



## A co-occurrence region based Bayesian network stepwise remote sensing image retrieval algorithm

Rui Zeng<sup>1,2,\*</sup>, Yingyan Wang<sup>1</sup>, Wanliang Wang<sup>2</sup>

<sup>1</sup>School of Electro-mechanical and Information Technology, Yi Wu Industrial & Commercial College, YiWu 322000, China

<sup>2</sup>College of Computer Science & Technology, Zhejiang University of Technology, Hangzhou 310014, China

\* Email of Corresponding author: jamse007@126.com

### ABSTRACT

Although scholars have conducted numerous researches on content-based image retrieval and obtained great achievements, they make little progress in studying remote sensing image retrieval. Both theoretical and application systems are immature. Since remote sensing images are characterized by large data volume, broad coverage, vague themes and rich semantics, the research results on natural images and medical images cannot be directly used in remote sensing image retrieval. Even perfect content-based remote sensing image retrieval systems have many difficulties with data organization, storage and management, feature description and extraction, similarity measurement, relevance feedback, network service mode, and system structure design and implementation. This paper proposes a remote sensing image retrieval algorithm that combines co-occurrence region based Bayesian network image retrieval with average high-frequency signal strength. By Bayesian networks, it establishes correspondence relationships between images and semantics, thereby realizing semantic-based retrieval of remote sensing images. In the meantime, integrated region matching is introduced for iterative retrieval, which effectively improves the precision of semantic retrieval.

*Keywords: the Bayesian network; Co-occurrence region; Remote sensing image retrieval.*

## Región de coocurrencia basada en una red bayesiana gradual hacia un algoritmo de recuperación de imágenes por teledetección

### RESUMEN

A pesar del gran número de trabajos en recuperación de imágenes basadas en contenido, poco se ha avanzado en la recuperación de imágenes por teledetección. Tanto la teoría como la aplicación de los sistemas son inmaduros. Ya que las imágenes por teledetección se caracterizan por un gran volumen de información, amplia cobertura, temas difusos y semántica abundante, los resultados de las investigaciones en imágenes naturales e imágenes médicas no pueden ser usados directamente en la recuperación de imágenes por teledetección. Incluso en una consulta perfecta de imágenes basadas en contenido, los sistemas tienen muchas dificultades con la organización de información, almacenamiento y manejo, descripción de características y extracción, medición de similitudes, retroalimentación relevante, modo de servicio de red y diseño e implementación del sistema estructural. Este artículo propone un algoritmo de recuperación de imágenes por teledetección que combina la co-ocurrencia local de una red bayesiana de recuperación de imágenes con el promedio de potencia de la señal de alta frecuencia. Por las redes bayesianas, se establecen las relaciones de correspondencia entre imágenes y semántica, además de permitir la recuperación de imágenes de teledetección a través de la semántica. Mientras tanto, se desarrolló el módulo de región integrada para la recuperación repetitiva, lo que mejora efectivamente la precisión de la recuperación semántica.

*Palabras clave: red bayesiana; región de coocurrencia; recuperación de imágenes por teledetección.*

### Record

Manuscript received: 05/07/2017

Accepted for publication: 19/01/2018

*How to cite item Zeng, R., Wang, Y., & Wang, W. (2018). A co-occurrence region based Bayesian network stepwise remote sensing image retrieval algorithm. Earth Sciences Research Journal, 21(1), 29-35*

DOI: <http://dx.doi.org/10.15446/esrj.v22n1.66107>

## 1. Introduction

Studies on natural and medial image retrieval have achieved remarkable results. Meanwhile, there have been numerous research projects designed specifically for remote sensing image retrieval. However, both theoretical and application systems are immature and need to be further researched (Samadzadegan et al., 2012; Hejazi et al., 2017). For example, the KIM system developed by DLR explores implied semantic features of images by Bayesian networks. Nanyang Technological University conducts multi-sensor (RS)2I thoroughly research feature description and extraction, multidimensional indexing technology and system structure design (Mountrakis, 2012). The study compares existing image retrieval systems and focuses on image storage, network transmission model, feature extraction and description, segmentation algorithm, reasonable segmentation, similarity measurement and relevance feedback. The results show that studies on remote sensing image retrieval have some deficiencies, which are mainly manifested as follows: (1) Although studies on content-based image retrieval have achieved brilliant results, they are mostly focused on natural and medical images. Rarely do scholars study remote sensing images (Han, 2012; Xu, 2015; Radan et al., 2017). (2) Currently, remote sensing image retrieval systems are mainly based on texts, strip numbers, latitudes, and longitudes, rather than contents (Abdi, 2013). (3) Scholars have made great use of Bayesian networks to study semantic-based remote sensing image retrieval. Retrieval precision of existing algorithms can be further improved. (4) The retrieval precision is inversely related to time efficiency for most image retrieval algorithms. In other words, high-precision image retrieval tends to consume lots of time (Zhai, 2014; Simon et al., 2017).

This paper proposes a stepwise Bayesian network algorithm for retrieving remote sensing images, aim to address these problems; the schema combines co-occurrence region-based Bayesian network image retrieval with average high-frequency signal strength, and adopts integrated region matching for iterative retrieval, thereby efficiently improving the precision of semantic retrieval and significantly reducing the retrieval time. In the meantime, semantic-based feature vectors are introduced to enhance the precision of image retrieval.

