



HAL
open science

Hierarchical Unsupervised Partitioning of Large Size Data and Its Application to Hyperspectral Images

Jihan Alameddine, Kacem Chehdi, Claude Cariou

► **To cite this version:**

Jihan Alameddine, Kacem Chehdi, Claude Cariou. Hierarchical Unsupervised Partitioning of Large Size Data and Its Application to Hyperspectral Images. *Remote Sensing*, 2021, 13 (23), pp.4874. 10.3390/rs13234874 . hal-03552768

HAL Id: hal-03552768

<https://hal.science/hal-03552768>

Submitted on 2 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



Article

Hierarchical Unsupervised Partitioning of Large Size Data and Its Application to Hyperspectral Images

Jihan Alameddine, Kacem Chehdi * and Claude Cariou

Institut d'Électronique et des Technologies du numÉrique, University of Rennes 1, IETR (UMR CNRS)/TSI2M, Enssat, 6 rue de KÉrampont, 22305 Lannion, France; jihan.alameddine@univ-rennes1.fr (J.A.); claude.cariou@univ-rennes1.fr (C.C.)

* Correspondence: kacem.chehdi@univ-rennes1.fr

Abstract: In this paper, we propose a true unsupervised method to partition large-size images, where the number of classes, training samples, and other *a priori* information is not known. Thus, partitioning an image without any knowledge is a great challenge. This novel adaptive and hierarchical classification method is based on affinity propagation, where all criteria and parameters are adaptively calculated from the image to be partitioned. It is reliable to objectively discover classes of an image without user intervention and therefore satisfies all the objectives of an unsupervised method. Hierarchical partitioning adopted allows the user to analyze and interpret the data very finely. The optimal partition maximizing an objective criterion provides the number of classes and the exemplar of each class. The efficiency of the proposed method is demonstrated through experimental results on hyperspectral images. The obtained results show its superiority over the most widely used unsupervised and semi-supervised methods. The developed method can be used in several application domains to partition large-size images or data. It allows the user to consider all or part of the obtained classes and gives the possibility to select the samples in an objective way during a learning process.



Citation: Alameddine, J.; Chehdi, K.; Cariou, C. Hierarchical Unsupervised Partitioning of Large Size Data and Its Application to Hyperspectral Images. *Remote Sens.* **2021**, *13*, 4874. <https://doi.org/10.3390/rs13234874>

Academic Editor: Paul Scheunders

Received: 17 October 2021

Accepted: 25 November 2021

Published: 30 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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/>).

Keywords: unsupervised classification; estimation of classes number; large size datasets; hyperspectral image; unsupervised learning; objective decision making; remote sensing

1. Introduction

The partitioning of large-size images, in order to exploit their rich information content for the construction of reliable decision-making systems, is of growing interest in many application fields. Indeed, hyperspectral images are today among the most widely used data thanks to their large spectral range (several hundred spectral bands covering the visible and infrared domains) and their fine spatial resolution (a few tens of centimeters). Due to this richness of information, which allows for better discrimination of objects, interest in hyperspectral images has increased during the last years in many application fields. These fields include geology [1], medicine [2–7], industrial production [8,9], safety [10], and the environment [11–20]. In this latter field, hyperspectral imagery has been of great interest and several applications have been treated. These applications include the inventory of vegetation species [11,12], the early detection of vegetation diseases [13,14] and invasive species [15,16], the identification of marine algae [17,18], and the human and animal impacts on the environment [19,20], etc. Thus, to efficiently meet the needs of all these areas, any partitioning method of hyperspectral images must be conducted while respecting the physical nature of the information provided by these images. A plethora of published methods do not respect the precise information given by these images, where the ground truth (GT) data used is simplified [21,22] and the classification of existing methods is sometimes ambiguous and not consistent. For example, in [23], the authors propose a method using learning samples by defining it as semi-supervised. Other authors introduce in [24], the number of classes in the partitioning process and qualify their method as

unsupervised. In general, it is wrong to consider all algorithms as unsupervised because the introduction of the number of classes by the user to run an algorithm is a supervised operation [25].

