Article

Alike Scene Retrieval from Land-Cover Products Based on the Label Co-Occurrence Matrix (LCM) †

Jun Liu, Bin Luo *, Qianqing Qin and Guopeng Yang

State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan 430079, China; liujunand@whu.edu.cn (J.L.); qqqin@lmars.whu.edu.cn (Q.Q.); yangguopeng@hotmail.com (G.Y.)
* Correspondence: luob@whu.edu.cn; Tel.: +86-186-2785-3175
† This paper is an extended version of our paper published in 2014 IEEE Geoscience and Remote Sensing Symposium, a Content Based Map Retrieval System for Land Cover Data, Jun Liu, Bin Luo, and Liangpei Zhang.

Abstract: The management and application of remotely sensed data has become much more difficult due to the dramatically growing volume of remotely sensed imagery. To address this issue, content-based image retrieval (CBIR) has been applied to remote sensing image retrieval for information mining. As a consequence of the growing volume of remotely sensed imagery, the number of different types of image-derived products (such as land use/land cover (LULC) databases) is also increasing rapidly. Nevertheless, only a few studies have addressed the exploration and information mining of these products. In this letter, for the sake of making the most use of the LULC map, we propose an approach for the retrieval of alike scenes from it. Based on the proposed approach, we design a content-based map retrieval (CBMR) system for LULC. The main contributions of our work are listed below. Firstly, the proposed system can allow the user to select a region of interest as the reference scene with variable shape and size. In contrast, in the traditional CBIR/CBMR systems, the region of interest is usually of a fixed size, which is equal to the size of the analysis window for extracting features. In addition, the user can acquire various retrieval results by specifying the corresponding parameters. Finally, by combining the signatures in the base signature library, the user can acquire the retrieval result faster .

Keywords: content-based map retrieval; land-cover datasets; similarity

1. Introduction

Alongside the rapid development of remote sensing platforms and sensors, the volume of remotely sensed imagery has also tremendously increased. Because of the large data volumes, the exploration and information mining from remote sensing archives is becoming increasingly difficult. Content-based image retrieval (CBIR) is one of the most active research areas in computer vision. CBIR is also widely used in remote sensing image retrieval, as it does not require the presence of semantic tags, which are rarely available and expensive to assign. In order to improve the situation of the low degree of utilization of the existing large archives of remotely sensed imagery, many studies [1–6] have addressed the application of CBIR to remotely sensed imagery. Li et al. [7] proposed a remote sensing image retrieval approach by adopting convolutional neural networks to extract unsupervised features. Demir and Bruzzone [8,9] introduced the hashing methods for large-scale remote sensing (RS) retrieval problems to provide highly time-efficient and accurate search capability within huge data archives. Aptoula [10,11] applied global morphological texture descriptors to the problem of content-based remote sensing image retrieval. Local description strategies and visual vocabularies were widely adopted by remote sensing content-based retrieval and scene classification [12–20].

To better manage remotely sensed data, Alonso et al. [21,22] designed a system and tools for effectively querying and understanding the available data archives. To improve Earth observation (EO) image retrieval results, Research projects like EOLib [23] and TELEIOS [24] have introduced the use of EO image metadata and linked data as query parameters. Along with EO image analysis data, an information layer extracted from OpenStreetMaps was used in the learning stage of the retrieval system [25]. There are also systems for the retrieval and analysis of remote sensing images time series, such as knowledge-driven information mining (KIM) [2], image information mining in time series (IIM-TS) [26], and the PicSOM system based on self-organizing maps (SOM) [27]. We should note that with the rapid growth of imagery databases, the number of image-derived products such as land use/land cover (LULC) products derived from high spatial resolution satellite or aerial images has also increased. These derived products allow us to retrieve remotely-sensed scenes with a higher efficiency and accuracy, when compared to retrieval based directly on imagery. Recently, a CBIR-like system specifically designed for category valued geospatial databases in general and for LULC databases in particular was proposed in [28]. The authors designed a pattern signature using only two features: the distribution of the labels and the size of the patches (which refer to the connected regions of the same label) [29]. Furthermore, a scene pattern signature was computed, which is the probability distribution of a 2-D variable with respect to the label and patch size. Based on the approach proposed in [28], LandEx (Landscape Explorer)—a content-based map retrieval (CBMR) system—was designed [30]. LandEx shares almost the same characteristics as the CBIR systems. In some cases, the distribution of the labels is enough for the retrieval task. However, for more complicated scenes with specific spatial patterns of different labels, the spatial relationships also need to be taken into consideration.

In this letter, we propose a CBIR-like remotely sensed scene retrieval system for use with LULC databases. To consider the spatial relationships of the labels in a scene, we propose a scene signature extraction method motivated by the grey level co-occurrence matrix (GLCM), which is utilized to compute the texture feature of the image. The similarity measurement we adopt is the cosine similarity, which is appropriate for our requirements even though it is simple. The scene signature can describe not only the probability distribution of the labels, but also the spatial relations. The CBMR system could also benefit from this approach.

In this study, we built a base signature library for the National Land Cover Database (NLCD) for 2006 [31,32]. The library consisted of the signature of every base block (100 × 100 pixels). By using the label co-occurrence matrix (LCM), the similarity between scenes of various sizes can be computed. In contrast, in most of the traditional retrieval systems, only the similarity between scenes of the same size can be computed. More specifically, signatures of the scenes with larger sizes (e.g., 200 × 200, 300 × 200, etc.) can be rapidly computed from the signature library, instead of re-computing the signatures for the whole database. In addition, the output of the retrieval results is more flexible. In addition to the similarity map measured between the reference scene and the whole database, we can also output the best-matching scene tiles, with the sizes not limited to the size of the analysis window.
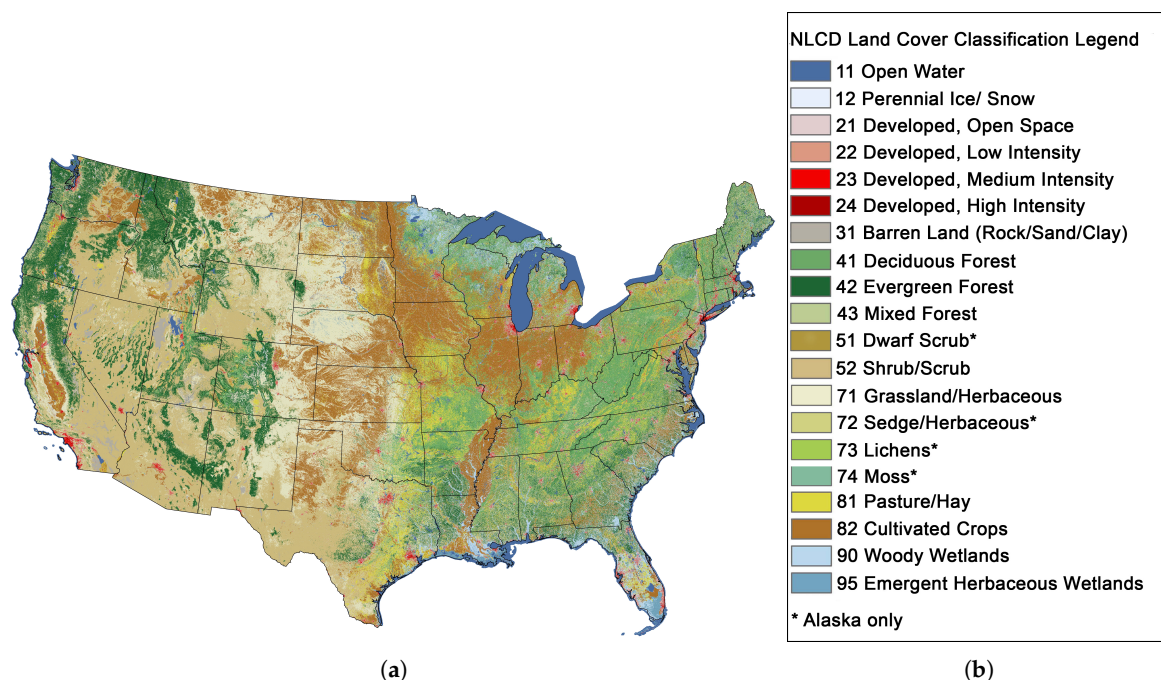
The rest of this letter is organized as follows. Section 2 briefly describes the proposed scene retrieval approach and introduces the proposed CBMR system in detail. In Section 3, we describe the series of experiments carried out to compare the scene retrieval performance of the proposed method with that of previous methods. A summary is provided in Section 4, along with a discussion of the achieved results.
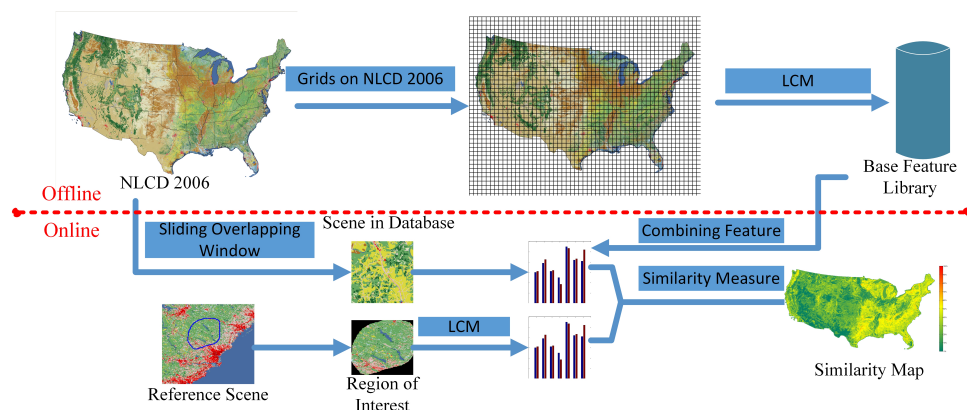
2. Methodology

This section provides a summary of the proposed approach for the retrieval of alike land-cover scenes from the NLCD 2006 database [31,32] (see Figure 1), which is shown by a flow chart in Figure 2. NLCD 2006 is a 16-class land cover classification product that covers the United States at a spatial resolution of 30 m.

The process can be divided into two main parts: (1) the off-line building of the signature library from the original LULC database; and (2) computation of the similarity between the selected reference

scene and the scenes in the database with a given size (width, height, and offset). Constructing the signature library is the foundation of the retrieval system.

Figure 1. (**a**) The National Land Cover Database (NLCD) 2006 map [33]; (**b**) The legend to NLCD 2006 map [34].

Figure 2. Basic framework of the proposed content-based map retrieval (CBMR) system. LCM: label co-occurrence matrix.

2.1. Label Co-Occurrence Matrix

The scene signature that we propose is inspired by the grey level co-occurrence matrix (GLCM) [35], and is utilized to extract the texture feature of the image. The principle of the GLCM has also been used in other applications and data. Barnsley et al. [36] designed a kernel-based procedure referred to as SPARK (spatial reclassification kernel) to examine both the frequency and the spatial arrangement of the class (land-cover) labels within a square kernel. To improve the classification of buildings, Zhang et al. [37] adopted modified co-occurrence matrix-based filtering. In the method proposed in [38], pairs of neighboring cells are regarded as primitive features. The primitive features are then extracted and counted to form a histogram to represent the pattern of a tile. The GLCM has

also been used to describe landscape structure [39]. In this letter, based on the same idea, we propose the label co-occurrence matrix for a LULC scene, which is defined as:

$$
\begin{aligned}
P(i,j;\sigma_x,\sigma_y) =& card\{(x,y)|I(x,y)=i; \\
& I(x+\sigma_x,y+\sigma_y)=j\}
\end{aligned}
\tag{1}
$$

where $card\{A\}$ represents the number of elements contained in set A, and i and j correspond to the labels of pixel (x,y) and $(x+\sigma_x,y+\sigma_y)$ in the scene. For a scene I, the co-occurrence matrices are computed at four orientations:

$$
P_{0,d}(i,j) = P(i,j;d,0)
\tag{2}
$$

$$
P_{45,d}(i,j) = P(i,j;d,d)
\tag{3}
$$

$$
P_{90,d}(i,j) = P(i,j;0,d)
\tag{4}
$$

$$
P_{135,d}(i,j) = P(i,j;-d,d)
\tag{5}
$$

where d is the distance parameter, which we set to 1. It can be seen that the LCM contains not only the first-order statistics of the labels in a scene (e.g., the histogram), but also the second-order statistics (i.e., the spatial relationships between adjacent pixels). In order to have a (relatively) rotation-invariant feature (because orientation information is not important in remote sensing), we sum the four co-occurrence matrices. For a scene I, its LCM should then be:

$$
H = \begin{pmatrix} H(1,1) & \cdots & H(1,n) \\ \vdots & \ddots & \vdots \\ H(n,1) & \cdots & H(n,n) \end{pmatrix}
\tag{6}
$$

$$
\begin{aligned}
H(i,j) =& P_{0,d}(i,j) + P_{45,d}(i,j) + P_{90,d}(i,j) \\
& + P_{135,d}(i,j)
\end{aligned}
\tag{7}
$$

where n is the amount of labels of the given scene, and H is a symmetrical matrix of which the dimension is n. Due to the symmetry of H we can only use matrix elements of the lower triangle to form a vector V (the dimension of V is $n*(n+1)/2$).

