Alignment of Hyperspectral Images Using KAZE Features

Álvaro Ordóñez [1],* , **Francisco Argüello [2]** and **Dora B. Heras [1]**

[1]   Centro Singular de Investigación en Tecnoloxías da Información (CiTIUS), Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain; dora.blanco@usc.es
[2]   Departamento de Electrónica e Computación, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain; francisco.arguello@usc.es
*   Correspondence: alvaro.ordonez@usc.es; Tel.: +34-8818-16474

check for updates

Abstract: Image registration is a common operation in any type of image processing, specially in remote sensing images. Since the publication of the scale–invariant feature transform (SIFT) method, several algorithms based on feature detection have been proposed. In particular, KAZE builds the scale space using a nonlinear diffusion filter instead of Gaussian filters. Nonlinear diffusion filtering allows applying a controlled blur while the important structures of the image are preserved. Hyperspectral images contain a large amount of spatial and spectral information that can be used to perform a more accurate registration. This article presents HSI–KAZE, a method to register hyperspectral remote sensing images based on KAZE but considering the spectral information. The proposed method combines the information of a set of preselected bands, and it adapts the keypoint descriptor and the matching stage to take into account the spectral information. The method is adequate to register images in extreme situations in which the scale between them is very different. The effectiveness of the proposed algorithm has been tested on real images taken on different dates, and presenting different types of changes. The experimental results show that the method is robust achieving image registrations with scales of up to $24.0\times$.

Keywords: hyperspectral data; image registration; KAZE features; remote sensing

1. Introduction

The advances in sensor development in the last decades allow us to obtain hyperspectral images (HSI) at a lower cost than before. Each pixel of these images contains a continuous spectrum that is formed by hundreds of narrow bands. As a result of this high spectral resolution, objects, plant species, and land–cover classes among others, can be distinguished with higher accuracy. Thanks to this large amount of available information, the use of hyperspectral remote sensing images has been extended to a multitude of applications such as vegetation science [1], land use classification [2], geology [3], quality control [4], and change detection [5] among others.

A previous fundamental task in many of these applications is the registration of images of the same scene which have been taken at different times from different viewpoints and which, furthermore, present changes in objects, in illumination, etc. The goal of the registration is to determine the geometric transformation that aligns the images. There are different types of registration algorithms due to the variety of images and the conditions that the method has to handle. They can be classified into two categories according to their nature [6]: area–based and feature–based methods. In the first group, the methods based on correlation between images [7], mutual information (MI) [8], and Fourier transform [9] stand out. These methods work directly with image intensity unlike feature–based methods which seek to detect distinctive features in objects or interest points at a

higher level. Feature–based methods extract significant regions, lines or points that must have certain characteristics: invariant to geometric transformations, good localization accuracy, and be insensitive to image degradation effects. This high level representation makes feature–based methods more resilient to illumination changes, intensity changes introduced by noise, and changes introduced by the use of different sensors. On the other hand, area–based methods are computationally more efficient and work better on images that are not so rich in details, e.g., medical images.

Due to the fact that remote sensing images are rich in details such as lines, edges, corners, regions, etc., and they normally present illumination changes, feature–based methods are the preferred approach when dealing with this type of images. The scale–invariant feature transform (SIFT) [10] is the most popular feature–based algorithm. It consists of four stages: scale–space extrema detection, keypoint localization, orientation assignment, and keypoint description. In the first step, a Gaussian scale space is created performing Gaussian convolutions and interpolations. The original images are smoothed at different levels using Gaussian filters and at different scales. The result is a multi–resolution pyramid of the images. This way, the characteristic of invariance to image scaling is achieved. Afterwards, a difference of Gaussian (DoG) is obtained to detect the keypoints. A point is considered keypoint if it is the local minimum or maximum compared to its 26 neighbours (8 neighbours at the same scale and 9 at the adjacent scales) of the DoG scale space. Once candidate keypoints are obtained, the second step carries out a selection to discard low–contrast keypoints, and their location and scale are accurately determined. In the third step, one or more consistent orientations are assigned to each keypoint to achieve the rotation invariance. It is calculated from the orientation histogram formed from the gradient orientations of points within the region of the keypoint. The size of this region depends on the scale and the gradient magnitude. Finally, in the last step, a 128 value descriptor for each keypoint is generated from a set of weighted histograms of the gradient orientation computed on different regions of the normalized neighbourhood. Then, the keypoints of each image are matched according to a ratio of the Euclidean distance between descriptors. The matched keypoints are used as control points to recover the image transformation in order to align the images.

Some approaches were proposed following the SIFT scheme. The Speeded Up Robust Features (SURF) [11] was proposed to be computed much faster than SIFT. This method exploits integral images to build the scale space. This is an approximation of the Gaussian scale space that reduces the computation time. Regarding the descriptor, it is formed only by 64 dimensions and the orientation is estimated using the Haar wavelet. The problem of the techniques to build the scale space in SIFT and SURF is that they do not respect the edges of the images, degrading important details at the same level as noise. Therefore, we are losing precision to detect keypoints and distinctiveness. KAZE [12] proposes the use of a nonlinear diffusion filter to build the scale space applying a selective blurring. This way, the noise is blurred but the details and edges are preserved. Moreover, it uses the M–SURF descriptor [13] which is faster and handles the boundaries better than the original SURF descriptor.

These feature–based methods are commonly well–known due to their effectiveness but they are not suitable for some applications. The binary descriptors are an alternative to the SIFT and SURF descriptors in cases where the speed–up is more important than the effectiveness. This type of descriptors provide a more efficient matching and are more compact to store thanks to their binary descriptor. The computation of the Hamming distance is used to match the keypoints. It can be done in a single instruction, i.e., its computation is cheaper than the Euclidean distance saving computation time in large sets of features. In the last decade, many feature–based methods including a binary descriptor have been proposed in the literature: BRIEF [14], ORB [15], LATCH [16], and A–KAZE [17], among others. Many of these also accelerate the scale space building process, e.g., A–KAZE uses a pyramidal Fast Explicit Diffusion (FED) scheme for building it much faster.

However, all these algorithms were developed to work with greyscale or RGB images, and not to deal with the spectral information available in multi and hyperspectral images. In order to register multispectral images, some authors proposed modifications to SIFT. In this line, Ref. [18]

suggested adding a scale restriction criteria to remove incorrect matching points in multispectral remote image registration. In addition, Ref. [19] also proposed applying an orientation restriction. A similar approach was followed by [20], who proposed a descriptor vector with 4 orientation bins instead of 8. These restrictions were also adapted to SURF in [21]. Moreover, Ref. [22] presented an approach that uses the spectral information of neighbouring keypoints in addition to the descriptor to discard false matches. However, these methods only use one band of each image to perform the registration, i.e., they do not use all the available spectral information. Ref. [23] introduced an interest point detector for hyperspectral images that uses principal component analysis (PCA) in order to reduce the dimensionality. For each PCA component, a scale space is built. Ref. [24] presented a method where the scale space is created using a nonlinear diffusion equation taking into account the spectral information. The keypoint localization consists in comparing each pixel vector to its neighbourhood according to their spectral signature. None of these methods is based on KAZE.

In this article, we present an automatic algorithm, called HSI–KAZE, to register hyperspectral remote sensing images taking into account the spectral information. It consists of a trade–off solution between KAZE and A–KAZE for hyperspectral images (HSI), and it is based on a pyramidal nonlinear scale space and on M–SURF descriptors. The algorithm is oriented towards extreme situations in which the images are very different in terms of scale, rotation, and other variations. An example of the registration problem considered in this work can be seen in Figure 1. Section 2 presents a review on the original KAZE. The proposed algorithm is described in Section 3. In Section 4, the results obtained under different images and under different conditions are discussed. Finally, some conclusions are given in Section 5.


| (a) | (b) | (c) |
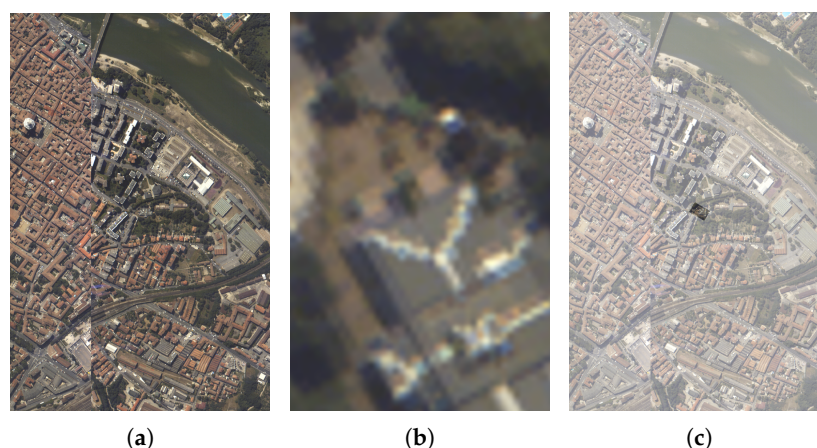
Figure 1. Example of registration considered in this work: (**a**) Reference image (size 1096×715), (**b**) Target image, and (**c**) Result of the registration process showing the correctly registered superposition of the reference and target registered image (scale $23.0\times$ and rotation angle $60°$).

2. Description of the KAZE Algorithm

The KAZE algorithm was presented by [12]. It is characterized by building the scale space using a nonlinear diffusion filter, along with the utilization of the M–SURF descriptors.

The method follows a similar approach to that of SIFT, building a scale space organized into octaves and, in turn, into scales. The difference lies in the approach to build this scale space. KAZE replaces the Gaussian approach proposed by Lowe [10] with a nonlinear diffusion filtering operation. The Gaussian approach blurs both noise and details, losing the natural boundaries of objects, buildings, etc. The nonlinear diffusion filtering method allows us to realize an adaptive blur keeping the details unaltered, as shown in Figure 2a. Specifically, KAZE builds the scale space using a semi–implicit scheme, one of the possible discretizations of the diffusion equations. This scheme is based on an Additive Operator Splitting (AOS) and is totally stable for any step size.

(a)

(b)
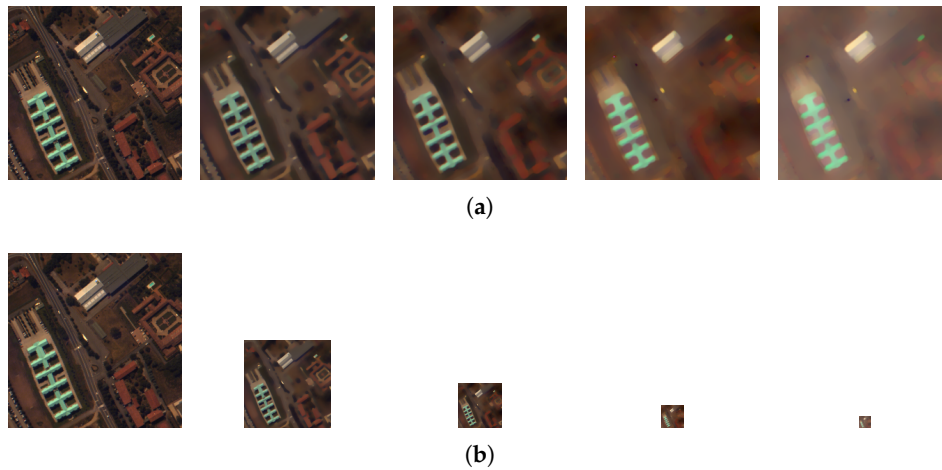
Figure 2. (**a**) Scale space used by KAZE (**b**) Pyramidal scale space used by A–KAZE. Each image represents the first sublevel of the next octave.

The classical nonlinear diffusion for an image X with spatial coordinates (x, y) and time t is expressed as

$$\frac{\partial X}{\partial t} = \text{div}(c(x, y, t) \cdot \nabla X), \tag{1}$$

where div and ∇ are the divergence and gradient operators, respectively, and c is the conductivity function that allows applying the diffusion to the image structure. Ref. [25] proposes that this function be dependent on the gradient in order to blur the inside of regions while preserving sharp edges, i.e.,

$$c(x, y, t) = g(|\nabla X_\sigma(x, y, t)|), \tag{2}$$

where function ∇X_σ is the gradient of a Gaussian smoothed version of the original image X with a standard deviation σ. The conductivity function g is chosen to promote wide regions,

$$g = \frac{1}{1 + \frac{|\nabla X_\sigma|^2}{k^2}}, \tag{3}$$

where constant k is the contrast factor that controls the level of diffusion. Following [12,26], the discretized version of Equation (1) can be expressed in matrix–vector notation as

$$\frac{X^{i+1} - X^i}{\tau} = \sum_{l=1}^{m} A_l(X^i) X^{i+1}, \tag{4}$$

where X^i and X^{i+1} are the smoothed images at the current and next level, respectively, m is the number of dimensions of image X (two in our case), τ is the time step size, and A_l is the matrix that encodes the diffusivities for each image dimension (derivatives along the l-th coordinate axis). This way, N_L levels of filtered images are computed, $i = 0, \ldots, N_L - 1$.