## 2. Material and Methods

In content-based image retrieval, scientific and reasonable similarity measurement models are the keys to achieving the precision of the entire retrieval system (Zhong et al., 2012; Bata et al., 2017). Bayesian network image retrieval systems consist of three parts (Boser et al., 1992; Roslee et al., 2017): (1) Segment images with a simple image segmentation algorithm and extract color or texture features of each sub-image to describe image contents, which is referred to as image segmentation and feature extraction. (2) Employ unsupervised classification to classify the extracted features, generate codebooks and encode images following the codebooks, thereby producing encoded image library. (3) Obtain the final retrieval results by selecting training samples and calculating related probabilities, which is known as semantic inference. The paper proposes two concepts: code co-occurrence matrix and semantic score function to simplify the process of semantic inference.

### 2.1 Co-occurrence Region Based Bayesian Network Algorithms for Remote Sensing Images Retrieval

#### 2.1.1 Features of IKONOS Images

IKONOS is the world's first high-resolution commercial satellite, with an orbit altitude of 681km, a revisit period of 3 days and a nadir image swath width of 11.3km. It has one panchromatic band and four multi-spectral bands (Pedergrana, 2013; Rahman et al., 2017). Wherein, the spectral resolution of the panchromatic band and multi-spectral bands is 1m and 4m respectively. Fundamental parameters of wave bands are shown in Table 1.

**Table 1.** Wave Bands of IKONOS Images

Wave Band	Name	Wavelength ( $\mu\text{m}$ )
1	Blue	0.045-0.516
2	Green	0.506-0.595
3	Red	0.632-0.698
4	Near Infrared	0.757-0.853
Panchromatic	~	0.526-0.929

#### Extraction

This study adopts a 256-dimensional color histogram to describe spectral features of images. In fact, Swain et al. proposed to describe color features of images with color histograms in 1991. Since color histogram, which presents the number of image pixels of each color after image colors is quantified, has transition invariance and strong robustness to rotation and scale changes, it is extensively used (Stumpf, 2014; Ruiz, 2014; Maleki-Ghelichi and Sharifi, 2017). Normalization processing shall be adopted, and color features shall be described by the proportion of

$$H(k) = n_k / N, k = 0, 1, 2, \dots, L-1 \quad (1)$$

different colors to keep the scale invariant; as shown in Formula (1).

In the formula, N is the total number of image pixels; L is the dimension of the color histogram;  $n_k$  is the number of image pixels in the k-th color dimension.

#### 2.1.3 Image Segmentation

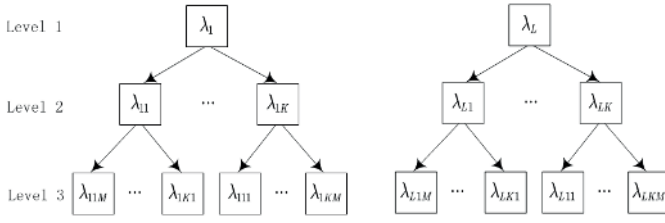
In the experiment, multi-spectral KONOS images are segmented into non-repetitive sub-images, and 256-dimensional color histograms of the sub-images are extracted and considered as their spectral features. However, the sub-images are in large quantities, and objects with similar characteristics are segmented into different sub-images, which is not conducive to image retrieval (Amiri et al., 2017). To better describe major features of images and reduce the number of feature vectors, the k-means algorithm is introduced to cluster segmented sub-images. The clustering results are deemed as the final segmentation results. Sub-image is taken as the minimum unit to judge whether image regions are spatially connected, preventing segmented regions from being spatially unconnected (Tariq et al., 2017). If not, the clustering results shall be segmented so that the images are spatially combined after segmentation.

## 2.2 Codebook and Image Coding

### 2.2.1 Codebook

Image features are accurately extracted for quantification and classified by the k-means algorithm to realize many-to-many matching between image features and semantic meanings. Assuming that all image areas in the database are expressed as  $D = \{R_1^1, \dots, R_d^1, \dots, R_d^N\}$ , N is the total number of image segmentation regions,  $F(D)$  is the region feature extraction function and  $F(D)$  represents features of all image areas in the database.  $F(D)$  is classified into L categories by the k-means algorithm, and each cluster  $\lambda_u (1 \leq u \leq L)$  is a set of features. Let  $C_u$  be the cluster center of  $\lambda_u$ , then there is a mapping relation  $M: R^n \rightarrow CB_i = \{c_1, \dots, c_u, \dots, c_L | c_u \in R^n\}$ . Wherein,  $CB_i$  is referred to as codebook and the index number between  $C_u$  and  $\lambda_u$  is referred to as code. For the image  $I_i$ , its region code  $R_i$  can be expressed as  $C_i(R_i)$ . Wherein,  $C_i$  is the code function of the region. Since codebooks generated by the k-means algorithm are all of the fixed length and appropriate codebook sizes are difficult to be identified, this paper adopts tree structured coding to solve the problem. Tree-structured codebooks are composed of different levels of  $CB_i^n$  (n represents the number of codebook levels). Figure 1 demonstrates a schematic drawing

of structured tree coding.



**Figure 1.** A Schematic Drawing of Three-Level Tree Structured Codebooks

According to Figure 1, top-level codes have more regional features, which flexibly and efficiently allows users to select codebooks required at different levels of  $CB_i^n$ . Based on the features of image areas, the chosen codebooks shall be used to encode images, thereby generating image library (encoded image library).