$$
\begin{aligned}
V =& \{(H(1,1),H(2,1),...,H(i,j),..., \\
& H(n,n)\}, where(1 \le i \le j \le n)
\end{aligned}
\tag{8}
$$

Then, a final vector F which we can have after a normalization process for V is used as the descriptor for the scene.

$$
\begin{aligned}
F =& \{\frac{H(1,1)}{C},\frac{H(2,1)}{C},...,\frac{H(i,j)}{C},...,\frac{H(n,n)}{C}\} \\
& where(1 \le i \le j \le n)
\end{aligned}
\tag{9}
$$

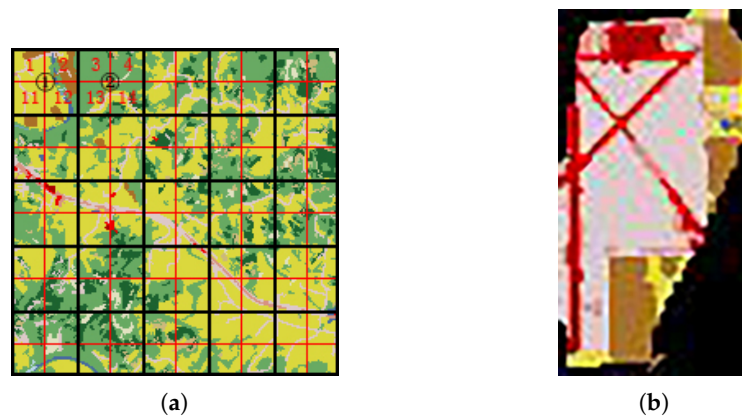where $C = \sum\limits_{H(i,j) \in V} H(i,j)$.

2.2. Similarity Measure

A number of different similarity measurements, including the Manhattan distance, the Hausdorff distance, the Mahalanobis distance, Kullback–Leibler divergence [40], Jeffrey divergence [41], etc., have been proposed for CBIR. In this letter, we chose cosine similarity for the similarity measure to be always within the range of 0 to 1 due to the elements in the signature vectors being positive.

The scenes are described by a signature vector computed by the LCM. Given two scenes k and l, their corresponding signature vectors are F_k and F_l, computed using Equation (9). The similarity between the two scenes is defined as:

$$D_{k,l} = \frac{F_k^{\mathrm{T}} F_l}{\|F_k\| \|F_l\|} \tag{10}$$

2.3. Determining the Signature of the Candidate Scene by Combining the Signatures in the Base Signature Library

In contrast to the traditional retrieval systems (for which the reference scene can only be of fixed size), we enable the user to select the reference scene more flexibly (see Figure 3b). The sizes of the retrieved scenes can also be varied according to the requirement of the user during each retrieval.

(**a**)　　　　　　　　　　　　　　　　　　(**b**)

Figure 3. (**a**) Illustration of determining the signature of the candidate scene by combining the signatures in the base signature library; (**b**) Reference scene region of interest (ROI).

The ability to return scenes with varied sizes relies on the computation of the scene signature. The construction process will be explained below in detail. Given that the database is a land cover scene in size, with width and height both 100 pixels. We first divided the whole map into small patches of 10×10 grids without overlap. The LCM was then computed for each patch, and was stored as the base signature library. Suppose that a ROI of 20×20 pixels is selected as the reference scene, and the user wants to obtain scenes of the same size with similar content. We first compute the LCM for the selected reference scene. For the signatures of the scenes in the database, we combine the adjacent 2×2 base signatures in the library as the signature of the candidate scene. The signature of the reference scene and the combined signatures in the database are then compared by the similarity measurement to return retrieval results. A sketch of this retrieval procedure is shown in Figure 3a.

$$V_{①}^1 = V_1 + V_2 + V_{11} + V_{12} \tag{11}$$

where V_1, V_2, V_{11}, and V_{12} correspond to the features of base blocks 1, 2, 11, and 12. $V_{①}^1$ omits the features of adjacent regions of each block. We now describe how to obtain the features of these regions in detail. As shown in Figure 3a, the adjacent regions can be divided into three parts: two columns C with the top left and bottom right coordinates of (1, 10) and (20, 11), two rows R with the top left and bottom right coordinates of (10, 1) and (11, 20), and a square S with four pixels (10, 10), (10, 11), (11, 10), and (11, 11).

$$V_{①}^2 = V_R + V_C - V_S \tag{12}$$

where V_R, V_C, and V_S correspond to the LCM features of C, R, S. Then, the complete LCM feature of block ① should be:

$$V_① = V_①^1 + V_①^2 \tag{13}$$

The above example explains the principle of combining the signature in the signature library to get the new signature of the scene explicitly. Suppose that the size of the base scene is $m \times m$ pixels. Based on the base signature library, we can set the parameters of the sliding window with width, height, and offset to be k, l, and t (where $k\&l = m, 2\,m, ...; t = m, 2\,m, ..., \& \ t \leq min(k,l)$).

2.4. Retrieval of Alike Scenes

In [28], the authors used an overlapping sliding window to search for alike scenes in the whole database. In this letter, we adopt a similar strategy. However, it can often occur that several scenes which are very close to each other are all in the retrieval results, due to the fact that the size of the retrieved scene is larger than the sliding window. In this case, we propose to merge nearby scenes when we display the retrieval results, in order to show the overall look of the retrieved scenes. More specifically, for each query, we first retrieve p similar scenes with the same size as the sliding window from the database (hereafter called "small scenes"). Among the p small scenes, scenes with overlapping areas are then merged. The merged scenes (hereafter called "larger scenes") have a larger size than the sliding window, and are not necessarily square. We attempt to retrieve six large scenes for the display of the retrieved results. If the p small scenes are not enough to form six large scenes, we retrieve more small scenes from the database. For comparison, the same operation was undertaken on the retrieval result returned by the approach proposed in [28]. In order to obtain the similarity map, we used a simple normalization operation (see Equation (14)) on the cosine similarity:

$$Sim = 1 - \frac{d_{cos} - min(d_{cos})}{max(d_{cos}) - min(d_{cos})} \tag{14}$$

where d_{cos} is the cosine distance between the reference scene and the scenes in the database, and *Sim* is the similarity after normalization.
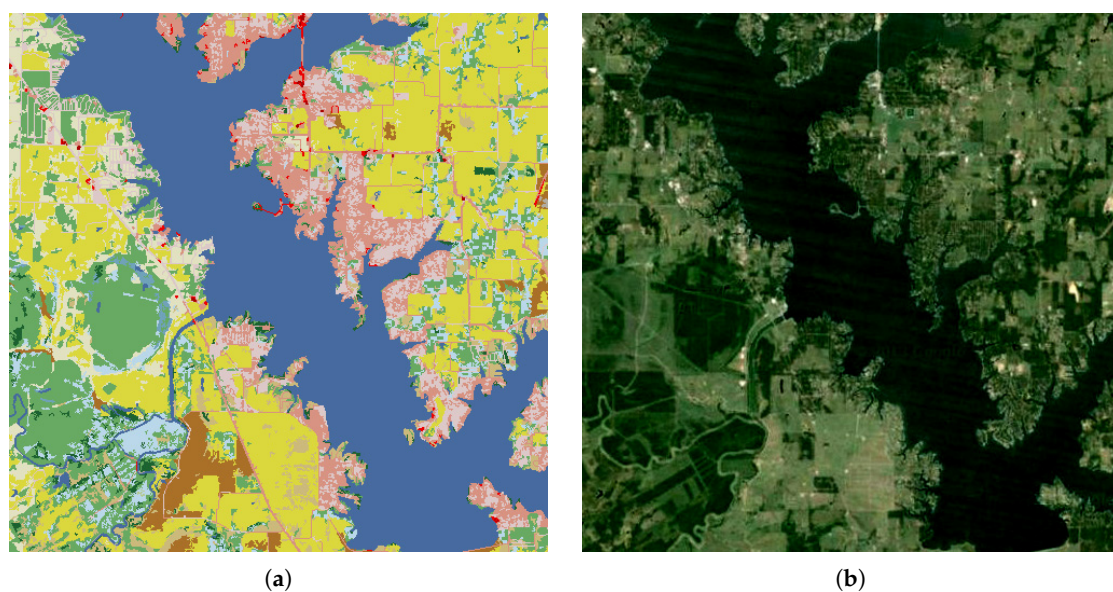
2.5. Retrieval Performance

Since there exists no ground truth data in the dataset we used, it is difficult to use the common ways of evaluating the retrieval performance—even though it could be determined whether the retrieved scenes are similar to the reference scene or not. Here, we try to evaluate the retrieval performance from two aspects: (1) the number of similar scenes appearing in the top retrievals; and (2) the ranking order of the similar scenes.

3. Experiments and Analysis

To demonstrate the validity of the proposed method, a number of comparison experiments were carried out. The whole database consists of a large landcover map with size $161,190 \times 104,424$ pixels. We first divided the whole map into small patches of 100×100 pixels without overlap. For the first three comparison experiments, two different scenes from the NLCD 2006 data set were chosen as the reference scenes. We adopted an overlapping sliding window with a size 500×500 pixels, where each time the window slid 100 pixels in either the horizontal or vertical direction. Since the NLCD database is of 30 m/pixel spatial resolution, each window covers 15 km × 15 km on the ground. With regard to the final experiments, we tried to retrieve the scenes containing an airport.

