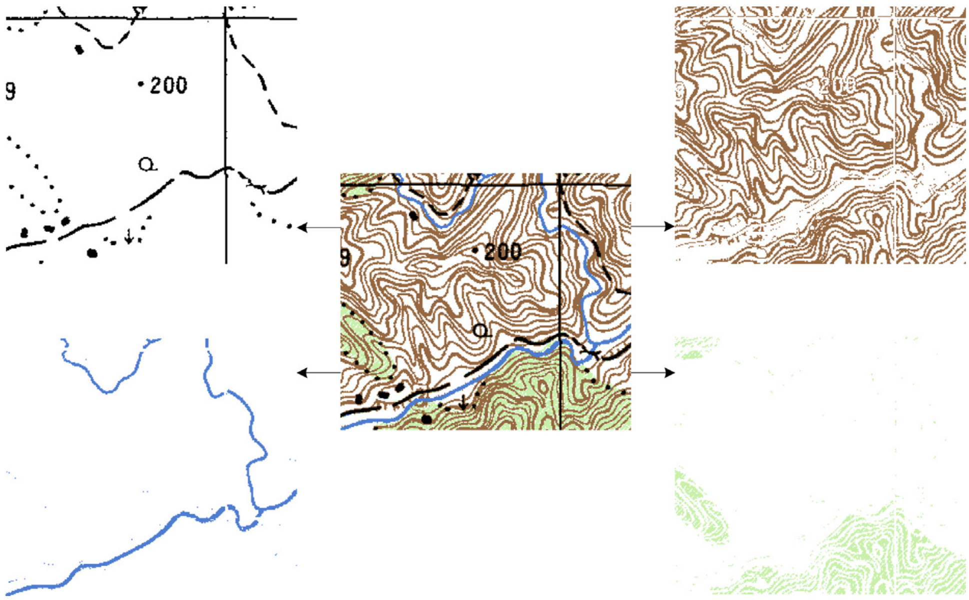

Fig. 1. The purpose of color image segmentation in color map. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

dimension geographical information. Therefore, the contour-line extraction is necessary and it has to be done accurately and efficiently. Automatic contour-line extraction techniques have great significance in constructing Geographical Information Systems.

Although many types of sequence steps in extracting contour-lines from topographic map have been shown in (Salvatore and Guitton, 2004; Samet and Hancer, 2012; Wu et al., 2009), there are two main necessary steps: (1) Color Image Segmentation (CIS): separate the contour-line color layer from the original map; (2) contour-line reparation: solve the problems of gap and conglutination. All the subsequent procedures strongly depend on the results of CIS (Chiang et al., 2014; Leyk, 2010). For the low quality maps, the result of CIS is not accurate and it is hard to resolve disconnection and conglutination of contour-lines. The most obvious difference between topographic map and natural image is that colors in topographic map are mainly used to distinguish different feature categories. The segmentation of topographic map can also be defined as segmentation of the map into different color layers, which represent different categories of elements. Fig. 1 shows the goal of color map segmentation and four color layers are separated from the original color map. Each layer represents a type of geographic elements, such as contour-line layer (brown), vegetation layer (green), rivers layer (blue) and other thematic objects layer (black).

Some researchers (Pouderoux et al., 2007; San et al., 2004; Salvatore and Guitton, 2004; Samet and Hancer, 2012; Xin et al., 2006) focused on the contour-line reparation step which deals with the CIS by color space transformation and threshold segmentation or other simple algorithms. Although these methods achieved good results in high quality maps, there are many low quality topographic maps that cannot be segmented well.

Numerous algorithms of topographic maps color segmentation were proposed in recent years(Chen et al., 2006; Khotanzad and Zink, 2003; Feng and Song, 1996; Wu et al., 1994; Zheng et al., 2003). Chen and Khotanzad used a color key set to solve the problem of color aliasing and false colors in maps for segmenting the color topographic maps (Chen et al., 2006; Khotanzad and Zink, 2003). Unfortunately, this method could only be applied to maps with high quality. Feng et al. proposed a method for feature separation based on color clustering (Feng and Song, 1996). Considering the existence of color aliasing and false colors, Wu et al. (1994) proposed a method which combined fuzzy clustering and neural networks to extract the lines and characters of the map. Zheng et al. presented a CIS method of fuzzy clustering based on two-dimensional histogram (Zheng et al., 2003). Although the above algorithms can segment maps automatically, these unsupervised methods cannot overcome the shortcomings of false color and color aliasing because they do not consider the planar spatial relationship.

In order to make full use of the information of distribution in color space, local homogeneity and connected regions in color map, Leyk proposed a segmentation method based on Seeded Region Growing (SRG) which employs the information from the local image plane, the frequency domain and the color space (Leyk and Boesch, 2010). G–K fuzzy clustering algorithm (Gustafson and Kessel, 1978), which is a transformation of Fuzzy c-means (FCM) algorithm, uses Mahalanobis distance (MD) as the distance measure. The results based on Euclidean distance show that the segmentation is good only when the data set contains clusters that are well separated or clusters of roughly the same shape, but G–K algorithm overcomes the this defect of FCM. Both these two methods have good performance in dealing with low-quality map segmentation, but they still have some limitations. Leyk's method requires extensive parameterization which needs to be turned through prior knowledge, such as the prototype initialization parameter will sensitively influence the results. It also has the limitation of overcoming the disadvantage of order dependencies required by SRG. On the other hand, mixed color and color aliasing problems could still cause inaccurate segmentation results in G–K algorithm since it cannot fix the problem of imbalance date sets clustering.

This kind of segmentation methods are designed to segment a map into different color layers. It is note worthy that it is hard to generate a segmentation framework suitable for all kinds of maps as well as can separate all layers accurately. Thus, we think improvement in map layers extraction can only be achieved at the cost of highly specialized algorithm. Our goal is to develop a robust and accurately contour-line color layer separation algorithm to extract the color layer of contour-lines in color topographic maps. We proposed a method which is slightly different from traditional information extraction method of topographic maps. Instead of general color image segmentation, the new method focuses on a specialized direction which is contour-line layer segmentation. The proposed method is direct at processing topographic maps whose contour-lines are abundantly and densely distributed. In such maps (like hilly area and mountainous region), the contour-lines always occupy the largest proportion in linear features and the contour-line separation is the most difficult task. In this method, the most-similar prototype and the second similar prototype of each pixel are obtained using fuzzy cluster. Based on fuzzy clustering results, we adopt a Single-prototype Region Growing algorithm for Contour-line Layer (SRGCL) to extract the color layer of contour-lines. Afterwards, a high quality contour-line layer can be obtained. The proposed method is on the basis of fuzzy clustering and a special SRG. It overcomes the disadvantages of both fuzzy clustering and SRG, which could make a significant improvement in separating contour-line layer.

The remaining of this paper is arranged as follows. In the Section 2, we analyze the color features of topographic map. In Section 3, the proposed method will be described in detail with regard to lines extraction, fuzzy clustering and single-prototype region growing. The experimental performance of our method will be shown in Section 4 and the conclusions will be discussed in Section 5.

2. Color analysis in scanned topographic map

Topographic map mainly consists of linear features and area features. The elements of contour-lines, roads, rivers and so on are linear features, and the elements of green field, bodies of water, background and etc. are area features. In many topographic maps like hilly area and mountainous region which include lots of variations in terrain, contour-lines occupy the largest proportion among linear features. In this kind of topographic maps, contour-lines are always distributed densely and abundantly. It makes the contour-lines very difficult to extract. Thus how to achieve the contour-line layer accurately from this kind of maps is a very important task for scanned topographic map digitalization.

The color space of CIE Lab, or named as Lab, specified by the International Commission on Illumination, which is designed to approximate human vision that aspires to perceptual uniformity (Ke et al., 2004). The original purpose of drawing map is to show geographic features to humans, different geographic features should be distinguished by different colors in human vision. Thus, all processes in this paper are based on the Lab color space.

Lots of challenges exist in contour-lines extraction process because of color aliasing and false colors derived from the scanning process. Meanwhile, closely spaced and intersecting/overlapping features inherent to the map would further increase the degree of challenges (Khotanzad and Zink, 2003). Fortunately, there exist some regular patterns of color in topographic maps which can be utilized for segmentation. Alireza et al. analyzed the difficulty of separating the brown contour-lines from a scanned image of a topographic map and arrived at the conclusion that color aliasing appears between the two colors which adjoin each other (Khotanzad and Zink, 2003). But the challenges of lines overlapped by area features are not discussed in detail, i.e. mixed color, which also is a big challenge on CIS. Fig. 2 shows this phenomenon. Fig. 2 (a1) and (b1) are two small pieces of original maps which only

