An Efficient Eigen-Space Approach
for Management of Satellite Image Databases

Lena Chang\textsuperscript{1}, S. J. Chang\textsuperscript{2}, S. W. Leu\textsuperscript{2} and J. D. Chen\textsuperscript{2}
\textsuperscript{1}Department of Communications and Guidance Engineering,
\textsuperscript{2}Department of Electrical Engineering,
National Taiwan Ocean University, Taiwan, R.O.C

\textbf{Abstract} - In the study, we propose an efficient approach, which consists of an algorithm to speed up the multi-dimensional image segmentation and an efficient automatic mechanism to manage satellite image databases. Concerning image segmentation, we develop a one-dimensional (1D) segmentation technique by eigen-subspace projection approach, which can perform the image segmentation efficiently by transforming the multi-dimensional image data into 1D projection length. In addition, we develop a data model of eigen-index by probabilistic approach. Based on the maximum a posterior probability (MAP) information criterion, the management scheme may determine the accuracy of the user added data entry. Therefore, the DB archive can be managed efficiently. Simulation results performed on SPOT images have demonstrated the proposed approach is suitable for management of satellite image databases.

\section{I. INTRODUCTION}

Multispectral satellite image has been used in a myriad of applications covering areas in global environments monitoring, land use planning, mapping and natural resource management. With such a diverse set of applications, building a satellite image database (DB) has become white hot in this era of Internet and the World Wide Web. Traditional search and retrieval systems for remotely satellite image, for example, the National Oceanic and Atmospheric Administration’s (NOAA) Satellite Active Archive Web Interference, provide facilities to query by sensor, temporal and location characteristics. For theses repositories, after retrieving a collection of records in response to a query, the user must browse the intensive retrieved images and decide whether a certain image is useful for a particular application.

Some intelligent remote sensing information systems attempting to provide novel access and retrieval techniques are designed for the respective applications such as automatic forest inventories and sea ice analysis of SAR images. Recently, based on Bayesian networks \cite{1} proposed an interactive learning and probabilistic retrieval in remote sensing archives. And \cite{2} developed a system to determine automatically whether an region-of-interest is visible in the image. In our previous study \cite{3}, we developed an eigen-index technique for an interactive content-based satellite image retrieval. In \cite{3}, the image is first partitioned into some eigen-regions according to local terrain characteristics by the eigen-region based segmentation technique \cite{4}. Then, an eigen-index is extracted for each region. Using the index differences between the specified cover type of the query image and that of each object image in the DB, the access of the archive could be facilitated.

In this study, to improve the overall archive access performance of \cite{3}, we propose an efficient approach, which consists of an algorithm to speed up the multi-dimensional (MD) image segmentation and an efficient automatic mechanism to manage the image DB. We first develop a one-dimensional (1D) segmentation technique by eigen-subspace projection approach, which can perform the image segmentation efficiently by transforming the MD image data into 1D projection length. After the transformation, any efficient 1D segmentation technique, such as moment-preserving method \cite{5}, can be applied in the segmentation. Then using the segmentation results, we extract the image feature of each region by the eigen-indexing technique \cite{3}, and allow an interactive DB retrieval. Furthermore, to develop an automatic management mechanism to maintain and update the DB archive regularly, we develop a data model of eigen-index by probabilistic approach. Based on the maximum a posterior probability (MAP) information criterion, the management scheme may determine the accuracy of the user added data entry. Therefore, the DB archive can be managed efficiently.

This paper is organized as follows. In Section II, we develop an novel segmentation technique by eigen-subspace projection approach. Then, the proposed segmentation technique is applied to interactively retrieve multi-spectral images database and an automatic management mechanism is proposed in Section III. Finally, some experiment results and discussions are shown in Section IV.

\section{II. IMAGE SEGMENTATION USING EIGEN-SUBSPACE PROJECTION APPROACH}

A typical satellite image exhibits a number of different terrains and each terrain has unique spectral signature. \cite{4} proposed the eigen-region-based segmentation technique which determines the similarity of local terrain characteristics of adjacent image data blocks by comparing the separation between image signature vector of a chosen image data block and those of its nearest 8 blocks. The eigen-region-based segmentation technique can divide the image into proper regions according to their local terrain characteristics. However, the segmentation procedure in \cite{4}
performs MD search and consumes intensive computation time to do the image partition.