The first scene selected was in the region located in Kaufman Country, Texas (see Figure 4). As we can see, the first reference scene is traversed by the Cedar Creek Reservoir, which is classified as "open water". This could be the most noticeable information to determine the similarity between the reference and the retrieved scenes. Along the reservoir, there are a few classes belonging to the
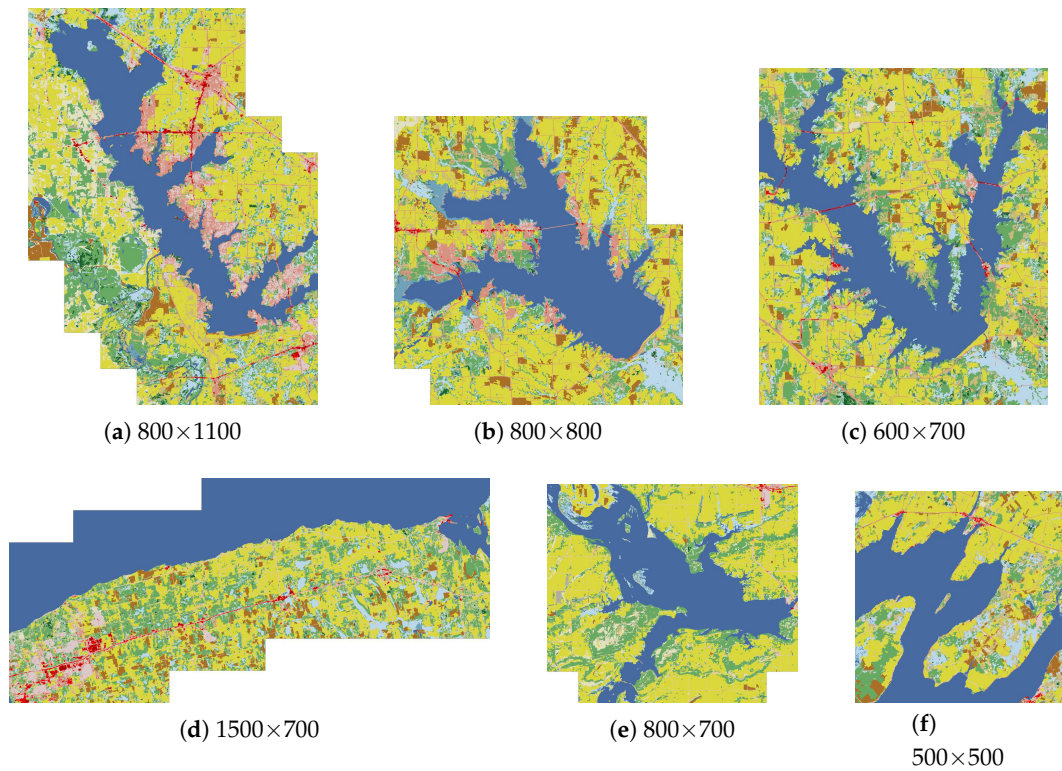
"developed" category. In addition to the "open water" and "developed" categories, the other two land-cover types of "pasture/hay" and "deciduous forest" occupy most of the remaining regions of this scene. The results of the retrieval experiment acquired by the use of the proposed method and the method proposed in [28] are shown in Figures 5 and 6, respectively. The scenes ranking first acquired by both methods are the same, and are in the same area where the reference scene is located. Figures 5 and 6 reveal that the ranking order of the proposed method is more reasonable than the of the method proposed in [28]. Judged only from the visual aspect, the last two scenes retrieved by the method proposed in [28] are quite different from the reference ones. The first four scenes (including the first retrieved one) retrieved by the proposed method are very similar to the reference ones. In Figures 7 and 8, the remotely sensed images correspond to the retrieved scenes shown in Figures 5 and 6. The images were acquired from Google Earth, using the geographical coordinates shown in Table 1.
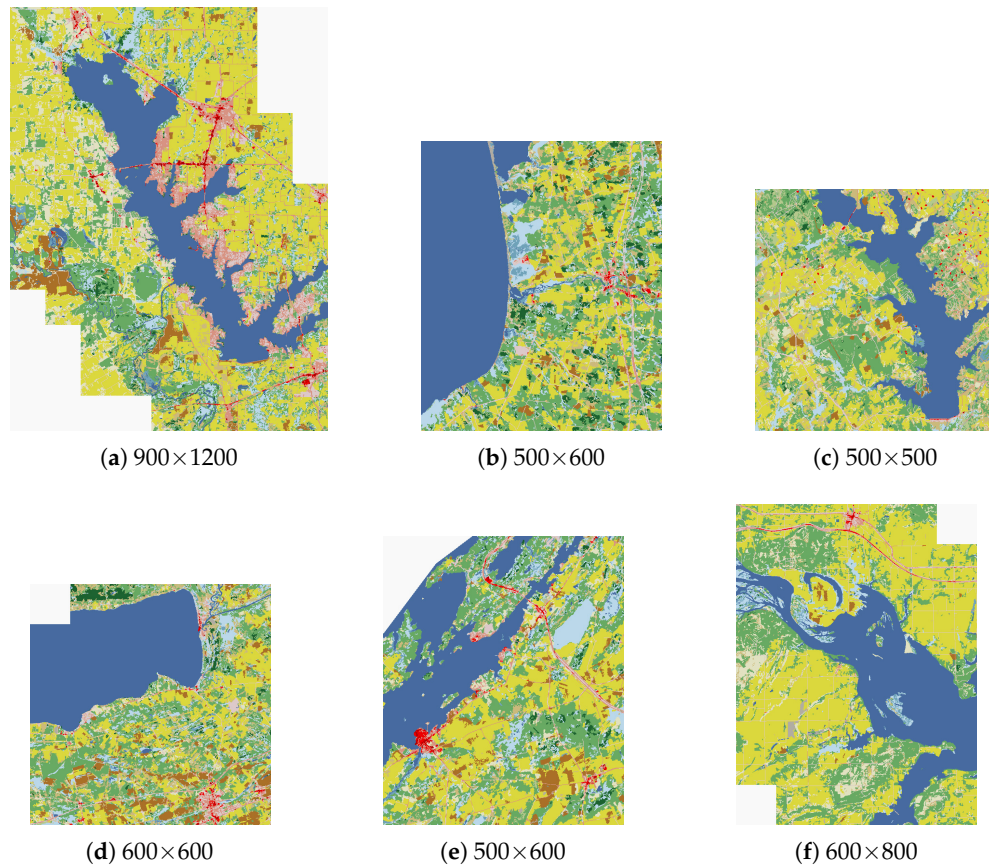

(**a**)　　　　　　　　　　　　　　　　　　　　　　　　　　　(**b**)

Figure 4. Reference scene I over Cedar Creek Reservoir in Texas. (**a**) The scene in NLCD 2006; (**b**) The remotely sensed image taken from Google Earth on the same region of the reference scene shown in (**a**) (for visualization, the histogram of the image has been stretched).

Table 1. The geographic coordinates of the center pixels in the top six returned scenes with reference scene I. Coordinates included in the brackets were acquired by [28], and the others are the results of the proposed method.
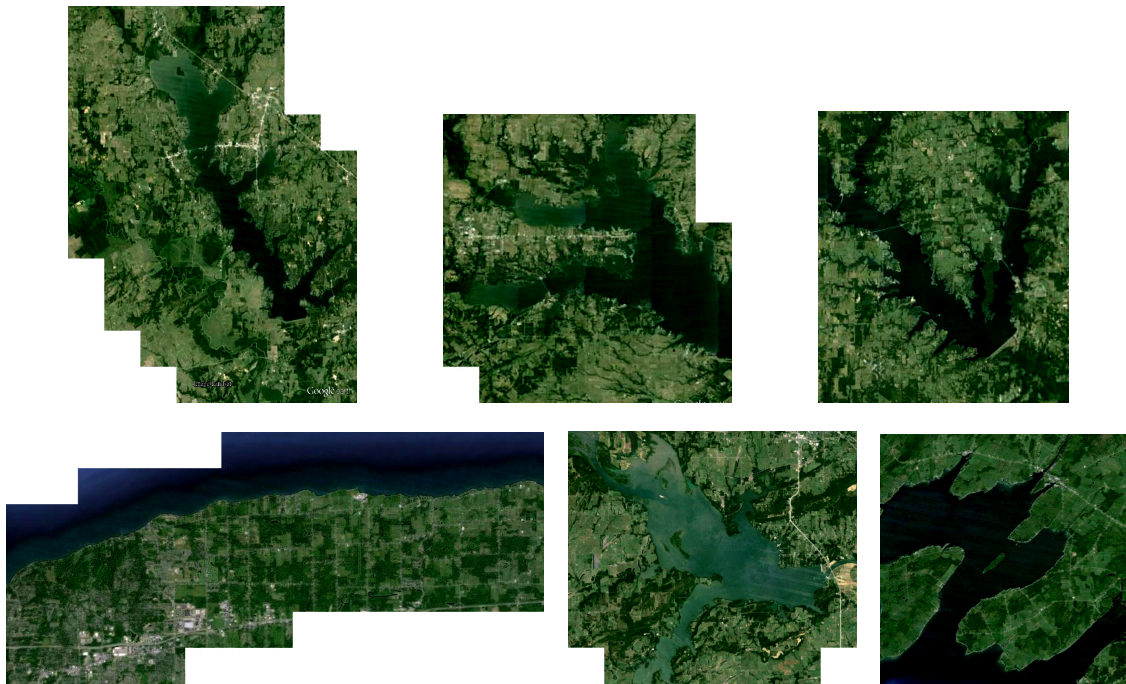
	Reference Scene I	Top Six Retrieval Results					
		Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6
Latitude (degree)	32.2619	32.2756 (32.2891)	32.8809 (43.5703)	32.8667 (31.3863)	43.2452 (43.1528)	35.3586 (44.2599)	44.0211 (35.3994)
Longitude (degree)	−96.1352	−96.1285 (−96.1446)	−96.0005 (−76.1792)	−95.5812 (−96.3645)	−77.2218 (−75.7893)	−94.9016 (−76.0132)	−76.1611 (−94.9677)

(**a**) 800×1100   (**b**) 800×800   (**c**) 600×700

(**d**) 1500×700   (**e**) 800×700   (**f**) 500×500

Figure 5. Top six retrieval results acquired by the use of the proposed method, for which the reference scene is shown in Figure 4a.

(**a**) 900×1200   (**b**) 500×600   (**c**) 500×500

(**d**) 600×600   (**e**) 500×600   (**f**) 600×800

Figure 6. Top six retrieval results of reference I acquired by the use of the method proposed in [28].
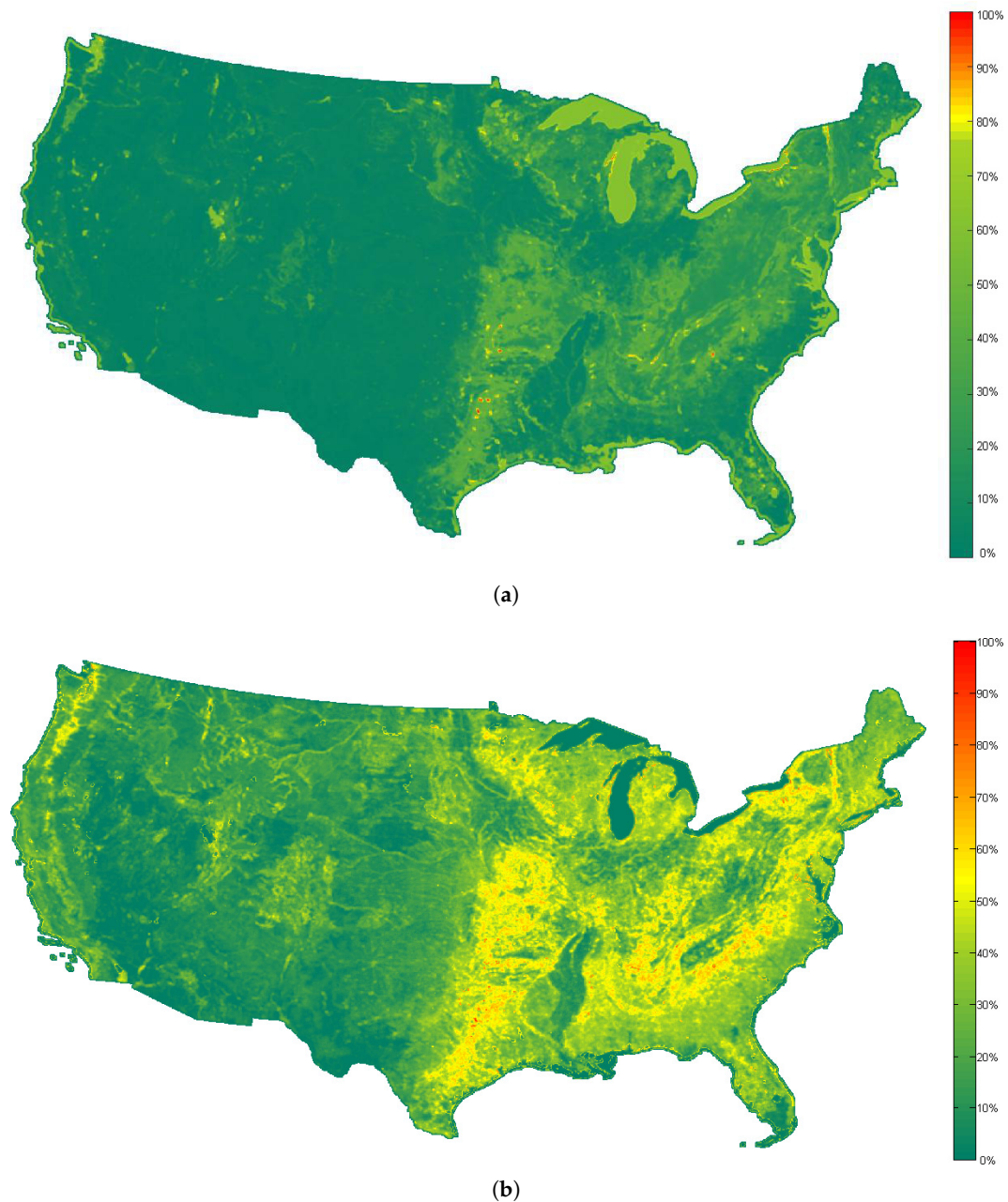
Figure 7. The remotely sensed images taken from Google Earth corresponding to the scenes shown in Figure 5 (for visualization, the histogram of the image has been stretched).
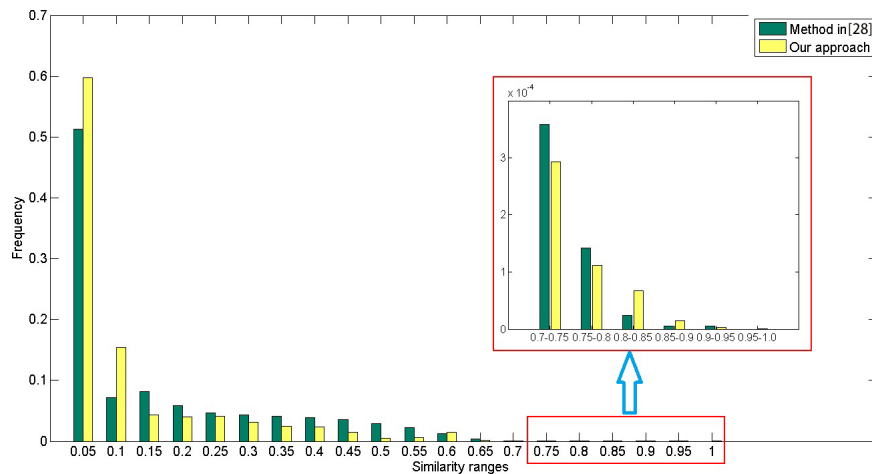

Figure 8. The remotely sensed images taken from Google Earth corresponding to the scenes shown in Figure 6 (for visualization, the histogram of the image has been stretched).