### 2.2.2 Image Coding

It is necessary to select proper codebooks from tree-structured codebooks to encode images. The coding function of image areas is defined as  $C_i^n(\cdot)$  ( $n$  represents the number of tree-structured codebook level), and the process of image coding is described as follows:

Step 1: select a  $m$ -level tree-structured codebook  $CB_i^n$  for  $R_i^j$  in  $I_i$  and its region features  $F_i(R_i^j)$ , and let the optimal codebook level  $l$  be 1.

Step 2: traverse each cluster  $\lambda_u$ , identify the cluster center  $C_u$  that contains clusters of regional features  $F_i(R_i^j)$  and encode the region  $R_i^j$  as  $u$ , then,  $l = l + 1$ .

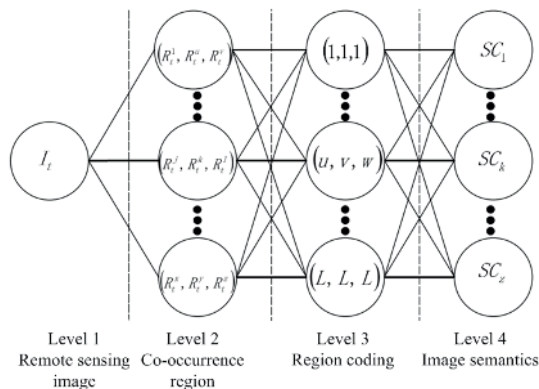
Step 3: if  $l > m$ , the program stops. Otherwise, it will go to Step 2.

Step 4: encode  $R_i^j$  into a code strand (each code represents clusters of image areas at different codebook levels), encode each image area and obtain encoded images, thereby forming encoded image library.

## 2.3 Semantic Inference

### 2.3.1 Mathematical Models

$R_i^j$ ,  $R_i^k$  and  $R_i^l$  coexist in the image  $I_i$ ,  $R_i^k$  and  $R_i^l$  are referred to as co-occurrence regions of  $R_i^j$ . Semantic inference is based on codes of  $R_i^j$  and its co-occurrence regions ( $R_i^k$  and  $R_i^l$ ). Spectral or texture features of the triad constituted by  $R_i^j$ ,  $R_i^k$  and  $R_i^l$  are described by corresponding code triad ( $C_i^n(R_i^j), C_i^n(R_i^k), C_i^n(R_i^l)$ ) (region code). Assuming that the codebook has  $L$  codes, it may generate  $L^3$  code combinations from the perspective of the combination. In this way, spectral or texture feature symbols of the region increased from  $L$  to  $L^3$ , thereby improving the precision of semantic inference. When inferring semantic meanings of images, the correlation between  $I_i$  and  $SC_k$  is measured by the posterior probability  $P(SC_k|I_i)$  of co-occurrence region based Bayesian networks. Co-occurrence region based Bayesian networks are shown in Figure 2:



**Figure 2.** Co-occurrence Region Based Bayesian Networks

Formula (2) is the formula for calculating the posterior probability of co-occurrence region based Bayesian networks.

$$P(SC_k|I_i) = \sum_{\substack{R_i^j, R_i^k, R_i^l \\ R_i^j, R_i^k \in cor(R_i^l)}} P[(R_i^j, R_i^k, R_i^l)|I_i] \sum_{u,v,w \in CB_i^n} P[SC_k|(u,v,w)]P[(u,v,w)|R_i^j, R_i^k, R_i^l] \quad (2)$$

According to Bayesian formula,

$$P[SC_k|(u,v,w)] = \frac{P[(u,v,w)|SC_k]P(SC_k)}{P[(u,v,w)]} \quad (3)$$

Formula (2) can be expressed as:

$$P(SC_k|I_i) = P(SC_k) \sum_{\substack{R_i^j, R_i^k, R_i^l \\ R_i^j, R_i^k \in cor(R_i^l)}} P[(R_i^j, R_i^k, R_i^l)|I_i] \sum_{u,v,w \in CB_i^n} \frac{P[(u,v,w)|SC_k]P[(u,v,w)|(R_i^j, R_i^k, R_i^l)]}{P[(u,v,w)]} \quad (4)$$

Where in,  $P[(u,v,w)]$  is the prior probability of  $(u,v,w)$ ; the conditional probability  $P[(u,v,w)|SC_k]$  reflects random links between codes and semantic concepts and can be obtained from user feedback or training samples. Assuming that the prior probability  $P(SC_k)$  follows a uniform distribution, it can be ignored when calculating  $P[(u,v,w)|(R_i^j, R_i^k, R_i^l)]$ . The calculation formula is as follows:

$$P[(u,v,w)|(R_i^j, R_i^k, R_i^l)] = \begin{cases} 1, & C_i^n(R_i^j) = u \wedge C_i^n(R_i^k) = v \wedge C_i^n(R_i^l) = w \\ 0, & \text{other} \end{cases} \quad (5)$$

When interpreting the image  $I_i$ , importance of the code triad ( $R_i^j, R_i^k, R_i^l$ ) can be described with  $P[(R_i^j, R_i^k, R_i^l)|I_i]$  and calculated by the product of region importance function  $RI(\cdot)$  and co-occurrence region importance function  $RCI(\cdot, \cdot)$ . Assuming that  $R_i^j$  and its co-occurrence regions ( $R_i^k, R_i^l$ ) are equally important to interpreting  $I_i$ , then  $RI(R_i^j) = 1/|I_i|$  and  $RCI(R_i^j, R_i^k, R_i^l) = 1/|cor(R_i^j)|$ . Where in, the function  $|\cdot|$  is element data in the set. If all co-occurrence regions are taken into account, the following formula is valid:

$$\sum_{R_i^j, R_i^k \in cor(R_i^l)} P[(R_i^j, R_i^k, R_i^l)|I_i] = \frac{1}{|I_i|} \quad (6)$$