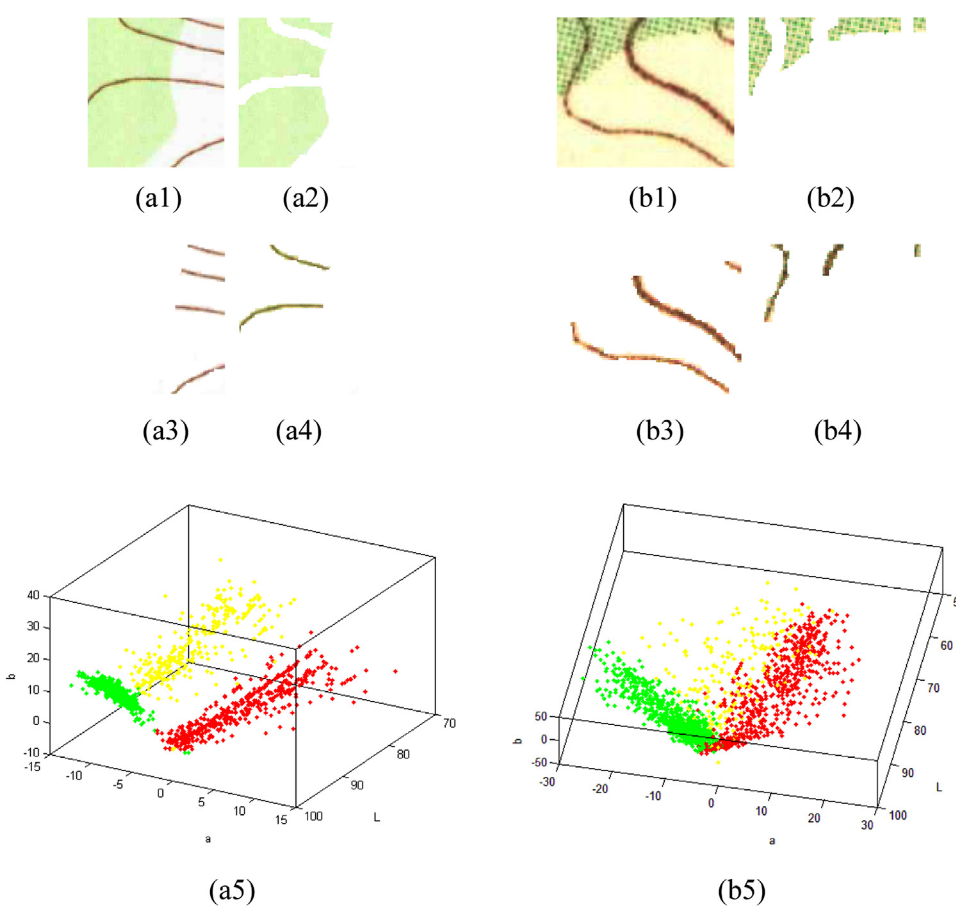

Fig. 2. Illustrations for showing color distribution of overlapped contour-lines. (a1) and (b1) are original maps. (a2–a4) and (b2–b4) are green field, contour-lines without overlapping and overlapped contour-lines respectively. (a5) and (b5) show their distribution in color space, where green points and red points represent the green fields and contour-lines without overlapping respectively, and yellow points represent overlapped contour-lines. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

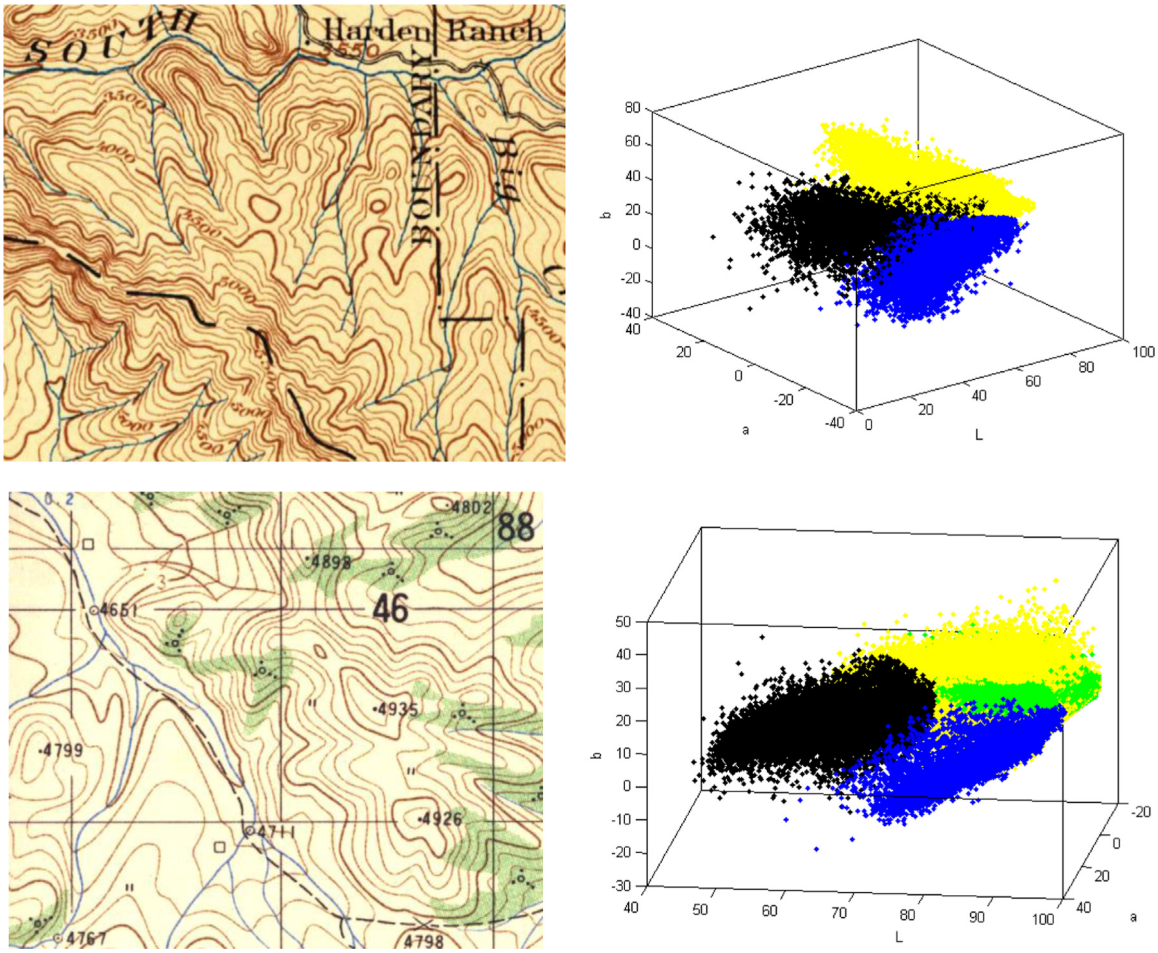

Fig. 3. The left column shows original scanned maps. The right column shows distribution of color layers in the Lab color space. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

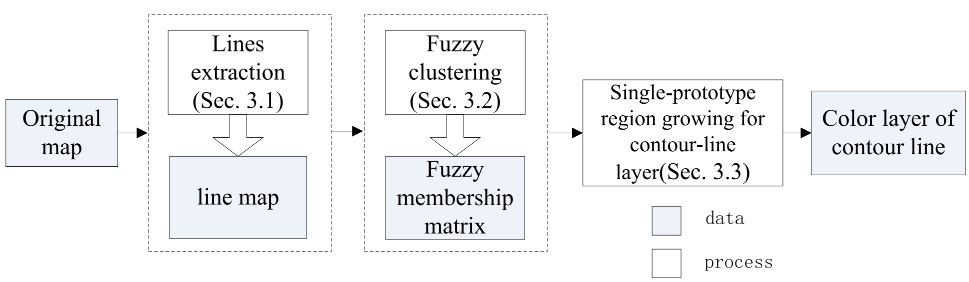

Fig. 4. The framework of proposed method.

consist of contour-lines and green field; (a2–a4) and (b2–b4) are the segmentation results obtained manually; (a5) and (b5) are the distribution of these pixels in the Lab color space. From the distribution we can see the color of the overlapped contour-lines is distributed between the color of the contour-lines without overlapping and the color of the green field in the Lab color space.

As a result of false colors, color aliasing and mixed colors, there is no clear boundary between different prototypes in color space. Meanwhile, the shapes of color layers in color space are not sphere but elliptical, as shown in Fig. 3. Two topographic maps are segmented into different geographic element layers manually (excluding background). Layers are mapped into the Lab color space with different color labels. Although a few pixels maybe lost during the process of manual segmentation, these mapping models show the distribution of topographic map in color space distinctively. In addition, they help to choose an appropriate clustering algorithm as shown in the following Section 3.

3. Proposed method

In this section, a new contour-line color layer separation method is proposed. This method, initial seeds can be automatically obtained from fuzzy clustering without prior knowledge. Additionally, the potentially growable seeds can also be discovered from fuzzy clustering process. Based on these clustering results, a region growing method SRGCL, which is specific to contour-line, is applied to grow the contour-line color layer. As discussed, this method is designed for only the contour-line color layer separation and works well when the maps show very dense and abundant contour-line. This method can obtain consecutive contour-line results without prior knowledge and sensitive parameters. The framework of this method is shown in Fig. 4.

Linear feature extraction method proposed by Miao et al. Qiguang et al. (2013) is first employed to get the line map which contains only linear features from scanned original map. Then, fuzzy clustering algorithm is applied on the line map. After fuzzy clustering, a fuzzy membership matrix can be obtained, which records the membership of each pixel to each prototype. Then the most-similar prototype and the second-similar prototype of each pixel are achieved from this matrix, which are the basic of the subsequent region growing process. Based on the previous fuzzy clustering results, the contour-line layer will iteratively grow from the initial seeds.

3.1. Line extraction

Although line features contain a large amount of geographic information, area features occupy the largest share of pixels in topographic maps. Moreover, color aliasing and mixed color are mainly caused by the overlapping of these geographic features. Removing the area features from topographic maps is necessary for the following reasons.

1) Area features reduces the calculation efficiency because of the large amount of pixels in them. Removing them could significantly reduce computational time.
2) Color aliasing and mixed color, caused by the overlapping of lines and areas, are distributed between the two original colors in the Lab color space as shown in Section 2. After line features extraction, one of the two original colors will be removed, which helps to ensure the accuracy of fuzzy clustering.
3) In our method, the characteristic of fuzzy clustering in dealing with imbalanced data set (will be introduced in Section 3.2) is used to obtain initial seeds. Thus, contour-lines prototype needs to contain the largest amount of pixels in all linear feature prototypes. Fortunately, in the contour-line densely distribution maps, after area features removing, contour-lines always occupy the largest proportion.

The line extraction method proposed by Miao et al. Qiguang et al. (2013), is based on energy density and the shear transform. The energy density in a grayscale image, is defined as the average energy in an area, can be described by the following formula.

$$E = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} (f(i,j))^2 \tag{1}$$

where E is the energy density, $M \times N$ is the size of the area, and $f(i,j)$ is the gray-value of a pixel in the negative image. Then templates in the horizontal and vertical directions are built to separate lines from background based on energy density. The linear information loss problem, due to the directional limits of lines only separated in one direction image, needs to be addressed. Hence, the shear transform is an affine transform, similar to the rotation transform (Xu et al., 2012), is employed to add the directional characteristics of the lines. Finally, a topographic map only contains lines obtained as shown in Fig. 5(b).