The most useful scene retrieval results are the similarity maps of the whole dataset compared to the reference scene. In Figure 9, the similarity maps computed by the use of the proposed method and the method proposed in [28] are shown. Figure 10 shows the histograms of the similarity maps in

Figure 9. It can be seen that although the visual aspects of the two similarity maps are different, high values of similarity measurements are always rare in both maps. In addition, the spatial distributions of both maps are correlated (i.e., the locations which have high values in Figure 9a also have high values in Figure 9b). However, according to Figures 5 and 6, the rank order of the most similar scenes obtained by the proposed method seems more reasonable, which can help the user to obtain the desired scenes from the large database more efficiently.
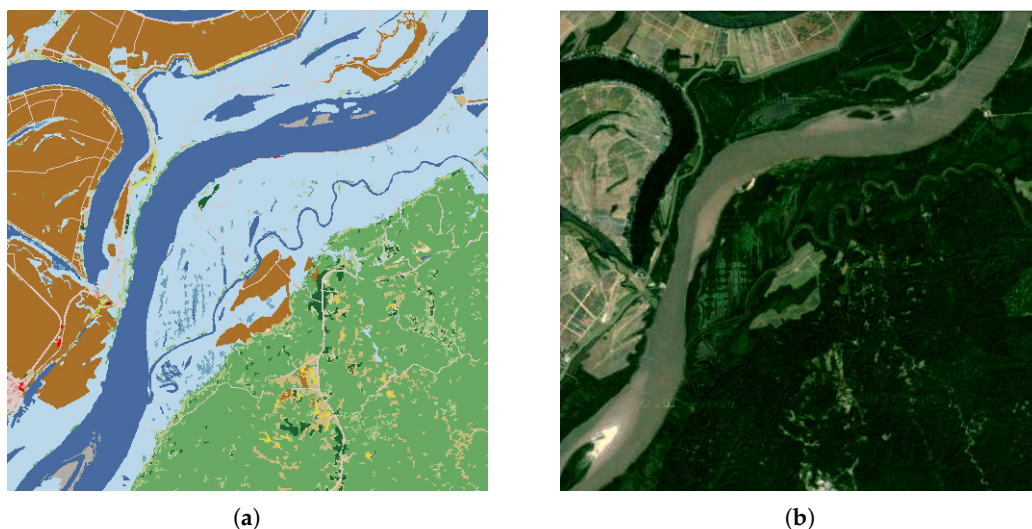

(**a**)

(**b**)

Figure 9. Similarity maps (with a resolution of 3 km) of reference scene I. (**a**) The proposed method; (**b**) The method proposed in [28].
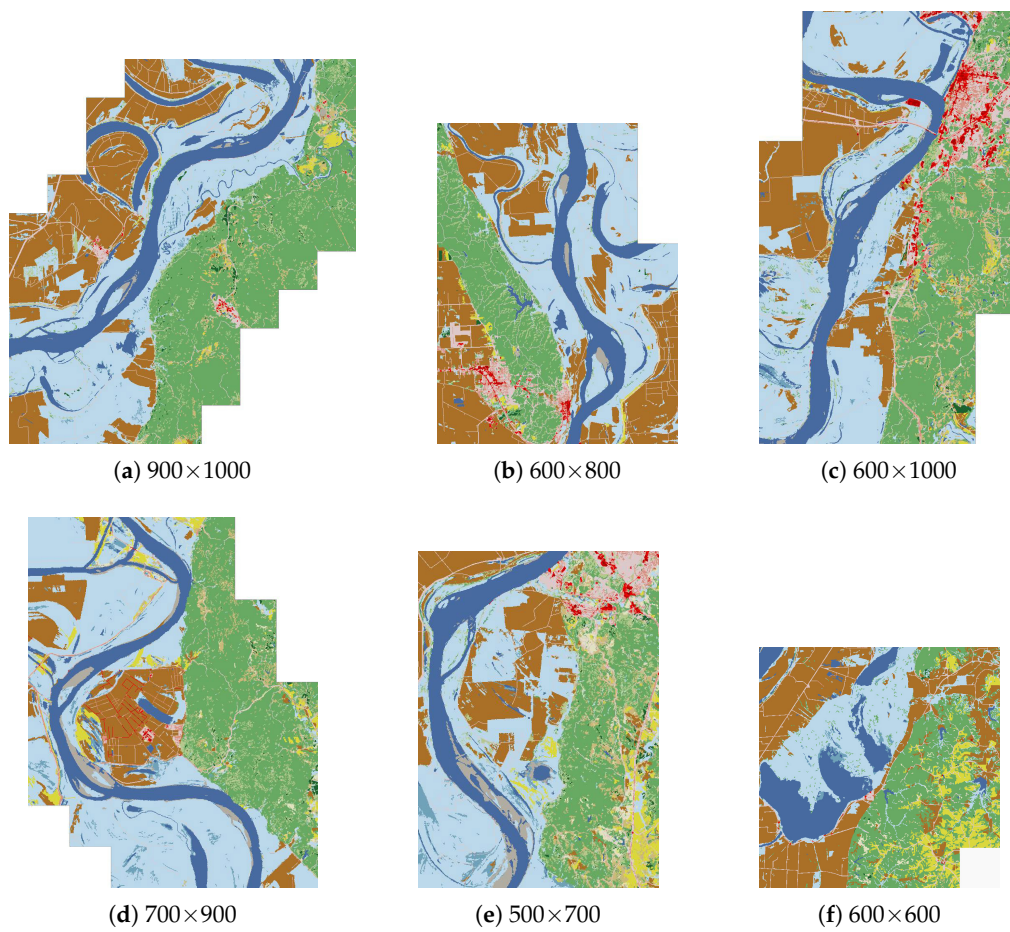
Figure 10. Comparison chart of the similarity histograms for the reference scene I.

The second reference scene considered in our experiment covers Yucatan Lake, which is located in the northeastern part of the U.S. state of Louisiana (Figure 11). The center of this scene is crossed by the Mississippi River, and the western part is divided into two parts by Lake Bruin—both of which are classified as "cultivated crops". Both the Mississippi River and Lake Bruin are classified as "open water". Both sides of the Mississippi River are covered with "woody wetlands", and "deciduous forest" occupies almost the whole southeastern part of the reference scene. Figures 12 and 13 show the results of this retrieval. The geographical coordinates of the retrieved scenes are presented in Table 2. The top five scenes given by the method proposed in [28] are all included in the results of the proposed method, except that the ranking order of the scenes is different, and the size of the scenes also varies. Again, the similarity maps obtained by the two methods (see Figure 14) are spatially correlated. Figure 15 shows the histograms of the similarity maps in Figure 14. However, the ranking order of the results is better when using the proposed method. The last scene obtained by the method proposed in [28] (Figure 13f) has different content.
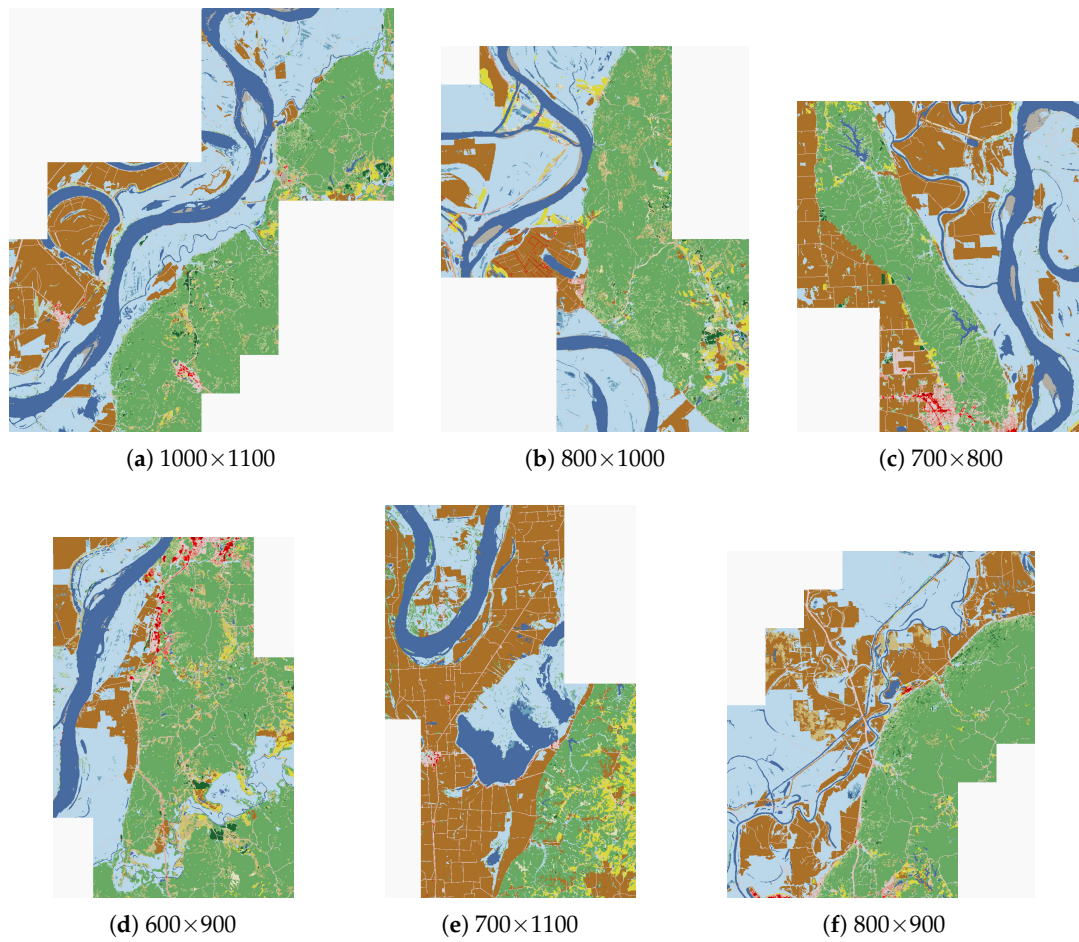

(a)　　　　　　　　　　　　　　　　　　　　　　　　　　(b)

Figure 11. Reference scene II over Yucatan Lake in Louisiana. (**a**) The scene in NLCD 2006; (**b**) The remotely sensed image taken from Google Earth on the same region of the reference scene shown in (**a**) (for visualization, the histogram of the image has been stretched).

