



## Article

# A Novel Semantic Content-Based Retrieval System for Hyperspectral Remote Sensing Imagery

Fatih Ömrüüzun <sup>1,\*</sup>, Yasemin Yardımcı Çetin <sup>1</sup>, Uğur Murat Leloğlu <sup>2</sup> and Begüm Demir <sup>3,4</sup>

- <sup>1</sup> Department of Information Systems, Graduate School of Informatics, Middle East Technical University, 06800 Ankara, Turkey; yyardim@metu.edu.tr
- <sup>2</sup> Faculty of Aeronautics and Astronautics, Turkish Aeronautical Association University, 06790 Ankara, Turkey; umleloglu@thk.edu.tr
- <sup>3</sup> BIFOLD—Berlin Institute for the Foundations of Learning and Data, Ernst-Reuter Platz 7, 10587 Berlin, Germany; demir@tu-berlin.de
- <sup>4</sup> Faculty of Electrical Engineering and Computer Science, Technische Universität Berlin, 10587 Berlin, Germany
- \* Correspondence: omruuzun@metu.edu.tr

**Abstract:** With the growing use of hyperspectral remote sensing payloads, there has been a significant increase in the number of hyperspectral remote sensing image archives, leading to a massive amount of collected data. This highlights the need for an efficient content-based hyperspectral image retrieval (CBHIR) system to manage and enable better use of hyperspectral remote-sensing image archives. Conventional CBHIR systems characterize each image by a set of endmembers and then perform image retrieval based on pairwise distance measures. Such an approach significantly increases the computational complexity of the retrieval, mainly when the diversity of materials is high. Those systems also have difficulties in retrieving images containing particular materials with extremely low abundance compared to other materials, which leads to describing image content with inappropriate and/or insufficient spectral features. In this article, a novel CBHIR system to define global hyperspectral image representations based on a semantic approach to differentiate foreground and background image content for different retrieval scenarios is introduced to address these issues. The experiments conducted on a new benchmark archive of multi-label hyperspectral images, which is first introduced in this study, validate the retrieval accuracy and effectiveness of the proposed system. Comparative performance analysis with the state-of-the-art CBHIR systems demonstrates that modeling hyperspectral image content with foreground and background vocabularies has a positive effect on retrieval performance.

**Keywords:** content-based hyperspectral image retrieval; hyperspectral imaging; remote sensing; semantic retrieval



**Citation:** Ömrüüzun, F.; Yardımcı Çetin, Y.; Leloğlu, U.M.; Demir, B. A Novel Semantic Content-Based Retrieval System for Hyperspectral Remote Sensing Imagery. *Remote Sens.* **2024**, *16*, 1462. <https://doi.org/10.3390/rs16081462>

Academic Editors: Jaime Zabalza, Jinchang Ren and Yijun Yan

Received: 15 March 2024

Revised: 9 April 2024

Accepted: 16 April 2024

Published: 20 April 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Hyperspectral images consist of many (hundreds in some cases) observation channels acquired at consecutive wavelengths. This virtue of hyperspectral imaging enables precise recognition and discrimination of matter in a scene. As such, hyperspectral imaging has become a prominent passive optical remote sensing technology utilized to solve various problems in diverse fields ranging from environmental monitoring to precision agriculture [1–3]. Consequently, a continuous increase in the deployment of hyperspectral imaging systems leads to a significant growth in the diversity and volume of hyperspectral remote sensing image collections. Furthermore, dense spectral information provided in hyperspectral imagery results in more data being processed than other optical imaging techniques [4]. Hence, the excessive amount of data emerging in imaging campaigns complicates the interpretation and management of the hyperspectral images. Accordingly, one of the critical tasks in remote sensing is the accurate and fast retrieval of hyperspectral

images from image collections in the context of spectral properties of the matter. Since spectral information provided in hyperspectral imagery leads to a very high capability for the identification and discrimination of the objects [5], content-based hyperspectral image retrieval (CBHIR) is the process of querying hyperspectral image collections in the context of matter through the dense information in the spectral domain.

Hyperspectral imaging is utilized in various fields to identify the composition of a scene through the exceptional spectral information provided. Thus, a proper CBHIR system should allow accurate access to desired hyperspectral imagery in an archive using a query that embodies/represents spectral features of similar content. In this context, accessing hyperspectral images with critical content requires fast and accurate retrieval in some applications. For instance, given the large expanse of land covered in a hyperspectral image archive, a precise CBHIR system can potentially enhance the effectiveness of hyperspectral imagery in various fields such as precision agriculture, forestry, mining, and defense. It can be beneficial in detecting and locating infected plants, specific types of trees, minerals, or targets that exhibit similar spectral characteristics in a given query image.

This study addresses content-based retrieval of hyperspectral imagery from different perspectives and proposes a promising semantic retrieval system, which is established on novel hyperspectral image descriptors that achieve both high accuracy and low computational complexity.

This article is organized as follows. In Section 2, a comprehensive literature review of CBHIR systems is presented. Section 3 explains the problem formulation and elaborates on the proposed system. Section 4 introduces the multi-label hyperspectral image archive used in the experiments. Section 5 elaborates on the experimental setup. In Section 6, comparative performance results are discussed. Finally, Section 7 concludes the study and criticizes the proposed CBHIR system.

## 2. Related Literature

Hyperspectral remote sensing imagery contains highly redundant information, and extracting proper features to model the image content sufficiently requires dedicated methods. CBHIR systems proposed in the literature, except [6], adopt endmember-based strategies to model hyperspectral images for two primary purposes: (1) to reveal spectral characteristics of the matter that constitute the scene and (2) to eliminate information redundancy in the hyperspectral imagery.

Spectral unmixing is a common and very crucial step for CBHIR systems available in the literature. It aims to find pure spectral signatures of the matter, so-called endmembers, in an image and decompose mixed pixel signatures, considering endmembers to calculate the abundances of those matter at a given pixel. Linear unmixing methods assume that mixed pixel signatures measured by hyperspectral imaging systems are composed of (i) a combination of pure material signatures (endmembers) in proportion to their abundances in a pixel and (ii) additive noise at each spectral band. On the other hand, since pure endmembers may not exist in a hyperspectral image due to insufficient spatial resolution of the imaging system or any other reason, specific linear unmixing methods utilize auxiliary endmember signature archives during the unmixing process.

CBHIR systems proposed in [7–9] model hyperspectral images with endmembers obtained via Pixel Purity Index (PPI), N-FINDR, and Automatic Pixel Purity Index (A-PPI) linear unmixing algorithms, respectively. In the retrieval phase, all three systems utilize a one-to-one endmember matching-based Spectral Signature Matching Algorithm (SSMA) to assess the similarity between the hyperspectral images. Unlike [7,8], the CBHIR system proposed in [9] employs the SSMA with Spectral Information Divergence-Spectral Angular Distance (SID-SAD)-based hybrid distance. In [10], an updated version the CBHIR system proposed in [7] is introduced that implements a distributed hyperspectral imaging repository on a cloud computing platform. In [11], an endmember matching-based distance for content-based hyperspectral image retrieval is proposed. This distance metric mutually maps each individual endmember that belongs to one image to an endmember of the other

image by considering SAD between them. Finally, the sum of the L-2 norm of vectors arising from minimum SAD between matched endmember pairs gives the Grana Distance between two hyperspectral images. The study evaluates the proposed hyperspectral image distance retrieval performance with the Endmember Induction Heuristic Algorithm (EIHA) and N-FINDR linear unmixing algorithms. In [12], the same research group introduces an alternative CBHIR system that utilizes both endmembers and their abundances. The proposed system assesses the similarity of two hyperspectral images by calculating the sum of SAD between each endmember pair arising from the Cartesian product of two endmember sets. In [6], yet another CBHIR approach is proposed that copes with spectral and spatial information redundancy in hyperspectral imagery with a data compression strategy. To this end, each hyperspectral image is converted to a text stream (either pixel-wise or band-wise) and then encoded with the Lempel–Ziv–Welch (LZW) algorithm to obtain a dictionary that models the image. In the retrieval phase, the level of similarity between two hyperspectral images is assessed by the dictionary distances that consider common and independent elements in the corresponding dictionaries. In [13], a hyperspectral image repository with retrieval functionality is introduced. The repository catalogs the hyperspectral images with endmembers obtained via either N-FINDR or Orthogonal Subspace Projection (OSP) linear unmixing algorithms in conjunction with their abundances. The user interacts with the system by choosing one or more spectral signatures from the library, already available in the repository, as a query. In the retrieval phase, the repository evaluates the level of similarity between query endmember(s) and cataloged image endmembers, considering the SAD. The CBHIR system proposed in [14] constructs a feature extraction strategy on sparse linear unmixing. This approach, which utilizes the SunSAL algorithm, aims to obtain image endmembers through spectral signatures already available in a library within the system. However, this CBHIR approach requires a large built-in library that accommodates spectral signatures of all possible materials for a proper feature extraction phase. In the retrieval phase, the proposed system evaluates the similarity of two images considering the SAD between image endmembers. In [15], hyperspectral images are characterized with two descriptors. The spectral descriptors corresponding to endmembers are obtained via N-FINDR algorithm. In addition, the proposed system uses Gabor filters to compute a texture descriptor to model the image. In the retrieval phase, the system considers the sum of spectral and texture descriptor distances to assess the similarity between two hyperspectral images. To this end, the distance between spectral and textural descriptors of two images is calculated by adopting the Significance Credit Assessment method introduced in [12] and squared Euclidean distance between Gabor filter vectors, respectively. Similar to [15], the CBHIR system proposed in [16] characterizes hyperspectral images with two descriptors: spatial and spectral. The spatial descriptor is computed with a saliency map that combines four features: the first component of the PCA, orientation, spectral angle, and visible spectral band opponent. On the other hand, the spectral descriptor corresponds to a histogram of spectral words obtained by clustering endmembers extracted from all the images in the archive. In the retrieval phase, the similarity between feature descriptors is calculated with squared Euclidean distance to assess the similarity between two images. In [17], a CBHIR system is proposed to secure hyperspectral imagery retrieval by encrypting the image descriptors. The system characterizes hyperspectral images with spectral and texture descriptors. To obtain the spectral descriptor, Scale-Invariant Feature Transform (SIFT) key-point descriptors of the RGB representation of the image and the endmembers extracted by the A-PPI linear unmixing algorithm are clustered with the k-means algorithm. This step defines spectral words that correspond to cluster centers. The proposed system employs the GLCM method to compute the texture descriptor to obtain contrast, correlation, energy, and entropy values. In the retrieval phase, these two descriptors are combined to model the images, and the Jaccard distance is used to assess the similarity between two images. Yet another CBHIR system that models the images with spectral and texture descriptors is introduced in [18]. The system obtains the spectral descriptors with endmembers extracted with the A-PPI unmixing algorithm. The system

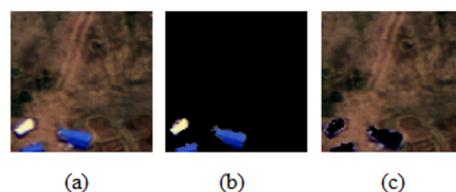
adopts the GLCM-based method introduced in [17] to obtain the texture descriptors. In the retrieval phase, the proposed system uses SID-SAM-based distance and Image Euclidean Distance to evaluate the similarity of spectral and texture descriptors, respectively. A bag-of-endmembers-based strategy for CBHIR is proposed in [19]. The proposed strategy aims to represent hyperspectral image content with a global spectral vocabulary obtained by clustering bag-of-endmembers from all endmembers extracted from the archive. In addition to the methods mentioned above, there is also a method that utilizes artificial neural networks. The method proposed in [20] suggests a model that provides pixel-based retrieval using a Deep Convolutional Generative Adversarial Network (DCGAN). For this purpose, an artificial neural network model is trained with a combination of spectral and spatial vectors obtained using manually selected pure material signatures from hyperspectral images and neighboring pixel signatures.