3.2. Fuzzy clustering

Fuzzy c-means (FCM) has many applications in image segmentation. Due to the introduction of fuzzification to every pixel,

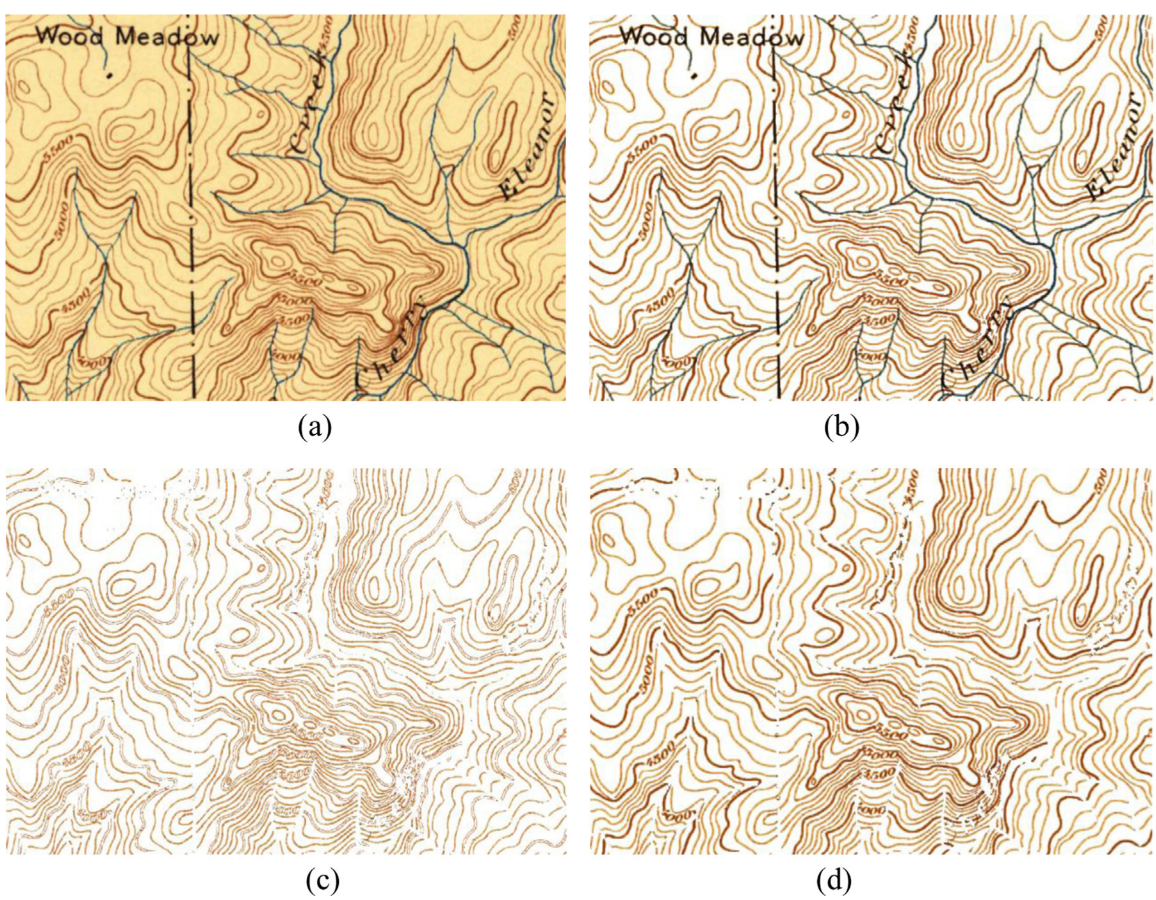

Fig. 5. (a) The original topographic map. (b) The line map obtained using lines extraction method. (c) The contour-line layer result of fuzzy cluster. (d) The final result of our proposed method.

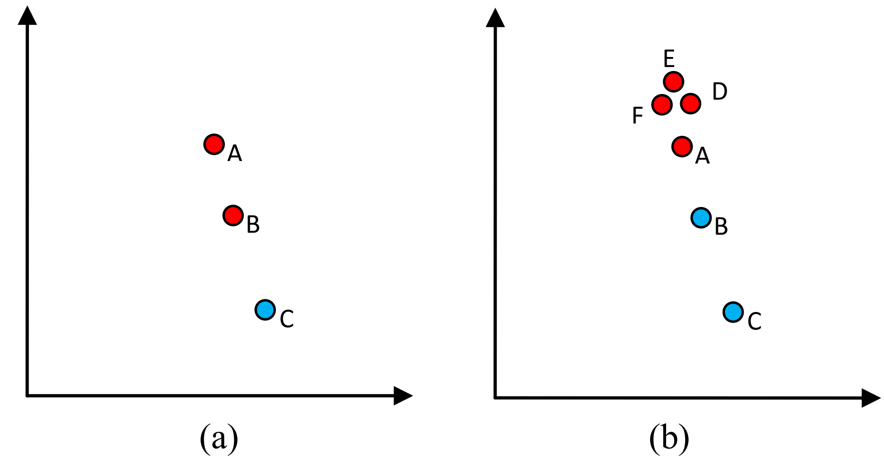

Fig. 6. Points with similarity from MND. (a) Before adding points, B and A are the same class. (b) After adding points, B and C are the same class. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

FCM achieved good performance (Cai et al., 2007; Krinidis and Chatzis, 2010). Compared with hard cluster, FCM keeps more original information of images (Sun et al., 2008). Because the Euclidean distance is employed as distance measure, the segmentation is good only when the data set contains clusters that are well separated or clusters of roughly the same size and shape (Krishnapuram and Kim, 1999). To overcome this shortcoming, Gustafson and Kessel presented the G–K algorithm with the Mahalanobis distance (MD) as the distance measure (Gustafson and Kessel, 1978). G–K algorithm preserves the volume and, hence, is suitable for cases where the data set contains ellipsoidal clusters of similar volume (Krishnapuram and Kim, 1999). As mentioned in Section 2, colors of topographic maps distribute as spheroids in the Lab color space. So, G–K algorithm is adopted in our method.

Assume that $X=(x_1, x_2, \cdots x_n)$ is a data set of p-dimensional feature vectors. Let $V=(v_1, v_2, \cdots v_c)$ represent a c-tuple centers of clusters and $U=(u_{jk})_{c\times n}$ be a fuzzy membership matrix with u_{jk} denoting the grade of membership of feature point x_k in cluster j. The objective function of G–K algorithm can be described as formula (2).

$$\begin{cases} \min J(V, U; X) = \sum_{j=1}^{c}\sum_{k=1}^{n} u_{jk}{}^{m}(x_k - v_j)^T A_j (x_k - v_j) \\ s.\ t.\ \ u_{jk} \in [0, 1]\ \ \forall\, j, k \\ 0 < \sum_{k=1}^{n} u_{jk} < n\ \ \forall\, j \\ \sum_{j=1}^{c} u_{jk} = 1\ \ \forall\, k \end{cases} \tag{2}$$

where m is any real number greater than 1 which determines the fuzziness of the resulting clusters. X contains all pixels of line features in the line map with 3-dimensional color feature vectors, c equals the number of feature layers in the line map, n is the size of X, and m is set as 2 which is widely accepted as a good choice (Hathaway and Bezdek, 2001; Xu and Wunsch, 2005). A_j is a positive definite symmetric matrix defined as formula (3).

$$A_j = \det(\rho_j F_j)^{1/p} F_j^{-1},\ \ \rho_j > 0 \tag{3}$$

where ρ_j is a cluster volume and is simply fixed as 1 for each cluster, which makes the clusters with approximately equal volumes. p is the dimension of feature vectors of data (in our new method the dimension of the Lab color space is 3. F_j is a fuzzy covariance matrix defined as formula (4).

$$F_j = \frac{\sum_{k=1}^{n} u_{jk}{}^{m}(x_k - v_j)(x_k - v_j)^T}{\sum_{k=1}^{n} u_{jk}{}^{m}} \tag{4}$$

The minimization of the objective functional in formula (2) is achieved by the alternating optimization method according to formulas (5) and (6) until the objective functional converge to a stable state.

$$u_{jk} = \frac{[(x_k - v_j)^T A_j (x_k - v_j)]^{1/(1-m)}}{\sum_{i=1}^{c}[(x_k - m_i)^T A_i (x_k - m_i)]^{1/(1-m)}},\ \ k = 1, 2, \cdots, n;\ j = 1, 2, \cdots, c \tag{5}$$

$$v_j = \frac{\sum_{k=1}^{n} u_{jk}{}^{m} x_k}{\sum_{k=1}^{n} u_{jk}{}^{m}},\ \ \mathrm{j} = 1, 2, \cdots, c \tag{6}$$

After clustering, fuzzy membership matrix U is achieved. Meanwhile, classifying each pixel to its most similar prototype, the contour-line layer result of fuzzy clustering is also available as shown in Fig. 5(c).

G–K algorithm is capable of detecting ellipsoidal cloud clusters of dissimilar sizes and orientations to various degrees (Krishnapuram and Kim, 1999), but the distribution of pixels in color space is rather complex. The boundaries between prototypes are fuzzy as analyzed in Section 2. Otherwise, in the map, the amount of pixels in contour-lines is always larger than that in other linear features. All these facts would cause misclassification in clustering results. To explain how this misclassification appears, we use mutual neighbor distance (MND) (Chidananda Gowda and Krishna, 1978) as distance measure and illustrate this situation in Fig. 6. MND is calculated based two values: $NN(a, b)$ is the neighbor number of b with respect to a and $NN(b, a)$ is the neighbor number of a with respect to b, where a and b are two points in the same space. The value of MND is given as formula (7).

$$MND(a, b) = NN(a, b) + NN(b, a) \tag{7}$$

In Fig. 6(a), the nearest neighbor of A is B, and B's nearest neighbor is A. So, $MND(A, B)$ is 2 and, for the same reason, $MND(B, C)$ is 3. Therefore, in Fig. 6(a), point B is more similar to A than C. However, in Fig. 6(b) by adding points D, E and F, $MND(A, B)$ comes to 5, but $MND(B, C)$ is still 3. In this case, point B is more similar to C than A. Similarly, in a line map the contour-line layer could be considered as red group of Fig. 6(b). The pixels of contour-line layer has larger relative frequency compare to that of other layers, thus in color space the contour-line prototype has bigger volume. B in Fig. 6 could represent the mixed color or color aliasing pixels obtained from the intersection or overlapping between contour-line and other features. Although these pixels ought to be part of contour-line layer, they are easily classified into other layers in clustering results.