As we mentioned and following the SIFT approach, this set of levels is mapped over octaves and sublevels (scales in the SIFT notation) using the octave index o and the sublevel index s. Therefore, in a scale space of $N_L = N_{oct} \cdot N_{sub}$ filtered images the octaves and sublevels are mapped to their corresponding scales σ_i, as follows

$$\sigma_i(o, s) = \sigma_0 2^{\frac{o+s}{S}}, \ i \in [0, \ldots, N_L - 1], \ o \in [0, \ldots, N_{oct} - 1], \ s \in [0, \ldots, N_{sub} - 1], \tag{5}$$

where σ_0 is the base scale level. It is necessary to note that KAZE does not perform a subsampling in each octave as SIFT does. Nonlinear diffusion filtering is defined in terms of time, so the scale units σ_i are mapped to time units t_i,

$$t_i = \frac{1}{2}\sigma_i^2, \ i \in [0, \ldots, N_L - 1]. \tag{6}$$

Summarizing, to build the scale space, KAZE smooths the original image using a Gaussian kernel of a standard deviation σ_0 to reduce noise. Next, the k parameter is obtained as the 70th percentile of the gradient histogram of this smoothed version of the image. In order to compute X^{i+1} from X^i in Equation (4) a tridiagonal linear system must be solved which can be efficiently done using Thomas algorithm.

Once the scale space has been built, it is necessary to detect the keypoints, which are points with certain characteristics (independence of position, robustness against image transformations, and scale independence). A pixel will be considered as keypoint if it is the maximum of its neighbourhood when the Hessian matrix is calculated. The determinant of the Hessian matrix at the different σ_i scale levels is computed as follows

$$X_{Hessian}^i = \sigma^2 (X_{xx} X_{yy} - X_{xy}^2), \tag{7}$$

where X_{xx} and X_{yy} are the second order horizontal and vertical derivatives, respectively, and X_{xy} is the second order cross derivative. This set of first and second order derivatives are approximated by 3×3 Scharr filters which provide better rotation invariance than other popular filters [27]. The maximum is computed at all the scale levels excluding the first and the last scales as well as the image borders. The neighbourhood is considered in a 3D space considering the 8 neighbours on the same scale and the 9 neighbours on the upper and lower scales. Therefore, we have 26 neighbours per pixel. To end with the feature detection step, the location of the extremas, keypoints from now on, is refined in position and scale using the local interpolation model proposed for SIFT in [10].

The next step of the KAZE algorithm is the orientation assignment for each keypoint. KAZE calculates the possible orientation through the derivatives around each point within a radius of $6\sigma_i$, being σ_i the scale of the keypoint. Similar to SURF, each derivative in the circular area is weighted with a Gaussian centred at the keypoint. For each point, the dominant orientation is calculated by summing the derivatives within a sliding circle window covering an angle of $\pi/3$. The sums are represented as vectors. The dominant orientation is selected from the longest vector. To save computation time, these derivatives are reused from the location refinement in the previous step.

With the main orientation estimated, the next step is the descriptor construction. KAZE uses the M–SURF descriptor [13] adapted to its nonlinear scale space. Firstly, for each keypoint, the first derivatives, X_x and X_y, are computed over a region of size $24\sigma_i \times 24\sigma_i$. This region is split into 4×4 subregions of size $9\sigma_i \times 9\sigma_i$ with an overlap of $2\sigma_i$ among them in order to avoid the boundary effects. Each subregion is weighted using a Gaussian ($\sigma = 2.5\sigma_i$) centred in the subregion centre and the derivative responses are summed up into a descriptor vector. Next, each subregion vector is weighted using another Gaussian ($\sigma = 1.5\sigma_i$) defined over a mask of size 4×4 and centred on the keypoint. Both, the samples of the $24\sigma_i \times 24\sigma_i$ region and the computation of derivatives, are taken into account to determine the main orientation of the keypoint. Finally, to achieve invariance to contrast, the descriptor vector is converted into an unit vector.

An accelerated method called Accelerated–KAZE (A–KAZE) was presented in [17]. A–KAZE is based on a pyramidal nonlinear scale space and on binary descriptors that makes it more efficient in terms of computation at the cost of accuracy loss. Figure 2b illustrates this scale space. Moreover, the scale space is built using a Fast Explicit Diffusion (FED) embedded in the pyramidal approach which is faster than any other discretization scheme [28], e.g., the AOS scheme. A–KAZE proposes an improvement of the Local Difference Binary (LDB) [29] called the Modified–LDB (M–LDB) that uses binary tests between the average of areas instead of single pixels, and adds the mean of the derivatives to the bits for the comparison.

3. HSI–KAZE: KAZE for Hyperspectral Remote Sensing Images

Hyperspectral remote sensing images are usually rich in edges, corners, boundaries, i.e., a large set of keypoints is obtained. In addition, they normally have a lot of repeated structures e.g., buildings, roads, fields, etc. These two conditions make harder registration, as it will need more computation time for the keypoint matching process, and several keypoints will probably be very similar, making it difficult to obtain correct matches. In this section, we present an automatic method to register two hyperspectral remote sensing images called HSI–KAZE that overcomes these limitations. It is based on KAZE and efficiently uses the spectral information available in the hyperspectral images. The spectral information is considered by searching for features in pairs of selected bands and by adding a spectral part to the descriptor. The main contributions of the method are the following:

- Band selection. A feature extraction method is proposed to select a set of most representative bands according to their entropy and their spectral distance.
- Keypoint detection. An interpolation is applied to the original images in order to highlight details and extract a large number of keypoints. The scale space is built using a nonlinear diffusion filter to blur noise and preserve details such as edges. In addition, it follows a pyramid scheme that improves the robustness to scale invariance. The discretization is carried out using a FED scheme.
- Keypoint description. HSI–KAZE uses a descriptor formed by a spatial and a spectral part: the M–SURF descriptor and the spectral signature.
- Keypoint matching using spectral information. The matching is based on the distance between the M–SURF descriptors and the cosine similarity between the spectral signatures. The spectral information allows refining the matching process discarding outliers.
- Band combination. Each selected band has unique characteristics that are not present in the other bands. The matched keypoints for the different bands are considered together in order to achieve a more accurate registration and use all the image information.
- Exhaustive search for registration. All the possible pairs of matched keypoints are considered and the outliers are discarded using a histogram–based approach.

Figures 3 and 4 show the outline of the proposed algorithm. The first stage performs a band selection in order to keep only the relevant spectral information and reduce the dimensionality of the hyperspectral images. In the second and third stages, keypoint detection and keypoint description, the features of each band are extracted and described. The fourth stage, keypoint matching, performs the matching of the extracted keypoints of both hyperspectral images. Then, in the fifth stage, band combination, all the matched keypoints are joined. Finally, in the last stage, registration, an exhaustive search based on histograms is performed to register the images. The pseudocode of the algorithm is presented in Figure 5. The main stages of the method are explained in the next subsections.
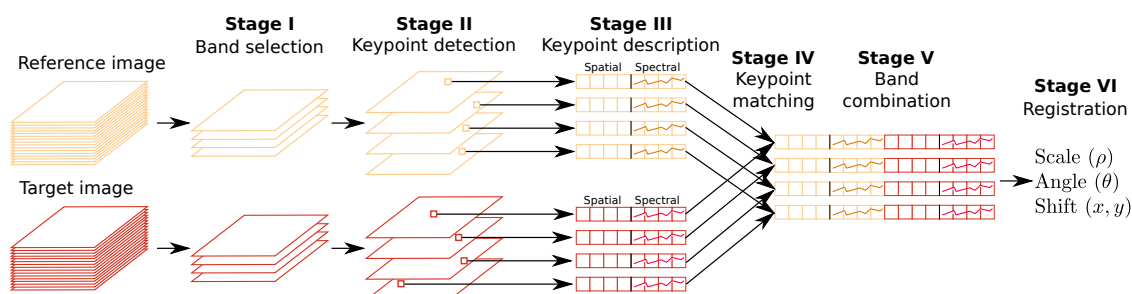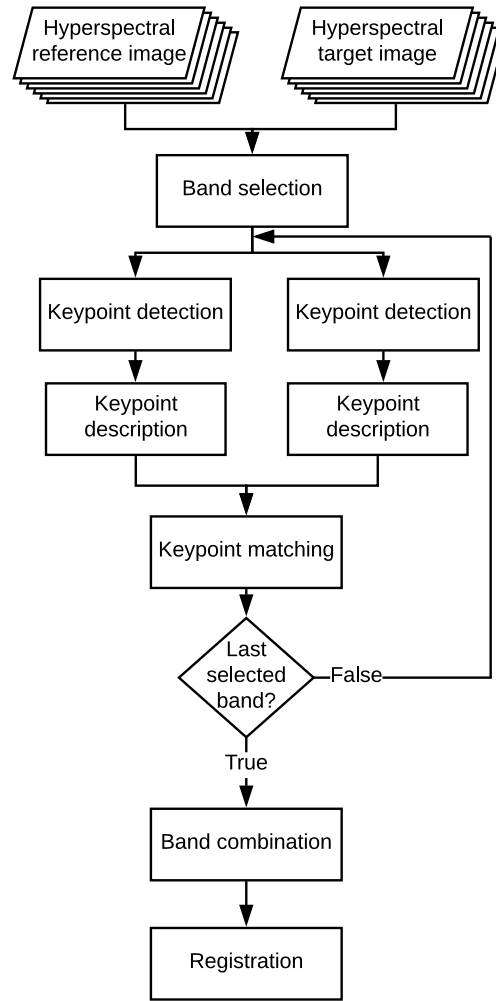

Figure 3. HSI–KAZE algorithm to register two hyperspectral images.

Figure 4. Flow chart of the proposed HSI–KAZE algorithm to register two hyperspectral images.

HSI–KAZE

Input: Hyperspectral reference image $\mathbf{I}_1$ and hyperspectral target image $\mathbf{I}_2$ with N_T bands.
Output: Scale factor ρ, rotation angle θ, and translation (x, y).

1: Perform feature selection over both images $\rightarrow \mathbf{B}_1$ and $\mathbf{B}_2$ ⊳ Band selection

2: **for each** band b **in** images $\mathbf{B}_1$ and $\mathbf{B}_2$ **do**
3:　　Extract keypoints of $\mathbf{B}_1^b \rightarrow \mathbf{P}_1^b$ ⊳ Keypoint detection
4:　　Extract keypoints of $\mathbf{B}_2^b \rightarrow \mathbf{P}_2^b$ ⊳ Keypoint detection
5:　　Calculate the M–SURF descriptor of each keypoint in $\mathbf{P}_1^b$ ⊳ Keypoint description
　　　and append the spectral signature $\rightarrow \mathbf{K}_1^b$
6:　　Calculate the M–SURF descriptor of each keypoint in $\mathbf{P}_2^b$ ⊳ Keypoint description
　　　and append the spectral signature $\rightarrow \mathbf{K}_2^b$
7:　　Match keypoints in $\mathbf{K}_1^b$ and $\mathbf{K}_2^b \rightarrow \mathbf{M}_b$ ⊳ Keypoint matching
8: **end for**

9: Combine all the matched keypoints $\mathbf{M}_b \rightarrow \mathbf{M}$ ⊳ Band combination
10: Perform an exhaustive search to recover the registration ⊳ Registration
　　parameters $\rightarrow \rho, \theta, (x, y)$

Figure 5. HSI–KAZE pseudocode.

3.1. Band Selection

The redundancy of information among bands in hyperspectral images is a well known problem. The different methods proposed in the literature can be classified into two groups: feature extraction and feature selection [30].

The first group, feature extraction methods, consists in combining the different bands of the image into a small set of new features. Some examples are Principal Component Analysis (PCA), Independent Component Analysis (ICA), or wavelet transforms, among others. These methods present the drawbacks of performing different transformations for each image and not preserving the original spectral information, making the feature–based registration methods are unable to identify the same spatial structures present in both images. This effect has a higher influence when both images present a large difference in scale, small overlapping areas, or changes in spatial structures due to the time difference.

On the other hand, the second group, feature selection methods, does not modify the original data and are essentially limited to choosing a band or a set of distinctive bands. Therefore, it is better to perform feature selection instead of feature extraction. Many methods choose the bands following a band selection criterion such as entropy, mutual information and the spectral angle mapper (SAM) among others. Some others apply a clustering method to group similar bands according different correlation measures.