### 3. Proposed CHBIR System

Unlike the existing CBHIRs reviewed in Section 2, which dominantly measure the similarity between two hyperspectral images by employing endmember matching-based methods, the system proposed in this study addresses content-based hyperspectral image retrieval with a semantic approach. The proposed system assumes that a hyperspectral remote sensing image archive comprises two types of content: (i) foreground and (ii) background.

It is worth noting that, to avoid terminological confusion, two definitions are used within the scope of this article: hyperspectral remote sensing payload data product and hyperspectral image. The hyperspectral remote sensing data product represents hyperspectral data obtained by the payload on the air or space platform covering an area on the Earth, and the hyperspectral image represents the patches that form the benchmark archive by dividing the data product into manageable small pieces.

The claim being made in this article is that when modeling hyperspectral remote sensing images, it is important to consider the varying prevalence of different types of materials that make up the land cover in a territory covered by the data product. Specifically, certain types of materials are much more common than others. These include cultivated or uncultivated lands, terrestrial barren lands, and water bodies. In contrast, material classes such as artificial surfaces, urban areas, mining areas, and areas of materials with semantically remarkable spectral features are less prevalent (see Figure 1). Failing to consider the prevalence of these material classes when creating content-based models for hyperspectral remote sensing images can have significant consequences. For example, it can result in errors in accurately modeling certain content types that are relatively less common. This fact also makes it challenging to access related images due to the limitations of the models that are being used. Therefore, it is crucial to consider the prevalence of different material classes when modeling hyperspectral remote sensing images to ensure accurate and reliable results.



**Figure 1.** Pseudo-color representation of a remote sensing hyperspectral image  $X_{1323}$  (a), illustration of foreground (b) and background (c) image contents.

The proposed method is constructed on this semantic approach to overcome the following shortcomings of existing CBHIR methods in the literature.

1. Poor retrieval performance issues caused by spectral information redundancy due to the relatively high abundance of background content in the archive images.
2. CBHIR methods that model hyperspectral images by only endmembers may not accurately extract the endmembers from the images, or pure material signatures

may not exist in the scene. These issues may lead to describing image content with inappropriate and/or insufficient spectral features.

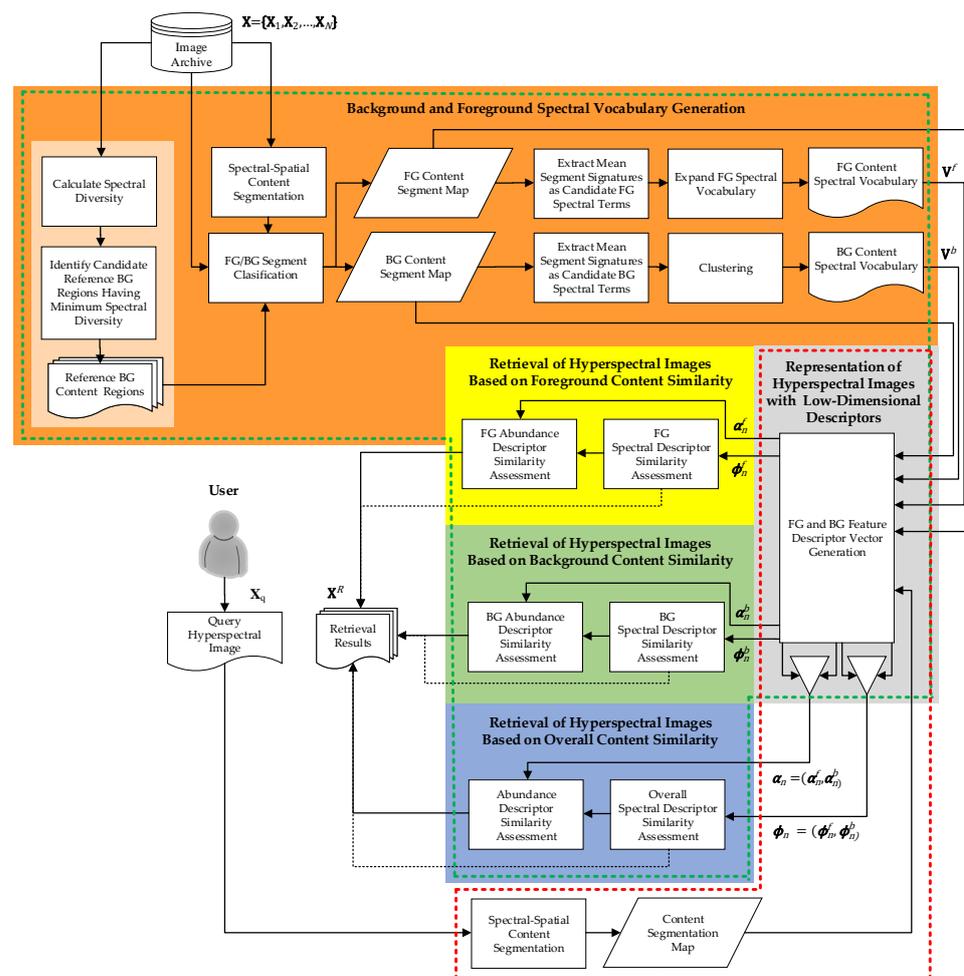
3. Strategies that combine and cluster all endmembers to generate a global spectral vocabulary to model hyperspectral images may ignore spectral signatures (endmember) of rarely seen content in the cases of using an inappropriate clustering method or setting parameters of clustering method inaccurately.

### 3.1. Problem Formulation and Notation

Let  $\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N\}$  be an archive of  $N$  hyperspectral images, where  $\mathbf{X}_n$  is the  $n$ -th image in the archive. The proposed CBHIR system aims at efficiently retrieving a set  $\mathbf{X}^R \subset \mathbf{X}$  of  $R$  hyperspectral images that contain similar content depicted by a query image  $\mathbf{X}_q$  provided by the user. (A list of all mathematical symbols used throughout the article is given in Appendix A.)

The proposed CBHIR system has two main modules: (1) an offline module to represent hyperspectral images with low-dimensional descriptors and (2) an online module to retrieve hyperspectral images using a computationally efficient hierarchical algorithm.

As illustrated in Figure 2, the proposed CBHIR system performs semantic feature extraction and representation of hyperspectral images with low-dimensional descriptors in the background offline. In contrast to existing CBHIR systems in the literature, the proposed CBHIR system allows for online retrieval of hyperspectral images through the low-dimensional descriptors obtained in this offline module. These novel feature representation and retrieval approaches are elaborated in the following subsections.



**Figure 2.** Block diagram of the proposed CBHIR system: green dashed lines represent offline processes that run in the background, while red dashed lines represent online processes.

### 3.2. Building Spectral Vocabularies

Spectral vocabulary generation and representing hyperspectral images with low-dimensional descriptors steps of the proposed CBHIR system aim at representing each hyperspectral image  $\mathbf{X}_n$  in  $\mathbf{X}$  with four low-dimensional descriptor vectors: two binary spectral descriptors  $\phi_n^f$  and  $\phi_n^b$  to represent the spectral characteristics of foreground and background content, respectively, and two abundance descriptors  $\alpha_n^f$  and  $\alpha_n^b$  to hold fractional abundance of corresponding content in the image  $\mathbf{X}_n$ . In addition to  $\phi_n^f$  and  $\phi_n^b$ , the proposed system uses descriptor  $\phi_n = (\phi_n^f, \phi_n^b)$  to represent spectral features of overall image content. Similarly, descriptor  $\alpha_n = (\alpha_n^f, \alpha_n^b)$  represents fractional abundance of corresponding content in the image  $\mathbf{X}_n$ . A new unsupervised spectral vocabulary generation method is introduced to calculate these descriptors.

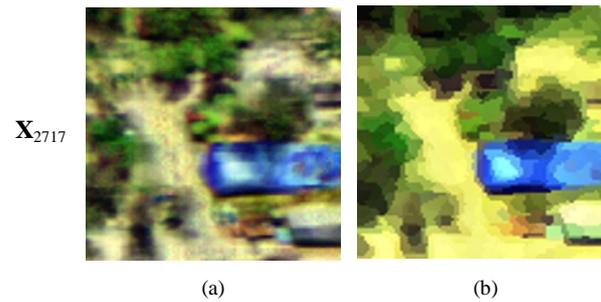
#### 3.2.1. Superpixel-Based Content Segmentation

The proposed CBHIR system benefits from spectral content vocabularies to retrieve hyperspectral images from the archive effectively in an online manner. Accordingly, discovering material diversity in the archive to generate the foreground and background content vocabularies is a crucial step for the proposed CBHIR system. To this end, a superpixel-based segmentation is performed on each hyperspectral image  $\mathbf{X}_n$  in  $\mathbf{X}$  to group image pixels with similar spectral features and spatial relations that belong to a phenomenon in the scene. However, an effective method is required to perform such a segmentation that can handle high-dimensional spectral information with low computational complexity.

To overcome this, the proposed CBHIR system benefits from a novel superpixel-based segmentation algorithm dedicated to hyperspectral imagery [21], which is a derivative of the Simple Linear Iterative Clustering (SLIC) method [22]. This superpixel-based segmentation algorithm, namely hyperSLIC in this study, is designed to cluster pixels in local regions rather than globally, which means that spatial correlation and spectral similarity are naturally considered during the segmentation process. There are three main reasons for using the hyperSLIC method in the proposed system. The first is the combined use of spectral and spatial relationships in segmentation. The second reason is the low complexity of segmentation performed in local regions. The third reason is the adaptability of the local neighborhood parameter to the spatial resolution of remote sensing images. Details of the hyperSLIC algorithm are given below.