In general, G–K algorithm could get a good performance in detecting the shape of geographic elements in color space. However, the result prototypes with larger amount of pixels will lose some parts of the boundary pixels. To overcome this shortcoming, a single-prototype region growing method is proposed for contour-line layer extraction.

3.3. Single-prototype region growing for contour-line layer

Seeded Region Growing algorithm was proposed by Adams and Bischof in 1994 (Adams and Bischof, 1994). The application of SRG is in the image segmentation in gray level, and the initial seeds input are needed (Mehnert and Jackway, 1997). Researchers later improved this method and expanded it to color image segmentation (Fan et al., 2001; Shih and Cheng, 2005), medical image segmentation (Hojjatoleslami and Kittler, 1998; Wu et al., 2008), video object analysis (Grinias and Tziritas, 2001) and 3-D image segmentation (Revol-Muller et al., 2002). Based on SRG, Leyk applied homogeneity in color topographic maps segmentation (Leyk and Boesch, 2010). This method makes full use of plane relation information, frequency of homogeneous pixels and color space information to segment map and produces acceptable results. It is dynamic in solving discrimination problems and it can prevent over-segmentation. However, as prior described, this method is directed at all color layers extraction. Thus, accuracy segmentation results in all layers are extremely difficult in this method, especially when the geographic element distributes complexity in the situation of contour-line in the maps which we are processing. To solve this issue, we apply a specific SRG to obtain accuracy contour-line layer as follows.

Based on fuzzy membership matrix U, the SRGCL consists of four steps of contour-line layer automatic identification, initialization of seeds, single-prototype region growing and post-processing. U denotes the grades of membership of each pixel in each cluster. Generally, the cluster that has the biggest grade is regarded as the ascription of each pixel. However, as mentioned before, color aliasing and mixed color will cause inaccurate segmentation in clustering. Hence, the second-similar prototype is employed in region growing. Every pixel x_k should hold a most similar prototype pm_k and a second similar prototype ps_k, which are defined in formulas (8) and (9) respectively.

$$pm_k = \{p|u_{pk} = \max(u_{jk}), j = 1, \cdots c\} \tag{8}$$

$$ps_k = \{q|u_{qk} = \max(u_{jk}), j = 1, \cdots c \text{ and } j \neq p\} \tag{9}$$

where c is the number of prototypes. Thus a most similar prototype set $PM = \{pm_1, pm_2, \cdots pm_n\}$ as well as a second-similar prototype set $PS = \{ps_1, ps_2, \cdots ps_n\}$ are obtained, where n is the number of pixels. As analyzed previously, most similar prototype alone could lose some mixed color and color aliasing pixels because of imbalance data distribution. Thus, we employ the second-similar prototype which is used to discover potential contour-line pixels. These potential pixels may lose their original membership of contour-line prototype during clustering. Hypothetically in Fig. 6, B represents a contour-line pixel (red color in Fig. 6a) which is influenced because of color aliasing or mixed color. With the increasing amount of contour-line pixels, B is easier to be classified into other layer (blue color in Fig. 6b). But it still has some similarity to contour-line prototype as described in Fig. 2. Thus, its second-similar prototype should be contour-line. Based on all of the above analysis, along with the combination of plane relation information contained in region growing process, these potential contour-line pixels could return to their original prototype which is the contour-line layer.

a) Contour-line layer automatic identification
For the random value of initial cluster centers and the complex computing, it is difficult to point out which prototype corresponds to the contour-line layer. However, for contour-lines of the maps which we focused on, due to their wide distribution they occupy the largest proportion in a line map, an automatic contour-line layer identify method is introduced here. The main idea of this method to first choose n rows and n column of line map randomly; then check corresponding pm_k of pixels lie on these lines to find out the prototype cp with the largest proportion, which is the prototype of contour-line layer. This entire ideal can be described by formula (10).

$$\begin{aligned} &PM' = \{pm_k\} \\ &cp = \{i| \neg \exists \ \ j,\ crad(PM' = j) > crad(PM' = i)\ i, j = 1, 2\cdots c\} \end{aligned} \tag{10}$$

where k equals the indexes of pixels lie on the random chosen lines, $crad(A)$ equals the number of elements in set A.

b) Initialization of seeds
This step has to have a sufficient number of initial seeds that SRGCL can successfully work over the maps. Meanwhile, the initial seeds must be pure enough that it only contains pixels which belong to the contour-line prototype. If a pixel k and the majority of the neighbors around it show the same most similar prototype as cp, then k is registered as an initial seed S_{cp} of contour-line prototype, thus it represents one starting pixel for SRGCL. The initial seed S_{cp} can be achieved by formula (11).

$$S_{cp} = \{k|Pm_k = cp \text{ and } crad(Pm_{\bar{k}} = cp) > \delta\} \tag{11}$$

where $\bar{k}$ is the neighbor of k, δ describes the threshold of a majority. As a majority, δ should be larger than 4 to guarantee the reliability of initial seeds. Otherwise, the selected seeds might not belong to contour-line layer. A over strict threshold like 7 or 8 could cause the quantity of initial seeds to be smaller, which then leads to higher iterations. However, this parameter has limited influence to the contour-line separation result. That is because the growing process will offset the missing of initial seeds by costing a few more iterations.

c) Single-prototype region growing method
Based on the contour-line prototype cp and initial seeds mentioned above, a single-prototype region growing method is introduced here.

Besides PM and PS, a label set $\{mr_k\}$ for denoting whether a pixel has been classified is needed. If a pixel has been classified to a prototype, mr_k is set to 1, otherwise, set mr_k to 0. An iterative process is described as follow.

Step 1 Starting at the initial seeds S_{cp} and then search the connected $m \times m$ neighbors for the pixels have the same pm. If the neighbors have not been classified and they have the same pm with S_{cp}, these pixels are registered as new seeds of the contour-line layer, as described in formula (12).

$$k \in S_{cp} \Leftarrow \ \forall\ k\ (pm_k = cp,\ mr_k = 0) \text{ and } (\ \exists\ \bar{k},\ pm_{\bar{k}} = cp) \tag{12}$$

Step 2 Beginning with the seeds S_{cp} of contour-line layer and search their 3×3 neighbors for the pixels whose second similar prototype is cp. These detected pixels are considered as new seeds of contour-line layer.

$$k \in S_{cp} \Leftarrow \ \forall\ k\ (ps_k = cp) \text{ and } (\ \exists\ \bar{k},\ \bar{k} \in S_{cp}) \tag{13}$$

Step 3 Repeat Step 1 and Step 2 until there is no change in S_{cp}.

Most similar prototype information will guide the region growing process in Step 1. This step allows the pixels with high similarity to contour-line prototype to become a part of contour-

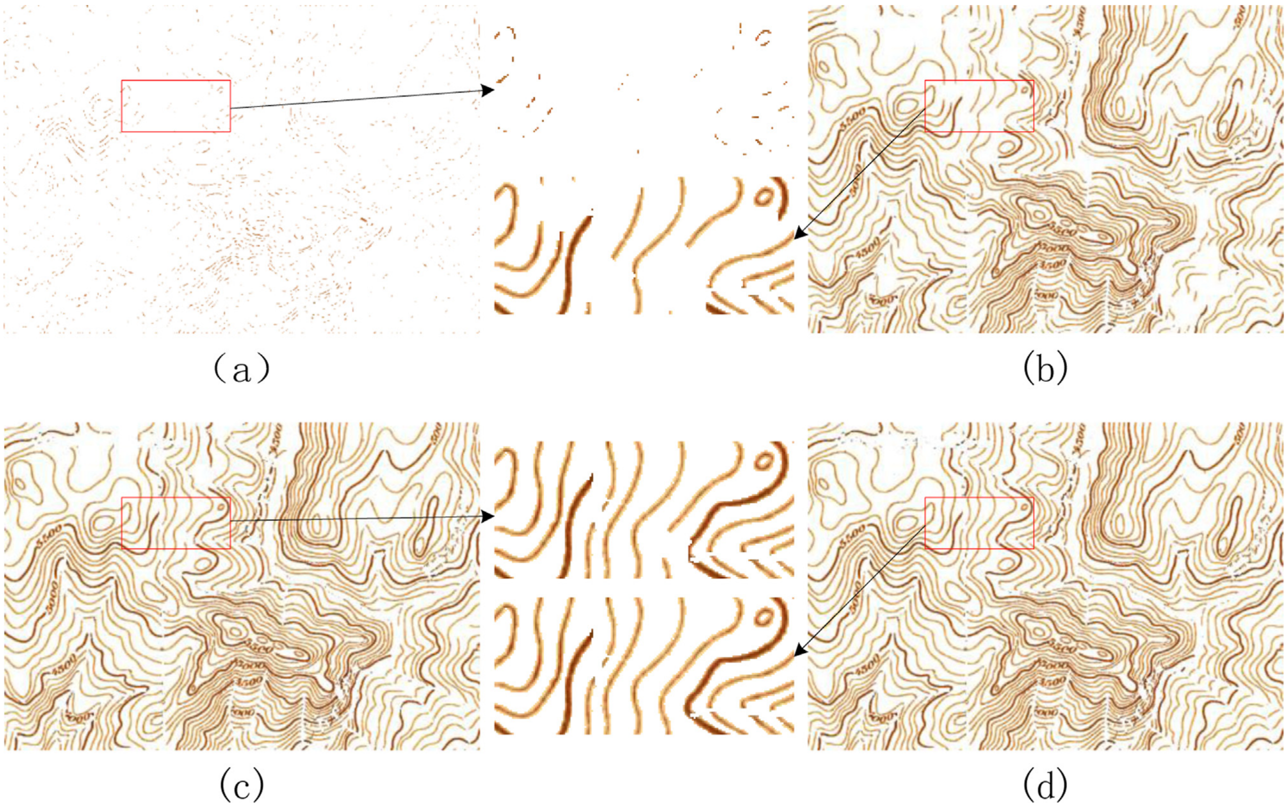

Fig. 7. (a) The initial seeds. (b) The iteration number equals 3. (c) The iteration number equals 6. (d) The final result of region growing and the iteration number equals 12. The middle column enlarges the same part of contour-line layer in different iteration.