In this work, we propose a feature selection method based on entropy and inter–band distance which is called Entropy–based Band Selection (EBS). A study of different feature reduction methods according to their effectiveness in order to produce a successfully registration was carried out. In the feature extraction category, PCA [31] and BandClust [32] are evaluated, and in the feature selection one, Ward's Linkage strategy using Mutual Information (WaluMI) [33] and our proposal (EBS) are considered.
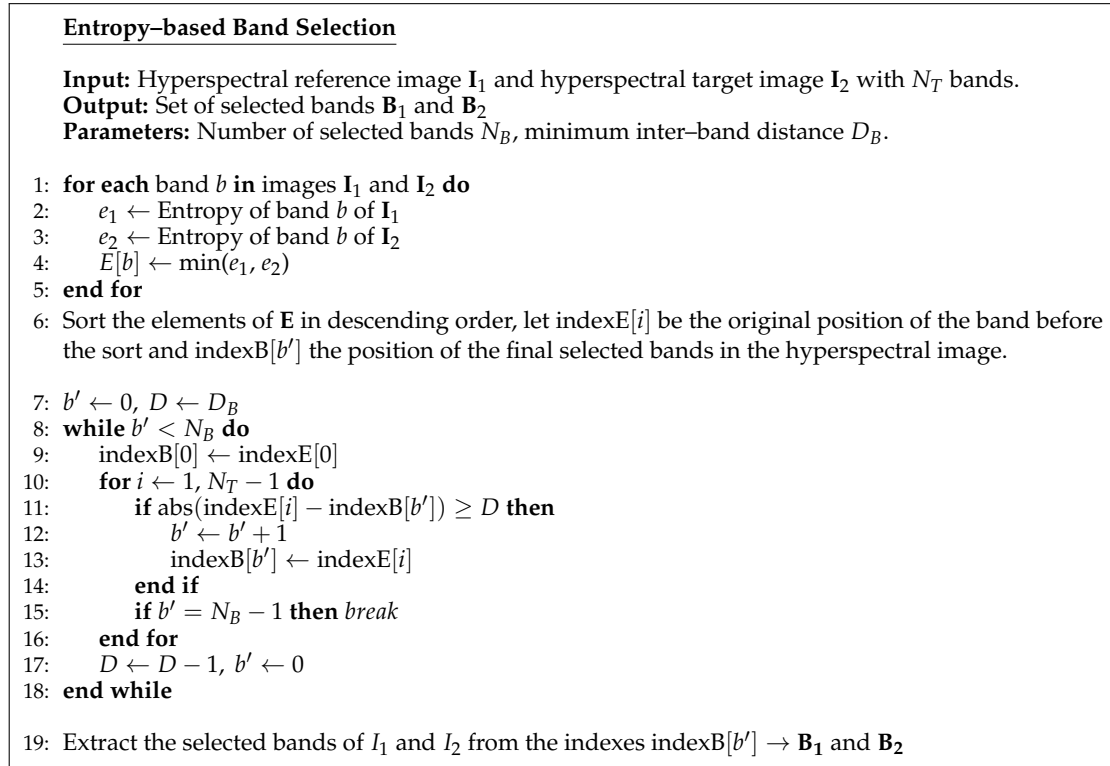
PCA is a well known statistical method to reduce the dimensionality in different types of problems. The main idea is to eliminate data redundancy present in the bands by means of a high correlation analysis. PCA generates a new set of linearly uncorrelated variables where the first few retain most of the variation present in all original variables.

BandClust follows a different approach since it is based on unsupervised clustering. It consists in splitting the initial set of bands into disjoint clusters according to a mutual information criterion. Each iteration splits the original set of bands into two clusters. The method stops when the minimum for the criterion is found. At the end, the average of each cluster is computed to obtain the final reduced image.

On the other hand, WaluMI performs an agglomerative clustering strategy based on Ward's linkage method [34]. At the beginning, each band forms a single cluster. Next, the algorithm searches for the two clusters with the minimum dissimilarity difference. In order to do that, the dissimilarity matrix is built based on the mutual information between each pair of bands. In the next steps, this matrix is updated using Ward's linkage strategy. Moreover, being feature selection method, WaluMI chooses the most representative band of each cluster at the end. These representative bands define the final compressed representation of the image.

The proposed method in this paper, EBS, performs a band selection based on entropy and inter–band distance. In contrast to the other methods, EBS uses both hyperspectral images. The method is detailed in the pseudocode shown in Figure 6. First, the entropy of each band of each image is computed (Figure 6, lines 2–3). Next, the minimum entropy of each band between the two images is selected (Figure 6, line 4). Finally, the N_b bands of highest entropy with an inter–band distance greater or equal than D_B between consecutive pairs are selected (Figure 6, lines 6–19). Specifically, this check is performed in line 11, where the candidate to be the next selected band (with index indexE[i]) must have, at least, an inter–band distance of D_B with respect to the previous selected band (with index indexB[b']). If it is not possible to select a set of bands with an inter–band distance D_B, D is decreased

until it becomes possible (Figure 6, line 17). N_b is fixed to 8 and the band separation D_B to 20 after experimental analysis.

Entropy–based Band Selection

Input: Hyperspectral reference image $\mathbf{I}_1$ and hyperspectral target image $\mathbf{I}_2$ with N_T bands.
Output: Set of selected bands $\mathbf{B}_1$ and $\mathbf{B}_2$
Parameters: Number of selected bands N_B, minimum inter–band distance D_B.

```
 1:  for each band b in images I₁ and I₂ do
 2:      e₁ ← Entropy of band b of I₁
 3:      e₂ ← Entropy of band b of I₂
 4:      E[b] ← min(e₁, e₂)
 5:  end for
 6:  Sort the elements of E in descending order, let indexE[i] be the original position of the band before
     the sort and indexB[b'] the position of the final selected bands in the hyperspectral image.

 7:  b' ← 0, D ← D_B
 8:  while b' < N_B do
 9:      indexB[0] ← indexE[0]
10:      for i ← 1, N_T − 1 do
11:          if abs(indexE[i] − indexB[b']) ≥ D then
12:              b' ← b' + 1
13:              indexB[b'] ← indexE[i]
14:          end if
15:          if b' = N_B − 1 then break
16:      end for
17:      D ← D − 1, b' ← 0
18:  end while

19:  Extract the selected bands of I₁ and I₂ from the indexes indexB[b'] → B₁ and B₂
```

Figure 6. Pseudocode for the band selection stage of the HSI–KAZE.

3.2. Keypoint Detection and Description

In order to detect a higher number of features and use the spectral information, the proposed HSI–KAZE presents some modifications in the feature extractor and in the descriptor as compared with the original KAZE.
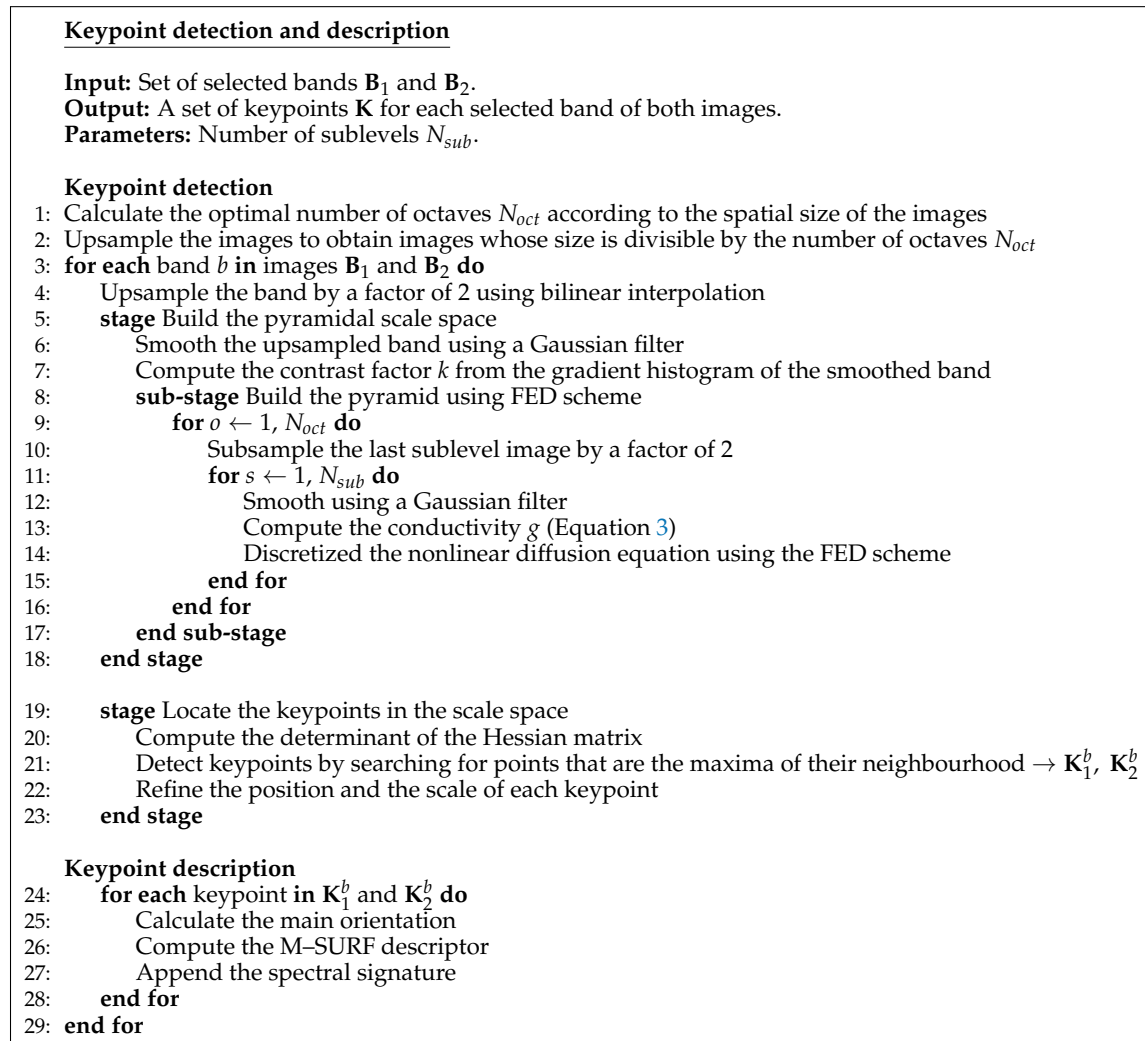
The pseudocode of these two stages is shown in Figure 7. The parameter N_{sub} is fixed to 4 as in the original KAZE [12]. The optimal number of octaves N_{oct} is calculated as follows [35]

$$N_{oct} = \min\left(8, \left\lfloor \log_2\left(\frac{\min(w, h)}{D_{sub}}\right) + 1 \right\rfloor\right) + 1, \tag{8}$$

where D_{sub} is the distance of initial upsampling applied to the image, in our case 2 which corresponds to a 2× interpolation, and w and h are the width and height of the interpolated image, respectively. First, the upsampling of the original image (Figure 7, line 4) is carried out using bilinear interpolation. It minimizes aliasing artefacts and allows extracting a higher number of keypoints. Second, the image data is used without normalization to avoid loss of precision. Third, regarding the scale space, it is built following a pyramidal scheme as in SIFT (Figure 7, lines 5–18), i.e., the image is subsampled for each octave using a bilinear interpolation. Moreover, HSI–KAZE uses FED for the discretization of the diffusion equations such as in A–KAZE (Figure 7, lines 8–17). FED schemes are characterized by their faster computation, ease of implementation, and higher accuracy than AOS approaches.

Finally, HSI–KAZE uses a descriptor consisting in two parts: the M–SURF descriptor as in the original KAZE and the spectral signature of the keypoint (Figure 7, lines 24–29). Thus, thanks to the use of spectral information, a more robust matching is performed. This spectral information corresponds to the components of the selected bands which provide enough information to discard false matches.

This strategy allows saving computation time in comparison with the use of all the bands of the image. At the end of this stage, a set of keypoints is obtained for each selected band of both images.

Keypoint detection and description

Input: Set of selected bands $\mathbf{B}_1$ and $\mathbf{B}_2$.
Output: A set of keypoints $\mathbf{K}$ for each selected band of both images.
Parameters: Number of sublevels N_{sub}.

Keypoint detection
1: Calculate the optimal number of octaves N_{oct} according to the spatial size of the images
2: Upsample the images to obtain images whose size is divisible by the number of octaves N_{oct}
3: **for each** band b **in** images $\mathbf{B}_1$ and $\mathbf{B}_2$ **do**
4:     Upsample the band by a factor of 2 using bilinear interpolation
5:     **stage** Build the pyramidal scale space
6:         Smooth the upsampled band using a Gaussian filter
7:         Compute the contrast factor k from the gradient histogram of the smoothed band
8:         **sub-stage** Build the pyramid using FED scheme
9:             **for** $o \leftarrow 1, N_{oct}$ **do**
10:                 Subsample the last sublevel image by a factor of 2
11:                 **for** $s \leftarrow 1, N_{sub}$ **do**
12:                     Smooth using a Gaussian filter
13:                     Compute the conductivity g (Equation 3)
14:                     Discretized the nonlinear diffusion equation using the FED scheme
15:                 **end for**
16:             **end for**
17:         **end sub-stage**
18:     **end stage**

19:     **stage** Locate the keypoints in the scale space
20:         Compute the determinant of the Hessian matrix
21:         Detect keypoints by searching for points that are the maxima of their neighbourhood $\rightarrow \mathbf{K}_1^b, \mathbf{K}_2^b$
22:         Refine the position and the scale of each keypoint
23:     **end stage**

Keypoint description
24:     **for each** keypoint **in** $\mathbf{K}_1^b$ and $\mathbf{K}_2^b$ **do**
25:         Calculate the main orientation
26:         Compute the M–SURF descriptor
27:         Append the spectral signature
28:     **end for**
29: **end for**

Figure 7. Pseudocode for the keypoint detection and description stages of the HSI–KAZE.