In order to remove all these confusions, we classify according to criteria related to the nature of the knowledge introduced by the users this plethora of methods into three main categories instead of two, as often proposed in the literature (supervised and unsupervised). The confusion is due to the fact that methods that require a priori knowledge of the number of classes are in the same category as those that estimate the number of classes. Thus, we subdivided the category of unsupervised methods into two categories (semi-supervised and unsupervised methods) [26]. It is important to recall that the methods of each category can be hierarchical or non-hierarchical and parametric or non-parametric. Consequently, the three main categories of classification or partitioning methods opted here are supervised, semi-supervised, and unsupervised, which we will specify after the definition of “partitioning”.

Definition 1. “Data partitioning” is a method of data analysis that consists of dividing a set of data described by quantitative features into different homogeneous “sub-sets” or classes, according to a similarity criterion in the sense that the data in each sub-set share common characteristics. The obtained classes form a partition.

Data partitioning can be performed with or without learning. In the latter case, all classes are located and the unidentified classes as in the case of several domains (environment, medicine, security, biology, etc.) can be analyzed and listed to add them as new known classes for possible learning or not.

We specify that in this article we use the term “partitioning” because it is both more practical and more appropriate to describe the developed data analysis method. In addition, it is more convenient to avoid using a long term, such as “classification of image pixels,” and a false term, such as “image classification,” which is another problem.

Partitioning an image or data consists in creating a partition formally defined as follows:

Let $X = \{x_1, x_2, \dots, x_N\}$ be a set of N objects x_i , where each object is characterized by B quantitative features.

The process of dividing X into K classes, C_i , consists in creating a partition, P , according to one or more objective optimization criteria with: $P = \{C_1, C_2, \dots, C_K\}$.

This partition, P , will therefore highlight the different classes in the dataset, X .

The quality of a partition will depend on the degree of homogeneity of the classes formed and consequently on their number. Thus, in order to obtain a correct partitioning result, three requirements must be verified simultaneously:

Completeness: all objects in the dataset must be associated with a class.

$$\forall x_i \in X, \exists C_j \text{ such that } x_i \in C_j$$

Separability: the classes must be sufficiently differentiable so that an object can only be associated with one class.

$$\bigcup_{i=1}^K C_i = X \text{ and } \bigcap_{i=1}^K C_i = \emptyset$$

Relevance: the association of an object to a class must be carried out according to an objective optimization criterion.

The different partitioning methods can be grouped into three categories [26] which we define as follows:

-*Supervised methods* require training samples to accomplish the partitioning task. Methods such as the Maximum Likelihood [27] and Support Vector Machines [28] are the most common methods used in this category.

We note that the information required by the methods of this category is not always available for many applications, and even if it exists, it is not always reliable [21,22].

-*Semi-supervised methods*, considered unsupervised in the literature, require the number of classes and threshold values or other empirical parameters from the operator. *K*-means [29] is a basic semi-supervised algorithm that classifies each object according to their similarity/dissimilarity, requiring the number of classes to be fixed by the user. Furthermore, it is sensitive to the random initialization of class centers, which makes it unstable [30–32]. Several extensions of *K*-means have been developed [33–36] but are still affected by the initial choice of class centers.

The Fuzzy C-Means (FCM) [37–40] derived from the *K*-means algorithm is another semi-supervised algorithm. This algorithm is also sensitive to the initial class centers and the choice of the fuzzification parameter.

In practice, the knowledge of the number of classes required by supervised and semi-supervised methods is imprecise and not often accessible, in particular for hyperspectral aerial images with large landscape areas. As stated in [21,22,32], prior knowledge of the number of classes is a subjective process in nature, which precludes an absolute judgment as to the relevance of all data analysis. Consequently, these methods do not allow for the discovery of novel relevant classes. In this case, the knowledge of the number of classes in semi-supervised and supervised methods can be considered as a constraint and does not often reflect reality.