line layer. With a broad neighbor size m, almost all the pixels have high similarity to contour-line prototype could be contained except false color pixels and noise pixels which are far away from contour-lines in map. The second similar prototype information in Step 2 makes the pixels with lower similarity have the opportunity to become a part of contour-line layer. Meanwhile, small neighbor size in Step 2 ensure the grown pixels are surrounding the seeds, which can get missed mixed color pixels and color aliasing pixels back to contour-line layer without bringing in other layers' pixels. In order to explain the growing process more clearly, we use the intersection region of contour-line and river as an example. This intersecting patch contains mixed color and color aliasing pixels. As previous analysis, the mixed color and color aliasing will show similarities to both contour-line and river but in different degrees. If a pixel in this patch shows the most similarity to contour-line, it will grow into contour-line layer in Step 1. Otherwise, its second-similar prototype should be contour-line, which will cause it grow into contour-line layer in Step 2. Thus, this growing method can successfully solve the mixed color and color aliasing problem even in intersection region and regardless of the size of intersecting patches.

Illustration of the SRGCL is shown in Fig. 7. Fig. 7(a) shows the initial seeds come from Fig. 5(c). Starting with these seeds, the iteration of growing method continues until the whole growing process finishes. As the enlarged portion shown, contour-lines are growing step by step.

a) Post processing

The region growing is specific to contour-line and it makes the contour-line layer complete. Thus, the unlabeled pixels, with the most similar prototype is cp, can be regarded as incorrect classified pixels in the fuzzy clustering. Therefore, these pixels can be ignored in this step. For the size of neighborhoods in region growing set by prior knowledge, some noise could emerge because of the false growing which resulting from a big neighborhood size m. So in the post processing, small connected regions are removed from the contour-line layer to fix this problem. The way to judge whether a region is a small connected region is counting its pixel amount S_n. According to our experiments, an over large threshold could cause missing information of the contour-line and a small one could not fix the noise problem in most cases. Thus, we set it as 5 in our method.

Finally, the final result of contour-line layer is obtained as shown in Fig. 5(d). It can be seen that after several iterations, most contour-lines are growing to complete lines.

4. Experimental results and analysis

In this Section, 5 different pages of topographic maps from USGS historical maps[1] are used to examine the proposed method. We use G–K method and Leyk's method as benchmarks to examine how much the new method could gain in accuracy of contour-line extracted through specializing on the contour-line layer. Both basic methods apply the same post processing algorithm to improve their results.

4.1. Measurement

In these experiments, we adopt the classic evaluation metrics, precision-recall and F1-measure (Martin et al., 2004), to measure the separated contour-line layer quality. Precision is the fraction of segmented results that are true positives rather than false positives, while recall is the fraction of true positives that are obtained by segmented results rather than missed. In other word, precision is the percent of valid pixels in segmented layers, and recall is the

[1] http://nationalmap.gov/historical/.

Table 1
Iteration number and evaluated results of different majority thresholds.

Majority threshold	5	6	7	8
Iteration number	13	14	15	15
F1-measure	0.8153	0.8153	0.8153	0.8153

percent of detected ground truth pixels in all ground truth pixels. F1-measure is a harmonic mean of precision and recall which could comprehensively assesses the segmentation quality of each method. The definition of F1-measure is shown in formula (14).

$$F1 - \text{measure} = 2\cdot\frac{\text{precision}\cdot\text{recall}}{\text{precision} + \text{recall}} \tag{14}$$

4.2. Parameters

In the fuzzy clustering step of SRGCL, all the parameters are set as default introduced in Section 3.2. In the seed initialization, the majority threshold δ is set as 5. To varify the good robustness of this parameter visualized, several values are tested on Fig. 5(a). Iteration number and evaluated results are shown in Table 1. It can be seen that F1-measure has no changes with different δ from 5 to 8.

In the region growing process, the neighborhood size m needs to be adjusted which will be indicated in each experimental group. And in the post processing, the noise size S_n is set as 5. The sensibility of the above two parameters will be discussed later.

4.3. Experiments

Continuation is a very important request in contour-line extraction. Gaps that exist in contour-line layer after segmentation will bring huge difficulties to contour-line extraction. Taking full account of relationships in color space and plane space among pixels, the contour-line color layer obtained by SRGCL has good performance in continuation, which can be verified in Fig. 8. As a result of the color aliasing and mixed color, parts of pixels on contour-line are lost in the result of G–K method. But after adding the special region growing procedure in the SRGCL algorithm, most of them grow back thus to insure the continuation of contour-lines, which is indicated in red rectangles in Fig. 8(c) and (e). In addition, SRGCL has big improvements in continuation relative to the SRG algorithm based on homogeneity, which is indicated with red rectangles in Fig. 8(d) and (e). In Fig. 8(d), huge gaps appear in contour-lines which almost lead to some of the contour-lines disappearing. This is caused by inaccuracy initial seeds and growing order dependencies problem. However, through using fuzzy clustering to initial seeds and focus on contour-line layer, these problems are overcome in SRGCL.

Besides gaps, noise is another serious difficulty for the extraction of contour-lines. The SRG algorithm based on homogeneity gets initial seeds by peak finding, which can consider all layers but is highly depended on the threshold. Meanwhile, the sphere shape of initial seeds in color space could easily cause the inaccuracy of prototype. So, some noise exists in the contour-line layer as indicated with blue ellipses in Fig. 8(d). SRGCL changes the way of initializing seeds by using fuzzy clustering result, which can fully consider all pixels in topographic map. In addition, ellipsoidal shape of initial seeds in color space could achieve accurate prototype. Thus, the results of SRGCL have less noise which can be seen from the corresponding area surrounded by blue ellipses in Fig. 8(e). In order to quantitatively evaluate the improvement of SRGCL relative to G–K method and Leyk's method, we used these basic methods as benchmark to show the increase of SRGCL in

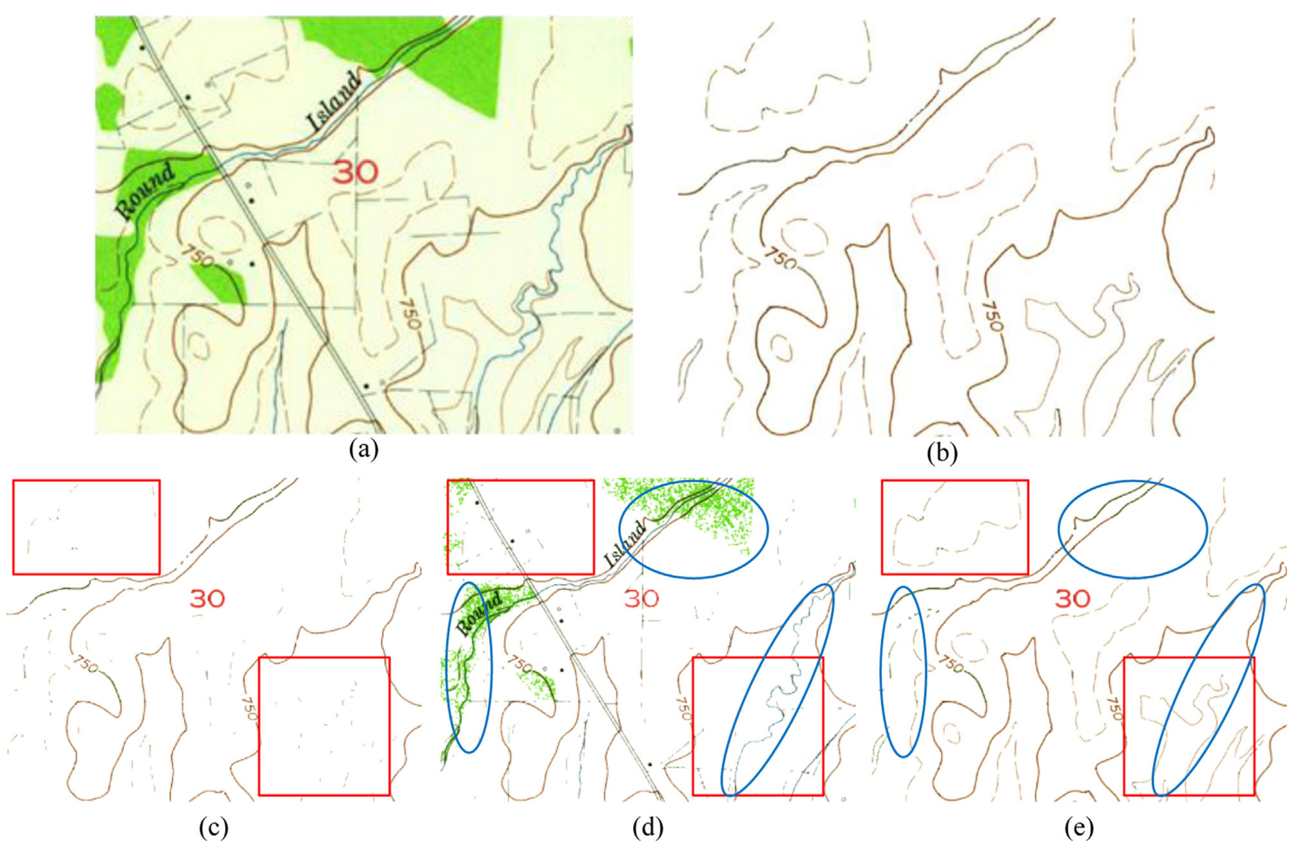

Fig. 8. (a) Original image of a map form 1995. (b) Ground truth of contour-line layer. (c) The contour-lines segmentation results achieved by the G–K algorithm. (d) The contour-lines segmentation results achieved by the SRG algorithm based on homogeneity. (e) The contour-lines segmentation results achieved by SRGCL which set m as 11. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