3.3. Keypoint Matching Using Spectral Information and Band Combination

As mentioned in the previous section, HSI–KAZE uses a descriptor is made up of a spatial part and a spectral one. The M–SURF descriptor is used for the spatial part while the spectral signature of the point forms the spectral one. The matching process is based on the calculation of the Euclidean distances for the spatial part and the cosine similarity for the spectral one.
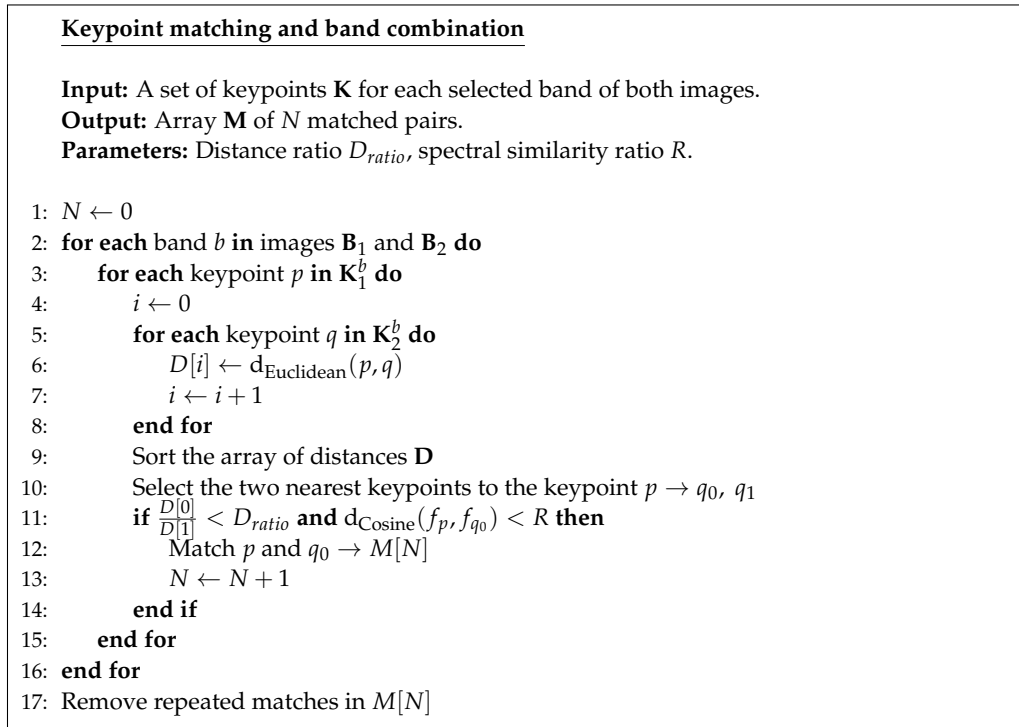
The pseudocode of these two stages can be shown in Figure 8. For each band, we have two sets of keypoints, one for each image. In this stage, the keypoints of each band are independently matched. As proposed in [10], a 2–nearest neighbour search using the Euclidean distance is used (Figure 8, lines 3–10). Two conditions must be fulfilled in order to consider the matching of a pair of keypoints (Figure 8, lines 11–14). First, the ratio of the Euclidean distances between keypoint p of the reference image and the two nearest keypoints q and r of the target image must be smaller than a distance ratio D_{ratio},

$$\frac{\mathrm{d}_{\text{Euclidean}}(p,q)}{\mathrm{d}_{\text{Euclidean}}(p,r)} < D_{ratio},\tag{9}$$

where the Euclidean distance is calculated using the M–SURF descriptor of each keypoint and D_{ratio} was fixed to 0.6 after experimental analyses. It allows us to discard false matches. Second, the cosine similarity between the spectral signature of keypoints p and q, f_p and f_q respectively, must be lower than a value R,

$$\mathrm{d}_{\mathrm{Cosine}}(f_p, f_q) < R, \tag{10}$$

where R is fixed to 0.9 when the images are taken by the same sensor, and to 0.8 when the sensor is different. Both values were chosen experimentally. The use of the spectral information allows us to refine the keypoint selection discarding false matches.

Keypoint matching and band combination

Input: A set of keypoints **K** for each selected band of both images.
Output: Array **M** of N matched pairs.
Parameters: Distance ratio D_{ratio}, spectral similarity ratio R.

```
 1: N ← 0
 2: for each band b in images B₁ and B₂ do
 3:     for each keypoint p in K₁ᵇ do
 4:         i ← 0
 5:         for each keypoint q in K₂ᵇ do
 6:             D[i] ← d_Euclidean(p, q)
 7:             i ← i + 1
 8:         end for
 9:         Sort the array of distances D
10:         Select the two nearest keypoints to the keypoint p → q₀, q₁
11:         if D[0]/D[1] < D_ratio and d_Cosine(f_p, f_q₀) < R then
12:             Match p and q₀ → M[N]
13:             N ← N + 1
14:         end if
15:     end for
16: end for
17: Remove repeated matches in M[N]
```

Figure 8. Pseudocode for the keypoint matching and band combination.

Once the pairs of keypoints have been obtained for each band, all the pairs are joined at the end of this stage and. This way, it is possible to detect features that are only present in some bands, as illustrated in Figure 9. Finally, the repeated matches, which are the result of detecting the same keypoints in the different scale space levels or bands, are removed.

3.4. Registration

An exhaustive search is used to recover the scale, the rotation angle, and the translation parameters that solve the registration problem. The method is detailed in the pseudocode shown in Figure 10. The input is the set of matched pairs of keypoints. First, for each combination of two matched pairs (4 keypoints), a scale factor, an rotation angle, and translation parameters are calculated (Figure 10, lines 1–7). Next, a histogram is created using the angle values with a bin size of 5° and 2.5° overlap between bins (Figure 10, line 8). The elements of the maximum bin are sorted by their scale factor in order to obtain the median (Figure 10, lines 9–10). The registration parameters of the median are finally selected to register the images (Figure 10, line 11).
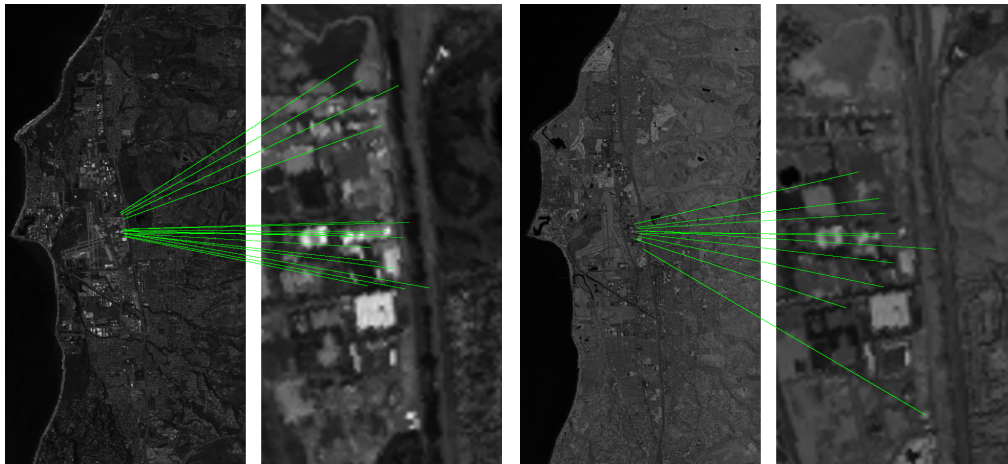
Figure 9. Matched keypoints detected in two pairs of bands belonging to the Santa Barbara Front scene (bands 26 and 93 with scale 9.5×).

Registration stage

Input: Array **M** of N matched pairs.
Output: Registration parameters (scale factor ρ, rotation angle θ, and translation (x, y)).

```
 1: n ← 0
 2: for i ← 1, N do
 3:     for j ← i + 1, N do
 4:         Calculate scale factor ρ_n, rotation angle θ_n, and translation (x_n, y_n)
            for the pairs M[i] and M[j]
 5:         n ← n + 1
 6:     end for
 7: end for

 8: Compute the histogram of the set of n angles
 9: Find the maximum bin of the histogram
10: Sort the elements of the maximum bin by their scale value
11: Recover the registration parameters from the median of the sorted elements → ρ, θ, (x, y)
```

Figure 10. Pseudocode for the registration stage of the HSI–KAZE.

4. Results and Discussion

This section presents some experimental results obtained by the HSI–KAZE algorithm using different hyperspectral remote sensing images. First, the test images and the experimental conditions are described in Section 4.1. Then, different feature reduction methods are evaluated in Section 4.2. In Section 4.3, HSI–KAZE is compared with the original KAZE and with other methods of the literature in terms of number of successful registered cases. Finally, in Section 4.4, a comparison in terms of different common measures of the literature is presented and discussed. Finally, the limitations of the method and future work are introduced.

4.1. Test Images and Experimental Conditions

Seven hyperspectral images are used to evaluate the method proposed in this article. Table 1 provides detailed information for each image: sensor, size, number of spectral bands, spatial resolution, and location. A false–colour composite of each image is displayed in Figures 11 and 12.

Table 1. Sensor, size, number of spectral bands, resolution (m/pixel), and location of the test hyperspectral images.

Image	Sensor	Size	Bands	Spatial Resolution	Location
Pavia University	ROSIS–03	610×340	103	1.3	N 45°12′13.176″, E 9°8′11.108″
Pavia Centre	ROSIS–03	1096×715	102	1.3	N 45°11′07.4″, E 9°8′34.8″
Indian Pines	AVIRIS	145×145	220	20	Refer to [36]
Salinas Valley	AVIRIS	512×217	204	3.7	Refer to [36]
Jasper Ridge 2006	AVIRIS	1286×588	224	3.3	N 37°24′8.5″, W 122°14′42.2″
Jasper Ridge 2007	AVIRIS	1286×588	224	3.4	N 37°24′8.5″, W 122°14′42.2″
Santa Barbara Box 2013	AVIRIS	1024×769	224	15.2	N 35°18′14.3″, W 118°50′22.6″
Santa Barbara Box 2014	AVIRIS	1024×769	224	15.2	N 35°18′14.3″, W 118°50′22.6″
Santa Barbara Front 2009	AVIRIS	900×470	224	16.4	N 34°26′24.3″, W 119°50′40.0″
Santa Barbara Front 2010	AVIRIS	900×470	224	11.3	N 34°26′24.3″, W 119°50′40.0″

(**a**)　　　　　　　　　　　　　　　(**b**)