In order to partition objects or form learning classes, unsupervised partitioning methods present many advantages with respect to supervised and semi-supervised methods, as defined below.

-*Unsupervised methods* do not require the number of classes, associated learning samples, or any other prior knowledge. The number of classes is objectively estimated following a given optimization criterion or several optimization criteria. This category is called exploratory data analysis by Xu and Wunsch [32]. It is more adapted to explore data analysis or to learn a novel object and discover a novel phenomenon. Indeed, the identification of the nature of each class is defined after the partitioning process and not before. This category leads to a fine and complete analysis of the observed data and does not limit the analysis to known classes. For these reasons, it meets a real need generated by some application domains, such as Earth observation, where the areas to be analyzed are very large or difficult to access. It provides accurate, objective, and consistent solutions that translate real information content of images independently of the GT data or learning samples which can be biased or simplified [21,22].

The definition of unsupervised methods given above ensures that no subjective knowledge is introduced by the user.

Several unsupervised methods have been proposed for data partitioning. In [41], an unsupervised version of the *K*-means algorithm, named Modified Linde–Buzo–Gray (MLBG) is proposed. This method improves the different stages of the Linde–Buzo–Gray (LBG) algorithm and is able to automatically determine the number of classes; however, it requires a long computing time. In [26], an optimized version of the baseline FCM algorithm (FCMO) was presented in order to make it stable and deterministic. Stability is achieved by initializing the class centers through an adaptive incremental procedure. In addition, an unsupervised evaluation criterion, based on the within- and between-class disparities, is introduced to estimate the optimal number of classes. This method can be used in unsupervised or semi-supervised mode.

One of the most elaborate unsupervised partitioning methods is Affinity Propagation (AP) [42]. This method has six advantages: (i) no a priori knowledge is required, especially the non-introduction of training samples, to give the user the possibility to detect and locate the known classes and new classes called “discovery classes”; (ii) stability of the results thanks to its deterministic character; (iii) possibility to objectively select the samples of the classes in a learning system in order to be able to detect them afterward; (iv) applicable to several domains without learning constraints; (v) it can be used in unsupervised or semi-supervised mode, and; and (vi) it is insensitive to initialization. Due to these advantages,

it is widely used in several application areas, such as environmental monitoring and safety [43–47], multimedia data management, and pattern recognition [48–57].

The main drawback of the standard AP method is that it is not applicable to partitioning large-size hyperspectral images. Indeed, the memory space required for each of the four main matrices has a quadratic relationship with the number of pixels to be partitioned. In [25], Chehdi et al. suggested a reduction in the number of pixels to allow the application of AP to partition large-size images. To reduce this number, the hyperspectral image is divided into blocks and the reduction step is applied independently within each block. Then, the AP method is applied only on exemplars of the duplicated pixels and non-duplicated pixels. However, in its current version, AP does not take into account the presence of identical objects after the reduction step in the calculation of the criteria used. Other drawbacks of the standard AP method are: (i) some parameters are not computed adaptively; and (ii) the criterion of identification of class exemplars may create not perfectly homogeneous classes, as demonstrated in Section 3.1.2.

To extend the AP method to large-size data, such as hyperspectral aerial images, and to optimize it by adaptively calculating the criteria and parameters used, we propose here a new version without any prior knowledge introduced by the user. We recall that a relevant partitioning method, whatever its nature (supervised, semi-supervised, or unsupervised), must provide results to the end-user, taking into account the physical characteristics provided by measuring instruments [21,22].

The main contributions of this work are: (1) modification and optimization of several steps of the standard AP algorithm; (2) introduction of a new approach to partition large size images and proposal of a fusion procedure of the classes obtained in the blocks with stable results regardless of the chosen block size; and (3) generation of hierarchical partitioning. For AP optimization, the improvements are the estimation of the preference parameter value for each object; taking into account the presence of identical objects in the calculation of similarity, responsibility, and availability criteria; updating procedure of estimating the values of the responsibility and the availability criteria; and finally, modification of the decisional criterion used to estimate the number of classes and identify their exemplars. The proposed hierarchical method allows the user to obtain several partitions while indicating the most optimal one. These partitioning results are very important and necessary in several application fields to perform a fine analysis and interpretation of the formed classes.