Formula (6) prevents the bias toward regions with more co-occurrence regions when calculating  $P(SC_k|I_i)$ . Based on Formula (5) and the calculation formula of  $P[(R_i^j, R_i^k, R_i^l)|I_i]$ , Formula (4) can be simplified as:

$$P(SC_k|I_i) \propto \frac{1}{|I_i|} \sum_{\substack{R_i^j, R_i^k \\ R_i^j, R_i^k \in cor(R_i^l)}} \frac{1}{|cor(R_i^j)|} \sum_{u,v,w \in CB_i^n} \frac{P[(u,v,w)|SC_k]P[(u,v,w)|(R_i^j, R_i^k, R_i^l)]}{P[(u,v,w)]} \quad (7)$$

### 2.3.2 Code Co-occurrence Matrix

The codes are defined to describe the importance and generate code co-occurrence matrix to simplify the calculation of  $P(SC_k|I_i)$ ,

The correlation between  $(u,v,w)$  and  $SC_k$  is reflected by the code triad importance function  $CPI_k(\cdot)$ . It is possible to determine the importance of the code triad to  $(u,v,w)$  by virtue of prior probability of  $SC_k$ . The calculation formula is as follows:

$$CPI_k(u,v,w) = P[SC_k|(u,v,w)] = \frac{P[(u,v,w)|P(SC_k)]}{P(u,v,w)} \quad (8)$$

Wherein,  $u$ ,  $v$  and  $w$  are codes that jointly constitute code co-occurrence matrix  $Mk$ . Assuming that the prior probability  $P(SC_k)$  follows a uniform distribution, Formula (7) can be simplified as:

$$P(SC_k | I_i) \propto \frac{1}{|I_i|} \sum_{\substack{R_i^j \in I_i \\ R_i^k, R_i^l \in \text{cor}(R_i^j)}} \frac{1}{|\text{cor}(R_i^j)|} \sum_{R_i^k, R_i^l \in \text{cor}(R_i^j)} M_k[C_i^n(R_i^j), C_i^n(R_i^k), C_i^n(R_i^l)] \quad (9)$$

### 2.3.3 Semantic Score Function

Semantic scores can be divided into region semantic scores and semantic image scores. Region semantic score function is used to calculate the correlation between  $SC_k$  and  $R_i^j$ :

$$SR(R_i^j) = \frac{1}{|\text{cor}(R_i^j)|} \sum_{R_i^k, R_i^l \in \text{cor}(R_i^j)} M_k[C_i^n(R_i^j), C_i^n(R_i^k), C_i^n(R_i^l)] \quad (10)$$

The normalization coefficient  $|\text{cor}(R_i^j)|$  prevents bias of semantic scores toward regions with more co-occurrence regions. According to the region semantic score function, the image semantic score function  $SI_k(\cdot)$  can be defined as:

$$SI_k(I_i) = \frac{1}{|I_i|} \sum_{R_i^j \in I_i} SR_k(R_i^j) \quad (11)$$

The image semantic score function is used to measure the correlation between  $I_i$  and  $SC_k$ . If  $SI_k(I_i)$  is greater than the predefined thresholds,  $I_i$  is deemed to contain the semantic concept  $SC_k$  at least. According to Formula (9) and Formula (11),

$$P(SC_k | I_i) \propto SI_k(I_i) \quad (12)$$

It should be noted that this is a special form of Formula (3), which indicates that the image semantic score function is consistent with co-occurrence region based Bayesian networks.

### 2.3.4 Learning with Bayesian Networks

The prior probability  $P[(u, v, w) | SC_k]$  can be obtained from a group of training samples that reflect semantic concept  $SC_k$  and consist of  $R_i^j$  and its co-occurrence regions to calculate the semantic score of images. The probability of the code triad can be calculated through Formula (13):

$$CTF_k(u, v, w) = \sum \mathcal{E}[R_k^i, R_k^j, R_k^m, (u, v, w)] \quad \forall R_k^i \in T_k \wedge R_k^j, R_k^m \in \text{cor}(R_k^i) \quad (13)$$

Since  $R_k^j$  and  $R_k^m$  may not belong to  $T_k$ , the choice function  $\mathcal{E}$  is defined as:

$$\mathcal{E}[R_k^i, R_k^j, R_k^m, (u, v, w)] = \begin{cases} 1, & C_i^l(R_k^i) = u \wedge C_i^l(R_k^j) = v \wedge C_i^l(R_k^m) = w \\ 0, & \text{other} \end{cases} \quad (14)$$

The prior probability  $P[(u, v, w) | SC_k]$  can be approximately calculated through Formula (15):

$$P((u, v, w) | SC_k, T_k) = \frac{CPF_k(u, v, w)}{\sum_{u, v, \xi} CPF_k(u, v, \xi)} \quad (15)$$

If users submit a new training sample  $T'_k$  and merge it with the original training sample  $T_k$ , the prior probability  $P[(u, v, w) | SC_k]$  shall be recalculated through Formula (16):

$$P((u, v, w) | SC_k, T'_k, T_k) = \frac{CPF_k(u, v, w) + CPF'_k(u, v, w)}{\sum_{u, v, \xi} [CPF_k(u, v, \xi) + CPF'_k(u, v, \xi)]} \quad (16)$$

If there are intersections between the new training sample and the original training sample, the calculation formula of the prior probability is  $P[(u, v, w) | SC_k, T'_k \cup T_k]$ .