(**c**)　　　　　　　　　　　　　　　(**d**)
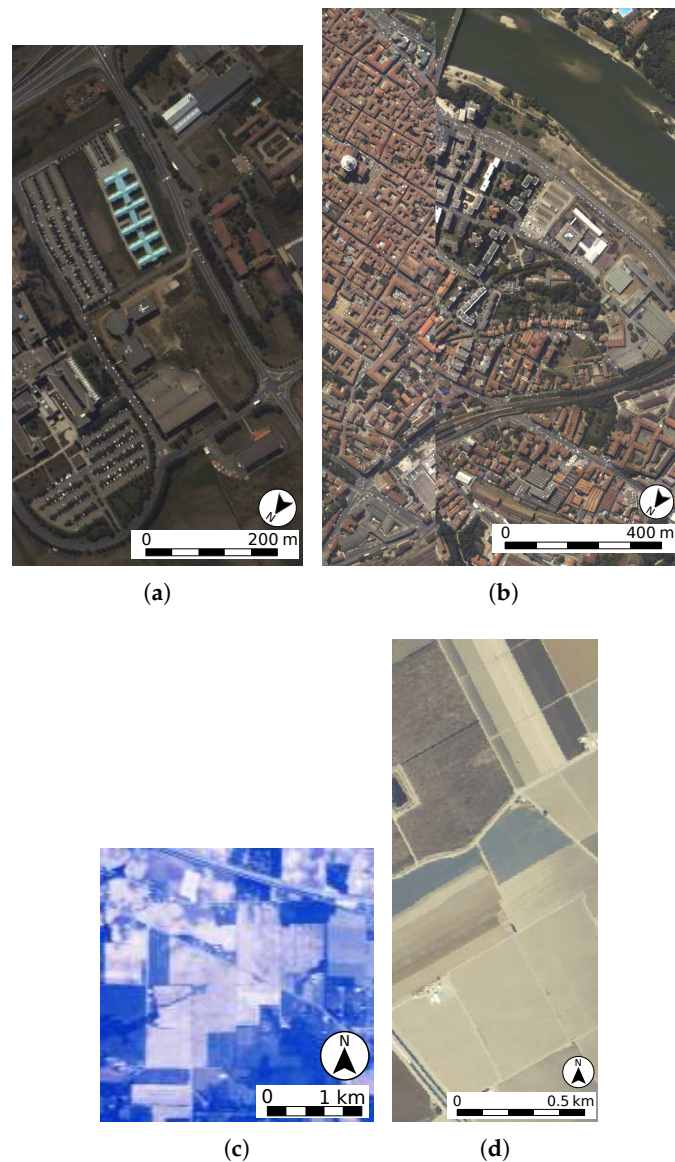
Figure 11. Hyperspectral images commonly used for testing in remote sensing: (**a**) Pavia University, (**b**) Pavia Centre, (**c**) Indian Pines, and (**d**) Salinas.
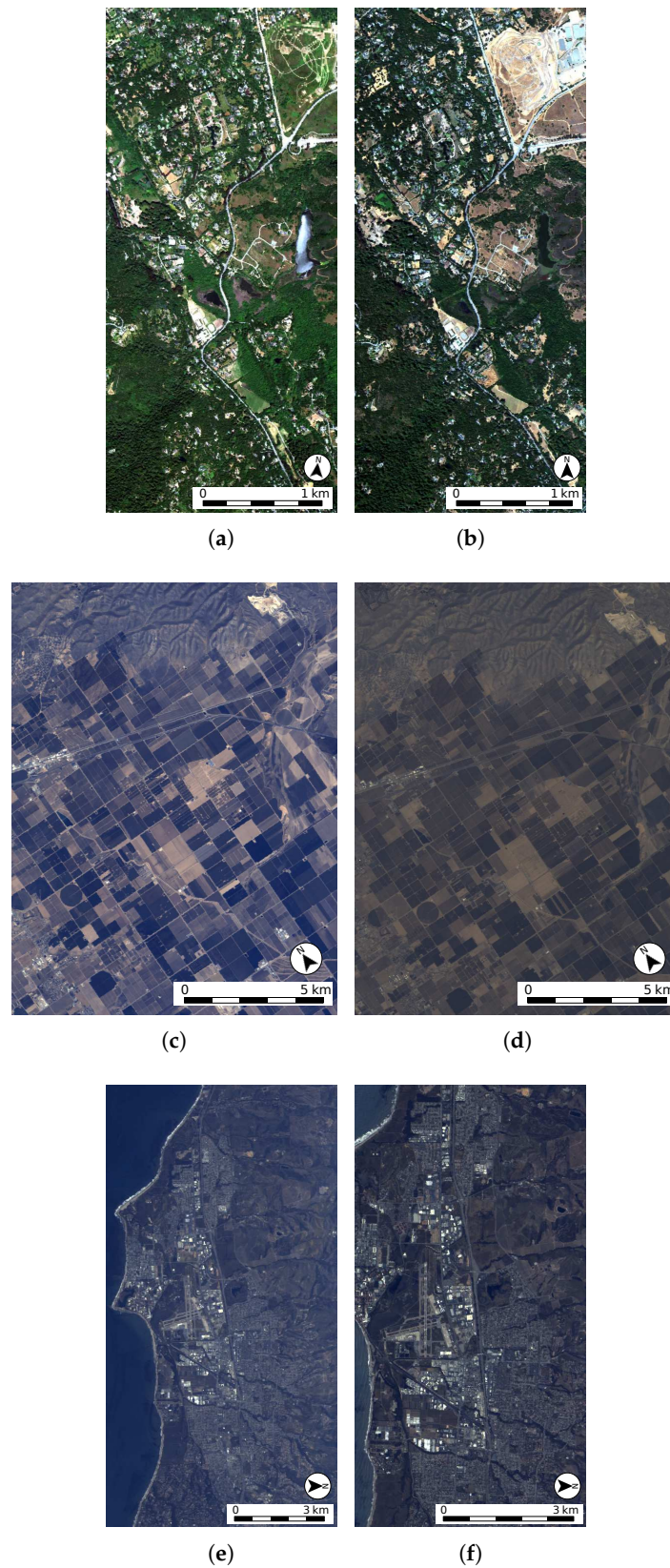
Figure 12. Second group of the test hyperspectral images: (**a**) Reference Jasper Ridge image taken on 5 December 2006, (**b**) Target Jasper Ridge image taken on 13 August 2007, (**c**) Reference Santa Barbara Box image taken on 11 April 2013, (**d**) Target Santa Barbara Box image taken on 16 April 2014, (**e**) Reference Santa Barbara Front image taken on 30 March 2009, and (**f**) Target Santa Barbara Front image taken on 30 April 2010.

The first four images (see Figure 11) are commonly used scenes for testing in the area of hyperspectral remote sensing [37]. The Pavia images were taken by the ROSIS–03 (Reflective Optics System Imaging Spectrometer) sensor while Indian Pines and Salinas Valley scenes were taken by the AVIRIS (Airbone Visible/Infrared Imaging Spectometrer) sensor. The set of target images is created by applying different scale factors and rotation angles to the original image. This way, we can investigate all the details in a controlled environment.

The other six images (see Figure 12) are pairs of scenes taken by the AVIRIS sensor at different dates [38]. Each pair of images presents changes in infrastructures, in buildings, vegetation, objects, and illumination, among others, as well as different scale factors, orientation and translation. Moreover, in order to increase the experimental range, different scale factors and rotation angles are applied to the target image of each pair.

The complete dataset is available in https://wiki.citius.usc.es/hiperespectral:hsi-kaze.

The experiments were carried out on a PC with a quad–core Intel Core i7–4790 CPU at 3.60 GHz and 24 GB of RAM. The code was compiled using the gcc and the g++ 5.4.0 version under Ubuntu 16.04. Furthermore, OpenCV 3.2.0.8 and Python were used to obtain results of registration effectiveness for the original KAZE and A–KAZE methods.

4.2. Evaluation of Feature Reduction Methods

Initially, we have evaluated HSI–KAZE using the feature selection method proposed in this paper, EBS, and compared with WaluMI [33], another feature selection method, and with both feature extraction methods: PCA [31] and BandClust [32]. All these methods were described in Section 3.1.

PCA transforms a number of possible variables into a new set of uncorrelated variables where the first few retain most of the variation of the original data. A different approach is followed by BandClust which performs an unsupervised clustering and, at the end, the reduced image is the combination of the bands of the same cluster. On the other hand, WaluMI carries out an agglomerative clustering strategy based on Ward's linkage method. The final image is obtained after choosing the most representative band of each cluster. EBS performs a band selection according to the entropy between the bands of both images and by complying with a minimum inter–band distance.

As explained in Section 4.1, we carry out an exhaustive set of tests by applying different scales and rotations to the target image. These exhaustive tests allow evaluating the method in the most unfavourable situations. The scale ranges go from $1/16\times$ to $25.5\times$ in increments of $0.5\times$ (65 scale factors). In all cases, the target images are trimmed on the central region to keep the same size as the original images, as shown in Figure 1b. For each scale factor, 72 rotation angles are applied, from $0°$ to $360°$ in increments of $5°$. Therefore, we have carried out an exhaustive evaluation of 4680 cases for each scene. For all the methods, 8 bands have been selected in order to make a fair comparison among proposals, with the exception of BandClust for which the band limits of each cluster for our datasets are shown in Table 2. In the PCA method the first 8 principal components were selected, while for WaLuMI the agglomerative clustering is set to stop in 8 clusters.

Table 2. Band limits of each cluster produced by BandClust for all the scenes.

Scene	Band Limits of Each Cluster
Pavia University	1–64, 65–103
Pavia Centre	1–23, 24–60, 61–102
Indian Pines	1–38, 39–60, 61–65, 66–79, 80–102, 103–116, 117–149, 150–160, 161–220
Salinas	1–11, 12–23, 24–33, 34–44, 45–61, 62–68, 69–78, 79–83, 84–104, 105–109, 110–121, 122–150, 151–155, 156–204
Jasper Ridge	1–14, 15–37, 38–88, 89–107, 108–111, 112–146, 147–154, 155–162, 163–175, 176–204, 205–209, 210–214, 215–224
Santa Barbara Front	1–15, 16–35, 36–63, 64–87, 88–114, 115–120, 121–174, 175–224
Santa Barbara Box	1–6, 7–36, 37–44, 45–65, 66–71, 72–85, 86–99, 100–125, 126–134, 135–138, 139–152, 153–157, 158–160, 161–165, 166–224

Table 3 summarizes the cases that were correctly registered for each scene using HSI–KAZE considering different feature reduction methods. The best results are achieved using EBS, 30.71 cases are correctly registered on average. WaLuMI registers 28.43 cases versus 25.71 achieved by BandClust.

PCA provides worse results because a different transformation is applied to the reference and the target images, leading to a smaller number of common keypoints present in both images, or even none, being recovered.

Table 3. Successfully registered cases for each scene using HSI–KAZE with different feature reduction methods. The number in parentheses summarizes the number of scales that were correctly registered for all angles. If an angle is incorrectly registered, the whole scale factor is considered incorrect, i.e., this case is not included in the table. The percentage values are calculated over 65, the number of scales considered.

Scene	PCA	BandClust	WaLuMI	EBS
Pavia University	$1/4\times$ to $4.5\times$ (26)	$1/9\times$ to $12.5\times$ (32)	$1/12\times$ to $13.0\times$ (36)	$1/11\times$ to $13.0\times$ (35)
Pavia Centre	$1/15\times$ to $18.5\times$ (50)	$1.0\times$ to $22.5\times$ (44)	$1/14\times$ to $25.5\times$ (63)	$1/16\times$ to $24.0\times$ (62)
Indian Pines	$1.5\times$ to $3.0\times$ (4)	$1/4\times$ to $5.5\times$ (13)	$1/3\times$ to $5.0\times$ (11)	$1/4\times$ to $5.5\times$ (13)
Salinas	$1/4\times$ $4.5\times$ (11)	$1/7\times$ to $6.5\times$ (18)	$1/7\times$ to $6.5\times$ (18)	$1/7\times$ to $6.0\times$ (17)
Jasper Ridge	$1/4\times$ $4.0\times$ (10)	$1/9\times$ to $11.0\times$ (29)	$1/9\times$ to $8.5\times$ (24)	$1/12\times$ to $12.5\times$ (35)
Santa Barbara Front	$1/4\times$ $3.0\times$ (8)	$1/8\times$ to $9.0\times$ (24)	$1/8\times$ to $8.5\times$ (23)	$1/9\times$ to $9.5\times$ (26)
Santa Barbara Box	$1/2\times$ $4.0\times$ (8)	$1/5\times$ to $8.5\times$ (20)	$1/9\times$ to $8.5\times$ (24)	$1/12\times$ to $8.5\times$ (27)
Number of scalings (average)	(16.71)	(25.71)	(28.43)	(30.71)
Number of scalings (percentage)	25.71%	39.56%	43.74%	47.25%

4.3. Registration Effectiveness

In this section, the evaluation of the HSI–KAZE algorithm is presented. Our proposal is compared with other methods in the literature. In particular, it is compared with two methods based on the Fourier transform, the Fourier–Mellin invariant symmetric phase–only matched filtering (FMI–SPOMF) [39] and the HYperspectral Fourier–Mellin algorithm (HYFM) [9], and on the other hand, with the original KAZE [12] and A–KAZE methods [17].

FMI–SPOMF uses the phase correlation and the log–polar grid to perform a translation, rotation, and scale-invariant grey–level image registration. However, HYFM was particularly designed for registering hyperspectral images. It exploits the information contained in different bands and is based on principal component analysis, multilayer fractional Fourier transform (MLFFT), combination of log-polar maps, and peak processing.

As explained in Section 2, KAZE is a feature–based method which uses the M–SURF descriptor and a nonlinear diffusion filter instead of Gaussian filters, as in SIFT, in order to build the scale space preserving sharp edges and smoothing noise. In contrast, A–KAZE uses a binary descriptor and builds a pyramidal scale space using a FED scheme which allows faster computation and higher accuracy. The binary descriptor also allows a faster computation because the matching stage is computed using the Hamming distance instead of the Euclidean distance.

The evaluation procedure is the same as the one introduced in the previous section. An exhaustive scaling range from $1/16\times$ to $25.5\times$, with 72 angles per scale, is applied to the target image. For KAZE and A–KAZE, the band of highest entropy is selected to perform the registration. On the other hand, FMI–SPOMF uses the first PCA component to perform the registration, while HYFM uses the first 8 PCA components [9]. The random sample consensus (RANSAC) algorithm was used in KAZE and A–KAZE in order to compute the registration parameters from the keypoints and is available in the OpenCV Library.

Table 4 summarizes the cases that were correctly registered for each scene and algorithm. The proposed method HSI–KAZE provides the best results on average, specifically, 30.71 cases as compared with 11.43 cases achieved by HYFM. FMI–SPOMF, KAZE and A–KAZE do not recover a high number of cases because they were developed to work with grey–level images and not to deal with spectral information. The bests results are achieved for the Pavia Centre and Jasper Ridge scenes in which a range of $1/16\times$ to $24.0\times$ and $1/12\times$ to $12.5\times$ are successfully registered, respectively.

Table 4. Successfully registered cases for each scene. The number in parentheses summarizes the number of scales that were correctly registered for all angles. If an angle is incorrectly registered, the whole scale factor is considered incorrect, i.e., this case is not included in the table. KAZE and A–KAZE are applied to the band with the highest entropy. KAZE and A–KAZE use RANSAC. The percentage values are calculated over 65, the number of scales considered.