The remainder of the paper is organized as follows: Section 2 provides an overview of the standard AP algorithm and related studies. Section 3 describes the proposed method, named “Unsupervised Partitioning by Optimized Affinity Propagation” (UP-OAP) and its hierarchical version, named “Hierarchical Unsupervised Partitioning approach by OAP” (HUP-OAP). Section 4 presents the assessment results of the proposed method on a synthetic hyperspectral image constructed from a real aerial image acquired by our platform. It is also evaluated on a real aerial large-size hyperspectral image. The target application of these images is the identification of marine algae species. To assess the efficiency of the proposed unsupervised method, each image is provided with validated GT data. Finally, Section 5 concludes this paper and provides some perspectives.

2. Review of the Standard Affinity Propagation Algorithm and Related Studies

Unsupervised partitioning methods, as defined in the introduction, have many advantages over supervised and semi-supervised methods: (i) they do not require any prior knowledge to aggregate objects in classes (neither the number of classes to discriminate, nor learning samples). The number of classes is estimated following a given optimization criterion, and (ii) they respect the physical characteristics of objects in the formation of classes. Thus, unsupervised methods provide more relevant results because the decision criteria for objects aggregation is independent of the GT data or learning samples, which can be biased or simplified in some cases [21,22]. For these reasons, we are interested in the development of an unsupervised partitioning method that excludes user intervention.

2.1. Overview of the Standard AP Algorithm

In the standard AP algorithm developed by Frey and Dueck [42], two procedures of message transmission, called responsibility and availability, are used to exchange messages between objects. These messages are used to identify in an iterative manner the best exemplar of each class that may exist. The responsibility, $r(x_i, x_k)$, is sent from object x_i to candidate exemplar x_k and reflects how well-suited it would be for object x_k to be the exemplar of object x_i . The availability, $a(x_i, x_k)$, is sent from candidate exemplar x_k to object x_i , and reflects how appropriate it would be for object x_i to choose candidate exemplar x_k as its exemplar. To calculate both criteria, the similarity matrix is used as the opposite of the squared Euclidean distance $-d_2^2(x_i, x_k)$, where d_2 is the distance associated to L_2 -norm.

For set $X = \{x_1, x_2, \dots, x_N\}$ of N objects to be partitioned, where each object, x_i , is characterized by a set of B features, S , R , and A denote the similarity, responsibility, and availability matrices of size $N \times N$, respectively. $s(x_i, x_k)$, $r(x_i, x_k)$, and $a(x_i, x_k)$ are their respective elements for objects x_i and x_k . Mathematically, responsibility, $r(x_i, x_k)$, and availability, $a(x_i, x_k)$, are defined as follows [42]:

$$r(x_i, x_k)_{i \neq k} = s(x_i, x_k) - \max_{k', k' \neq k} [s(x_i, x_{k'}) + a(x_i, x_{k'})] \quad (1)$$

$$r(x_k, x_k) = p - \max_{k', k' \neq k} [s(x_k, x_{k'}) + a(x_k, x_{k'})] \quad (2)$$

with $s(x_i, x_k) = -d_2^2(x_i, x_k)$, for $i \neq k$ and $s(x_k, x_k) = p$, where p denotes a global preference parameter, whose value is generally set as the minimum or median value of the similarity matrix S , which implicitly controls the number of classes.