The hyperSLIC algorithm begins by assigning a pre-defined number of superpixel centers at equal distances. To streamline the clustering search process, hyperSLIC sets a defined local neighborhood around each cluster center. This neighborhood takes the shape of a rectangular region with a width of  $w$  and a height of  $h$ . Limiting the search to only the surrounding  $w \times h$  pixels for each cluster center significantly reduces the computational complexity compared to traditional clustering algorithms. During the main loop step, the algorithm employs the SID-SAM and Euclidean spectral and spatial distance criteria, respectively, to cluster each pixel in the local neighborhood for every cluster center. Following each iteration of the clustering algorithm, the cluster centers are updated to enhance the accuracy of subsequent iterations.

The sample image presented in Figure 3 was selected from the dataset described in Section 4. This image has undergone a segmentation process using the hyperSLIC algorithm. The minimum segment size for this process was set to  $4 \times 4$  pixels, meaning that the image was divided into smaller segments, with each segment being at least  $4 \times 4$  pixels. This step helps to identify the content segments within the image, which will be further analyzed in the feature extraction process.



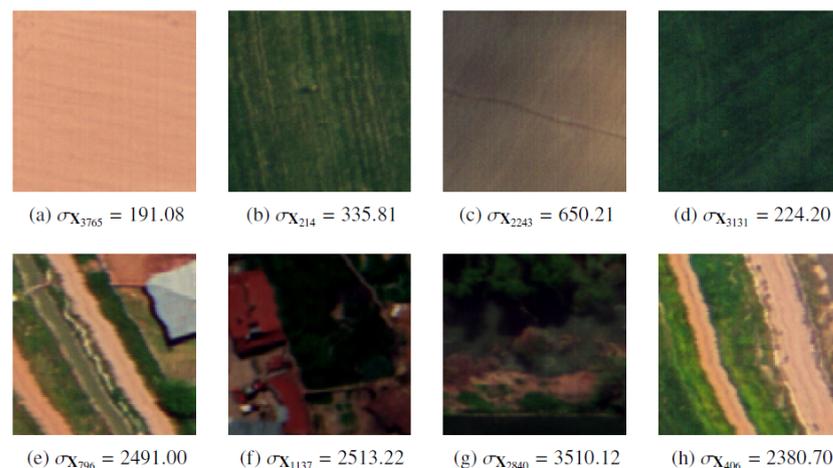
**Figure 3.** Sample superpixel-based content segmentation with hyperSLIC. (a) False-color original image. (b) False-color segmented image.

### 3.2.2. Background Suppression

Segmentation of hyperspectral imagery with a proper algorithm (i.e., hyperSLIC) results in identifying semantically (both spectral and spatial) related content pixels. This is a helpful step in dealing with highly redundant spectral information in hyperspectral imagery. On the other hand, the relatively high proportion of background content in the discovered segments poses a problem for efficient and quick retrieval of desired content. To overcome this problem, the proposed CBHIR system introduces a novel background suppression-based method to make foreground content more easily identifiable. This method examines each content segment in the images concerning spectral features of the territorial background content and identifies each segment's dissimilarity to spectral features of the territorial background regions.

#### Discovering Spectral Diversity of Candidate Territorial Background Content

The proposed CBHIR system benefits from two spectral diversity to identify territorial background regions in the data products to use these regions in the background suppression process. Hyperspectral images with relatively smaller intra-spectral diversity are more capable of representing background and can be used as reference background imagery for a territory as depicted in Figure 4. In the first step of the background suppression algorithm, spectral diversity  $\sigma_{X_i}$  for each individual hyperspectral image that has been created from the same hyperspectral remote sensing data product, which covers a specific region on the Earth, is calculated. The reason for adopting a regional approach in determining background contents is that hyperspectral images, which are spatially close to each other, tend to have similar hyperspectral background contents.



**Figure 4.** Sample hyperspectral images with low and high spectral diversity.

$$\sigma_{\mathbf{X}_n} = \frac{1}{P^2} \sum_{i=1}^P \sum_{j=1}^P \cos^{-1} \left( \frac{\mathbf{x}_n^i \cdot \mathbf{x}_n^j}{\|\mathbf{x}_n^i\| \|\mathbf{x}_n^j\|} \right) \quad (1)$$

where  $P$  is the total number of pixels in image  $\mathbf{x}_n$ .  $\mathbf{x}_n^i$  and  $\mathbf{x}_n^j$  represent  $i$ -th and  $j$ -th pixels of  $\mathbf{x}_n$ . Equation (1) was inspired by Spectral Angular Mapper (SAM) [23], and the non-linearity of the equation in calculating the dissimilarity of two spectral signatures enables better discrimination of low and high spectral diversity in image content.

#### Discovering Spectral Diversity of Candidate Territorial Background Content

After calculating intra-spectral diversity for each image created from the same data product covering a specific region on the Earth, a specific number of hyperspectral images are identified as reference background images in this step. To this end, hyperspectral image  $\mathbf{X}_n$  in the archive with minimum intra-spectral diversity is identified as the first reference background image. Later on, the next hyperspectral image with minimum intra-spectral diversity is chosen as a candidate reference background image. A hyperspectral image is labeled as a reference background image if the spectral dissimilarity between the mean spectral signature of this image and the previously identified background images is bigger than a threshold defined by the user. This process is terminated if the desired number of hyperspectral images are identified as reference background images. In this way, the proposed system scans through the images created from the same hyperspectral remote sensing data product and prevents identifying similar reference background images to model the background content better. Figure 5 demonstrates hyperspectral images identified as reference background images for each hyperspectral remote sensing data product introduced in Section 4.

#### Identifying Foreground–Background Content Segments

As illustrated in the block diagram of the proposed CBHIR system (please see Figure 2), foreground and background content in a hyperspectral image are discriminated based on a background suppression-based approach. Thus, this method requires a reliable method to distinguish foreground and background contents using the reference hyperspectral images with materials representing the regional spectral features of the background for that specific territory.

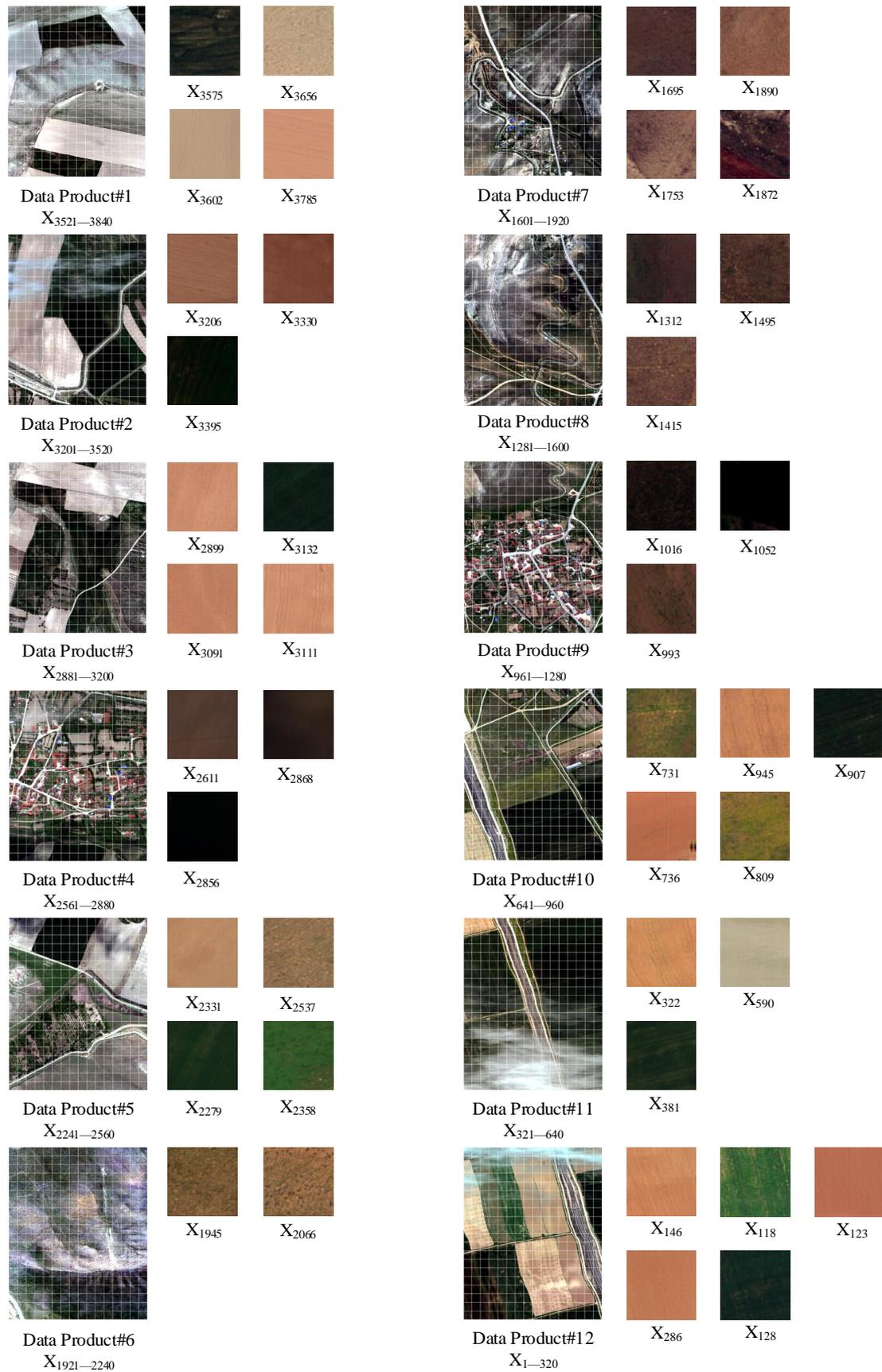
Mahalanobis distance is a measure used to quantify the dissimilarity between a sample and a distribution. It considers the correlations between variables, making it particularly useful when dealing with multivariate data such as hyperspectral imagery.

The proposed method calculates how closely a content segment resembles the spectral characteristics of reference background image contents using the Mahalanobis distance-based scoring approach. In other words, to determine whether a content segment belongs to the foreground or background class, the spectral signature of the segment is compared against a set of pre-defined reference background images. If the spectral features of the segment/pixel noticeably deviate from the spectral features of all the background images, it is classified as foreground content.