Scene	FMI–SPOMF	HYFM	KAZE (RANSAC)	A–KAZE (RANSAC)	HSI–KAZE
Pavia University	$1/5\times$ to $4.5\times$ (12)	$1/4\times$ to $5.5\times$ (13)	$1/3\times$ to $2.5\times$ (6)	$1/3\times$ to $2.0\times$ (5)	$1/11\times$ to $13.0\times$ (35)
Pavia Centre	$1/6\times$ to $6.0\times$ (16)	$1/5\times$ to $7.5\times$ (18)	$1/4\times$ to $5.0\times$ (12)	$1/4\times$ to $4.0\times$ (10)	$1/16\times$ to $24.0\times$ (62)
Indian Pines	$1/2\times$ to $3.0\times$ (6)	$1/2\times$ to $4.0\times$ (8)	$1.0\times$ to $2.0\times$ (3)	$1.0\times$ (1)	$1/4\times$ to $5.5\times$ (13)
Salinas	$1/2\times$ to $4.0\times$ (8)	$1/2\times$ to $4.5\times$ (9)	$1/2\times$ to $2.0\times$ (4)	$1/2\times$ to $2.0\times$ (4)	$1/7\times$ to $6.0\times$ (17)
Jasper Ridge	$1/3\times$ to $2.5\times$ (6)	$1/5\times$ to $3.0\times$ (9)	(0)	(0)	$1/12\times$ to $12.5\times$ (35)
Santa Barbara Front	$1/4\times$ to $2.5\times$ (7)	$1/4\times$ to $3.5\times$ (9)	$1.0\times$ to $2.0\times$ (3)	$1/2\times$ to $2.0\times$ (4)	$1/9\times$ to $9.5\times$ (26)
Santa Barbara Box	$1/2\times$ to $2.0\times$ (4)	$1/4\times$ to $6.0\times$ (14)	$1.0\times$ to $1.5\times$ (2)	$1/2\times$ to $1.5\times$ (3)	$1/12\times$ to $8.5\times$ (27)
Number of scalings (average)	(8.43)	(11.43)	(4.29)	(3.86)	(30.71)
Number of scalings (percentage)	12.97%	17.58%	6.59%	5.93%	47.25%

To make a deeper study, we have analysed the performance of KAZE and A–KAZE replacing the RANSAC method based on the random sampling theory by the registration method proposed in HSI–KAZE (see Section 3.4). Table 5 shows the correct registered cases for each scene and method. The average number of successful registered cases has been doubled for KAZE and A–KAZE for all scenes, specifically, 9.00 and 8.29 cases have correctly been registered versus 4.29 and 3.86 using the RANSAC version, respectively (see Table 4). In particular, for Pavia Centre, ranges of $1/7\times$ to $9.0\times$ and $1/7\times$ to $8.5\times$ have been recovered for KAZE and A–KAZE, respectively. Moreover, KAZE and A–KAZE outperform HYFM on Pavia Centre and Pavia University scenes but not in the other three scenes taken at different dates. Using our exhaustive registration method, KAZE and A–KAZE achieve registering 1 and 2 cases for Jasper Ridge images (Table 5), respectively, instead of 0 where RANSAC is used (Table 4). Even replacing RANSAC by the proposed registration method in these methods, HSI–KAZE achieves the best results thanks to exploitation of the spectral information.

Table 5. Successfully registered cases for each scene. The number in parentheses summarizes the number of scales that were correctly registered for all angles. If an angle is incorrectly registered, the whole scale factor is considered incorrect, i.e., this case is not included in the table. KAZE and A–KAZE are applied to the band with the highest entropy. In this case, KAZE and A–KAZE use the registration method proposed in HSI–KAZE. The percentage values are calculated over 65, the number of scales considered.

Scene	FMI–SPOMF	HYFM	KAZE	A–KAZE	HSI–KAZE
Pavia University	$1/5\times$ to $4.5\times$ (12)	$1/4\times$ to $5.5\times$ (13)	$1/5\times$ to $6.0\times$ (15)	$1/4\times$ to $4.5\times$ (11)	$1/11\times$ to $13.0\times$ (35)
Pavia Centre	$1/6\times$ to $6.0\times$ (16)	$1/5\times$ to $7.5\times$ (18)	$1/7\times$ to $9.0\times$ (23)	$1/7\times$ to $8.5\times$ (22)	$1/16\times$ to $24.0\times$ (62)
Indian Pines	$1/2\times$ to $3.0\times$ (6)	$1/2\times$ to $4.0\times$ (8)	$1/2\times$ to $2.5\times$ (5)	$1.0\times$ to $1.5\times$ (2)	$1/4\times$ to $5.5\times$ (13)
Salinas	$1/2\times$ to $4.0\times$ (8)	$1/2\times$ to $4.5\times$ (9)	$1/4\times$ to $3.5\times$ (9)	$1/3\times$ to $3.0\times$ (7)	$1/7\times$ to $6.0\times$ (17)
Jasper Ridge	$1/3\times$ to $2.5\times$ (6)	$1/5\times$ to $3.0\times$ (9)	$1.0\times$ (1)	$1.0\times$ to $1.5\times$ (2)	$1/12\times$ to $12.5\times$ (35)
Santa Barbara Front	$1/4\times$ to $2.5\times$ (7)	$1/4\times$ to $3.5\times$ (9)	$1/2\times$ to $2.5\times$ (5)	$1/4\times$ to $3.0\times$ (8)	$1/9\times$ to $9.5\times$ (26)
Santa Barbara Box	$1/2\times$ to $2.0\times$ (4)	$1/4\times$ to $6.0\times$ (14)	$1/2\times$ to $2.5\times$ (5)	$1/4\times$ to $2.0\times$ (6)	$1/12\times$ to $8.5\times$ (27)
Number of scalings (average)	(8.43)	(11.43)	(9.00)	(8.29)	(30.71)
Number of scalings (percentage)	12.97%	17.58%	13.85%	12.75%	47.25%

The total number of successfully registered cases considering both the scale factors and angles of rotation (from the all 65×72 test cases) is shown in Table 6. HSI–KAZE presents the best results, it correctly registers $17,153$ cases, i.e., 52.36% of the test cases. These results are very similar to those already presented in Table 5 in which only the scales in which the 72 angles have been successfully registered are included.

Table 6. Number and percentage of successfully registered cases for each scene and method for the entire test range (from $1/16\times$ to $25.5\times$, 65 scale factors and 72 angles per scale).

Scene	FMI–SPOMF	HYFM	KAZE	A–KAZE	HSI–KAZE
Pavia University	999 (21.35%)	1139 (24.34%)	1446 (30.90%)	856 (18.29%)	2781 (59.42%)
Pavia Centre	1268 (27.12%)	1472 (31.45%)	1859 (39.72%)	1829 (39.08%)	4647 (99.29%)
Indian Pines	586 (12.52%)	739 (15.79%)	393 (8.40%)	152 (3.25%)	954 (20.38%)
Salinas	876 (18.72%)	856 (18.29%)	907 (19.38%)	625 (13.35%)	1312 (28.03%)
Jasper Ridge	663 (14.17%)	697 (14.89%)	148 (3.16%)	352 (7.52%)	2920 (62.39%)
Santa Barbara Front	733 (15.66%)	859 (18.35%)	873 (18.65%)	799 (17.07%)	2404 (51.37%)
Santa Barbara Box	508 (10.58%)	948 (20.26%)	715 (15.28%)	720 (15.38%)	2135 (45.62%)
Total successful cases	5634 (17.20%)	6710 (20.48%)	6341 (19.36%)	5333 (16.28%)	17,153 (52.36%)

4.4. Measures of Registration Accuracy

In general, a small scale factor range is analysed in the literature, even in many cases, only one scale. Table 7 summarizes the scale factor range considered in the recent literature. It includes methods for registering both panchromatic images as well as multispectral (from up to 31 bands) and hyperspectral (from up to 242 bands) images. In our study, the considered scale factor range goes from $1/16\times$ to $25.5\times$ (65 scale factors) where for each scale 72 angles are evaluated, as shown in Tables 3–5. Table 7 also summarizes the most common measures used to evaluate registration methods in the literature: chequerboard, number of matches, number of correct matches, correct match rate, root-mean-square error (RMSE) and registration error. For that reason, in this section, a evaluation of the feature–based methods KAZE, A–KAZE and HSI–KAZE is presented according to these measures.

Table 7. The most common measures used to evaluate the effectiveness, accuracy and robustness of registration methods in the literature as well as the image type and the scale factor range considered.

Ref.	Image Type	Scale Factor Range	Measures
[18]	Multispectral	$0.7\times$ to $1.0\times$	Chequerboard, correct match rate, registration error
[19]	Multispectral	$0.5\times$ to $2.0\times$	Chequerboard, correct match rate, RMSE
[20]	Multispectral	$1.0\times$ to $2.0\times$	Correct match rate, number of matches
[21]	Multispectral	$1.0\times$ to $2.0\times$	Correct match rate, number of correct matches, number of matches
[22]	Multispectral	$1.0\times$	Number of correct matches
[40]	Multispectral and hyperspectral	$0.7\times$	RMSE
[41]	Hyperspectral	$1.0\times$	Registration error
[42]	Multispectral and hyperspectral	$1.0\times$ to $2.0\times$	Chequerboard, number of correct matches, RMSE
[24]	Hyperspectral	$8.0\times$	Number of keypoints, correct match rate
[43]	Multispectral	$1.0\times$	Chequerboard, correct match rate, number of correct matches, RMSE
[44]	Multispectral and hyperspectral	$2.0\times$	Correct match rate, number of correct matches, RMSE
[45]	Multispectral and panchromatic	$2.8\times$ to $4.0\times$	RMSE
[46]	Multispectral	$1.0\times$	Number of matches, RMSE
[47]	Multispectral	$6.0\times$	Chequerboard, number of correct matches, RMSE
[48]	Multispectral	$0.5\times$ to $1.5\times$	Registration error, chequerboard

The experimental results shown in this section correspond to the pairs of images of the second dataset described in Section 4.1, but in this case we only consider the original scale factor, the angle of rotation and the translation, i.e., no additional transformations will be apply to the images. The reference registration parameters for each scene are shown in Table 8.

Table 8. Reference registration parameters for the second group of the test hyperspectral images.

Scene	Scale Factor	Rotation Angle (Degrees)	Translation (x,y) (Pixels)
Jasper	0.97	-6.05	$(-12, 24)$
Santa Barbara Front	1.45	3.52	$(2, 6)$
Santa Barbara Box	1.00	0.00	$(33, 40)$

Figure 13 shows the chequerboard registered images for the Jasper Ridge and Santa Barbara scenes using HSI–KAZE. It can be seen that the borders and objects of the two images are correctly overlapped even when they were obtained in different dates.

(**a**)   (**b**)   (**c**)
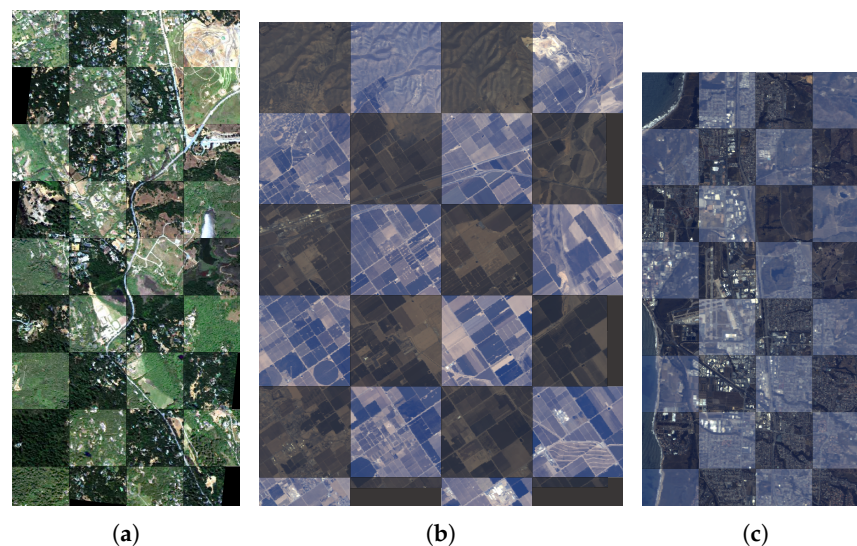
Figure 13. Chequerboard registered images for the scenes taken by the AVIRIS sensor at different dates: (**a**) Jasper Ridge, (**b**) Santa Barbara Box, and (**c**) Santa Barbara Front.

Although the number of keypoint matches found by a method does not influence the quality of the registration, it is interesting to compare the values obtained by the different methods in order to better explain the registration details. Table 9 compares the number of keypoint matches for each scene using the methods KAZE, A–KAZE and HSI–KAZE. It is important to note that HSI–KAZE discards false matches thanks to the use of the spectral signature as part of the keypoint descriptor (see Section 3.3), as can be seen in the second row for each scene. In the third row, it can be seen that the elimination of repeated matches performed by HSI–KAZE leads to even a higher reduction of the final number of used matches. To calculate the number of correct matches shown in this section, the Euclidean distance between the keypoint of the reference image and the keypoint of the target image after applying the original transformation is used as error measure. A match is considered incorrect if the error is higher than 2 pixels.

More detailed results of the registration process are shown in Table 10 in which the number of matches actually used to register the images is presented. For HSI–KAZE, these matches have been obtained using the exhaustive research method explained in Section 3.4 which calculates the registration parameters taken into account all the possible combinations between the matched keypoints. Then, the histogram of the obtained angles of rotation for all these matches is computed and the registration parameters of the bin with the maximum value are selected. The final registration parameters are obtained from the median of the scales of the parameters of this bin. It is important to note that the matches used for registering the images are entirely correct in the case of HSI–KAZE as can be seen in the third row of the table, i.e., the correct match ratio is 100%. This correct match ratio stands at around 83% on average in the case of KAZE and A–KAZE with RANSAC used to filter incorrect matches.