## 2.4 Stepwise Remote Sensing Image Retrieval

Traditional content-based remote sensing image retrieval systems tend to calculate the distance between low-level feature vectors. Due to semantic gaps, however, images with similar features and irrelevant semantics are always retrieved (Ali et al., 2017). In the process of image retrieval, time efficiency is also a key factor that determines the performance of the retrieval system. Users require not only high precision and recall but also demand high time efficiency.

### 2.4.1 Integrated Region Matching

Integrated region matching (IRM), which has the strong robustness to possible segmentation errors, determines the overall similarity of images by comprehensively calculating the similarity of all image regions. Mathematically speaking, image similarity can be determined by calculating the distance between two point sets in high-dimension space (feature space). Although there are many methods of calculating the distance, such as Euclidean distance, between points in high-dimension space, they cannot be used to calculate the distance between two point sets. The biggest problem of defining the distance between two point sets in feature space is to maximize the difference between overall similarity and people's subjective feelings of image similarity. By comprehensively considering significance and distance between image regions, integrated region matching allows one region to match multiple regions, thereby minimizing retrieval errors caused by inaccurate image segmentation (Yasin et al., 2017).

According to the research results of Wang et al. (2016), the algorithm is described as follows: Image 1 and Image 2 are represented by  $R_1 = (r_1, r_2, \dots, r_n)$  and  $R_2 = (r'_1, r'_2, \dots, r'_n)$ , and feature vectors of the segmentation region  $i$  and  $j$  are represented by  $r_i$  and  $r'_j$  respectively. Also, the distance between  $r_i$  and  $r'_j$  is  $d(r_i, r'_j)$ . When calculating the distance between  $R_1$  and  $R_2$ , it is necessary to match all regions of the two images, calculate the distance between all regions and give corresponding weights to significance factors  $S_{i,j}$ . Since significance factor represents the matching degree between  $r_i$  and  $r'_j$ ,  $S = \{s_{i,j}\}, 1 \leq i \leq m, 1 \leq j \leq n$  is referred to as significance matrix.

$$d(R_1, R_2) = \sum_{i,j} s_{i,j} d_{i,j} \quad (17)$$

The significance matrix  $S$  of integrated region matching is determined by significance factors  $s_{i,j}$  – the area ratio between one region and the whole image. Assuming that the area ratio of  $r_i$  in Image 1 and  $r'_j$  in Image 2 is  $p_i$  and  $p'_j$  separately, then

$$\sum_{j=1}^n s_{i,j} = p_i, i = 1, \dots, m \quad (18)$$

$$\sum_{i=1}^m s_{i,j} = p'_j, j = 1, \dots, n \quad (19)$$

Under standard conditions,  $\sum_{i=1}^m p_i = \sum_{j=1}^n p'_j = 1$ . A reasonable matching mechanism is required to involve all image regions and give priority to the matching of the most similar regions. Under this matching mechanism, regions in Image 1 will only match with similar regions in Image 2. In other words, the distance between the two images is 0. By distributing significance factors, integrated region matching connects the most similar regions in two images with the shortest distance, which is referred to as the most similar highest priority (MSHP).

It is assumed that  $d(i', j')$  is the minimum distance,  $s_{i',j'} = \min(p_{i'}, p'_{j'})$  and  $p_{i'} \leq p'_{j'}$  to obtain significance matrix with an iterative method. When  $s_{i',j'} = 0, j' \neq j'$ ,  $i'$  and  $j'$ , which are regions with the minimum matching distance, are successfully matched together by significance factors of  $i'$ . The significance factor of  $j'$  is calculated through  $p'_{j'} - p_{i'}$ . According to MSHP, the matching problem of  $s_{i',j'}, i' \neq i'$  can be solved when the following four formula requirements are met.

$$\sum_{j=1}^n s_{i,j} = p_i, 1 \leq i \leq m, i \neq i' \quad (20)$$

$$\sum_{i: 1 \leq i \leq m, i \neq i'} s_{i,j} = p'_j, 1 \leq j \leq n, j \neq j' \quad (21)$$

$$\sum_{i: 1 \leq i \leq m, i \neq i'} s_{i,j} = p'_j - p_{i'} \quad (22)$$

$$s_{i,j} > 0, 1 \leq i \leq m, i \neq i', 1 \leq j \leq n \quad (23)$$

The iteration stops when all significant factors are calculated.

By co-occurrence region based Bayesian networks, the algorithm selects a group of candidate images that are highly correlated with the query semantics from large-scale remote sensing image database and adopts integrated region matching to reorder candidate images and return top ranking images, thereby realizing rapid image retrieval. Since co-occurrence region based Bayesian network image retrieval is characterized by low time complexity, it can retrieve a group of candidate images that are highly correlated with the query semantics within a short time (Shahzad et al., 2017; Basarian and Tahir, 2017). Integrated region matching has high time complexity. In comparison with the entire database, candidate images are small in quantities, which effectively reduce image retrieval time. Flowchart of the whole process is shown in Figure 3.