The Mahalanobis distance between a segment spectral signature and a distribution is defined as follows:

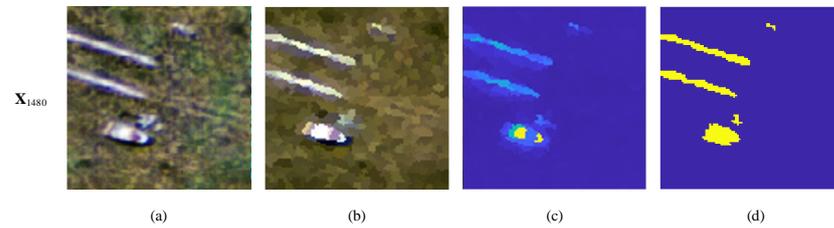
$$\delta(\mathbf{x}_n^s) = (\mathbf{x}_n^s - \boldsymbol{\mu}_B)^T \boldsymbol{\Gamma}_B^{-1} (\mathbf{x}_n^s - \boldsymbol{\mu}_B) \quad (2)$$

where  $\mathbf{x}_n^s$ ,  $\boldsymbol{\mu}_B$ , and  $\boldsymbol{\Gamma}_B^{-1}$  represent mean spectral signature vector of  $s$ -th content segment in image  $\mathbf{X}_n$ , sample mean, and sample covariance matrix of territorial background image  $\mathbf{B}$  that is a combination of reference background images identified for that specific geographical region, respectively.



**Figure 5.** Background content regions designated by the proposed CBHIR system for hyperspectral remote sensing payload products.

As a result, the similarities of content segments to the background within archive images can be measured unsupervised, as depicted in Figure 6.



**Figure 6.** Foreground–background content segment classification. (a) False-color original image, (b) segmented image, (c) Mahalanobis score map, (d) foreground–background segment classification map.

### Identifying Spectral Terms

Two distinct methods are used to create foreground and background content vocabularies to enhance the semantic significance of emphasized foreground contents in the study and minimize the redundant spectral information related to the background content. The foreground content vocabulary includes the spectral signatures of previously identified foreground content segments as is, while a clustering-based approach is used to create the background content vocabulary. This approach helps reduce the density of repeated background content information. By differentiating between foreground and background contents in the images within the archive, dedicated vocabularies related to each content type can be generated.

Creating the background content vocabulary through the clustering process is a meticulous procedure. Research conducted in the context of the article has revealed the advantages of utilizing the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [24] clustering method over other methods, including k-means and kernel k-means. DBSCAN offers the advantages of automatically detecting clusters of arbitrary shapes while being robust to noise (rarely seen material signatures), requiring minimal parameter tuning, and not being sensitive to initialization.

### 3.3. Representing Hyperspectral Images with Low-Dimensional Descriptors

The proposed CBHIR system represents each hyperspectral image  $X_n$  in  $X$  by four low-dimensional descriptor vectors: two binary partial spectral descriptors  $\phi_n^f$  and  $\phi_n^b$  to represent the spectral characteristics of foreground and background content, respectively, and two partial abundance descriptors  $\alpha_n^f$  and  $\alpha_n^b$  to hold fractional abundance of corresponding content in the image  $X_n$  as depicted in Figure 7. In addition to  $\phi_n^f$  and  $\phi_n^b$ , the proposed system uses the overall descriptor  $\phi_n = (\phi_n^f, \phi_n^b)$  to represent spectral features of overall image content. Similarly, descriptor  $\alpha_n = (\alpha_n^f, \alpha_n^b)$  represents fractional abundance of corresponding content in the image  $X_n$  as depicted in Figure 8.

	V <sub>1</sub>	V <sub>2</sub>	V <sub>3</sub>	V <sub>4</sub>	V <sub>5</sub>	V <sub>6</sub>	V <sub>7</sub>	V <sub>8</sub>
$\phi_n^f$	1	0	0	1	1	0	1	1
$\alpha_n^f$	0.03	0	0	0.08	0.11	0	0.06	0.01
$\phi_n^b$	1	0	1	1	0	1	0	1
$\alpha_n^b$	0.12	0	0.19	0.02	0	0.43	0	0.24

**Figure 7.** Illustration of low-dimensional foreground and background content descriptors (where  $|\mathbf{V}^f| = 8$  and  $|\mathbf{V}^b| = 8$ ).

At this point, it is essential to underline that while the first part of the low-dimensional descriptors describes the image content that defines materials having a significant difference compared to the background in terms of spectral characteristics (e.g., artificial materials, anomalies), the second part defines the background content commonly seen in archive images.

	V <sub>1</sub>	V <sub>2</sub>	V <sub>3</sub>	V <sub>4</sub>	V <sub>5</sub>	V <sub>6</sub>	V <sub>7</sub>	V <sub>8</sub>	V <sub>9</sub>	V <sub>10</sub>	V <sub>11</sub>	V <sub>12</sub>	V <sub>13</sub>	V <sub>14</sub>	V <sub>15</sub>	V <sub>16</sub>
$\phi_n$	1	0	0	1	1	0	1	1	1	0	1	1	0	1	0	1
$\alpha_n$	0.03	0	0	0.08	0.11	0	0.06	0.01	0.12	0	0.19	0.02	0	0.43	0	0.24

**Figure 8.** Illustration of low-dimensional overall content descriptors (where  $|\mathbf{V}^f| = 8$  and  $|\mathbf{V}^b| = 8$ ).

To compute foreground spectral image descriptors, initially a spectral distance matrix  $\mathbf{D}_{\phi_n^f, \mathbf{V}^f} = [d_{s,\psi}; s = 1, \dots, S; \psi = 1, \dots, \Psi]$  is constructed, where  $d_{s,\psi}$  denotes a spectral distance estimated between  $s$ -th foreground segment mean signature extracted from the image  $\mathbf{X}_n$  and  $\psi$ -th spectral term in foreground content vocabulary  $\mathbf{V}^f$ . Any distance measure can estimate this, whereas in this study, the well-known spectral angular distance is considered. Then, the distance matrix  $\mathbf{D}_{\phi_n^f, \mathbf{V}^f}$  is quantized by setting the minimum element of each row to 1 and the remaining elements to 0. In this way, each image foreground segment mean signature is associated with a spectral term in the vocabulary  $\mathbf{V}^f$  considering the degree of spectral similarity. Then,  $\mathbf{D}_{\phi_n^f, \mathbf{V}^f}$  is compressed into a fixed-size binary descriptor to obtain  $\phi_n^f$  by applying the Boolean OR operator along each column.

Similarly, a distance matrix  $\mathbf{D}_{\phi_n^b, \mathbf{V}^b} = [d_{s,\omega}; s = 1, \dots, S; \omega = 1, \dots, \Omega]$  is constructed, where  $d_{s,\omega}$  denotes a spectral distance estimated between  $s$ -th background segment's mean signature extracted from the image  $\mathbf{X}_n$  and  $\omega$ -th spectral term in background content vocabulary  $\mathbf{V}^b$ . Then, the distance matrix  $\mathbf{D}_{\phi_n^b, \mathbf{V}^b}$  is quantized in the same way to obtain fixed-size descriptor  $\phi_n^b$ .

Accordingly,  $\phi_n^f = [\phi_n^{f_1}, \dots, \phi_n^{f_\Psi}]$  and  $\phi_n^b = [\phi_n^{b_1}, \dots, \phi_n^{b_\Omega}]$  are defined as  $\Psi$  and  $\Omega$  dimensional binary spectral descriptors, where each element of the vector (i.e., descriptor) indicates existence of a unique material in hyperspectral image represented by the  $\psi$ -th and  $\omega$ -th spectral term in the spectral vocabularies  $\mathbf{V}^f$  and  $\mathbf{V}^b$ , respectively. Obtained binary spectral descriptors have two main advantages: (1) they enable real-time search and accurate retrieval, and (2) they reduce the memory required for storing hyperspectral image descriptors in the archives.

To calculate the foreground abundance descriptor  $\alpha_n^f$  for  $\mathbf{X}_n$ , normalized fractional abundance of each foreground spectral term in  $\mathbf{V}^f$  is computed as given in Equation (3).

$$\alpha_n^{V_\psi^f} = \frac{c_n^{V_\psi^f}}{P} \quad (3)$$

where  $c_n^{V_\psi^f}$  and  $P$  correspond to the cumulative number of pixels in the segments labeled as  $\psi$ -th spectral term in foreground content vocabulary  $\mathbf{V}^f$  and the total number of pixels in  $\mathbf{X}_n$ , respectively. Similarly, for the background abundance descriptor  $\alpha_n^b$  for  $\mathbf{X}_n$ , normalized fractional abundance of each foreground spectral term in  $\mathbf{V}^b$  is computed as given in Equation (4).

$$\alpha_n^{V_\omega^b} = \frac{c_n^{V_\omega^b}}{P} \quad (4)$$

In addition to  $\phi_n^f$  and  $\phi_n^b$ , the proposed system uses descriptor  $\phi_n = (\phi_n^f, \phi_n^b)$  (or  $\phi_q$ ) to represent spectral features of overall image content. Similarly, descriptor  $\alpha_n = (\alpha_n^f, \alpha_n^b)$  (or  $\alpha_q$ ) represents fractional abundance of corresponding content in the image  $\mathbf{X}_n$  as depicted in Figure 8.

### 3.4. Retrieving Hyperspectral Images with Low-Dimensional Feature Descriptors

The proposed novel CBHIR system allows users to perform hyperspectral retrieval with a hierarchical algorithm. Furthermore, the proposed hierarchical algorithm significantly reduces the image retrieval time since (1) it filters out a high number of irrelevant images (with respect to the spectral characteristics of distinct materials present in the query image) at the first step by considering simple bitwise operations on low-dimensional spectral descriptors, and (2) in the second step the reduced set  $\mathbf{X}^H$  of images is queried only to retrieve the set  $\mathbf{X}^R \subset \mathbf{X}^H$  of images with the highest similarities in terms of spectral characteristics of distinct materials and their fractional abundances present in the query image. It is worth noting that due to the considered two-step strategy, the proposed algorithm can be performed by either considering or neglecting the evaluation of the similarities among the abundances of materials. Accordingly, the proposed strategy meets the diverse needs of different types of CBHIR applications.

#### 3.4.1. Retrieving Hyperspectral Images Based on Overall Content Similarity

In this retrieval scenario, the user benefits from the proposed system to retrieve hyperspectral images concerning overall content similarity by utilizing spectral and abundance descriptors of foreground and background contents. To this end, concatenated spectral and abundance descriptors calculated for foreground and background content of each hyperspectral image  $\mathbf{X}_n$  in the archive and spectral and abundance descriptors calculated query image  $\mathbf{X}_q$  are employed to perform multiple material-based retrieval.