The RMSE and the registration error are measures frequently used in the literature to evaluate registration algorithms as seen in Table 7. The registration error is measured in pixels and is computed as the average Euclidean distance between the keypoints used in the registration of the reference image and the keypoints used in the registration of the target image after applying the reference transformation (see Table 8). The results obtained for KAZE, A–KAZE and HSI–KAZE are very similar

as shown in Table 11. However, it is important to point out that HSI–KAZE allows registering images for a range of scales much higher than that the obtained by KAZE and A–KAZE, as shown in Section 4.3 in Tables 4 and 5.

Table 9. Comparisons of KAZE, A–KAZE and the proposed method HSI–KAZE regarding the original number of matches obtained for each scene.

		KAZE	A–KAZE	HSI–KAZE
Jasper	Number of matches	17	23	21,437
	Number of matches after spectral discarding	-	-	21,105
	Number of matches after removing repeated matches	-	-	20,860
	Number of correct matches	13	18	20,535
Santa Barbara Front	Number of matches	207	328	44,550
	Number of matches after spectral discarding	-	-	44,490
	Number of matches after removing repeated matches	-	-	43,537
	Number of correct matches	176	282	20,741
Santa Barbara Box	Number of matches	307	230	37,185
	Number of matches after spectral discarding	-	-	37,001
	Number of matches after removing repeated matches	-	-	36,314
	Number of correct matches	234	171	18,502

Table 10. Comparisons of KAZE, A–KAZE and the proposed method HSI–KAZE regarding the matches used to register the images.

		KAZE	A–KAZE	HSI–KAZE
Jasper	Number of matches used in registration	17	22	4
	Number of correct matches used in registration	13	18	4
	Correct match ratio	0.76	0.82	1.00
Santa Barbara Front	Number of matches used in registration	185	262	4
	Number of correct matches used in registration	165	234	4
	Correct match ratio	0.89	0.89	1.00
Santa Barbara Box	Number of matches used in registration	241	221	4
	Number of correct matches used in registration	204	170	4
	Correct match ratio	0.85	0.77	1.00

**Table 11.** Results in terms of RMSE and registration error for KAZE, A–KAZE and HSI–KAZE.

		KAZE	A–KAZE	HSI–KAZE
Jasper	RMSE	0.98	1.54	1.35
	Registration error (pixels)	0.92	1.38	1.23
Santa Barbara Front	RMSE	1.33	1.37	1.44
	Registration error (pixels)	1.16	1.18	1.42
Santa Barbara Box	RMSE	1.48	1.68	0.72
	Registration error (pixels)	1.28	1.44	0.70

Figures 14 and 15 show all the matches obtained for the eight selected bands by HSI–KAZE of Santa Barbara Front and Jasper scenes. The matches which were discarded by the spectral information are in brown colour, the correct matches used in registration are green, the remainder of correct matches are yellow, and the incorrect matches are blue. In these figures it is easy to note that the use of different pair of bands allows detecting different keypoints due to the different information contained in each one. The final matches used to register these scenes are obtained from different bands, bands 126 and 146 for Santa Barbara Front, and bands 21, 135 and 207 for Jasper Ridge. Moreover, a high number of false matches are discarded thanks to the use of the spectral signature, as can be seen, for example, in Figure 14f.
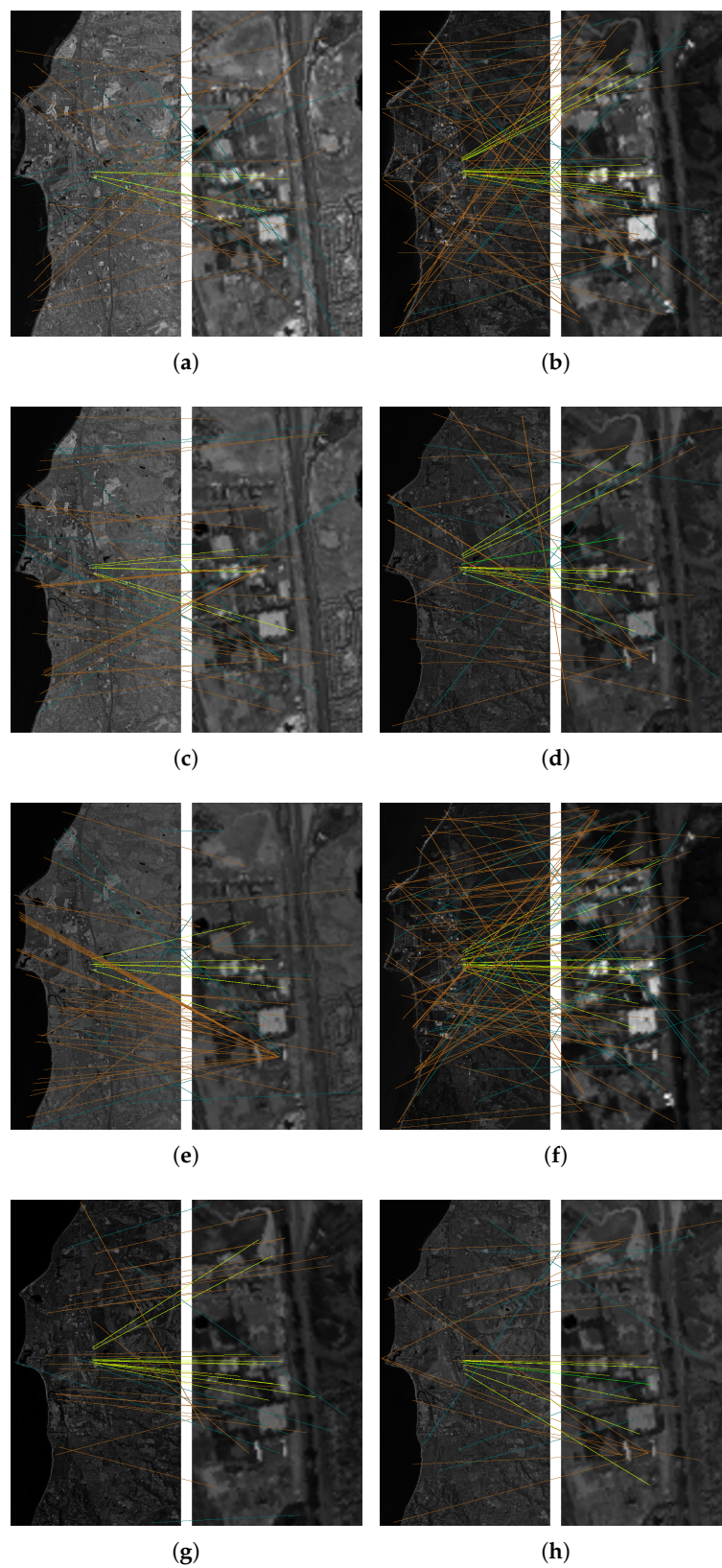
Figure 14. Matched keypoints detected in the eight selected bands ordered by decreasing entropy from Santa Barbara Front scene with scale 9.5×: (**a**) Band 46, (**b**) Band 26, (**c**) Band 71, (**d**) Band 126, (**e**) Band 93, (**f**) Band 6, (**g**) Band 185, and (**h**) Band 146. Matches discarded after considering spectral information (brown), incorrect matches (blue), correct matches (yellow), and correct matches used in registration (green).
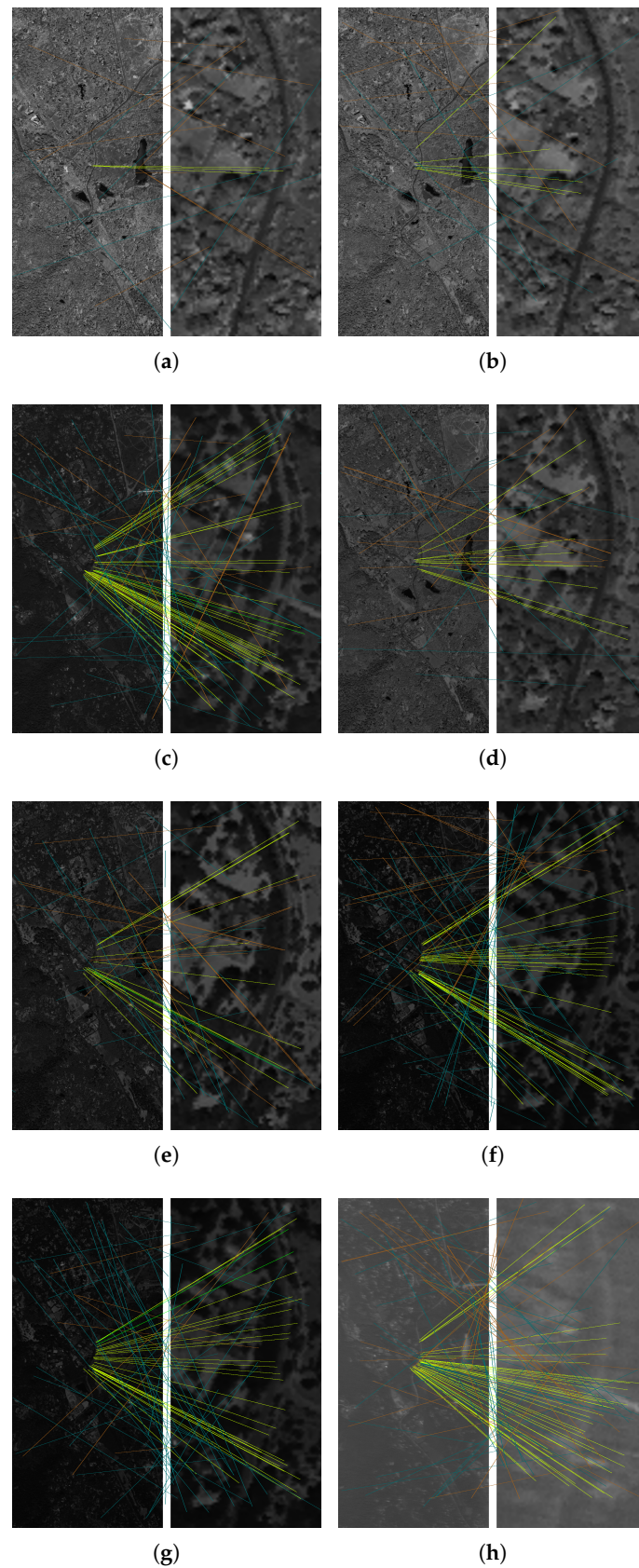
Figure 15. Matched keypoints detected in the eight selected bands ordered by decreasing entropy from Jasper Ridge scene with scale 9.0×: (**a**) Band 46, (**b**) Band 71, (**c**) Band 21, (**d**) Band 94, (**e**) Band 135, (**f**) Band 187, (**g**) Band 207, and (**h**) Band 1. Matches discarded after considering spectral information (brown), incorrect matches (blue), correct matches (yellow), and correct matches used in registration (green).

4.5. Limitations of the Method

HSI–KAZE has an upper limit in the scales that can register successfully. Nevertheless, this limit is higher than the previously proposed methods (see Table 7) thanks to the selection of different bands, the use of spectral information in the descriptor, and the proposed exhaustive registration method. An example of the extreme range of scales studied is shown in Figure 1. The higher scale factor, the harder registration due to the lack of overlapping between the images.

Table 12 shows the computational cost (execution times and memory requirements experimentally obtained) in CPU for each scene and method considering the last scale successfully registered in Table 5. Regarding computational time, HSI–KAZE is more costly than the other methods. The reason is the exploitation of the spectral information and the exhaustive search to recover the registration parameters considering all the possible pairs of matches that involves higher computational costs. This cost is reduced by discarding false and repeated keypoint matches, as it was explained in Section 3.3. It is important to note that FMI–SPOMF, KAZE and A–KAZE were developed to register greyscale or RGB images and not hyperspectral ones, so they perform the registration by using only a pair of bands of the images. HYFM uses different pairs of bands as HSI-KAZE, so their execution times are similar. These computation times can be additionally decreased by projecting HSI–KAZE onto HPC infrastructures or onto commodity devices such as general purpose GPU (GPGPU).

Regarding the memory requirements, HYFM needs an average of 233.49 MiB. The highest peak of memory is achieved during the keypoint detection and description stages due to the building of scale space, and during the keypoint matching stage where all the possible pairs of matches are considered. KAZE is the method that requires the highest amount of memory. On the other hand, FMI–SPOMF and HYFM present similar memory requirements due to the highest memory requirement is in the computation of the principal components (PCA). It is a common stage in both methods for which the whole hyperspectral image must be in memory.

Table 12. Execution times (in seconds) and memory requirements (in MiB) considering the last scale successfully registered in Table 5 for each scene and method in CPU.