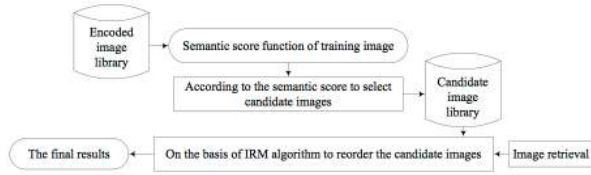


Figure 3. A Flowchart of Stepwise Region Retrieval

Stepwise region retrieval is mathematically described as follows:

Step 1: select a group of candidate images that are highly correlated with the query semantics. There are two alternatives:

(1) Users choose candidate images by searching images by the image semantic score function trained from co-occurrence region based Bayesian networks. Let  $I_q$  be the query image and  $\delta_k$  be the threshold of semantic image score; then the candidate images can be expressed as:

$$A = \{I_p | SI_k(I_p) \geq \delta_k\} \quad (24)$$

(2) Users select candidate images by searching semantic information of pictures by semantic score function. Assuming that  $I_q$  has two semantic meanings,  $SI_k^1(\cdot)$  and  $SI_k^2(\cdot)$  are the two semantic score functions, and  $\delta_k^1$  and  $\delta_k^2$  are thresholds of the two functions, then the candidate images selected can be expressed as:

$$A = \{I_p | SI_k^1(I_p) \geq \delta_k^1 \wedge SI_k^2(I_p) \geq \delta_k^2\} \quad (25)$$

Under normal circumstances, the threshold shall be set to 0, to ensure the precision of image retrieval.

Step 2: adopt integrated region matching and query  $I_q$  to reorder the candidate image gallery A, return top ranking images and finish the search. The more the returned images are similar to the query images in spectral and texture features, the more top-ranked the images will be.

### 3. Results

This experiment involves a total of 28 multi-spectral remote sensing images photographed by IKONOS and synthesizes standard true-color images with blue, green and red wave bands. Each image is segmented into 256x256 sub-images, and the maximum repeatability between two sub-images is 50%. After deleting unqualified images, the database has a total of 12,000 sub-images, covering nearly 40,000km<sup>2</sup> earth surfaces. Images in the database are classified into six categories following semantic meanings: farmland (644), city (3,547), water body (3,243), vegetation (2,937), bare earth (1,033) and rock (1,730). Wherein, the figures in brackets represent numbers of images related to corresponding semantic concepts, and one image may be correlated with multiple semantic concepts. This paper measures performance of the retrieval system with precision and recall.

In the experiment, the author segments images into 32x32 sub-images extracts color histograms of regions with spectral features and employs the k-means algorithm for image clustering. When generating encoded image library, the color feature vectors are classified into tree codes by the k-means algorithm, thereby generating codebooks (Hussin et al., 2017). Each image area is encoded into corresponding codes following its feature vectors, thereby making encoded image library. This experiment creates three-level codebooks: 150, 300 and 600.

Experiments are designed to verify the retrieval algorithm proposed, which focuses on co-occurrence region based Bayesian network stepwise remote sensing image retrieval, verifies precision and time efficiency of the algorithm.

#### 3.1 Experiment Design and Results

Considering that the retrieval time of integrated region matching may be influenced by the segmentation region numbers of query images, the experiment selects images with 4 to 13 image regions, with integrated region matching and stepwise image retrieval as controlled trails. Each retrieval experiment is repeated for three times, to ensure the accuracy of time, and the average of the retrieval time is deemed as the experiment result. Since stepwise image retrieval experiment records the time consumed in Step 1 and Step 2, it is possible to obtain the total retrieval time. This experiment measures performances of retrieval systems by retrieving time and precision. The retrieval time of integrated region matching and stepwise image retrieval is shown in Table 2, and the comparison results are shown in Figure 4(a). Since integrated region matching belongs to sorting algorithm, rather than classification algorithm, it selects the first 100 images for precision calculation. As for stepwise image retrieval, it will fetch the first 100 images if it retrieves more than 100 images; otherwise, it will be subject to the number of images returned. The final retrieval precision is demonstrated in Figure 4(b).

Table 2. A Comparison of Retrieval Time between Integrated Region Matching and Stepwise Image Retrieval

Image region	IRM(s)	150 yards			300 yards			600 yards		
		1st	2nd	Sum	1st	2nd	Sum	1st	2nd	Sum
4	2.934	1.309	0.028	1.337	1.414	0.01	1.424	2.377	0.01	2.387
5	3.437	1.306	0.037	1.343	1.434	0.012	1.446	2.442	0.002	2.444
6	4.034	1.311	0.071	1.382	1.446	0.052	1.498	2.353	0.031	2.384
7	4.595	1.308	0.072	1.38	1.439	0.049	1.488	2.381	0.031	2.412
8	5.18	1.311	0.213	1.524	1.432	0.07	1.502	2.386	0.036	2.422
9	5.696	1.312	0.115	1.427	1.439	0.105	1.544	2.366	0.1	2.466
10	6.244	1.306	0.317	1.623	1.427	0.158	1.585	2.401	0.05	2.451
11	6.75	1.316	0.511	1.827	1.463	0.288	1.751	2.356	0.211	2.567
12	7.373	1.294	0.567	1.861	1.432	0.236	1.668	2.345	0.117	2.462
13	7.734	1.322	0.145	1.467	1.438	0.081	1.519	2.381	0.078	2.459
Average	5.398	1.31	0.208	1.517	1.436	0.106	1.543	2.379	0.067	2.445