(**a**) 900×1000   (**b**) 600×800   (**c**) 600×1000

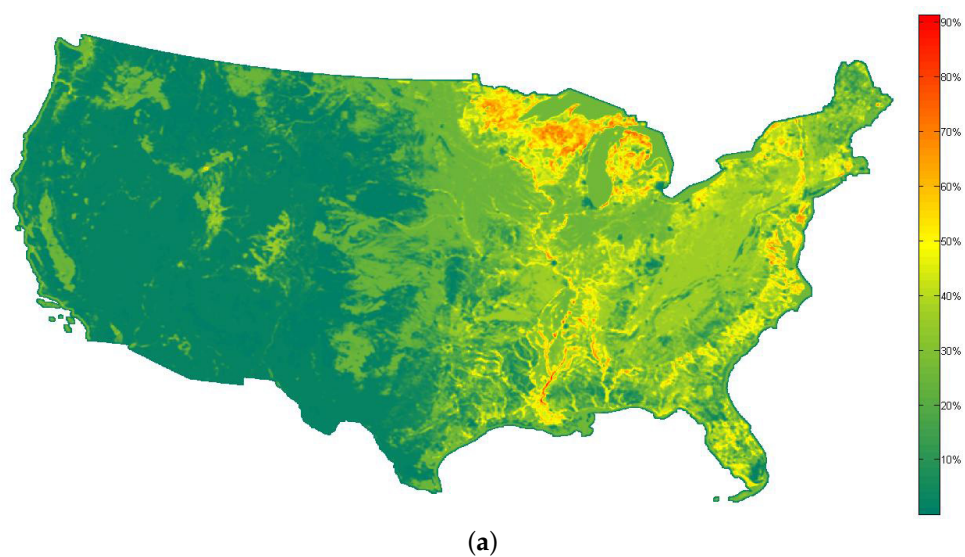(**d**) 700×900   (**e**) 500×700   (**f**) 600×600

Figure 12. Top six retrieval results of reference II acquired by the use of the proposed method.

Table 2. The geographic coordinates of the center pixels in the top six returned scenes with reference scene II. Coordinates included in the brackets were acquired by [28], and the others are the results of the proposed method.
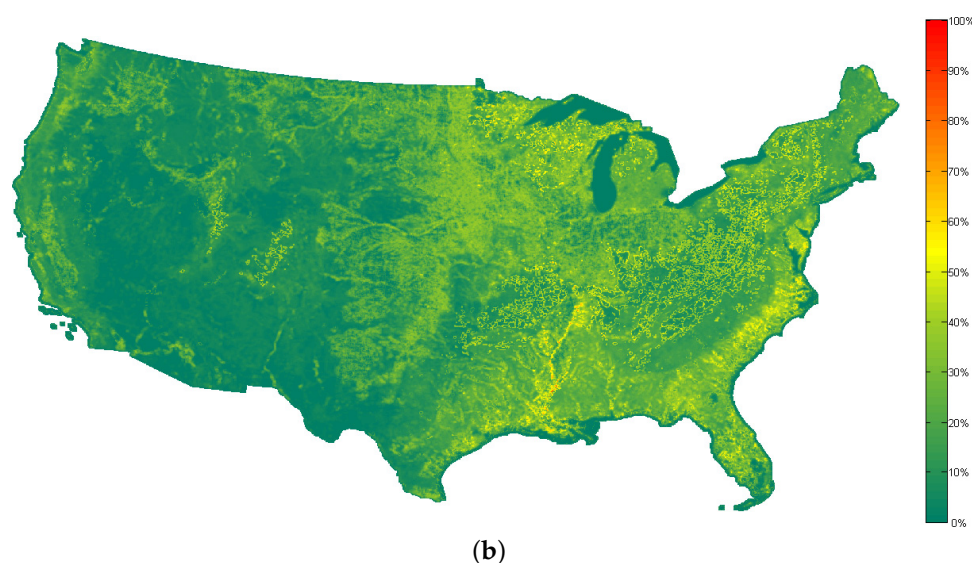
	Reference Scene II	Top 6 Retrieval Results					
		Rank 1	**Rank 2**	**Rank 3**	**Rank 4**	**Rank 5**	**Rank 6**
Latitude (degree)	31.9739	31.9163 (−31.9815)	34.6084 (−30.9879)	32.2567 (−34.6375)	30.975 (−32.1881)	31.456 (−36.4141)	36.3979 (−32.5061)
Longitude (degree)	−91.1757	−91.1705 (−91.1187)	−90.5897 (−91.5569)	−90.9416 (−90.6372)	−91.5734 (−90.914)	−91.4518 (−89.3943)	−89.3449 (−90.7972)

(**a**) 1000×1100　　　　　　(**b**) 800×1000　　　　　　(**c**) 700×800
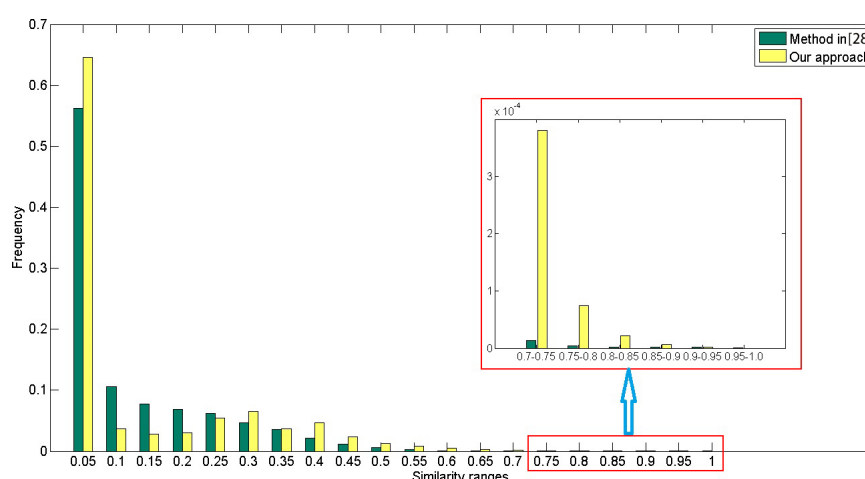
(**d**) 600×900　　　　　　(**e**) 700×1100　　　　　　(**f**) 800×900

Figure 13. Top six retrieval results of reference II acquired by the use of the method proposed in [28].

(**a**)

Figure 14. *Cont.*

(**b**)

Figure 14. Similarity maps (with a resolution of 3 km) for the reference scene II. (**a**) The proposed method. (**b**) The method proposed in [28].
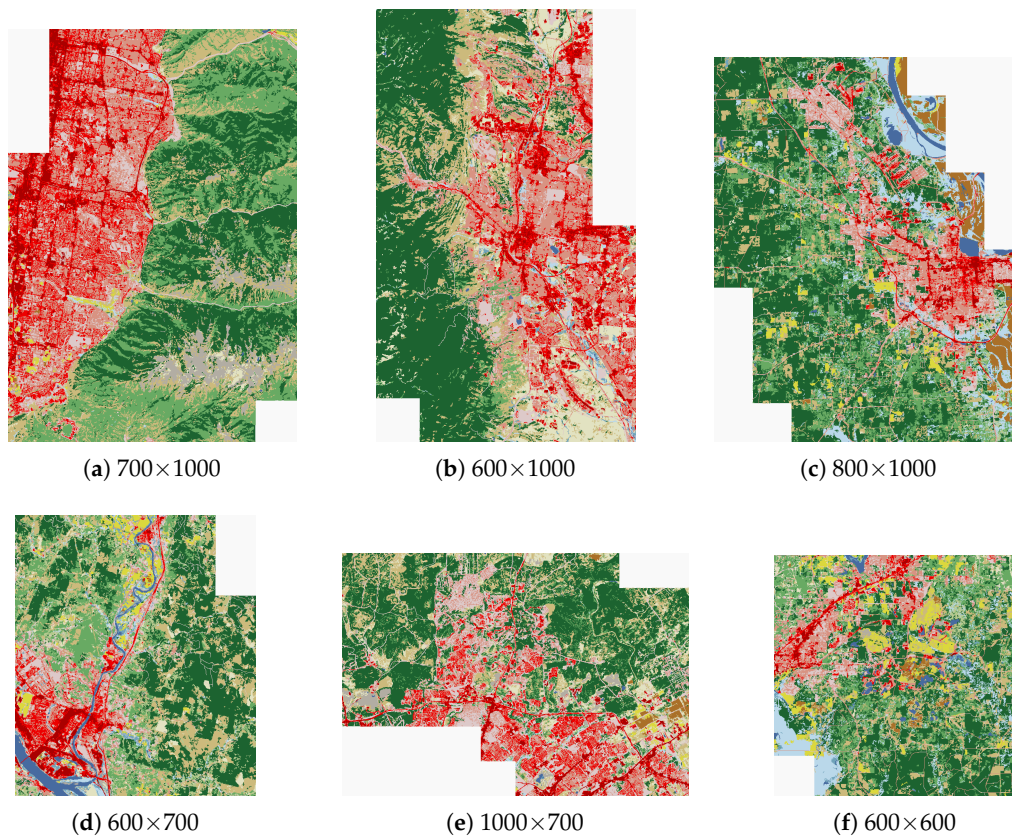

Figure 15. Comparison chart of the similarity histograms for reference scene II.

The third reference scene we selected is located in the northeastern part of the American state of Utah (Figure 16). The upper part of the reference scene is located within Springville, while the majority of the reference scene is located within Mapleton. Compared to the two previous scenes, the third reference scene is more heterogeneous. Seen from Figure 16, the western parts of the reference scene are urban areas, and the eastern parts are mountainous areas. The western parts are almost occupied by "developed" category with different intensities. Apart from the "developed" category, the other land-cover type "pasture/hay" occupies the remaining regions of the western parts. Within the eastern parts, the most common land-cover types are "evergreen forest", "deciduous forest", "shrub/scrub", and "barren land", which are typical landscape of the mountains. Figures 17 and 18 show the results acquired by different methods. The geographical coordinates of the retrieved scenes are presented in Table 3. The top four scenes given by the method proposed in [28] are the same as the results of our proposed method, and the ranking order of the scenes are also identical, except that the size of the scenes varies. For the last two scenes (see Figures 17 and 18) retrieved by each method, judged from the visual aspect, scenes acquired by the proposed method are much more similar than [28]. The distribution of areas with relative high similarity are identical in Figure 19a,b. So, the similarity
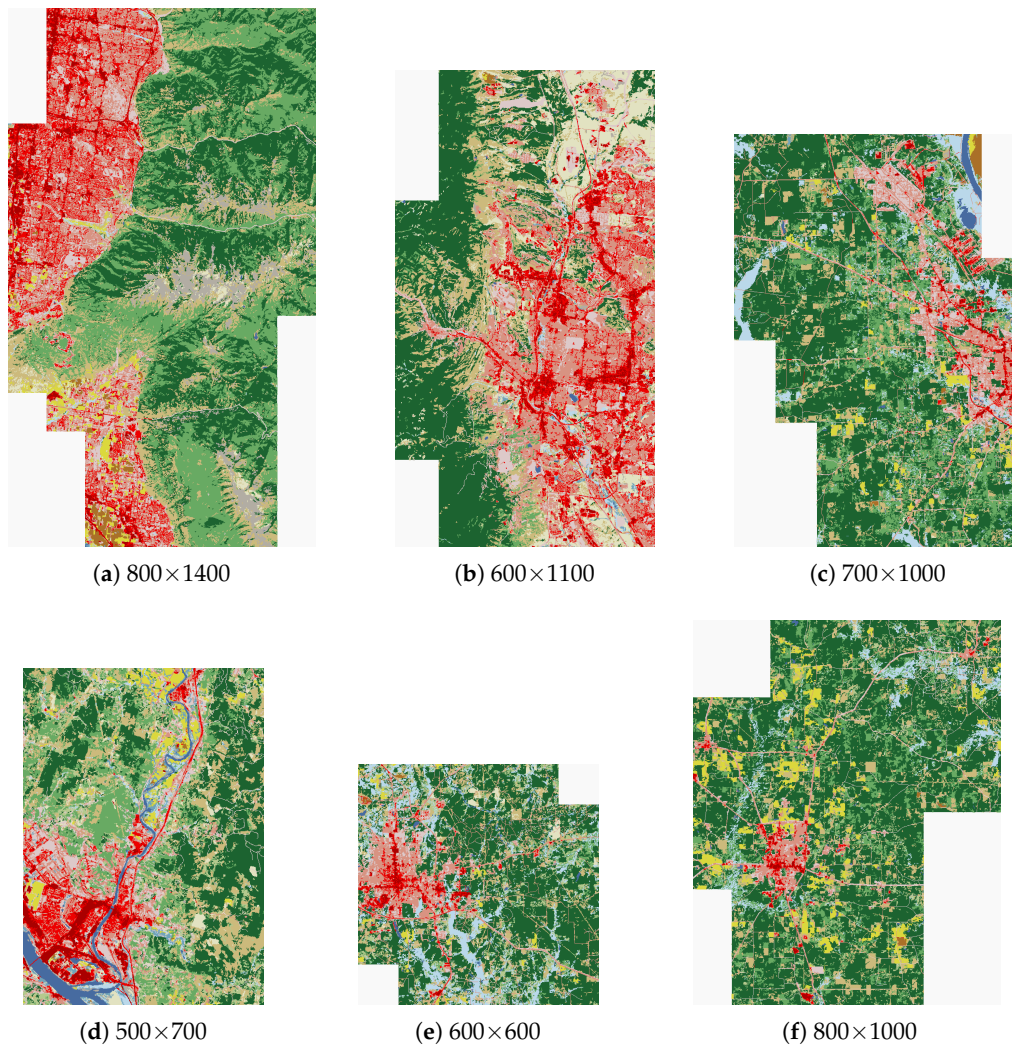
maps obtained by the two methods (see Figure 19) are spatially correlated. Figure 20 shows the histograms of the similarity maps in Figure 19.