In the first step, the similarity between  $\mathbf{X}_q$  and  $\mathbf{X}_n$  is computed concerning the binary spectral descriptors  $\boldsymbol{\phi}_n = (\boldsymbol{\phi}_n^f, \boldsymbol{\phi}_n^b)$  and  $\boldsymbol{\phi}_q$  by estimating the Hamming distance between them. Then, a set  $\mathbf{X}_H$  of  $H \leq R$  images having the lowest Hamming distances are selected, while the remaining images in the archive are filtered out. In the case of considering only spectral descriptor-based similarity between hyperspectral images for retrieval,  $\mathbf{X}^H$  is considered as the final set of retrieved images (i.e.,  $\mathbf{X}^H = \mathbf{X}^R$ ) and the algorithm stops at this step. If the abundance of materials is also considered for retrieval,  $\mathbf{X}^H$  is forwarded to the second step. In the second step, the similarity between abundance descriptor  $\boldsymbol{\alpha}_n = (\boldsymbol{\alpha}_n^f, \boldsymbol{\alpha}_n^b)$  of each image in  $\mathbf{X}^H$  and  $\boldsymbol{\alpha}_q$  of the query image is estimated by considering the Euclidean distance measure. Then, the set  $\mathbf{X}^R \subset \mathbf{X}^H$  of  $R$  images that have the highest similarity to the query image  $\mathbf{X}_q$  in terms of the fractional abundance of materials defined in abundance descriptors are chosen.

#### 3.4.2. Retrieving Hyperspectral Images Based on Foreground Content Similarity

Since the proposed system independently models foreground and background content, in this scenario, the user can configure the retrieval process by forcing the system to focus only on foreground content. To this end, spectral and abundance descriptors calculated for foreground content of each hyperspectral image  $\mathbf{X}_n$  in the archive and overall spectral and abundance descriptors calculated query image  $\mathbf{X}_q$  are employed to perform multiple material-based retrieval. It is critical to note that, in this retrieval scenario, overall spectral and abundance descriptors of  $\mathbf{X}_n$  are modified such that portions of the descriptors related to background content are discarded (set to zero) to perform the retrieval by focusing on foreground content only.

In the first step, the similarity between  $\mathbf{X}_q$  and  $\mathbf{X}_n$  is computed concerning the binary spectral descriptors, *modified*  $\boldsymbol{\phi}_n$  and  $\boldsymbol{\phi}_q$ , only by estimating the Hamming distance between them. Then, a set  $\mathbf{X}_H$  of  $H \leq R$  images having the lowest Hamming distances are selected, while the remaining images in the archive are filtered out. In the case of considering only spectral descriptor-based similarity between hyperspectral images for retrieval,  $\mathbf{X}^H$  is considered as the final set of retrieved images (i.e.,  $\mathbf{X}^H = \mathbf{X}^R$ ) and the algorithm stops at this step. If the abundance of materials is also considered for retrieval,  $\mathbf{X}^H$  is forwarded to the second step. In the second step, the similarity between *modified* abundance descriptor  $\boldsymbol{\alpha}_n$  of each image in  $\mathbf{X}^H$  and  $\boldsymbol{\alpha}_q$  of the query image is estimated by considering the Euclidean distance measure. Then, the set  $\mathbf{X}^R \subset \mathbf{X}^H$  of  $R$  images that have the highest similarity to

the query image  $X_q$  in terms of the fractional abundance of materials defined in abundance descriptors are chosen.

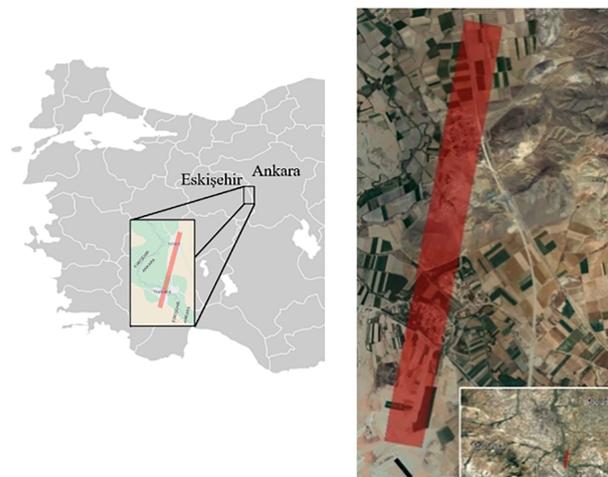
### 3.4.3. Retrieving Hyperspectral Images Based on Background Content Similarity

Similar to retrieving hyperspectral images concerning foreground content similarity, the proposed CBHIR system allows the user to query hyperspectral images by only considering the background content similarity. In contrast to foreground content-based retrieval, in this retrieval scenario, overall spectral and abundance descriptors of  $X_n$  are modified such that portions of the descriptors related to foreground content are discarded (set to zero) to perform the retrieval by focusing on background content only.

## 4. Dataset Description

### 4.1. Data Source

To evaluate the retrieval performance of the proposed CBHIR system and compare it with the state-of-the-art systems available in the literature, a multi-label benchmark hyperspectral image archive was created from very high-resolution hyperspectral imagery. The hyperspectral images used during archive generation were acquired over a flight line covering Yenice and Yeşilkaya towns (which are located on the border of the cities Eskişehir and Ankara in Turkey) by VNIR hyperspectral imager of a multimodal imaging system (see Figure 9).



**Figure 9.** Fingerprint of the area imaged during flight and used in benchmark archive generation.

The sensor components of the multimodal imaging system are composed of two co-aligned very high-resolution hyperspectral (VNIR + SWIR) imagers, one RGB multispectral imager, and one Fiber Optic Downwelling Irradiance Sensor (FODIS) to simultaneously measure the power of incident light during flight for atmospheric correction of VNIR hyperspectral images. The data acquisition flight was performed with a Cessna 206 aircraft on 4 May 2019. Details of flight parameters and corresponding ground resolution obtained with each sensor are given in Table 1.

**Table 1.** Flight parameters and corresponding ground resolutions obtained with the sensors.

<b>Aircraft</b>	Cessna 206		
<b>Flight Altitude (m)</b>	~3000 (AGL)/~3815 (ASL)		
<b>Flight Speed (knots)</b>	~90		
<b>Flight Polygon Size (m)</b>	8000 × 790		
	<b>VNIR HS</b>	<b>SWIR HS</b>	<b>RGB MS</b>
<b>FOV (m)</b>	445.17	276.48	1696.00 × 1356.80
<b>GSD (cm)</b>	27.86	72	1.32 × 1.32

#### 4.2. Data Pre-Processing

A set of pre-processing tasks was performed on the raw data to generate a coherent benchmark archive from large consecutive images acquired during the mission and prepare the patches for the labeling phase. The data pre-processing step consists of the following tasks: (1) digital number (raw image) to radiance conversion, (2) radiance to reflectance conversion, and (3) slicing images to obtain patches to be labeled. The first and second tasks were performed using commercial Headwall SpectralView (v3.2.0) software. In the last step of data pre-processing, twelve reflectance hyperspectral images with  $2000 \times 1600$  pixels were equally sliced into  $100 \times 100$  pixel square patches. By the end of this step, 3840 patches, each of which approximately covered  $7.8 \text{ km}^2$  on the ground, were obtained.

#### 4.3. Data Labeling

Accurate labeling of samples in any benchmark archive is a crucial task that explicitly affects performance analysis. Thus, the labeling of patches in the benchmark archive was performed through auxiliary VHR multispectral imagery, which provides a  $1.32 \text{ cm}$  ground sampling distance, which was acquired during the same flight. (see Figure 10).



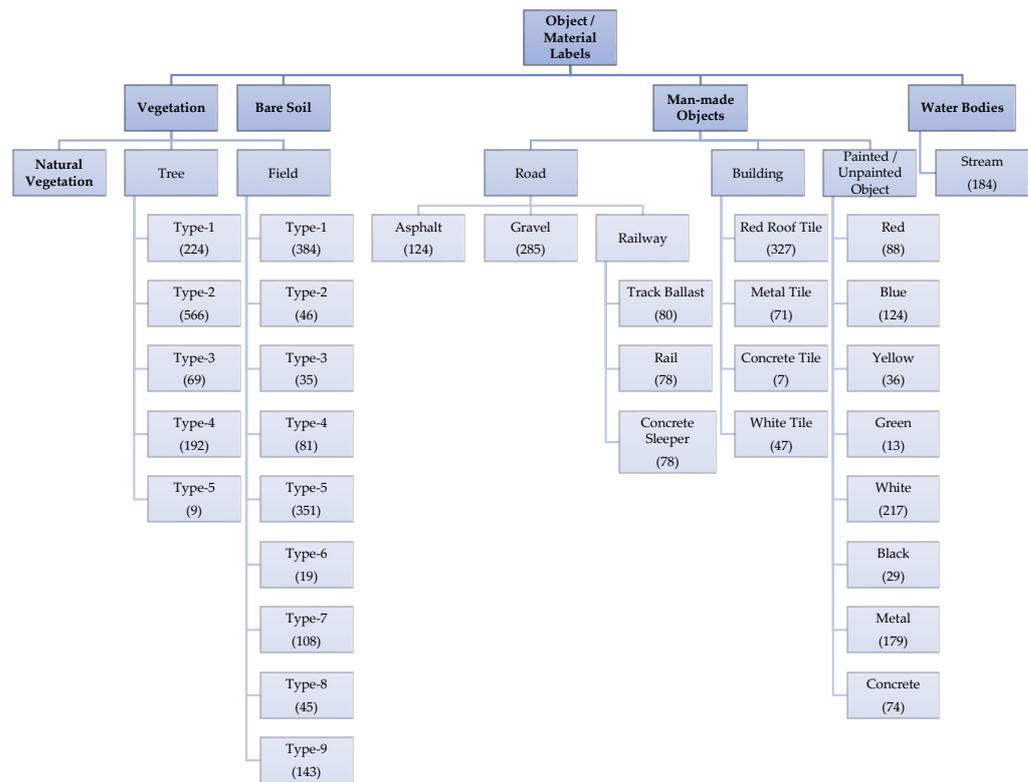
**Figure 10.** Utilizing VHR imagery for identifying hyperspectral image labels precisely; (a) a data product sliced to obtain hyperspectral images for benchmark archive, (b) corresponding VHR multispectral image section acquired during the same flight.

In addition to labeling each sample in the benchmark archive with VHR multispectral imagery, fieldwork was also performed on 31 October 2021 along the flight path to enhance the quality of labeling. In this fieldwork, objects in the hyperspectral image archive were photographed from the ground to obtain more information about them (see Figure 11).



**Figure 11.** Fieldwork to enhance the accuracy of the content labeling phase. The blue circle indicates the location where the ground-truth picture was taken.

Taxonomy of the benchmark hyperspectral image archive is presented in Figure 12.



**Figure 12.** Taxonomy of content labels and the corresponding number of images labeled under each individual sub-category.

## 5. Experimental Setup

This section of the article elaborates on the experimental setup designed for performing objective and comparative performance analysis between the proposed and other CBHIR systems available in the literature.