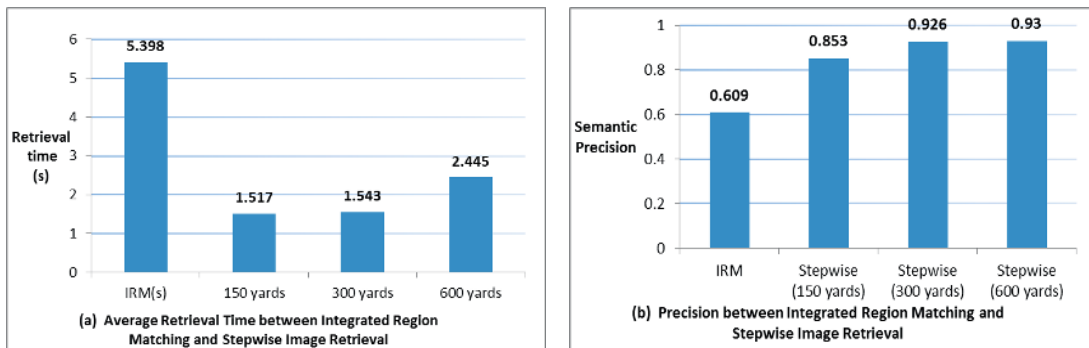


Figure 4. A Comparison of between Integrated Region Matching and Stepwise Image Retrieval

4. Discussion

4.1 Time Efficiency Analysis

According to Table 2, the retrieval time of integrated region matching increases gradually with the increase in region numbers, which is mainly because the number increase adds difficulties to the computation. Nevertheless, the increase in region numbers has little impact on the retrieval time of stepwise image retrieval. It fluctuates around 1.3s, 1.4s and 2.35s at 150, 300 and 600 yards respectively. With the increase of codebook numbers, however, the retrieval time increases accordingly. On average, the retrieve time rises by 0.126s from 150 yards to 300 yards and by 0.943s from 300 yards to 600 yards. Due to the increase of codebook numbers, code co-occurrence matrices appear. The codebooks at 300 yards are eight times those at 150 yards, and the codebooks at 600 yards are eight times those at 300 yards, which indicates that the ratio of retrieval time difference is 1:8. Through experiment, the ratio is proved to be 1:7.4841, which is consistent with the speculation. The retrieval time of Step 2 is related to the number of candidate images returned in Step 1. It can be seen from Figure 4(a) that the retrieval time of stepwise image retrieval is significantly shorter than that of

integrated region matching, which is because co-occurrence region based remote sensing image retrieval has lower time complexity.

4.2 Recall Analysis

According to Figure 4(b), the precision of integrated region matching is 0.609, which is significantly higher than that of stepwise image retrieval (0.853, 0.926 and 0.93 at 150, 300 and 600 yards respectively). The reasons are as follows: images retrieved by integrated region matching are merely similar to each other in respect to low-level features, without considering semantic similarity. Stepwise image retrieval selects candidate images highly correlated with query images and retrieves images with similar characteristics to query images by integrated region matching, which effectively improves the precision of image retrieval.

4.3 Retrieval Cases

This paper selects an image of city region for retrieval analysis and 600 yards for stepwise image retrieval. The first ten images retrieved are shown in Table 3: arranged from left to right and from top to bottom.








Image to be retrieved	Stepwise remote sensing image retrieval algorithm (600 yards)	Integrated region matching
		
		
		
		
		
		
		
		
		
		

Table 3. A Comparison of Retrieval Results between Stepwise Image Retrieval and Integrated Region Matching

## 5. Conclusions

By considering remote sensing image features and remote sensing image retrieval, the paper proposes a co-occurrence region based Bayesian network stepwise remote sensing image retrieval algorithm that takes into account both time efficiency and retrieval precision. This algorithm mainly draws reference from stepwise remote sensing image retrieval proposed by Moustakidis et al. (2012) and based on CSBN and integrated region matching. It consists of two parts: co-occurrence region based remote sensing image retrieval and integrated region matching. Select a group of candidate images that are highly correlated with query images in semantics by co-occurrence region based remote sensing image retrieval and reorder the candidate images with integrated region matching, thereby obtaining the final retrieval results. From the perspective of time efficiency, stepwise remote sensing image retrieval consumes 1.5171s, 1.5425s and 2.4454s at 150, 300 and 600 yards respectively. Regarding the precision, integrated region matching enjoys an accuracy of 0.609 while stepwise image retrieval significantly improves the precision to 0.853, 0.926 and 0.93 at 150, 300 and 600 yards respectively.

## Acknowledgments

This research was financially supported by Science and technology project of Zhejiang Province (Grant No. 2015C31171) and National education information technology research plan project (Grant No. 146241806) (Grant No. 136241559).