(**a**)

(**b**)

Figure 16. Reference scene III located in Mapleton, Utah, USA. (**a**) The scene in NLCD 2006; (**b**) The remotely sensed image taken from Google Earth on the same region of the reference scene shown in (a) (for visualization, the histogram of the image has been stretched).
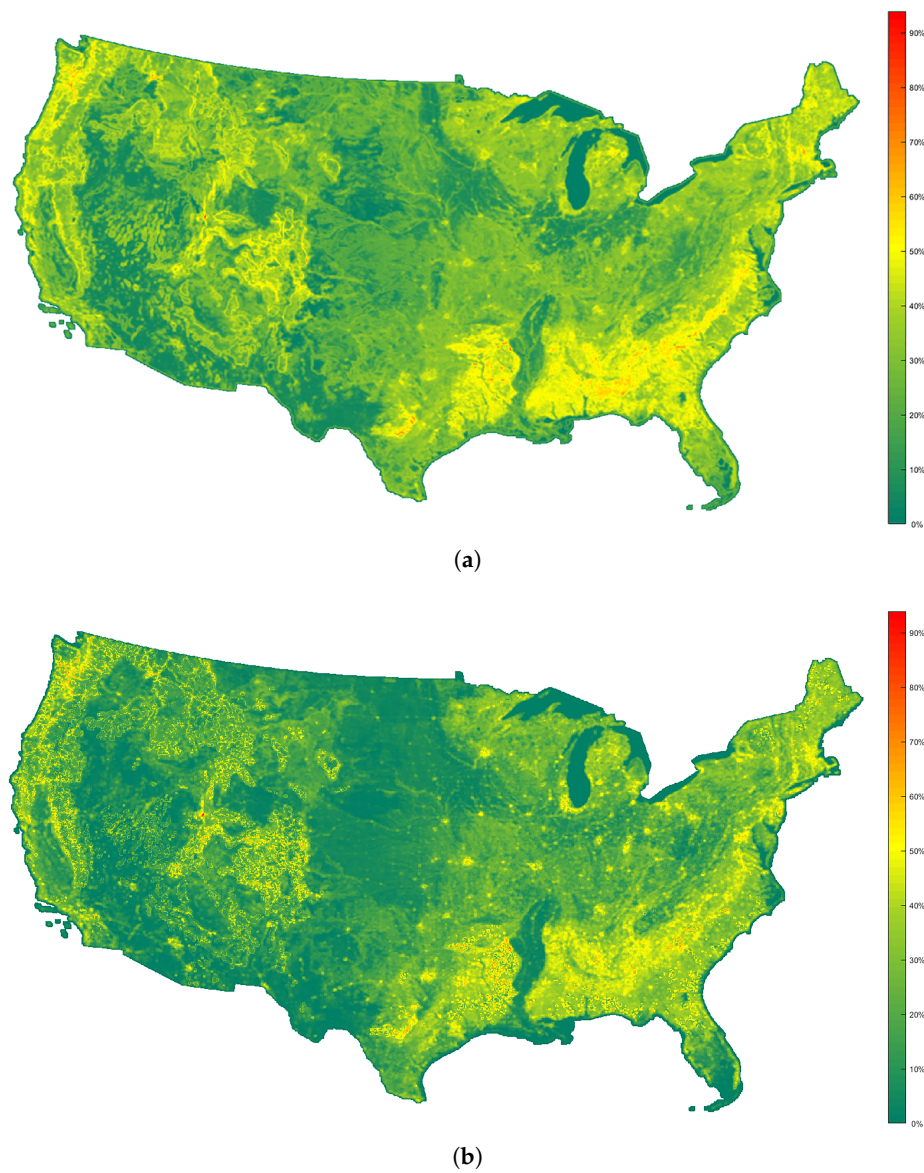

(**a**) 700×1000

(**b**) 600×1000

(**c**) 800×1000

(**d**) 600×700

(**e**) 1000×700

(**f**) 600×600

Figure 17. Top six retrieval results of reference III acquired by the use of the proposed method.

(**a**) 800×1400　　　　　　　(**b**) 600×1100　　　　　　　(**c**) 700×1000

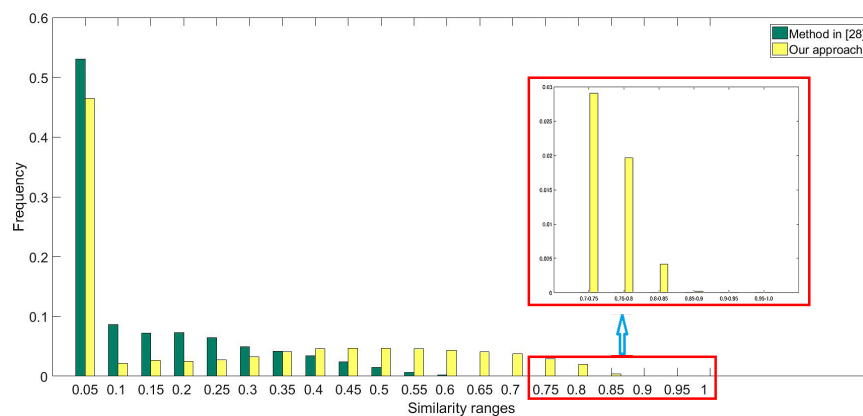(**d**) 500×700　　　　　　　(**e**) 600×600　　　　　　　(**f**) 800×1000

Figure 18. Top six retrieval results of reference III acquired by the use of the method proposed in [28].

Table 3. The geographic coordinates of the center pixels in the top six returned scenes with reference scene III. Coordinates included in the brackets were acquired by [28], and the others are the results of the proposed method.

	Reference Scene II	Top 6 Retrieval Results					
		Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6
Latitude (degree)	40.5858	40.6092 (40.5322)	38.8373 (38.8773)	34.2388 (34.2404)	46.1929 (46.1891)	29.6310 (33.2067)	34.5472 (33.3031)
Longitude (degree)	−111.7792	−111.7873 (−111.7520)	−104.8394 (−104.8443)	−92.0940 (−92.1432)	−122.8871 (122.9057)	−98.4487 (−92.5989)	−92.5071 (−93.1788)

(**a**)

(**b**)

Figure 19. Similarity maps (with a resolution of 3 km) for reference scene III. (**a**) The proposed method. (**b**) The method proposed in [28].

Figure 20. Comparison chart of the similarity histograms for reference scene III.

In the last experiment, we tried to extract scenes with concrete semantics from the NLCD database. Due to the coarse resolution of the NLCD dataset (30 m), few ground objects are visible. It was observed that the airstrips in the NLCD 2006 dataset are usually represented by developed areas of medium and high intensities. We also observed that the neighbourhood regions of the airstrips are always filled with open space. Based on this information, we could try to retrieve scenes containing an airport.