A series of experiments was conducted to assess the proposed CBHIR system's performance compared to other CBHIR systems in the literature. To this end, it is necessary to set specific variables and methods to perform the experiments presented in this section, including the proposed CBHIR system and other CBHIR systems from the literature. These parameters are essential for obtaining accurate experimental results. Therefore, preliminary experiments were conducted to determine the best values for these parameters. This section gives a detailed explanation of the values determined as a result of these preliminary experiments. First, the experimental setup of the CBHIR system proposed in this study and other studies in the literature are given.

Within the scope of the study, spectral–spatial segmentation is performed on hyperspectral images using the proposed system. In this segmentation step with the hyperSLIC algorithm, the local neighborhood parameter is set to  $4 \times 4$  pixels corresponding to an area of  $\sim 1 \text{ m}^2$  on the ground. Such an area is clear enough to observe spectral features of matter in the scene for the spatial resolution of the imager at the given flight altitude in Table 1.

Another parameter the proposed method requires is the maximum number of reference background images to be determined for each hyperspectral remote sensing data product. When examining the hyperspectral remote sensing data products that comprise the archive, this number was determined to be five. When determining reference background images, it has been observed that selecting the average spectral angular distance between images as 0.25 radians is suitable for different background image sets. The proposed CBHIR system uses Mahalanobis distance to regional reference background images to classify foreground and background content segments. At this stage, the threshold value is the

highest Mahalanobis distance to the regional background image pixels created by merging reference background images.

During the vocabulary creation stage, the spectral angular distance for foreground content dictionaries is set to 0.10 radians to eliminate the existence of repetition for the same material signature. To evaluate the performance of the proposed system, three state-of-the-art methods for comparison were considered: (1) the bag-of-endmember-based method (denoted as BoE), (2) the endmember matching algorithm based on the Grana Distance (denoted as EM-Grana), and (3) the endmember matching algorithm that weights the distances estimated by the SAD between each endmember pair by their abundances (denotes as EM-WSAD). Vertex Components Analysis (VCA) was used in the experiments for endmember-based methods to obtain the endmembers. HySime [25] was used in the experiments to estimate the number of endmembers.

In all experiments, CBHIR systems are requested to retrieve the 10 most similar images to a given query image, and each hyperspectral image in the benchmark archive is used as a query image. Beyond each system's retrieval performance, the retrieval time is also measured.

### 5.1. Computational Environment

The experiments were conducted in MATLAB R2023b environment installed on a Microsoft Windows 10 operating system computer with 3.6 GHz Intel® i7-9750H processor 2.6 GHz and 32 GB RAM.

### 5.2. Performance Metrics

Since this study performs performance evaluation on a multi-label benchmark archive, four compatible multi-label performance metrics were used: (i) accuracy, (ii) precision, (iii) recall, and (iv) Hamming Loss. Let  $\mathbf{L}_{\mathbf{X}_q}$  and  $\mathbf{L}_{\mathbf{X}_r}$  be the label sets for the query image  $\mathbf{X}_q$  and any particular image  $\mathbf{X}_r$  in the corresponding set of retrieved images  $\mathbf{X}^R$ , respectively.

Accuracy is the fraction of identical content labels of the query and retrieved images in the union of label sets of two images and is defined as:

$$Accuracy = \frac{|\mathbf{L}_{\mathbf{X}_q} \cap \mathbf{L}_{\mathbf{X}_r}|}{|\mathbf{L}_{\mathbf{X}_q} \cup \mathbf{L}_{\mathbf{X}_r}|} \quad (5)$$

Thus, accuracy is directly proportional to the cardinality of the intersection of label sets of query and retrieved images. The retrieval performance increases when accuracy approaches 1. Precision is the fraction of identical content labels of query and retrieved images in the content label set of the retrieved image and is defined as:

$$Precision = \frac{|\mathbf{L}_{\mathbf{X}_q} \cap \mathbf{L}_{\mathbf{X}_r}|}{|\mathbf{L}_{\mathbf{X}_r}|} \quad (6)$$

In comparison with accuracy, precision evaluates the retrieval performance of the system by mainly focusing on the content labels of the retrieved image. Accordingly, the content labels of the query image apart from the matched ones are ignored. The retrieval performance increases when precision approaches 1. Unlike precision, recall is the fraction of identical content labels of query and retrieved images in the content labels of the query image and is defined as:

$$Recall = \frac{|\mathbf{L}_{\mathbf{X}_q} \cap \mathbf{L}_{\mathbf{X}_r}|}{|\mathbf{L}_{\mathbf{X}_q}|} \quad (7)$$

Thus, the content labels of the retrieved image, apart from those of the matched ones, are ignored. The retrieval performance increases when precision approaches 1.

Hamming Loss evaluates the retrieval performance by calculating the symmetric difference ( $\Delta$ ) between two content label sets and is defined as:

$$\text{Hamming Loss} = \frac{|\mathbf{L}_{X_q} \Delta \mathbf{L}_{X_r}|}{|\mathbf{L}_{X_q}|} \quad (8)$$

According to Hamming Loss, the system is penalized for each item that is not in the intersection of query and retrieved image content label sets. The retrieval performance increases when Hamming Loss approaches zero.

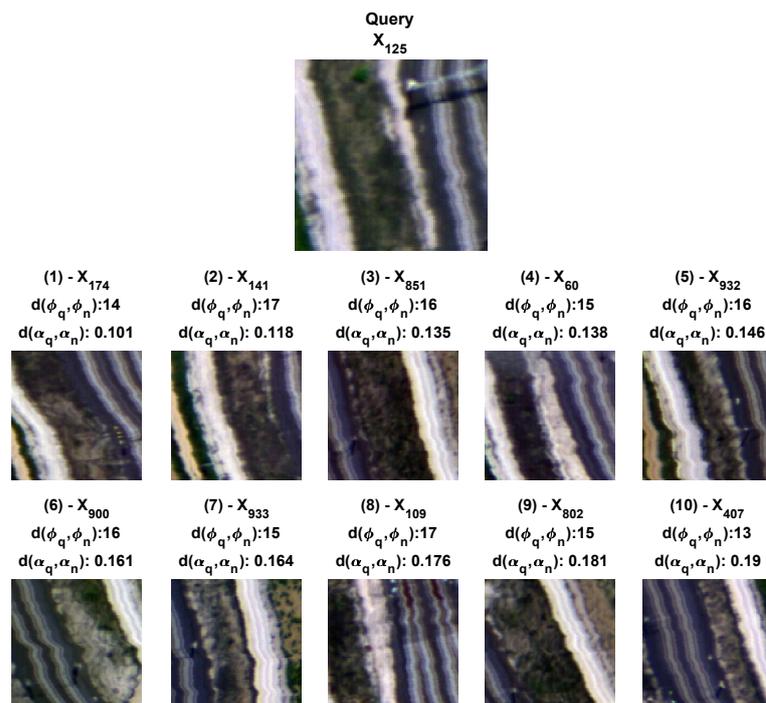
## 6. Experimental Results

In this section, the retrieval performance of the proposed CBHIR system is compared with state-of-the-art systems available in the literature detailed in Section 2.

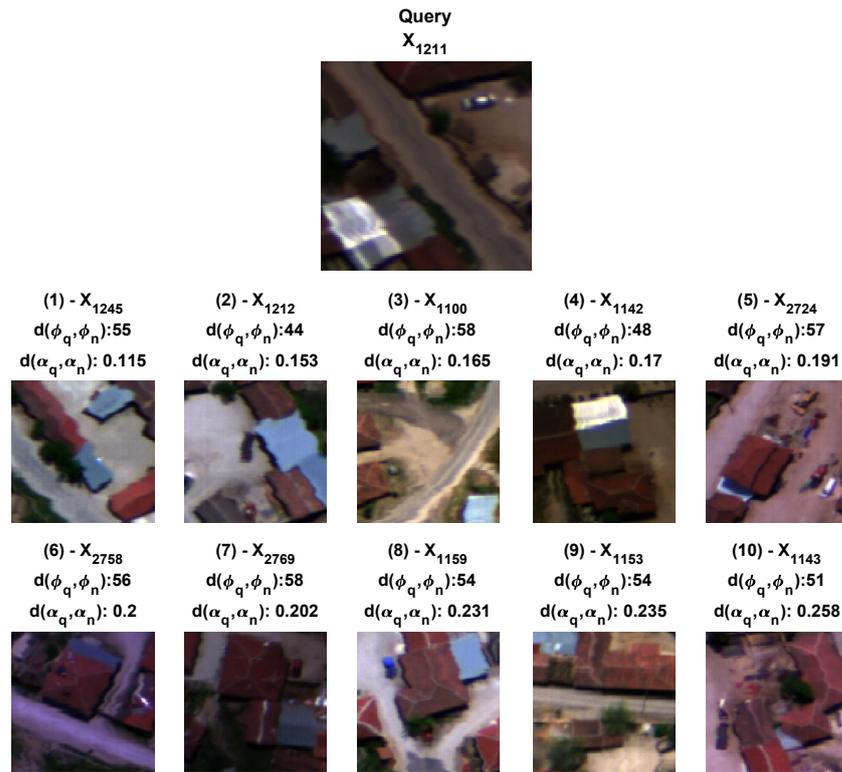
### 6.1. Sample Retrieval Results for the Proposed CBHIR System

In this subsection, the retrieval performance of the proposed CBHIR system within the scope of the article is demonstrated with visual examples using different query images. For this purpose, query hyperspectral images are selected from different regions of the hyperspectral image archive used in the study.