## References

- Abdi, M. J., & Giveki, D. (2013). Automatic detection of erythematous diseases using PSO-SVM based on association rules. *Engineering Applications of Artificial Intelligence*, 26(1), 603–608.
- Ali, S. S., Ijaz, N., Aman, N., Nasir, D. A., Anjum, D. L., & Randhawa, D. I. A. (2017). Clinical Waste Management Practices In District Faisalabad. *Earth Sciences Pakistan*, 1(2), 01–03.
- Amiri, M., Arabhosseini, A., Kianmehr, M. H., Mehrjerdi, M. Z., & Mirsaeedghazi, H. (2017). Environmental impact assessment of total alkaloid extracted from the *Atropa belladonna* L. using LCA. *Geology, Ecology, and Landscape*, 1(4), 259–263.
- Basarian, M. S., & Tahir, S. H. (2017). Groundwater Prospecting Using Geoelectrical Method at Kg Gana, Kota Marudu, Sabah. *Earth Science Malaysia*, 1(2), 7–9.
- Bata T., Samaila, N. K., Maigari, A. S., Abubakar, M. B., & Ikyoive, S. Y. (2017). Common Occurrences Of Authentic Pyrite crystals in Cretaceous Oil Sands as Consequence of Biodegradation Processes. *Geological Behavior*, 1(2), 26–30.
- Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). *A training algorithm for optimal margin classifiers*. In: Proceedings of the fifth annual workshop on Computational learning theory, Pittsburgh, pp. 144–152.
- Han, M., Zhu, X., & Yao, W. (2012). Remote sensing image classification based on neural network ensemble algorithm. *Neurocomputing*, 78(1), 133–138.
- Hejazi, S. M., Lotfi, F., Fashandi, H., & Alirezazadeh, A. (2017). Serishm: an eco-friendly and biodegradable flame retardant for fabrics. *Environment Ecosystem Science*, 1(2), 05–08.
- Hussin, H., Fauzi, N., Jamaluddin, T. A., & Arifin, M. H. (2017). Rock Mass Quality Effected by Lineament Using Rock Mass Rating (RMR) – Case Study from Former Quarry Site. *Earth Science Malaysia*, 1(2), 13–16.
- Maleki-Ghelichi, E., & Sharifi, M. (2017). Prioritize and choose the best process of anaerobic digestion to produce energy from biomass using analytic hierarchy process (AHP). *Geology, Ecology, and Landscapes*, 1(4), 219–224.
- Moustakidis, S., Mallinis, G., Koutsias, N., Theocharis, J. B., & Petridis, V. (2012). SVM-based fuzzy decision trees for classification of high spatial resolution remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 50(1), 149–169.
- Pedergnana, M., Marpu, P. R., Dalla-Mura, M., Benediktsson, J. A., & Bruzzone, L. (2013). A novel technique for optimal feature selection in attribute profiles based on genetic algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, 51(6), 3514–3528.
- Radan, A., Latifi, M., Moshtaghi, M., Ahmadi, M., & Omidi, M. (2017). Determining the Sensitive Conservative Site in Kolah Ghazi National Park, Iran, In Order to Management Wildlife by Using GIS Software. *Environment Ecosystem Science*, 1(2), 13–15.
- Rahman, N. A., Tarmudi, Z., Rosdy, M., & Muhiddin, F. A. (2017). Flood Mitigation Measures Using Intuitionistic Fuzzy Dematel Method. *Malaysian Journal Geosciences*, 1(2), 01–05.
- Roslee, R., Mickey, A. C., Simon, N., Norhisham, M. N. (2017). Landslide Susceptibility Analysis (LSA) Using Weighted Overlay Method (WOM) Along the Genting Sempah To Bentong Highway, Pahang. *Malaysian Journal Geosciences*, 1(2), 13–19.
- Ruiz, P., Mateos, J., Camps-Valls, G., Molina, R., & Katsaggelos, A. K. (2014). Bayesian active remote sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 52(4), 2186–2196.
- Samadzadegan, F., Hasani, H., Schenk, T. (2012). Simultaneous feature selection and SVM parameter determination in classification of hyperspectral imagery using ant colony optimization. *Canadian Journal of Remote Sensing*, 38(2), 138–156.
- Shahzad, A., Munir, M. H., Yasin, M., Umar, M., Rameez, S., Samad, R., Altaf, S., & Sarfraz, Y. (2017). Biostratigraphy of Early Eocene Margala Hill Limestone in The Muzaffarabad Area (Kashmir Basin, Azad Jammu And Kashmir). *Pakistan Journal of Geology*, 1(2), 16–20.
- Simon, N., Roslee, R., & Lai, G. T. (2017). Temporal Landslide Susceptibility Assessment Using Landslide Density Technique. *Geological Behavior*, 1(2), 10–13.
- Stumpf, A., Lachiche, N., Malet, J. P., Kerle, N., Puissant, A. (2014). Active learning in the spatial domain for remote sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 52(5), 2492–2507.
- Tariq, W., Hussain, S. Q., Nasir, D. A., Tayyab, N., Gillani, S. H., & Rafiq, A. (2017). Experimental Study on Strength And Durability Of Cement And Concrete By Partial Replacement Of Fine Aggregate With Fly Ash. *Earth Sciences Pakistan*, 1(2), 07–11.
- Wang, Q., Lin, J., & Yuan, Y. (2016). Salient band selection for hyper spectral image classification via manifold ranking. *IEEE Transactions on Neural Networks and Learning System*, 27(6), 1279–1289.
- Xu, L., Li, Y. P., Li, Q. M., Yang, Y. W., Tang, Z. M., & Zhang, X. F. (2015). Proportional fair resource allocation based on hybrid ant colony optimization for slow adaptive OFDMA system. *Information Sciences*, 293, 1–10.
- Yasin, M., Shahzad, A., Abbasi, N., Ijaz, U., & Khattak, Z. (2017). The Use of Stratigraphic Section in Recording Quagmire of Information For The Fluvial Depositional Environment – A Worked Example In District Poonch, Azad Jammu And Kashmir, *Pakistan. Pakistan Journal of Geology*, 1(2), 01–02.
- Zhai, S., & Jiang, T. (2014). A novel particle swarm optimization trained support vector machine for automatic sense-through-foilage target recognition system. *Knowledge-Based Systems*, 65, 50–59.
- Zhong, Y., & Zhang, L. (2012). An adaptive artificial immune network for supervised classification of multi-/hyperspectral remote sensing imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 50(3), 894–909.