To improve the efficiency of [4], a novel projection method is proposed for multi-spectral image in this section. The approach first divides the image into blocks with fixed block size, such as 2x2, 4x4, or 8x8. The image feature of each block is estimated by using the principal eigenvector corresponding to the correlation matrix of image data block, $\phi_t$, which is denoted as the image signature vector. To perform multi-spectral image segmentation efficiently, we project the multi-spectral image data onto one referenced subspace. The referenced subspace is formed by the image signature vector, $\phi_t$, of one referenced image block. Then we project the image signature vector of each image block, $\phi_t$, onto the referenced subspace, and get a sequence a sequence of 1D projection length,

$$d_j = \|P_r \phi_j\|$$  \hspace{1cm} (1)

where $P_r$ is the projection matrix of $\phi_j$. By the way, we transform the multi-spectral image data into 1D projection length which represents the separation between the image signature vector of each image block and that of one referenced block.

Next, we adopt a proper 1D segmentation method to perform image partition. In this paper, we use 1D binary moment preserving method [5] as an example. By utilizing the threshold $\xi$ calculated by [5], we may separate the image into 2 regions. One region formed by the image blocks with $d_i < \xi$ is referred as the referenced region, in which the image blocks have similar image characteristics to that of the referenced image block. The other formed by the image blocks with $d_i > \xi$, has significantly different image characteristics with that of the referenced block. The separation procedure may be performed iterative. In each iteration, one referenced image block is chosen form the region which will be further partition. By changing the referenced block, we project the image onto an alternative referenced subspace in each iteration. Thus, by using the projection approach, we transform the MD segmentation into a sequence of 1D binary segmentation. The proposed segmentation technique using projection approach is summarized as

1. Divide the image into blocks with fixed block size.
2. Compute the principal eigen-vector, $\phi_j$, as image signature vector, for each block.
3. Select one reference image block and employ its signature vector to form the referenced subspace.
4. Project the image signature vector $\phi_j$ onto the referenced subspace, and a sequence of projection length.
5. Determine a threshold $\xi$ by moment preserving method, and partition the image into 2 regions.
6. Repeat 3 to 5 until the image has been properly segmented.

The following simulations will validate the proposed method improving significantly in computations than the eigen-region-based segmentation technique.

III. AUTOMATIC DATABASE MANAGEMENT MECHANISM

In this section, we apply the proposed segmentation technique to interactively retrieve multi-spectral images DB. Furthermore, to manage the DB efficiently, we then propose an automatic management scheme using data modeling for image cover types.

(A) Interactive Database Retrieval

After the image has been properly segmented into some regions via the above segmentation method, we then extract the image feature for each region by the eigen-index technique [3]. The segmentation results and the eigen-index for each region are stored in the DB. Then, we set up an interactive interface. Through the interface, a user may select the cover type of his interest, named $C_j$, on the client computer via Web. According to the content of specified cover type, the server generates the associated eigen-index, $e_q$, in real time. By comparing the eigen-index of the query image, $e_q$, with those of candidate images, $e_j$, in the DB, we may retrieval the best match object image in the DB by searching the minimum separation angles between the eigen-index vector $e_j$ and $e_q$.

$$\min_{e_j} D(e_q, e_j) = \min_{e_j} \cos^{-1}(e_q^T, e_j)$$  \hspace{1cm} (2)

Furthermore, the user specifies cover type name and its associated eigen-index can be stored in the DB for the future query.

(B) Management Mechanism by Using Data Modeling

For the interactive nature of DB retrieval, it is essential to set up a management scheme for managing and updating DB regularly. To determine the accuracy of the user specified cover type $C_j$ and its associated eigen-index $e_j$, which are stored in the DB, we apply the decision rule maximum a posterior probability (MAP) to make a decision as

$$P(C_j \mid e_j) > P(C_i \mid e_j) \hspace{1cm} i \neq j$$  \hspace{1cm} (3)

where $C_i$, $i \neq j$, represents other image cover types, and $P(C_j \mid e_j)$ specifies the probability in believing the cover type $C_j$ under the condition that $e_j$ is known. If (3) is satisfied, then we accept the specified cover type $C_j$ and its corresponding eigen-index $e_j$. Otherwise, this entry will be deleted. For example, $C_j$ being that a user thinks of “sea”,

$$\min_{e_j} D(e_q, e_j) = \min_{e_j} \cos^{-1}(e_q^T, e_j)$$  \hspace{1cm} (2)

Furthermore, the user specifies cover type name and its associated eigen-index can be stored in the DB for the future query.