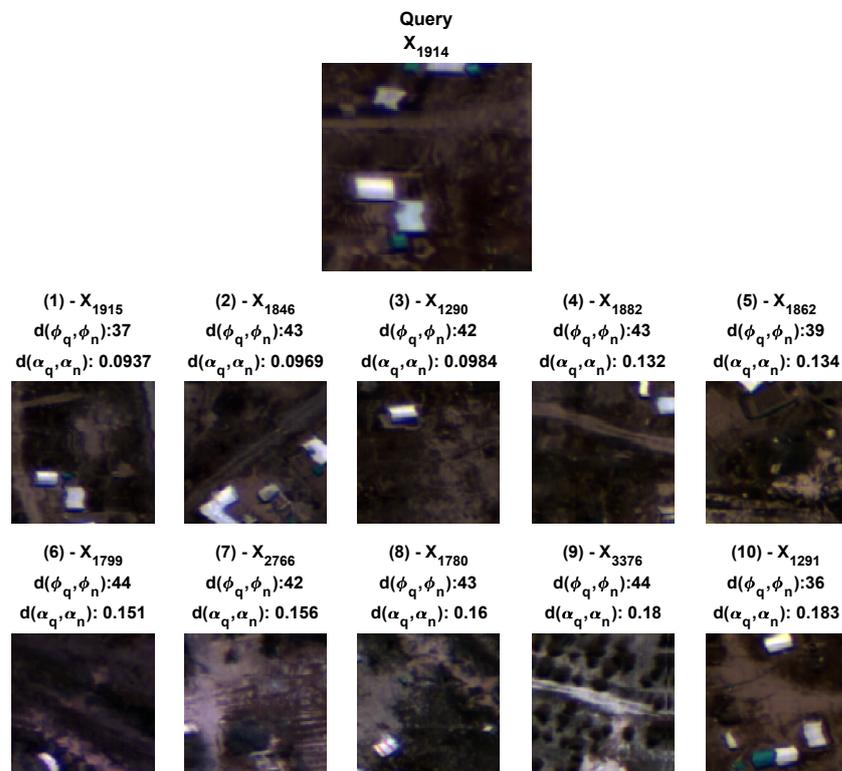
The retrieval results presented in Figure 13 consist of content predominantly related to railway ballast material, steel rail, natural vegetation cover, and stabilized road, using a query image. The proposed system has successfully retrieved other images from the archive containing materials with similar spectral characteristics. For the retrieval results presented in Figure 14, a query image was used with content primarily focused on a red-tiled roof, metal sheet roof, natural vegetation cover, and stabilized road. The proposed system retrieves other hyperspectral images from the archive containing materials with similar spectral characteristics. In Figure 15, retrieval results for a query hyperspectral image specifically containing white tent tarpaulin observed in rural regions are shown. Figure 16 presents the retrieval results of a query hyperspectral image that is dominantly composed of bare soil and a specific tree type.



**Figure 13.** Content-based retrieval results of the proposed CBHIR system,  $X_q = X_{125}$ . Content labels of each image are given in Table 2.



**Figure 14.** Content-based retrieval results of the proposed CBHIR system,  $X_q = X_{1211}$ . Content labels of each image are given in Table 3.



**Figure 15.** Content-based retrieval results of the proposed CBHIR system,  $X_q = X_{1914}$ . Content labels of each image are given in Table 4.



**Table 4.** Content labels for retrieval results,  $X_q = X_{1914}$ .

$X_q$	1	2	3	4	5	6	7	8	9	10
$X_{1914}$	$X_{1245}$	$X_{1212}$	$X_{1100}$	$X_{1142}$	$X_{2724}$	$X_{2758}$	$X_{2769}$	$X_{1159}$	$X_{1153}$	$X_{1143}$
white tent	white tent	white tent	white tent	white tent	white tent	white tent	white tent	white tent	white tent	water stream
metal sheet	metal sheet	metal sheet		metal sheet				gravel road	metal sheet	blue painted object
blue painted object								tree (Type-1)	bare soil	bare soil
natural vegetation	natural vegetation	natural vegetation	natural vegetation	natural vegetation	natural vegetation	natural vegetation	natural vegetation	natural vegetation	natural vegetation	natural vegetation
bare soil	bare soil	bare soil	bare soil	bare soil	bare soil	bare soil	bare soil	bare soil	natural vegetation	natural vegetation

**Table 5.** Content labels for retrieval results,  $X_q = X_{2440}$ .

$X_q$	1	2	3	4	5	6	7	8	9	10
$X_{2440}$	$X_{2456}$	$X_{2441}$	$X_{2454}$	$X_{2455}$	$X_{2453}$	$X_{2438}$	$X_{2425}$	$X_{992}$	$X_{2486}$	$X_{2424}$
tree (Type-3)										
bare soil										
natural vegetation										

## 6.2. Comparative Performance Analysis

The retrieval performance of the proposed CBHIR system is compared with three state-of-the-art systems available in the literature: (1) the bag-of-endmember-based method (denoted as BoE) [19], (2) the endmember matching algorithm based on the Grana Distance (denoted as EM-Grana) [11], and (3) the endmember matching algorithm that weights the distances estimated by the SAD between each endmember pair by their abundances (denoted as EM-WSAD) [12].

In the experiments, the proposed CBHIR system and BoE were examined in both retrieval scenarios defined in Section 3. When only spectral similarity is considered, single-stage retrieval (SSR) is applied to the images represented by the binary spectral content descriptors (BSDs). In the case of both spectral similarity and abundance of corresponding materials considered (BSAD), both spectral and abundance descriptors are considered with two-stage hierarchical retrieval (TSHR).

To measure the retrieval performance of the system in this regard, each hyperspectral image  $X_n$  in  $\mathbf{X}$  was used as the query hyperspectral image to retrieve 10 hyperspectral images that contain similar materials. It is worth noting that while the proposed system performs retrieval based on overall content, other CBHIR systems perform retrieval based on the strategy they built on.

Comparative performance results given in Table 6 show that the proposed system performs the retrieval with the highest accuracy (82.20%) in cases where both spectral and abundance descriptors are utilized by considering overall image content. Likewise, the proposed system has the highest precision (84.28%) and recall (85.54%) values. Similarly, the lowest Hamming Loss score also belongs to the proposed algorithm when the retrieval is performed concerning the spectral descriptor only.

**Table 6.** Performance evaluation of CBHIR systems.

CBHIR SYSTEM	Method	Accuracy (%)	Precision (%)	Recall (%)	Hamming Loss	Retrieval Time (ms)
BoE	BSD-SSR	64.82	76.03	74.17	6.02	<b>0.114</b>
	BSAD-TSHR	66.43	63.22	73.48	6.21	0.129
Proposed System	Overall-SSR	76.65	<b>84.28</b>	<b>85.54</b>	<b>4.48</b>	0.146
	Overall-TSHR	<b>82.20</b>	83.25	82.43	5.21	0.159
EM-Grana		58.47	61.26	64.25	7.03	83.442
EM-WSAD		51.47	54.18	57.18	9.44	18756.36

On the other hand, it has been observed that the proposed CBHIR system exhibits an increase in retrieval time compared to the previously suggested bag-of-endmember-based CBHIR system. This is because the descriptor vector lengths in the proposed CBHIR system are longer than those calculated in the previously suggested bag-of-endmember-based CBHIR system.

In observations made with different query images within the archive, it has been observed that in some cases, the results retrieved for images containing much more diverse and less prominent foreground material (e.g., urban areas) are negatively affected by this diversity compared to others. This phenomenon has been attributed to the Hamming distance criterion used in the first stage of the image retrieval process. Although Hamming distance significantly reduces computational complexity in the retrieval process with binary vectors, spectral differences within the same type of foreground content have been observed to lead to such results.

## 7. Conclusions and Future Work

This study proposes a novel content-based hyperspectral image retrieval (CBHIR) system to define global hyperspectral image representations based on a semantic approach to differentiate foreground and background image content. This approach significantly improves the performance at the expense of slightly increasing the retrieval time compared to the bag-of-endmembers method, whereas it is superior to the other methods in both aspects. It offers several advantages over the conventional approach of using only endmembers to retrieve hyperspectral images from an archive. The proposed system considers spatial and spectral relationships through obtained content segments, which enables more accurate modeling of content in hyperspectral imagery. It categorizes the content of hyperspectral images into two classes—foreground and background—and defines the content belonging to these two classes with different spectral vocabularies. This allows for considering less common materials than those typically seen in hyperspectral image archives during the modeling phase of images. Thus, the proposed CBHIR system enables accurate retrieval of hyperspectral imagery from an archive using a query representing spectral features of similar content, including rarely seen materials. This could be advantageous in various applications such as, but not limited to, precision agriculture, forestry, mining, and defense to detect and locate less abundant materials in an archive. Furthermore, the system allows hyperspectral images to be retrieved online by characterizing the hyperspectral image content using four low-dimensional global feature descriptors in the background. Therefore, it is a more effective and sophisticated approach to accessing hyperspectral images in remote sensing archives.

A multi-label benchmark hyperspectral image archive was created from high-resolution airborne hyperspectral remote sensing data products to evaluate the retrieval performance of the proposed CBHIR system and compare it with the state-of-the-art systems available in the literature. The experiments conducted on this benchmark archive of hyperspectral images demonstrate the effectiveness of the proposed system in terms of retrieval accuracy and time.

Although the proposed CBHIR system exhibits higher retrieval performance compared to other systems during the experimental process, it also has certain shortcomings observed.

The first of these is the input requirement from the user in modeling background content, even though this process is carried out semi-supervised. It is believed that fully unsupervised decomposition of foreground and background content would positively impact the system's performance. In future work, alternative methods, e.g., a neural network-based model, to decompose image content in an unsupervised manner will be taken.

Another observed limitation is the use of Hamming distance in comparing spectral descriptors. Hamming distance evaluates two spectral descriptor vectors in a binary manner, assigning a penalty score for each spectral term that is not common between the two vectors. As a future work, a weighted distance measure for descriptors considering spectral features of the terms could be developed.

**Author Contributions:** Conceptualization, F.Ö., Y.Y.Ç., U.M.L. and B.D.; methodology, F.Ö., Y.Y.Ç., U.M.L. and B.D.; software, F.Ö.; validation, F.Ö. and Y.Y.; formal analysis, F.Ö.; investigation, F.Ö.; resources, F.Ö.; data curation, F.Ö.; writing—original draft preparation, F.Ö.; writing—review and editing, Y.Y.Ç., U.M.L. and B.D.; visualization, F.Ö.; supervision, Y.Y.Ç., U.M.L. and B.D.; project administration, Y.Y.Ç. and B.D.; funding acquisition, F.Ö. and B.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the German Federal Ministry of Education and Research under the grant BIFOLD24B.

**Data Availability Statement:** The benchmark hyperspectral image archive, false-color representations, and content and category labels of each image can be downloaded at <https://www.doi.org/10.17605/OSF.IO/H2T8U> in MATLAB (.mat), bitmap (.bmp), and comma-separated values (.csv) formats, respectively.