A scene (see Figure 21a) containing Arkansas International Airport was selected as the reference scene. We set the size of the sliding overlapping window as 100 × 100 pixels and the step size as 100 pixels. Figures 21 and 22 show the top 12 similar scenes by each method. Apart from the similarity map (Figure 23) and histograms of the similarity maps (Figure 24), we also returned the ranking of the similarity degree. Therefore, we could also observe each scene from the top-ranking scene to the bottom. In order to give an authentic look to the scenes, rather than classified map, the geographic coordinates of each scene were returned. This allowed us to locate the scenes in Google Earth. Among the top 12 most similar scenes, there are six scenes (see Figure 21a–c,e,i,k) containing an airport, which respectively correspond to Arkansas International Airport, Indianapolis International Airport, Monroe Regional Airport, South Bend Airport, Salina Municipal Airport, and Valley International Airport. Meanwhile, by the use of the method proposed in [28], only three airports—Arkansas International Airport, Orangeburg Municipal Airport, and Fort Wayne Municipal Airport (Figure 22a,i,k)—were returned among the top 12 scenes. Furthermore, we compared the precision of the top 96 scenes retrieved by the use of the proposed method and the method proposed in [28]. The proposed method could obtain 29 airports, while the method proposed in [28] could only acquire eight.
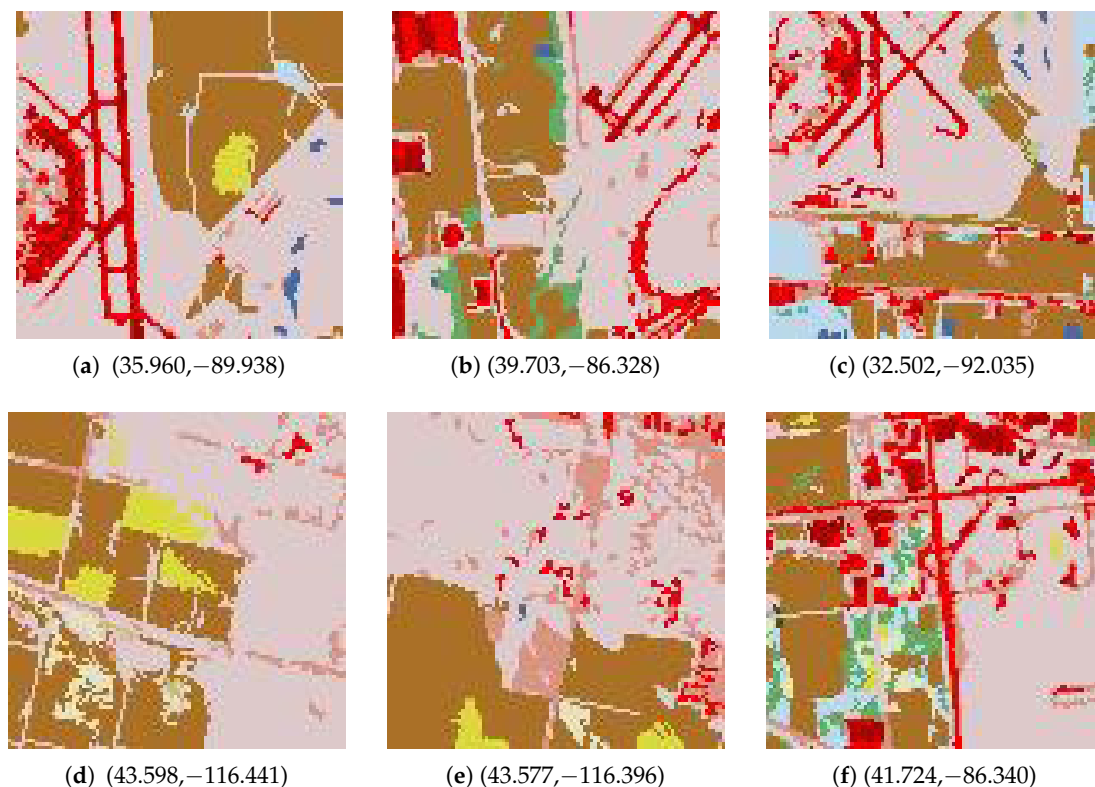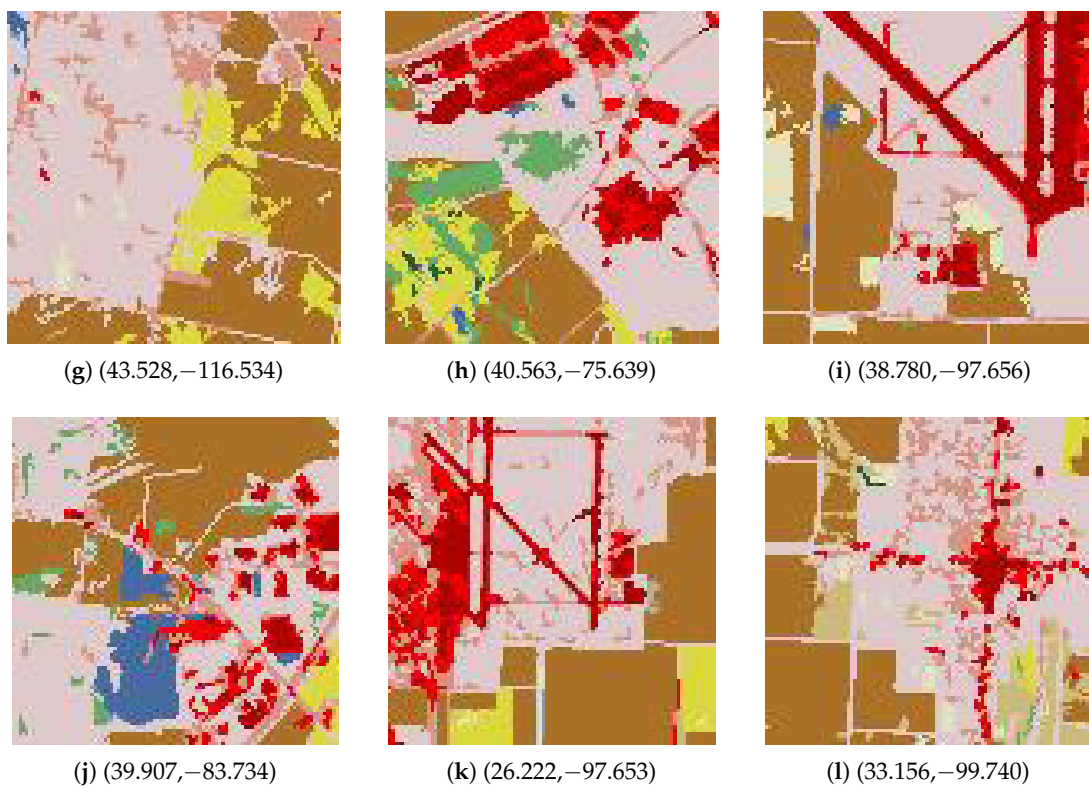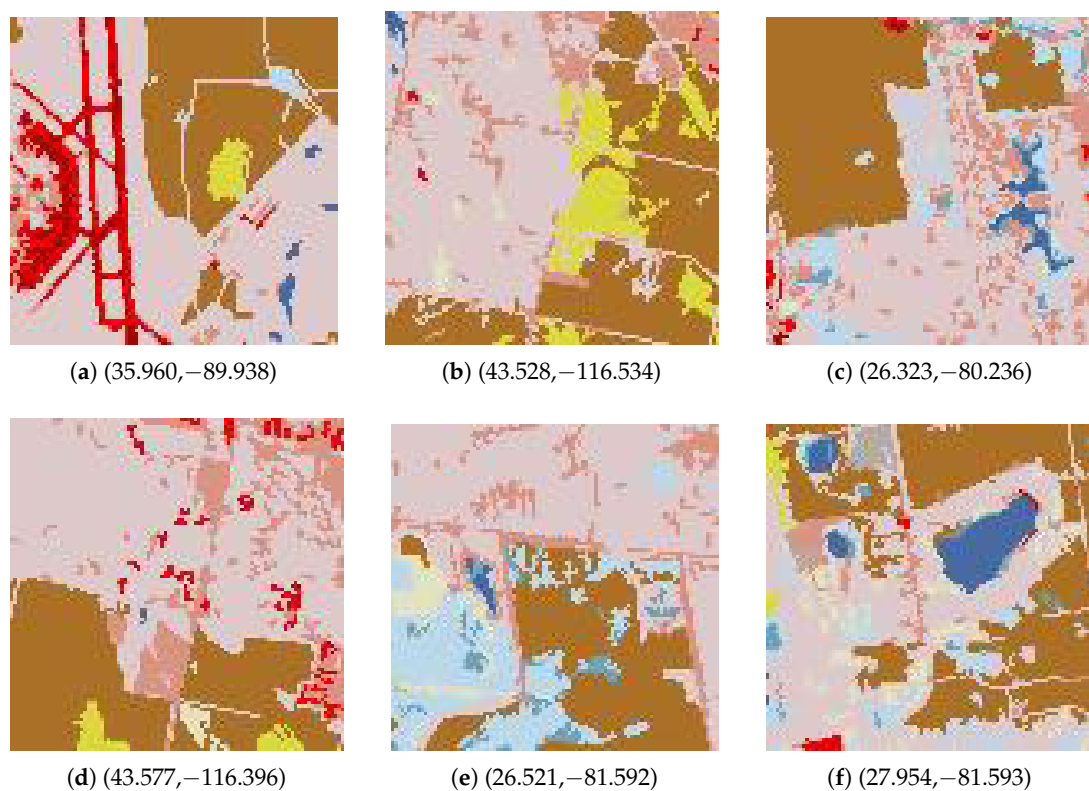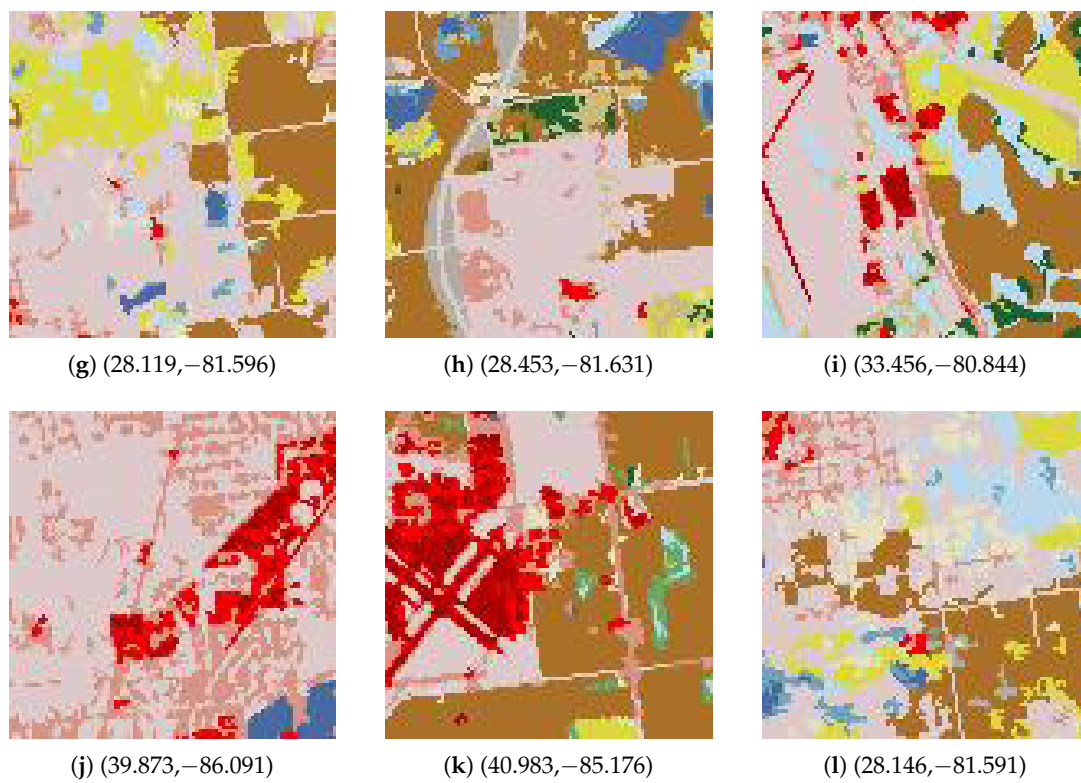

(**a**) (35.960,−89.938)   (**b**) (39.703,−86.328)   (**c**) (32.502,−92.035)

(**d**) (43.598,−116.441)   (**e**) (43.577,−116.396)   (**f**) (41.724,−86.340)

Figure 21. *Cont.*

(**g**) (43.528,−116.534)

(**h**) (40.563,−75.639)

(**i**) (38.780,−97.656)

(**j**) (39.907,−83.734)

(**k**) (26.222,−97.653)

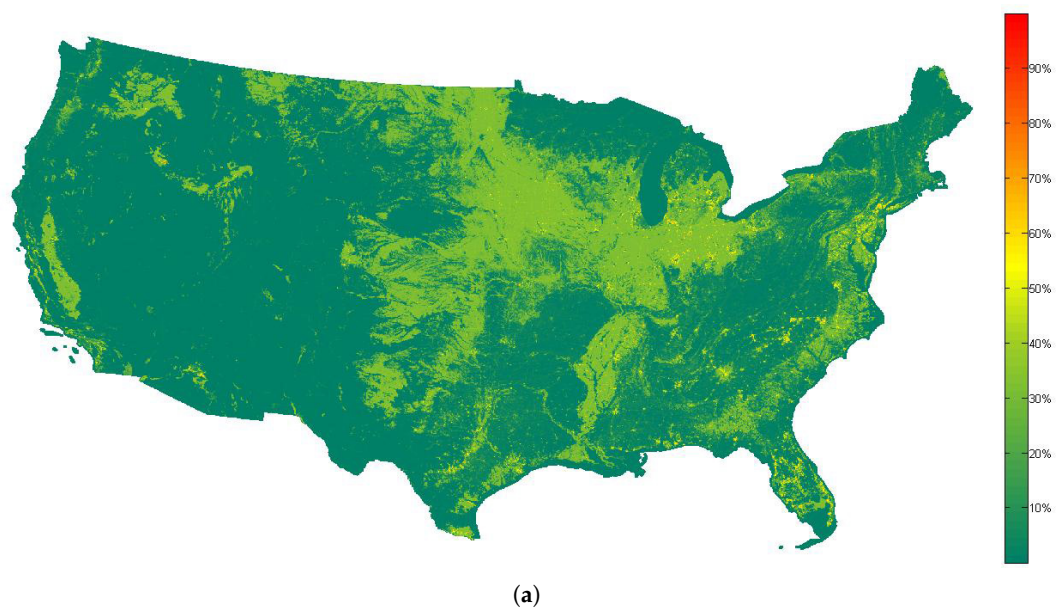(**l**) (33.156,−99.740)

Figure 21. Top 12 scenes most similar to the reference scene (see Figure 21a) obtained by the proposed method. The content in the brackets is the geographic coordinates of the scenes.

(**a**) (35.960,−89.938)

(**b**) (43.528,−116.534)

(**c**) (26.323,−80.236)

(**d**) (43.577,−116.396)

(**e**) (26.521,−81.592)

(**f**) (27.954,−81.593)

Figure 22. *Cont.*

(**g**) (28.119,−81.596)   (**h**) (28.453,−81.631)   (**i**) (33.456,−80.844)

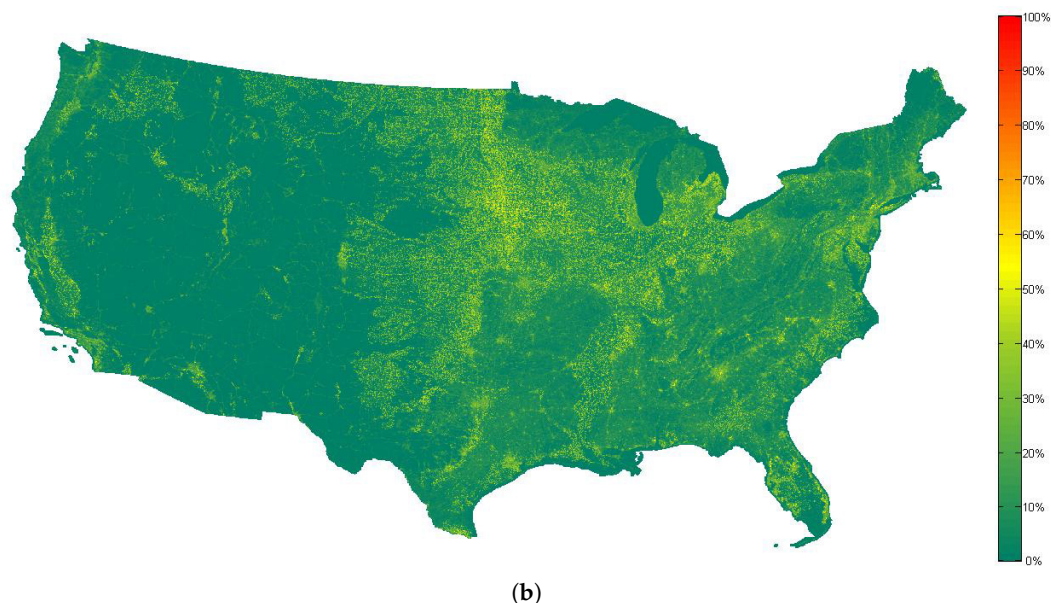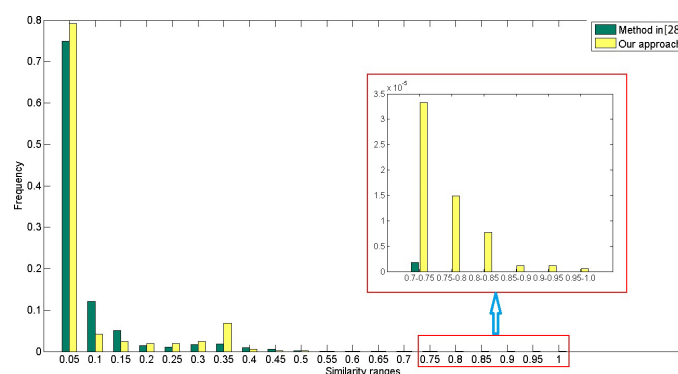(**j**) (39.873,−86.091)   (**k**) (40.983,−85.176)   (**l**) (28.146,−81.591)

Figure 22. Top 12 scenes most similar to the reference scene (see Figure 21a) obtained by the approach proposed in [28]. The content in the brackets is the geographic coordinates of the scenes.