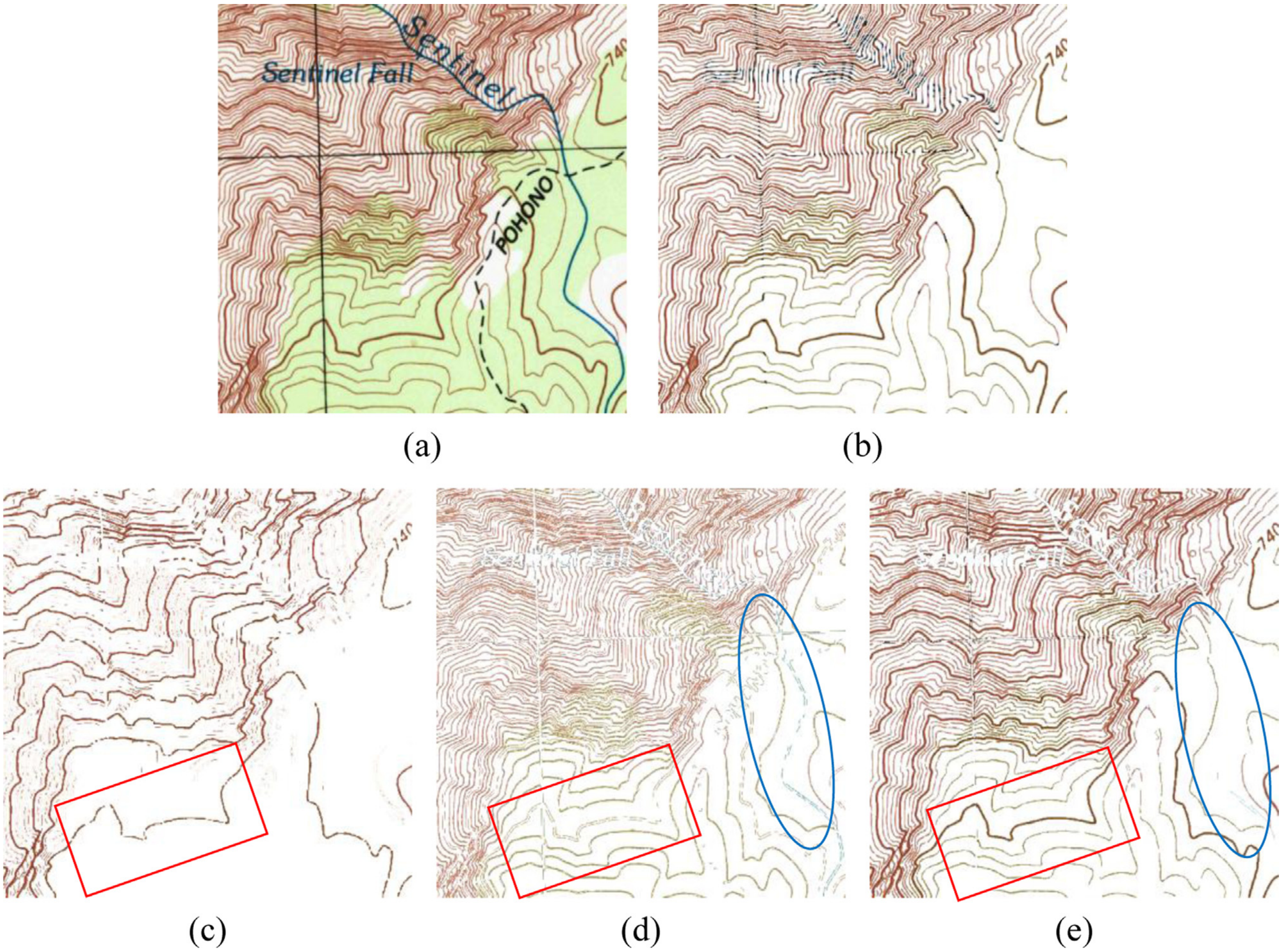

Fig. 9. (a) Original image of a map form 1997. (b) Ground truth of contour-line layer. (c) The contour-lines segmentation results achieved by the G–K algorithm. (d) The contour-lines segmentation results achieved by the SRG algorithm based on homogeneity. (e) The contour-lines segmentation results achieved by SRGCL which set m as 11. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

precision, recall and F1-measure. The quantified improvements of separated results are shown in Table 2.

These four groups of experiments (Figs. 9–12) above further show the improvements of SRGCL in continuation (red rectangles) and with freedom from noise (blue ellipses). The quantified improvements of separated results are shown in Table 2. Although SRGCL does not improve all precisions and recalls in all cases, it shows obvious improvement in F1-measures.

Since SRGCL uses the G–K results as the basis of initial seeds, it has stable improvement in recall relative to G–K method. This is because the growing process could keep causing the boundary pixels into contour-line layer. Thus the results of SRGCL have higher degree in completeness and continuity. Since the contour-line layers of SRGCL have more true positive pixels than that of G–K method, the ratio of true positive pixels is increased which leads to SRGCL improvement in precision. In some cases, the boundary pixels of other geographic elements happen to have high levels of similarity contour-line layer. As in Fig. 10, the background has very similar color to contour-line. Meanwhile the rivers are drawn in dark blue, which leads to the mixed color and color aliasing pixels of these two elements have high similarity to contour-lines. The boundary pixels adjacent to contour-lines are falsely segmented into contour-line layer and the precision of SRGCL is decreased relative to G–K algorithm. However, the amount of falsely segmented pixels is very small, which barely influence the separated results of SRGCL. This point can be seen from both the image and F1-measure of Fig. 10 in Table 2. In Figs. 8 and 9, the mixed color and color aliasing problem are serious because of low map quality and close space between elements. This makes the G–K clustering results miss lots of contour-line pixels, consequently the improvement in recall of SRGCL appears a lot in these two figures.

SRGCL has stable improvement in precision relative to Leyk's method, which can be seen from the images and Table 2. This improvement demonstrates further significant in Fig. 8. The color

Table 2
Quantified improvements of SRGCL relative to basic methods relative to the contour-line portion produced by the two basic methods.

		Fig. 8 (%)	Fig. 9 (%)	Fig. 10 (%)	Fig. 11 (%)	Fig. 12 (%)
G–K	Precision	3.6	47.3	−4.5	183.6	0.2
	Recall	88.2	157.1	46.6	69.3	25.8
	F1-measure	62.2	106.6	21.7	123.3	14.5
Leyk's method	Precision	205.9	8.7	9.0	9.7	16.5
	Recall	49.3	24.1	27.5	−7.5	12.9
	F1-measure	97.3%	17.0	18.5	0.6	14.5

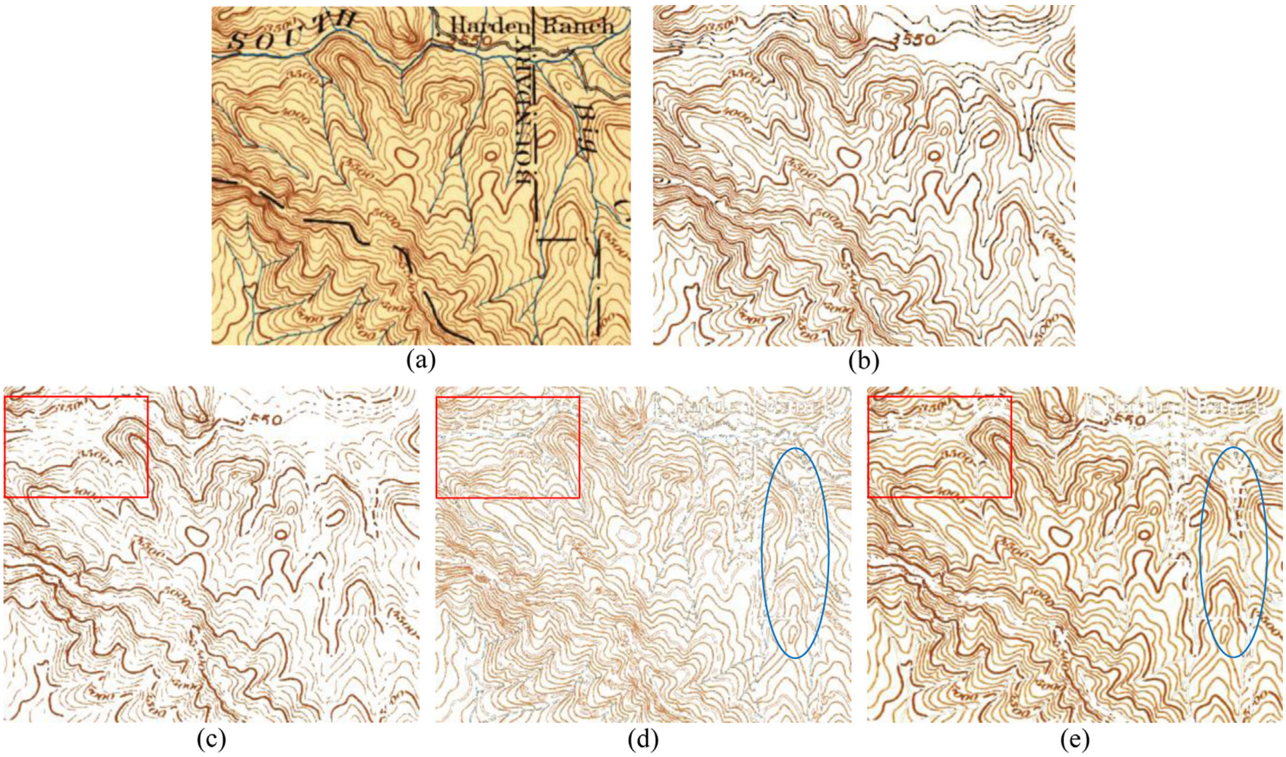

Fig. 10. (a) Original image of a map form 1897. (b) Ground truth of contour-line layer. (c) The contour-lines segmentation results achieved by the G–K algorithm. (d) The contour-lines segmentation results achieved by the SRG algorithm based on homogeneity. (e) The contour-lines segmentation results achieved by SRGCL which set m as 5. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