(B) Management Mechanism by Using Data Modeling

For the interactive nature of DB retrieval, it is essential to set up a management scheme for managing and updating DB regularly. To determine the accuracy of the user specified cover type $C_j$ and its associated eigen-index $e_j$, which are stored in the DB, we apply the decision rule maximum a posterior probability (MAP) to make a decision as

$$P(C_j \mid e_j) > P(C_i \mid e_j) \hspace{1cm} i \neq j$$  \hspace{1cm} (3)

where $C_i$, $i \neq j$, represents other image cover types, and $P(C_j \mid e_j)$ specifies the probability in believing the cover type $C_j$ under the condition that $e_j$ is known. If (3) is satisfied, then we accept the specified cover type $C_j$ and its corresponding eigen-index $e_j$. Otherwise, this entry will be deleted. For example, $C_j$ being that a user thinks of “sea”,

$$\min_{e_j} D(e_q, e_j) = \min_{e_j} \cos^{-1}(e_q^T, e_j)$$  \hspace{1cm} (2)

Furthermore, the user specifies cover type name and its associated eigen-index can be stored in the DB for the future query.
given the associated eigen-index \( e_j \) in DB. The management tool allows the user to make a decision of “sea”, versus another type, such as “city” or “vegetation”.

The posterior probability function in (3) could equivalently estimated by the priori probability function \( P(e_j | C_j) \), which is the probability of the eigen-index \( e_j \) under the condition that the image cover type \( C_j \) is known. Therefore, the MAP decision rule in (3) could be simplified as

\[
P(e_j | C_j)P(C_j) > P(e_j | C_i)P(C_i)
\]

(4)

where \( P(C_j) \) is the probability of cover type \( C_j \), and may be estimated by \( P(C_j) = N_j / N \). In this estimation, \( N_j \) is the image pixels belonging to cover type \( C_j \) and \( N \) represents the image block size.

By collecting a large of sampled data, we observed that the distribution of the eigen-index \( e_j \) given the cover type \( C_j \) could be approximated by Gaussian distribution with mean vector \( \mu_{e_j} \) and covariance matrix \( K_{e_j} \). By using the linear transformation of \( e_j \),

\[
Z_j = A(e_j - \mu_{e_j})
\]

(5)

we may get a normal random vector \( Z_j \) with zero mean and identity covariance matrix. In (5), \( A \) is the transformation matrix and satisfies \( AK_{e_j}A^T = I \). We may determine the confidence space \( S_{Z_j} \) for the cover type \( C_j \), in which the probability of \( Z_j \) is \( p \). That is

\[
P(Z_j \in S_{Z_j}) = p
\]

(6)

From (5) and (6), we may get the confidence space \( S_{e_j} \) in which eigen index have \( p \) probability falling inside. Thus by examining the stored eigen index \( e_j \) using the confidence space \( S_{e_j} \), we may quickly determine the accuracy the stored cover type \( C_j \) and its corresponding eigen index \( e_j \).

IV. EXPERIMENT RESULTS AND DISCUSSIONS

In this section, simulation are presented to show the performance of the proposed segmentation technique and DB management mechanism. We have conducted the experiments on the SPOT satellite images, which have 512x512 size and 3 bands coded at 8 bit/band. The test images are divided into blocks with fixed block size 2x2 in the following simulations. From the segmentation results shown in Figs. 1, we observe that the proposed technique can partition the image into 2 regions, the gray and black regions which illustrate “sea” and “land” image features, respectively. In Figs.1, we also include the segmentation result from [4]. As we see the segmentation results of the two methods are similar. Regarding the computation time, the proposed technique takes 113 sec., whereas the technique in [4] requires 355 sec. by using a Pentium-4 2.66G-clock-rate personal computer. These experiment results validate the efficiency of the proposed segmentation technique.

Then, the segmentation results obtained by the proposed segmentation technique are applied to set up an archive of SPOT images. A Web-based interface agent using the eigen-index for content-based access to this test archive is available over the internet [http://spl.int.ntou. edu.tw]. In the test archive, three cover types: sea, city and vegetation, could be specified. Additional cover types could be added by the user. User specified cover types can be interactively defined via Web and added to the public inventory. The archive can be accessed by the proposed approach which searches the minimal separation between the eigen-index of the user specified cover type of query image and that of each object image in the database. To show the performance of the proposed DB retrieval method, we display in the Fig. 2 the query results of five best-match candidates for cover type: vegetation, using the query image “Taipei Basin”. As we see, the proposed approach can retrieve images with vegetation feature correctly.

Finally, we have implemented the proposed DB management mechanism for above SPOT image archive according to following cases: (1) the user makes wrong decision in the selection of cover type, (2) the stored entry is correct but has already in the inventory and (3) the added entry is both correct and new. This mechanism also includes a Web-based management interface shown in Fig. 3. User feedback of the proposed management mechanism has been very positive, particularly regarding the convenience of management. However, the performance evaluation, such as accuracy comparison with other methods, is still open for future study.

REFERENCES


Figs. 1 Segmented results of SPOT image “Harbor” (a) original image (b) segmented results by the proposed technique (c) segmented results by the eigen-region-based technique [4].

Fig. 2 The query results for cover type: vegetation.

Fig. 3 The management interface for the proposed DB management mechanism.