(**a**)

Figure 23. *Cont.*

(**b**)

Figure 23. Similarity maps (with a resolution of 3 km) of the reference scene (Figure 21a). (**a**) The proposed method; (**b**) The method proposed in [28].

Figure 24. Comparison chart of the similarity histograms for the reference scene (Figure 21a).

4. Discussion

We propose a CBIR-like remotely sensed scene retrieval system for use with LULC databases in this study. Motivated by GLCM, we propose a scene signature extraction method for LULC scenes. To validate our method, four reference scenes are selected to conduct the comparison experiments. For the first three reference scenes, we evaluated the performance of our method in two aspects: the content of the retrieved scenes, and the ranking order. As depicted in Figures 5, 6, 12, 13, 17 and 18 that scenes returned by the proposed method seem much more similar compared to [28]. Additionally, the rank order of the most similar scenes obtained by the proposed method seems more reasonable, which can help the user to obtain the desired scenes from the large database more efficiently. For the last reference scene, we selected a scene containing an airport. Among the top twelve most similar scenes, our method can return six scenes containing an airport, while [28] had only three. Furthermore, we compared the precision of the top 96 scenes retrieved by the use of the proposed method and the method proposed in [28]. The proposed method could obtain 29 airports, while the method proposed in [28] could only acquire 8.

To achieve a finer retrieval result, we propose the LCM as scene signature. Through the designed CBMR, the user could obtain not only the alike scenes in varied scales, but also the similarity map. Moreover, the geographical coordinates of the alike scenes are also presented. Due to this, the scenes

could be easily located in Google Earth or other map tools. Per the results mentioned above in Section 3, the proposed method could achieve better performance. The retrieval result of the proposed method is more accurate and reasonable. Some future work is also expected. At first, the proposed method can be extended to ordinary remotely sensed images rather than land cover products. A rough classification of the images in order to convert the grey-valued (or multi-spectral) images into label images is required before the retrieval. In addition, we need to merge the feedback technique into our system. With relevance feedback technique involved, users would be able to indicate to the image retrieval system whether the retrieved results are "relevant", "irrelevant", or "neutral". Retrieval results would then be refined iteratively. This could help the user obtain much more accurate results. The next aspect we need to take into consideration is the complex semantic information; for example, a large open space with many cars is likely to be a parking lot.

5. Summary and Conclusions

In this letter, we have introduced an effective approach for retrieving alike scenes in large category-valued geospatial data sets. The LCM has been used as the scene signature which is able to not only describe the distribution of the labels in a scene, but also to extract the spatial relations of the different labels. The experimental results show the effectiveness of the proposed method. By constructing the base signature library, much time could be saved, as we do not need to recalculate the signature of the scenes. Further, our CBMR system can support reference scenes with irregular shape. In the end, benefiting from the LCM, we can compare the similarity of scenes of different size. Based on this principle, we can acquire retrieval results in varied scales, and not only the same size as reference scenes. This could increase the efficiency for the users to acquire desired scenes from a large database.

Author Contributions: Bin Luo and Jun Liu conceived and designed the experiments; Jun Liu performed the experiments, analyzed the data and wrote the paper. Bin Luo, Qianqing Qin and Guopeng Yang gave valuable comments and suggestions, and carefully revised the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CBIR	content-based image retrieval
CBMR	content-based map retrieval
EO	Earth observation
GLCM	Grey Level Co-occurrence Matrix
IIM-TS	image information mining in time series
KIM	knowledge-driven information mining
LandEx	Landscape Explorer
LCM	Label Co-occurrence Matrix
LULC	land use/land cover
NLCD	National Land Cover Database
RS	remote sensing
SOM	self-organizing maps
SPARK	spatial reclassification kernel

References

1. Seidel, K.; Datcu, M. Query by image content from remote sensing archives. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Seattle, WA, USA, 6–10 July 1998; pp. 393–396.

2. Datcu, M.; Daschiel, H.; Pelizzari, A.; Quartulli, M.; Galoppo, A.; Colapicchioni, A.; Pastori, M.; Seidel, K.; Marchetti, P.G.; Elia, S.D. Information mining in remote sensing image archives: System concepts. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 2923–2936.

3. Li, J.; Narayanan, R.M. Integrated spectral and spatial information mining in remote sensing imagery. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 673–685.

4. Daschiel, H.; Datcu, M. Information mining in remote sensing image archives: System evaluation. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 188–199.

5. Shyu, C.; Klaric, M.; Scott, G.; Barb, A.S.; Davis, C.; Palaniappan, K. GeoIRIS: Geospatial information retrieval and indexing system content mining, semantics modeling and complex queries. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 839–852.

6. Luo, B.; Jiang, S.; Zhang, L. Indexing of remote sensing images with different resolutions by multiple features. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 1899–1912.

7. Li, Y.; Zhang, Y.; Tao, C.; Zhu, H. Content-based high-resolution remote sensing image retrieval via unsupervised feature learning and collaborative affinity metric fusion. *Remote Sens.* **2016**, *8*, 709.

8. Demir, B.; Bruzzo, L. Kernel-based hashing for content-based image retrval in large remote sensing data archive. In Proceedings of the 2014 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Quebec City, QC, Canada, 13–18 July 2014; pp. 3542–3545.

9. Demir, B.; Bruzzone, L. Hashing-based scalable remote sensing image search and retrieval in large archives. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 892–904.

10. Aptoula, E. Remote sensing image retrieval with global morphological texture descriptors. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 3023–3034.

11. Aptoula, E. Bag of morphological words for content-based geographical retrieval. In Proceedings of the 2014 12th International Workshop on Content-Based Multimedia Indexing (CBMI), Klagenfurt, Austria, 18–20 June 2014; pp. 1–5.

12. Bosilj, P.; Aptoula, E.; Lefèvre, S.; Kijak, E. Retrieval of remote sensing images with pattern spectra descriptors. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 228.

13. Yang, Y.; Newsam, S. Geographic image retrieval using local invariant features. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 818–832.

14. Özkan, S.; Ateş, T.; Tola, E.; Soysal, M.; Esen, E. Performance analysis of state-of-the-art representation methods for geographical image retrieval and categorization. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 1996–2000.

15. Gueguen, L. Classifying compound structures in satellite images: A compressed representation for fast queries. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 1803–1818.

16. Negrel, R.; Picard, D.; Gosselin, P.H. Evaluation of second-order visual features for land-use classification. In Proceedings of the 2014 12th International Workshop on Content-Based Multimedia Indexing (CBMI), Klagenfurt, Austria, 18–20 June 2014; pp. 1–5.

17. Zhao, L.J.; Tang, P.; Huo, L.Z. Land-use scene classification using a concentric circle-structured multiscale bag-of-visual-words model. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 4620–4631.

18. Zhang, F.; Du, B.; Zhang, L. Saliency-guided unsupervised feature learning for scene classification. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 2175–2184.

19. Napoletano, P. Visual descriptors for content-based retrieval of remote sensing images. *arXiv* **2016**, arXiv:1602.00970.

20. Wang, Y.; Zhang, L.; Tong, X.; Zhang, L.; Zhang, Z.; Liu, H.; Xing, X.; Mathiopoulos, P.T. A three-layered graph-based learning approach for remote sensing image retrieval. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 6020–6034.

21. Alonso, K.; Espinozamolina, D.; Datcu, M. LUCAS Visual Browser: A tool for land cover visual analytics. In Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, Italy, 26–31 July 2015; pp. 1484–1487.

22. Alonso, K.; Espinoza-Molina, D.; Datcu, M. Multilayer architecture for heterogeneous geospatial data analytics: Querying and understanding EO archives. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 791–801.

23. Espinoza-Molina, D.; Manilici, V.; Dumitru, C.; Reck, C.; Cui, S.; Rotzoll, H.; Hofmann, M.; Schwarz, G.; Datcu, M. The Earth Observation Image Librarian (EOLIB): The data mining component of the TerraSAR-X Payload Ground Segment. In Proceedings of Big Data from Space (BiDS'16), Auditorio de Tenerife, Santa Cruz de Tenerife, Spain, 15–17 March 2016; pp. 228–231.

24. Espinoza-Molina, D.; Datcu, M. Earth-observation image retrieval based on content, semantics, and metadata. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 5145–5159.

25. Alonso, K.; Datcu, M. Accelerated probabilistic learning concept for mining heterogeneous earth observation images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 3356–3371.

26. Bovolo, F.; Bruzzone, L. *Image Information Mining in Time Series: Algorithms and Methods for Prototyping*; European Space Agency (ESA): Trento, Italy, 2007.

27. Molinier, M.; Laaksonen, J.; Hame, T. Detecting man-made structures and changes in satellite imagery with a content-based information retrieval system built on self-organizing maps. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 861–874.

28. Jasiewicz, J.; Stepinski, T.F. Example-based retrieval of alike land-cover scenes from NLCD 2006 database. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 155–159.

29. Alnuweiri, H.; Prasanna, V. Parallel architectures and algorithms for image component labeling. *IEEE Trans. Pattern Anal. Mach. Intell.* **1992**, *14*, 1014–1034.

30. Jasiewicz, J.; Netzel, P.; Stepinski, T.F. LandEx - A GeoWeb tool for query and retrieval of spatial patterns in land cover datasets. *IEEE J. Sel. Top. Appl. Earth Obs.* **2013**, *7*, 257–266.

31. National Land Cover Database 2006. Available online: http://www.mrlc.gov/nlcd2006.php (accessed on 15 January 2014).

32. Fry, J.; Xian, G.; Jin, S.; Dewitz, J.; Homer, C.; Yang, L.; Barnes, C.; Herold, N.; Wickham, J. Completion of the 2006 National Land Cover Database for the conterminous United States. *Photogramm. Eng. Remote Sens.* **2011**, *77*, 858–864.

33. National Land Cover Database 2006 Product. Available online: http://www.mrlc.gov/nlcd06_data.php (accessed on 15 January 2014).

34. Legend of National Land Cover Database 2006. Available online: http://www.mrlc.gov/nlcd06_leg.php (accessed on 15 January 2014).

35. Haralick, R.; Shanmugam, K.; Dinstein, I. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *SMC-3*, 610–621.

36. Barnsley, M.; Barr, S. Inferring urban land use from satellite sensor images using Kernel-based spatial reclassification. *Photogramm. Eng. Remote Sens.* **1996**, *62*, 949–958.

37. Zhang, Y. Optimisation of building detection in satellite images by combining multispectral classification and texture filtering. *ISPRS J. Photogramm. Remote Sens.* **1999**, *54*, 50–60.

38. Jasiewicz, J.; Netzel, P.; Stepinski, T. Content-based landscape retrieval using geomorphons. *Geomorphometry* **2013**. Available online: http://www.geomorphometry.org/system/files/Jasiewicz2013geomorphometry_0.pdf (accessed on 10 July 2014).

39. McGarigal, K. Landscape pattern metrics. *Encycl. Environmetrics* **2002**. Available online: https://www.umass.edu/landeco/pubs/mcgarigal.2002.pdf (accessed on 20 June 2014)

40. Do, M.N.; Vetterli, M. Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance. *IEEE Trans. Image Process.* **2002**, *11*, 146–158.

41. Puzicha, J.; Hofmann, T.; Buhmann, J. Non-parametric similarity measures for unsupervised texture segmentation and image retrieval. In Proceedings of the 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Juan, Puerto Rico, 17–19 June 1997; pp. 393–396.