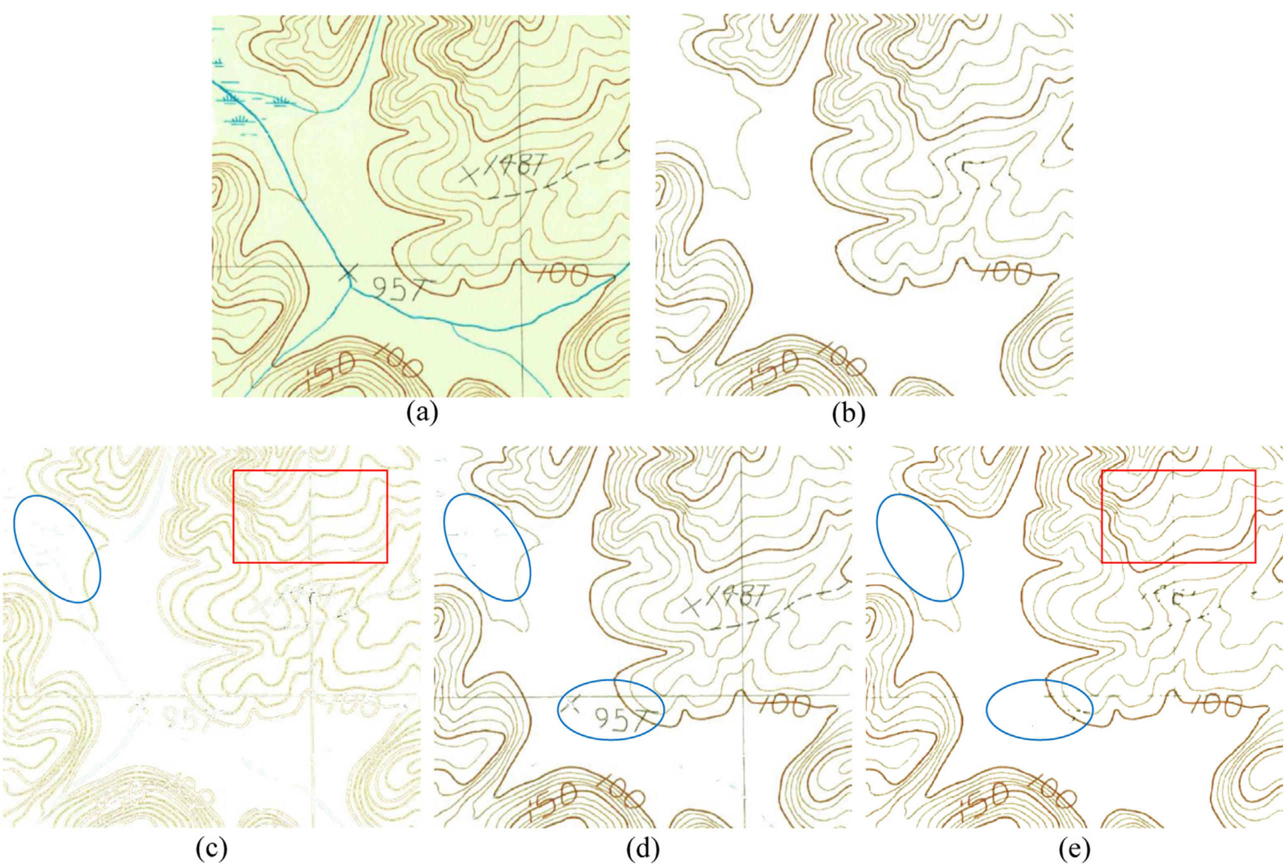

Fig. 11. (a) Original image of a map form 1983. (b) Ground truth of contour-line layer. (c) The contour-lines segmentation results achieved by the G–K algorithm. (d) The contour-lines segmentation results achieved by the SRG algorithm based on homogeneity. (e) The contour-lines segmentation results achieved by SRGCL which set m as 7. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

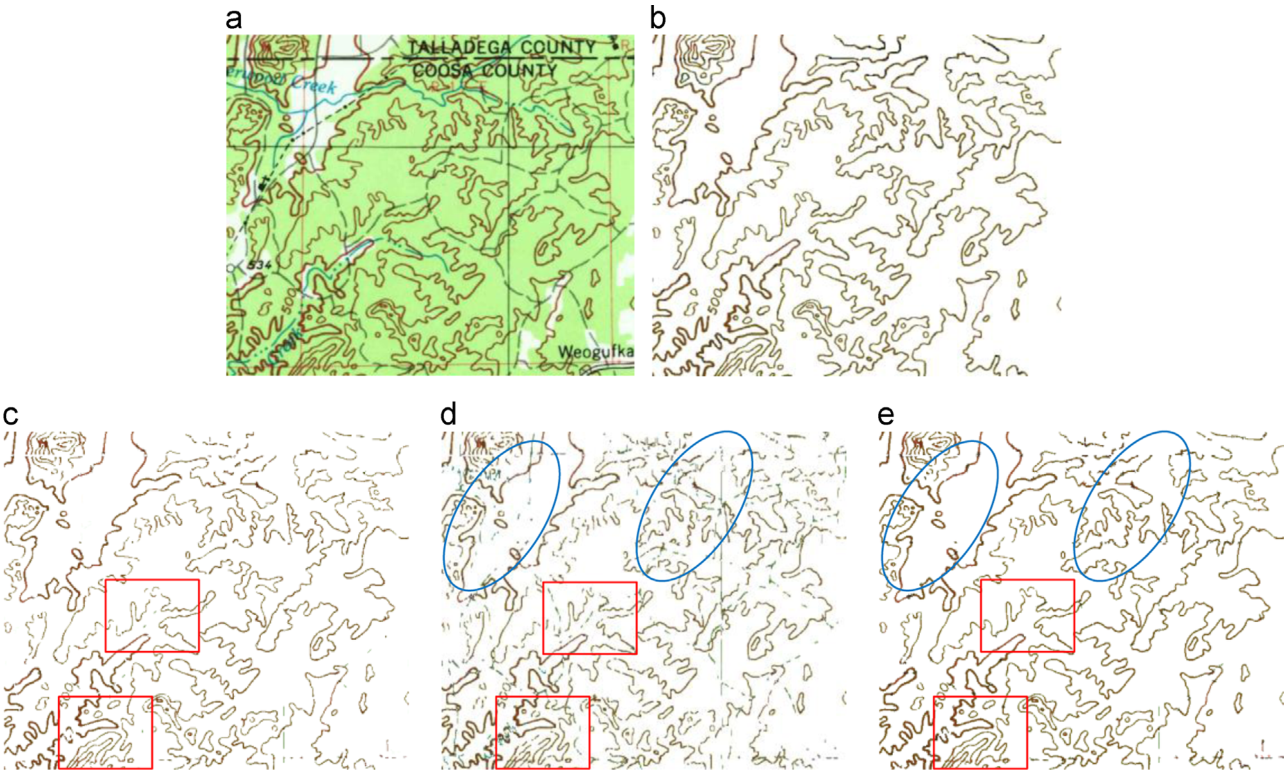

Fig. 12. (a) Original image of a map form 1969. (b) Ground truth of contour-line layer. (c) The contour-lines segmentation results achieved by the G–K algorithm. (d) The contour-lines segmentation results achieved by the SRG algorithm based on homogeneity. (e) The contour-lines segmentation results achieved by SRGCL which set m as 11. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3
Evaluation on different neighborhood sizes.

m	1	3	5	7	9	11	13	15	17	19
Precision	0.9773	0.9758	0.9753	0.9748	0.9738	0.9738	0.9736	0.9734	0.9734	0.9734
Recall	0.7392	0.7573	0.7628	0.7629	0.7639	0.7639	0.7639	0.7639	0.7639	0.7639
F1-measure	0.8417	0.8528	0.8560	0.8559	0.8562	0.8562	0.8561	0.8560	0.8560	0.8560

Table 4
Evaluation on different noise sizes.

S_n	1	2	3	4	5	6	7	8	9	10
Precision	0.9726	0.9733	0.9734	0.9737	0.9738	0.9738	0.9742	0.9749	0.9754	0.9764
Recall	0.7645	0.7643	0.7642	0.7641	0.7639	0.7639	0.7638	0.7631	0.7628	0.7622
F1-measure	0.8561	0.8563	0.8562	0.8563	0.8562	0.8562	0.8562	0.8561	0.8561	0.8561

distribution of Fig. 8 in color space is rather uniform, thus it is very difficult to select an ideal threshold for initializing seeds of all color layers in Leyk's method. This is the reason that large redundancy appears in its contour-line layer. In SRGCL, the initial seeds are selected based on the clustering results, which are determined by the data itself and show the essence of the data. Additionally, SRGCL focus on only contour-line layer, which could guarantee the accuracy of the separation results. All of the above lead to significant improvement in precision. Meanwhile, other figures can verify this improvement in the same way. SRGCL not only increases the ratio but also increases the amount of true positive pixels, which result in the increasing recall of contour-line layer separation results. It is worth of being noticed that the recall in Fig. 11 is decreased which is because Leyk's method indeed has good performance in this type of good quality maps. The color distribution of different layers in color space is well separated, which make it easy to select threshold. However, the image shows some texts of other layers are also included in the contour-line layer of Leyk's result, which can be reflected through precision and F1-measure as well.