Scene	FMI–SPOMF	HYFM	KAZE	A–KAZE	HSI–KAZE	Average
Pavia University	30.64 s 30.70 MiB	123.27 s 37.95 MiB	1.53 s 238.26 MiB	1.11 s 169.49 MiB	67.38 s 121.40 MiB	44.79 s 119.56 MiB
Pavia Centre	136.99 s 114.83 MiB	507.59 s 151.78 MiB	6.56 s 547.93 MiB	6.72 s 276.16 MiB	454.98 s 482.22 MiB	222.57 s 314.58 MiB
Indian Pines	1.65 s 6.83 MiB	6.59 s 6.83 MiB	0.19 s 140.78 MiB	0.10 s 130.05 MiB	4.72 s 19.93 MiB	2.65 s 60.88 MiB
Salinas	6.89 s 32.62 MiB	28.79 s 32.62 MiB	0.94 s 188.17 MiB	0.57 s 149.40 MiB	22.71 s 40.57 MiB	11.98 s 88.67 MiB
Jasper Ridge	140.17 s 242.83 MiB	508.71 s 242.83 MiB	11.04 s 541.02 MiB	12.80 s 272.57 MiB	628.25 s 482.24 MiB	260.19 s 356.30 MiB
Santa Barbara Front	33.66 s 170.50 MiB	120.87 s 184.50 MiB	4.61 s 147.02 MiB	4.50 s 209.09 MiB	226.62 s 163.27 MiB	78.05 s 174.88 MiB
Santa Barbara Box	32.47 s 252.87 MiB	121.93 s 252.87 MiB	10.29 s 556.77 MiB	11.00 s 278.06 MiB	616.88 s 324.79 MiB	158.51 s 333.07 MiB
Average	54.64 s 121.60 MiB	202.54 s 129.91 MiB	5.02 s 337.14 MiB	5.26 s 212.12 MiB	288.79 s 233.49 MiB	111.25 s 206.85 MiB

It is important to note that HSI–KAZE is similar to the other methods in terms of RMSE, as can be seen in Table 11 and discussed in Section 4.4, in spite of the wide range of scales that HSI–KAZE can recovered. The RMSE is only a measure of precision in the cases correctly registered.

Finally, HSI–KAZE could fail in the detailed alignment of the images. It searches for a consensus global affine transformation for the whole image. Due to images usually present different distortions inconsistently, it will be necessary to apply a fine grained registration in a second stage in order to correct these distortions with a higher level of detail and decrease the RMSE and the registration error.

5. Conclusions

In this work, a feature–based method for registering hyperspectral remote sensing images called HSI–KAZE is proposed. It has been designed to register images with large scale factors and any angle of rotation and translation. HSI–KAZE is based on KAZE but improving the registration effectiveness by exploiting the spectral information available in the images. The spectral information is considered when a set of representative bands of the image are selected based on their entropy and spectral distance. Each keypoint descriptor also incorporates spectral information because it consists of a M–SURF descriptor and a spectral signature. A combination stage considers together all the keypoints for the selected bands. HSI–KAZE carries out an exhaustive search for registration using a histogram-based approach where all pairs of matched keypoint are taking into account.

We have performed several experiments using four well–known hyperspectral images in the remote sensing area, and three pairs of hyperspectral images taken by the AVIRIS sensor on different dates. In addition to the changes that the images already present, we have applied additional similarity transformations to extend our study. Moreover, we have evaluated different band selection methods present in the literature regarding their effectiveness in terms of registration robustness.

The proposed method has been compared with the KAZE, A–KAZE, FMI–SPOMF, and HYFM registration methods in terms of successful registrations, number of matches, number of correct matches, correct match ratio, RMSE and registration error. The results show that the highest number of correctly recovered scales for all angles are obtained by HSI–KAZE. For example, a scale factor of $24.0\times$ has been recovered in Pavia Centre by the HSI–KAZE method, while for the KAZE method the maximum achieved scale is $9.0\times$.

As future work, HSI–KAZE can be used as a first stage of a fine grained multilevel registration method. HSI–KAZE would perform a high–level registration that would be refined in a second registration stage to correct geometric deformations at a lower level.

References

1. Thenkabail, P.S.; Lyon, J.G. *Hyperspectral Remote Sensing of Vegetation*; CRC Press: Boca Raton, FL, USA, 2016.
2. Garea, A.S.; Heras, D.B.; Argüello, F. GPU classification of remote-sensing images using kernel ELM and extended morphological profiles. *Int. J. Remote Sens.* **2016**, *37*, 5918–5935. [CrossRef] [CrossRef]
3. Kruse, F.A.; Boardman, J.W.; Huntington, J.F. Comparison of airborne hyperspectral data and EO-1 Hyperion for mineral mapping. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1388–1400. [CrossRef]
4. Sun, D.W. *Hyperspectral Imaging for Food Quality Analysis and Control*; Elsevier: London, UK, 2010.

5. López-Fandiño, J.; Heras, D.B.; Argüello, F.; Mura, M.D. GPU Framework for Change Detection in Multitemporal Hyperspectral Images. In Proceedings of the 10th International Symposium on High-Level Parallel Programming and Applications, Valladolid, Spain, 10–11 July 2017; pp. 115–132.

6. Zitova, B.; Flusser, J. Image registration methods: A survey. *Image Vis. Comput.* **2003**, *21*, 977–1000. [CrossRef]

7. Pratt, W.K. Correlation techniques of image registration. *IEEE Trans. Aerosp. Electron. Syst.* **1974**, *AES-10*, 353–358. [CrossRef]

8. Wells, W.M.; Viola, P.; Atsumi, H.; Nakajima, S.; Kikinis, R. Multi-modal volume registration by maximization of mutual information. *Med. Image Anal.* **1996**, *1*, 35–51. [CrossRef]

9. Ordóñez, Á.; Argüello, F.; Heras, D.B. GPU Accelerated FFT-Based Registration of Hyperspectral Scenes. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 4869–4878. [CrossRef]

10. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [CrossRef]

11. Bay, H.; Ess, A.; Tuytelaars, T.; Van Gool, L. Speeded-up robust features (SURF). *Comput. Vis. Image Underst.* **2008**, *110*, 346–359. [CrossRef]

12. Alcantarilla, P.F.; Bartoli, A.; Davison, A.J. KAZE features. In Proceedings of the European Conference on Computer Vision, Florence, Italy, 7–13 October 2012; Springer: Berlin, Germany, 2012; pp. 214–227.

13. Agrawal, M.; Konolige, K.; Blas, M.R. Censure: Center surround extremas for realtime feature detection and matching. In Proceedings of the European Conference on Computer Vision, Marseille, France, 12–18 October 2008; Springer: Berlin, Germany, 2008; pp. 102–115.

14. Calonder, M.; Lepetit, V.; Strecha, C.; Fua, P. Brief: Binary robust independent elementary features. In Proceedings of the Computer Vision–ECCV 2010, Crete, Greece, 5–11 September 2010; pp. 778–792.

15. Rublee, E.; Rabaud, V.; Konolige, K.; Bradski, G. ORB: An efficient alternative to SIFT or SURF. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 2564–2571.

16. Levi, G.; Hassner, T. LATCH: Learned arrangements of three patch codes. In Proceedings of the WACV 2016: IEEE Winter Conference on Application of Computer Vision, Lake Placid, NY, USA, 7–9 March 2016; pp. 1–9.

17. Alcantarilla, P.F.; Nuevo, J.; Bartoli, A. Fast explicit diffusion for accelerated features in nonlinear scale spaces. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *34*, 1281–1298.

18. Yi, Z.; Zhiguo, C.; Yang, X. Multi-spectral remote image registration based on SIFT. *Electron. Lett.* **2008**, *44*, 107–108. [CrossRef]

19. Li, Q.; Wang, G.; Liu, J.; Chen, S. Robust scale-invariant feature matching for remote sensing image registration. *IEEE Geosci. Remote Sens. Lett.* **2009**, *6*, 287–291.

20. Mehmet, F.; Yardimci, Y.; Temlzel, A. Registration of multispectral satellite images with orientation-restricted SIFT. In Proceedings of the 2009 IEEE International Geoscience and Remote Sensing Symposium, Cape Town, South Africa, 12–17 July 2009; Volume 3, p. III-243.

21. Teke, M.; Temizel, A. Multi-spectral satellite image registration using scale-restricted SURF. In Proceedings of the 20th International Conference on Pattern Recognition, Istanbul, Turkey, 23–26 August 2010; pp. 2310–2313.

22. Hasan, M.; Jia, X.; Robles-Kelly, A.; Zhou, J.; Pickering, M.R. Multi-spectral remote sensing image registration via spatial relationship analysis on sift keypoints. In Proceedings of the 2010 IEEE International Geoscience and Remote Sensing Symposium, Honolulu, HI, USA, 25–30 July 2010; pp. 1011–1014.

23. Mukherjee, A.; Velez-Reyes, M.; Roysam, B. Interest points for hyperspectral image data. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 748–760. [CrossRef]

24. Dorado-Muñoz, L.P.; Velez-Reyes, M.; Mukherjee, A.; Roysam, B. A vector SIFT detector for interest point detection in hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 4521–4533. [CrossRef]

25. Perona, P.; Malik, J. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.* **1990**, *12*, 629–639. [CrossRef]

26. Weickert, J.; Romeny, B.T.H.; Viergever, M.A. Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE Trans. Image Process.* **1998**, *7*, 398–410. [CrossRef] [PubMed]

27. Weickert, J.; Scharr, H. A scheme for coherence-enhancing diffusion filtering with optimized rotation invariance. *J. Vis. Commun. Image Represent.* **2002**, *13*, 103–118. [CrossRef]

28. Grewenig, S.; Weickert, J.; Bruhn, A. From Box Filtering to Fast Explicit Diffusion. In *DAGM-Symposium*; Springer: Berlin, Germany, 2010; pp. 533–542.

29. Alvarez, L.; Lions, P.L.; Morel, J.M. Image selective smoothing and edge detection by nonlinear diffusion. II. *SIAM J. Numer. Anal.* **1992**, *29*, 845–866. [CrossRef]

30. Jia, X.; Kuo, B.C.; Crawford, M.M. Feature mining for hyperspectral image classification. *Proc. IEEE* **2013**, *101*, 676–697. [CrossRef]

31. Richards, J.A. *Remote Sensing Digital Image Analysis*; Springer: Berlin, Germany; New York, NY, USA, 1999; Volume 3.

32. Cariou, C.; Chehdi, K.; Le Moan, S. BandClust: An unsupervised band reduction method for hyperspectral remote sensing. *IEEE Geosci. Remote Sens. Lett.* **2011**, *8*, 565–569. [CrossRef]

33. Martínez-Usó, A.; Pla, F.; Sotoca, J.M.; García-Sevilla, P. Clustering-based Hyperspectral Band selection using Information Measures. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 4158–4171. [CrossRef]

34. Ward, J.H., Jr. Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* **1963**, *58*, 236–244. [CrossRef]

35. Otero, I.R. Anatomy of the SIFT Method. Ph.D. Thesis, École Normale Supérieure de Cachan-ENS Cachan, Cachan, France, 2015.

36. Gualtieri, A.; Chettri, S.R.; Cromp, R.; Johnson, L. Support Vector Machine Classifiers as Applied to AVIRIS Data. In Proceedings of the 8th JPL Airborne Earth Science Workshop, Pasadena, California, 8–11 February 1999; pp. 217–227.

37. Hyperspectral Remote Sensing Scenes at the University of the Basque Country. Available online: http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes. (accessed on 9 February 2017).

38. AVIRIS Database. Available online: http://aviris.jpl.nasa.gov/data/index.html. (accessed on 9 February 2017).

39. Chen, Q.S.; Defrise, M.; Deconinck, F. Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **1994**, *16*, 1156–1168. [CrossRef]

40. Goncalves, H.; Corte-Real, L.; Goncalves, J.A. Automatic image registration through image segmentation and SIFT. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 2589–2600. [CrossRef]

41. Chen, S.; Guo, Q.; Leung, H.; Bosse, E. A maximum likelihood approach to joint image registration and fusion. *IEEE Trans. Image Process.* **2011**, *20*, 1363–1372. [CrossRef] [PubMed]

42. Zhang, H.; Liu, X.; Li, L.; Rao, J.; Li, Q.; Chen, S.; Wang, T. Accurate multi-spectral image registration based on scale invariant feature. In Proceedings of the 2012 2nd International Conference on Computer Science and Network Technology, Changchun, China, 29–31 December 2012; pp. 847–852.

43. Gong, M.; Zhao, S.; Jiao, L.; Tian, D.; Wang, S. A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 4328–4338. [CrossRef]

44. Li, Q.; Qi, S.; Shen, Y.; Ni, D.; Zhang, H.; Wang, T. Multispectral image alignment with nonlinear scale-invariant keypoint and enhanced local feature matrix. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 1551–1555.

45. Zhang, Q.; Cao, Z.; Hu, Z.; Jia, Y.; Wu, X. Joint image registration and fusion for panchromatic and multispectral images. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 467–471. [CrossRef]

46. Wu, Y.; Miao, Q.; Ma, W.; Gong, M.; Wang, S. PSOSAC: Particle Swarm Optimization Sample Consensus Algorithm for Remote Sensing Image Registration. *IEEE Geosci. Remote Sens. Lett.* **2017**, *15*, 242–246. [CrossRef]

47. Ma, W.; Wen, Z.; Wu, Y.; Jiao, L.; Gong, M.; Zheng, Y.; Liu, L. Remote sensing image registration with modified sift and enhanced feature matching. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 3–7. [CrossRef]

48. Chen, S.J.; Shen, H.L.; Li, C.; Xin, J.H. Normalized total gradient: A new measure for multispectral image registration. *IEEE Trans. Image Process.* **2018**, *27*, 1297–1310. [CrossRef]