**Acknowledgments:** Part of this research was supported by TÜBITAK-BİDEB 2214-A International Research Fellowship Programme for Ph.D. Students.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Abbreviations

A-PPI	Automatic Pixel Purity Index
AGL	Above Ground Level
ASL	Above Sea Level
BoE	Bag-of-Endmembers-Based Method Proposed in [19]
BSAD	Binary Spectral and Abundance Descriptors
BSD	Binary Spectral Descriptors
CBHIR	Content-Based Hyperspectral Image Retrieval
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
EIHA	Endmember Induction Heuristic Algorithm
EM-Grana	Grana Distance Proposed in [11]
EM-WSAD	Endmember Matching-Based Algorithm Proposed in [12]
FODIS	Fiber Optic Downwelling Irradiance Sensor
GLCM	Gray Level Co-occurrence Matrix
hyperSLIC	Hyperspectral Simple Linear Iterative Clustering
LZW	Lempel–Ziv–Welch
OSP	Orthogonal Subspace Projection
PPI	Pixel Purity Index
RGB	Red–Green–Blue
SAD	Spectral Angular Distance
SAM	Spectral Angle Mapper
SID-SAD	Spectral Information Divergence–Spectral Angular Distance
SIFT	Scale Invariant Feature Transform
SLIC	Simple Linear Iterative Clustering
SSMA	Spectral Signature Matching Algorithm

SSR	Single-Stage Retrieval
SWIR	Shortwave Infrared
TSHR	Two-Stage Hierarchical Retrieval
VCA	Vertex Component Analysis
VHR	Very High Resolution
VNIR	Visible–Near Infrared

## Appendix A

**Table A1.** Symbols and Their Descriptions

Symbol	Description
$\mathbf{X} = \{\mathbf{X}_n\}_{n=1}^N$	Archive of N hyperspectral images
$\mathbf{X}_n$	$n$ -th hyperspectral image in $\mathbf{X}$
$\mathbf{X}_q$	Query hyperspectral image
$\mathbf{X}^R$	The ranked set of R retrieved images that are most similar to $\mathbf{X}_q$
$\mathbf{X}_r$	$r$ -th retrieved hyperspectral image in $\mathbf{X}^R$
$W$	Number of spectral bands
$P$	Number of pixels
$\mathbf{x}_n^p \in \mathbb{R}^W$	Spectral signature vector of $p$ -th spatial pixel in $\mathbf{X}_n$ , where $1 \leq p \leq P$
$\mathbf{V}^f = [\mathbf{v}_1^f, \dots, \mathbf{v}_\Psi^f]$	Foreground content spectral vocabulary, where $\mathbf{v}_\psi \in \mathbb{R}^W$ and $\psi = 1, 2, \dots, \Psi$
$\mathbf{V}^b = [\mathbf{v}_1^b, \dots, \mathbf{v}_\Omega^b]$	Background content spectral vocabulary, where $\mathbf{v}_\omega \in \mathbb{R}^W$ and $\omega = 1, 2, \dots, \Omega$
$\phi_n^f$	Foreground content spectral descriptor of $\mathbf{X}_n$
$\phi_n^b$	Background content spectral descriptor of $\mathbf{X}_n$
$\phi_n$	Overall content spectral descriptor of $\mathbf{X}_n$
$\phi_q$	Overall content spectral descriptor of $\mathbf{X}_q$
$\alpha_n^f$	Foreground content abundance descriptor of $\mathbf{X}_n$
$\alpha_n^b$	Background content abundance descriptor of $\mathbf{X}_n$
$\alpha_n$	Overall content abundance descriptor of $\mathbf{X}_n$
$\alpha_q$	Overall content abundance descriptor of $\mathbf{X}_q$
$\sigma_{\mathbf{X}_n}$	Spectral diversity of $\mathbf{X}_n$
$S$	Number of content segments extracted from $\mathbf{X}_n$
$s$	$s$ -th content segment extracted from $\mathbf{X}_n$
$\mathbf{x}_n^s$	Spectral signature representing $s$ -th segment extracted from $\mathbf{X}_n$
$\mu_{\mathbf{B}}$	Sample mean for territorial background image $\mathbf{B}$
$\mathbf{B}$	Territorial background image
$\Gamma_{\mathbf{B}}^{-1}$	Covariance matrix for territorial background image
$\mathbf{L}_{\mathbf{X}}$	Set of associated category labels with archive $\mathbf{X}$
$\mathbf{L}_{\mathbf{X}_q}$	Set of associated category labels with $\mathbf{X}_q$
$\mathbf{L}_{\mathbf{X}_r}$	Set of associated category labels with $\mathbf{X}_r$

## References

- Lu, B.; Dao, P.D.; Liu, J.; He, Y.; Shang, J. Recent Advances of Hyperspectral Imaging Technology and Applications in Agriculture. *Remote Sens.* **2020**, *12*, 2659. [\[CrossRef\]](#)
- Jia, J.; Wang, Y.; Chen, J.; Guo, R.; Shu, R.; Wang, J. Status and application of advanced airborne hyperspectral imaging technology: A review. *Infrared Phys. Technol.* **2020**, *104*, 103115. [\[CrossRef\]](#)
- Sneha; Kaul A. Hyperspectral imaging and target detection algorithms: A review. *Multimed. Tools Appl.* **2022**, *81*, 44141–44206. [\[CrossRef\]](#)
- Ghamisi, P.; Yokoya, N.; Li, J.; Liao, W.; Liu, S.; Plaza, J.; Rasti, B.; Plaza, A. Advances in Hyperspectral Image and Signal Processing: A Comprehensive Overview of the State of the Art. *IEEE Geosci. Remote Sens. Mag.* **2017**, *5*, 37–78. [\[CrossRef\]](#)

5. Singh, P.; Pandey, P.C.; Petropoulos, G.P.; Pavlides, A.; Srivastava, P.K.; Koutsias, N.; Deng, K.A.K.; Bao, Y. 8-Hyperspectral remote sensing in precision agriculture: Present status, challenges, and future trends. In *Hyperspectral Remote Sensing*; Pandey, P.C., Srivastava, P.K., Balzter, H., Bhattacharya, B., Petropoulos, G.P., Eds.; Earth Observation; Elsevier: Amsterdam, The Netherlands, 2020; pp. 121–146. [[CrossRef](#)]
6. Veganzones, M.; Datcu, M.; Grana, M. Dictionary based Hyperspectral Image Retrieval. In Proceedings of the ICPRAM (1), Vilamoura, Algarve, Portugal, 6–8 February 2012; pp. 426–432.
7. Plaza, A.J.; Plaza, J.; Paz, A. Parallel heterogeneous CBIR system for efficient hyperspectral image retrieval using spectral mixture analysis. *Concurr. Comput. Pract. Exp.* **2010**, *22*, 1138–1159. [[CrossRef](#)]
8. Plaza, A.J. Content-based hyperspectral image retrieval using spectral unmixing. In Proceedings of the Image and Signal Processing for Remote Sensing XVII, Prague, Czech Republic, 19–21 September 2011; Volume 8180. [[CrossRef](#)]
9. Zhang, J.; Zhou, Q.; Zhuo, L.; Geng, W.; Wang, S. A CBIR System for Hyperspectral Remote Sensing Images Using Endmember Extraction. *Int. J. Pattern Recognit. Artif. Intell.* **2017**, *31*, 1–15. [[CrossRef](#)]
10. Zheng, P.; Wu, Z.; Sun, J.; Zhang, Y.; Zhu, Y.; Shen, Y.; Yang, J.; Wei, Z.; Plaza, A. A Parallel Unmixing-Based Content Retrieval System for Distributed Hyperspectral Imagery Repository on Cloud Computing Platforms. *Remote Sens.* **2021**, *13*, 176. [[CrossRef](#)]
11. Graña, M.; Veganzones, M.A. An Endmember-Based Distance for Content Based Hyperspectral Image Retrieval. *Pattern Recognit.* **2012**, *45*, 3472–3489. [[CrossRef](#)]
12. Veganzones, M.A.; Grana, M. A Spectral/Spatial CBIR System for Hyperspectral Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *5*, 488–500. [[CrossRef](#)]
13. Sevilla, J.; Plaza, A. A New Digital Repository for Hyperspectral Imagery With Unmixing-Based Retrieval Functionality Implemented on GPUs. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 2267–2280. [[CrossRef](#)]
14. Sevilla, J.; Jiménez, L.I.; Plaza, A. Sparse Unmixing-Based Content Retrieval of Hyperspectral Images on Graphics Processing Units. *IEEE Geosci. Remote. Sens. Lett.* **2015**, *12*, 2443–2447. [[CrossRef](#)]
15. Shao, Z.; Zhou, W.; Cheng, Q.; Diao, C.; Zhang, L. An effective hyperspectral image retrieval method using integrated spectral and textural features. *Sens. Rev.* **2015**, *35*, 274–281. [[CrossRef](#)]
16. Geng, W.; Zhang, J.; Zhuo, L.; Liu, J.; Chen, L., Creating Spectral Words for Large-Scale Hyperspectral Remote Sensing Image Retrieval. In *Advances in Multimedia Information Processing—PCM 2016: 17th Pacific-Rim Conference on Multimedia, Xi'an, China, 15–16 September 2016, Proceedings, Part II*; Chen, E.; Gong, Y.; Tie, Y., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 116–125. [[CrossRef](#)]
17. Zhang, J.; Geng, W.; Liang, X.; Zhuo, L.; Chen, L. Secure retrieval method of hyperspectral image in encrypted domain. *J. Appl. Remote Sens.* **2017**, *11*, 1035021. [[CrossRef](#)]
18. Zhang, J.; Geng, W.; Liang, X.; Li, J.; Zhuo, L.; Zhou, Q. Hyperspectral remote sensing image retrieval system using spectral and texture features. *Appl. Opt.* **2017**, *56*, 4785–4796. [[CrossRef](#)] [[PubMed](#)]
19. Ömrüuzun, F.; Demir, B.; Bruzzone, L.; Çetin, Y.Y. Content based hyperspectral image retrieval using bag of endmembers image descriptors. In Proceedings of the 2016 8th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Los Angeles, CA, USA, 21–24 August 2016; pp. 1–4. [[CrossRef](#)]
20. Zhang, J.; Chen, L.; Zhuo, L.; Liang, X.; Li, J. An Efficient Hyperspectral Image Retrieval Method: Deep Spectral-Spatial Feature Extraction with DCGAN and Dimensionality Reduction Using t-SNE-Based NM Hashing. *Remote Sens.* **2018**, *10*, 271. [[CrossRef](#)]
21. Xu, X.; Li, J.; Wu, C.; Plaza, A. Regional clustering-based spatial preprocessing for hyperspectral unmixing. *Remote Sens. Environ.* **2018**, *204*, 333–346. [[CrossRef](#)]
22. Achanta, R.; Shaji, A.; Smith, K.; Lucchi, A.; Fua, P.; Süsstrunk, S. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 2274–2282. [[CrossRef](#)] [[PubMed](#)]
23. Kruse, F.; Lefkoff, A.; Boardman, J.; Heidebrecht, K.; Shapiro, A.; Barloon, P.; Goetz, A. The spectral image processing system (SIPS)—Interactive visualization and analysis of imaging spectrometer data. *Remote Sens. Environ.* **1993**, *44*, 145–163. Airbone Imaging Spectrometry, [[CrossRef](#)]
24. Ester, M.; Kriegel, H.P.; Sander, J.; Xu, X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, Portland, OR, USA, 2–4 August 1996; AAAI Press: Washington, DC, USA 1996; pp. 226–231.
25. P Nascimento, J.M.; Bioucas-Dias, J.M. Hyperspectral signal subspace estimation. In Proceedings of the 2007 IEEE International Geoscience and Remote Sensing Symposium, Barcelona, Spain, 23–28 July 2007; pp. 3225–3228. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.