From the evaluations in both visualization and quantification we can see, SRGCL will significantly improve the continuation of lines and the ability of lessening noise in contour-line layer separation relative to the other two basic CIS methods.

In order to test the sensibility of neighborhood size m and noise size S_n, various values of these parameters are applied in SRGCL. The examined map is Fig. 9 and the resulting evaluations are shown in Tables 3 and 4. To test m as well as S_n, we fix other parameters as the default values. It can be seen from these tables that there are very small changes through very wide ranges in both m and S_n, which indicate the strong robustness of these parameters.

There also exist some shortages in our method. For instance, when some features are represented by the color which is similar to the color of contour-line, it is hard to separate these feature from contour-line in fuzzy clustering, and it will result in false segmentation just like '30' in Fig. 8 and '148T' in Fig. 11. Besides, the cost of achieving good contour-line extraction results in our method is, losing the ability of extracting other layers in a topographic map. In addition, in this paper we only test USGS series maps in which contour-lines are the largest proportion of line features. Thus, other series topographic maps and the maps in which contour-lines are not the largest proportion of line features should be considered in the future research.

5. Conclusions

This paper presents a new algorithm SRGCL for contour-line color layer separation in topographic maps which contain wide and dense contour-lines. Different from general topographic map segmentation methods, SRGCL only focus on contour-line layer extraction. However, this limitation on other layers segmentation has an exchange of more accurate contour-line layer segmentation results. This method possesses nice properties of fuzzy clustering and single-prototype region growing to separate contour-line color layer specifically. By introducing the planar spatial relationship into images, the proposed method can solve the problem of inaccuracy raised by mixed color, color aliasing and false color in fuzzy clustering. Moreover, by taking advantage of unsupervised cluster and SRGCL, both problems in initial seeds selection and in order dependencies of SRG are solved. The results obtained by our method have excellent performance in continuation and with freedom of noise, which have been proven by experiments. In addition, our method shows good robustness with only a few parameters.

Acknowledgments

The work was jointly supported by the National Natural Science Foundation of China under grant Nos. 61472302, 61272280, U1404620 and 41271447; The Program for New Century Excellent Talents in University under grant No. NCET-12-0919; The Fundamental Research Funds for the Central Universities under grant Nos. K5051203020, K5051303018, JB150313 and BDY081422; Natural Science Foundation of Shaanxi Province, under grant Nos. 2014JM8310 and 2010JM8027; The Creative Project of the Science and Technology State of xi'an under grant No. CXY1441(1); The State Key Laboratory of Geo-information Engineering under grant No. SKLGIE2014-M-4-4. We are grateful to T.Y. Wright-Kenney for her proof-reading and editing.

References

Adams, R., Bischof, L., 1994. Seeded region growing. Pattern Anal. Mach. Intell. IEEE Trans. 16 (6), 641–647.
Chiang, Y.Y., Leyk, S., Knoblock, C.A., 2013. Efficient and Robust Graphics Recognition from Historical Maps. Graphics Recognition. New Trends and Challenges. Springer, Berlin Heidelberg, pp. 25–35.
Chen, Y., Wang, R., Qian, J., 2006. Extracting contour-lines from common-conditioned topographic maps. Geosci. Remote. Sens. IEEE Trans. 44 (4), 1048–1057.
Chiang, Y.Y., Leyk, S., Knoblock, C.A., 2014. A survey of digital map processing techniques. ACM Comput. Surv. (CSUR) 47 (1), 1.
Chiang, Y.Y., Knoblock, C.A., Shahabi, C., et al., 2009. Automatic and accurate extraction of road intersections from raster maps. GeoInformatica 13 (2), 121–157.
Cao, R., Tan, C.L., 2002. Text/Graphics Separation in Maps. Graphics Recognition Algorithms and Applications. Springer, Berlin Heidelberg, pp. 167–177.
Chidananda Gowda, K., Krishna, G., 1978. Agglomerative clustering using the concept of mutual nearest neighbourhood. Pattern Recognit. 10 (2), 105–112.
Cai, W., Chen, S., Zhang, D., 2007. Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation. Pattern Recognit. 40 (3), 825–838.
Feng, Y., Song, E., 1996. An algorithm to gather colors from image of color maps. J. Softw. 7 (8), 466–470.
Fan, J., Yau, D.K.Y., Elmagarmid, A.K., Aref, W.G., 2001. Automatic image segmentation by integrating color-edge extraction and seeded region growing. Image Process. IEEE Trans. 10 (10), 1454–1466.
Gamba, P., Mecocci, A., 1999. Perceptual grouping for symbol chain tracking in digitized topographic maps. Pattern Recognit. Lett. 20 (4), 355–365.
Gustafson, D.E., Kessel, W.C., 1978. Fuzzy clustering with a fuzzy covariance matrix [C]. In: Proceedings of the 17th Symposium on Adaptive Processes, IEEE Conference on Decision and Control including. 17, pp. 761–766.
Grinias, Ilias, Tziritas, Georgios, 2001. A semi-automatic seeded region growing algorithm for video object localization and tracking. Signal Process.: Image Commun. 16 (10), 977–986.
Hathaway, R.J., Bezdek, J.C., 2001. Fuzzy c-means clustering of incomplete data. Systems, Man. Cybern. B: Cybern. IEEE Trans. 31 (5), 735–744.
Hojjatoleslami, S.A., Kittler, Josef, 1998. Region growing: a new approach. Image Process. IEEE Trans. 7 (7), 1079–1084.
Khotanzad, A., Zink, E., 2003. Contour-line and geographic feature extraction from USGS color topographical paper maps. Pattern Anal. Mach. Intell. IEEE Trans. 25 (1), 18–31.
Ke, W., Changde, L., Wande, Y., et al., 2004. Color harmony system based on lab perceptual uniform color space. J. Northwest. Polytech. Univ. 22 (6), 695–699.
Krinidis, S., Chatzis, V., 2010. A robust fuzzy local information C-means clustering algorithm. Image Process. IEEE Trans. 19 (5), 1328–1337.
Krishnapuram, R., Kim, J., 1999. A note on the Gustafson-Kessel and adaptive fuzzy clustering algorithms. Fuzzy Syst. IEEE Trans. 7 (4), 453–461.
Leyk, S., Boesch, R., Weibel, R., 2006. Saliency and semantic processing: extracting forest cover from historical topographic maps. Pattern Recognit. 39 (5), 953–968.
Leyk, S., 2010. Segmentation of colour layers in historical maps based on hierarchical colour sampling. Graphics Recognition. Achievements, Challenges, and Evolution. Springer, Berlin Heidelberg, pp. 231–241.
Leyk, S., Boesch, R., 2010. Colors of the past: color image segmentation in historical topographic maps based on homogeneity. GeoInformatica 14 (1), 1–21.
Mehnert, Andrew, Jackway, Paul, 1997. An improved seeded region growing algorithm. Pattern Recognit. Lett. 18 (10), 1065–1071.
Martin, D.R., Fowlkes, C.C., Malik, J., 2004. Learning to detect natural image boundaries using local brightness, color, and texture cues. Pattern Anal. Mach. Intell. IEEE Trans. 26 (5), 530–549.
Pouderoux, J., Gonzato, J.C., Pereira. A., et al., 2007. Toponym recognition in scanned color topographic maps[C]. In: Proceedings of the Ninth International Conference on Document Analysis and Recognition, IEEE. 1, pp. 531–535.
Qiguang, Miao, Xu, Pengfei, Liu, Tiange, Yang, Yun, Zhang, Junying, Li, Weisheng, 2013. Linear feature separation from topographic maps using energy density and the shear transform. Image Process. IEEE Trans. 22 (4), 1548–1558.
Revol-Muller, Chantal, Peyrin, Francoise, Carrillon, Yannick, Odet, Christophe, 2002. Automated 3D region growing algorithm based on an assessment function. Pattern Recognit. Lett. 23 (1), 137–150.
San, L.M., Yatim, S.M., Sheriff, N.A.M., et al., 2004. Extracting contour-lines from scanned topographic maps[C]. In: Proceedings of the International Conference on IEEE Computer Graphics, Imaging and Visualization, pp. 187–192.
Salvatore, S., Guitton, P., 2004. Contour-line recognition from scanned topographic maps. J. WSCG 12 (1–3), 419–426.
Samet, R., Hancer, E., 2012. A new approach to the reconstruction of contour-lines extracted from topographic maps. J. Vis. Commun. Image Represent. 23 (4), 642–647.
Shih, F.Y., Cheng, S., 2005. Automatic seeded region growing for color image segmentation. Image Vis. Comput. 23 (10), 877–886.
Sun, J.G., Liu, J., Zhao, L.Y., 2008. Clustering algorithms research. J. Softw. 19 (1), 48–61.
Wu, R.Q., Cheng, X.R., Yang, C.J., 2009. Extracting contour-lines from topographic maps based on cartography and graphics knowledge. J. Comput. Sci. Technol. 9 (2), 58–64.
Wu, J., Yan, H., Chalmers, A.N., 1994. Color image segmentation using fuzzy clustering and supervised learning. J. Electron. Imaging 3 (4), 397–403.
Wu, J., Poehlman, S., Noseworthy, M.D., Kamath, M.V., 2008. Texture feature based automated seeded region growing in abdominal MRI segmentation. Int. Conf. Biomed. Eng. Inform. BMEI 2, 263–267.
Xin, D., Zhou, X., Zheng, H., 2006. Contour-line extraction from paper-based topographic maps. J. Inf. Comput. Sci. 1 (5), 275–283.
Xu, P., Miao, Q., Shi, C., et al., 2012. An edge detection algorithm based on the multi-direction shear transform. J. Vis. Commun. Image Represent. 23 (5), 827–833.
Xu, R., Wunsch, D., 2005. Survey of clustering algorithms. Neural Netw., IEEE Trans. 16 (3), 645–678.
Zheng, Huali, Zhou, Xianzhong, Wang, Jianyu, 2003. Research and implementation of automatic color segmentation algorithm for scanned color maps. J. Comput. Aided Des. Comput. Graph. 15 (1), 29–33.