In this method, all diagonal elements $s(x_k, x_k)$ of the similarity matrix S are equal to the p value instead of zero.

$$a(x_i, x_k) = \min \{0, r(x_k, x_k) + \sum_{i', i' \neq \{i, k\}} \max[0, r(x_{i'}, x_k)]\} \quad (3)$$

$$a(x_k, x_k) = \sum_{k', k' \neq k} \max[0, r(x_{k'}, x_k)] \quad (4)$$

At each iteration, m , responsibility and availability are estimated as follows:

$$\hat{r}(x_i, x_k)_m = \lambda \hat{r}(x_i, x_k)_{m-1} + (1 - \lambda) r(x_i, x_k)_m \quad (5)$$

$$\hat{a}(x_i, x_k)_m = \lambda \hat{a}(x_i, x_k)_{m-1} + (1 - \lambda) a(x_i, x_k)_m \quad (6)$$

where λ is a damping factor ($\lambda \in [0, 1]$).

At any step of the iterative process, responsibilities and availabilities are combined to identify the exemplar of each class to be formed. The criterion that identifies object x_k as an exemplar of object x_i is:

$$E^*(x_i) = \operatorname{argmax}_k [\hat{r}(x_i, x_k) + \hat{a}(x_i, x_k)] \quad (7)$$

The standard AP algorithm requires as inputs (see Algorithm 1), the value of the damping factor, λ , and the choice between two possibilities of the value of the preference parameter, p , which can be fixed at the minimum or median value of the similarity matrix, whatever the object. The choice of the values of these two parameters conditions the partitioning results; therefore, the optimality of the results is not guaranteed.

Algorithm 1 Standard AP**Input:**

- **Data table** (N objects \times B features)
- **Parameters to be set by the user:**
 - Damping factor λ ($\lambda \in]0,1[$)
 - Preference parameter value p : fixed to the minimum or median value of the similarity matrix, S

Preliminary steps:

- Calculation of the similarity matrix S of size $N \times N$
 $s(x_i, x_k) = -d_2^2(x_i, x_k)$
- Identification of the preference parameter value, p , according to the fixed choice
- Initialization: $r(x_i, x_k) = 0$, $a(x_i, x_k) = 0$

Procedure:

1. Replacement of the diagonal elements of S by the value of p
2. Calculating all responsibilities given the availabilities according to Equations (1), (2), and (5)
3. Calculating all availabilities given the responsibilities according to Equations (3), (4), and (6)
4. Combining availabilities and responsibilities according to Equation (7) for each object x_i and identifying exemplars x_k that maximize $[\hat{r}(x_i, x_k) + \hat{a}(x_i, x_k)]$
5. **if** exemplars do not change, proceeding to the next step (6)
else repeat steps (2) to (4) until convergence
end if
6. Merging every object to its nearest exemplar and break

Output: Partition P of K classes and exemplar of each class

2.2. Related Studies

The AP algorithm has large success in numerous application areas by providing relevant partitioning results, as mentioned in the introduction [42–57], but its use remains limited to images of small sizes. This algorithm can also be improved to give better partitioning results by adapting the damping factor and the preference parameter.

In [43–45,48], several extensions of the AP algorithm are proposed, but the preference parameter, p , and the damping factor, λ , remain not adaptive.

To adapt the preference parameter, a solution is proposed in [47]. Each preference parameter, p_j , is automatically adjusted during the iteration process according to the data distribution by fixing two thresholds. With this method, the problem of partitioning large-size images remains. Furthermore, the introduction of these two thresholds makes this method parametric. In [58], an extension of the AP algorithm is presented in order to reduce the computing time and the memory space, but the values of the preference parameter and the damping factor are to be set by the user. Other propositions are described in [56,59,60], but they are also not adapted to partition large-size data. Furthermore, the damping factor and the preference parameter are fixed by the user. Finally, in [61,62], the AP algorithm was combined with other methods to improve the learning task. For these last two methods, prior knowledge is required for the learning task, the preference parameter is also not adaptive, and the damping factor must also be chosen by the user.

3. Proposed Hierarchical Unsupervised Partitioning Method

In this section, we first present a new unsupervised partitioning method based on an Optimization of Affinity Propagation [42], which we name “UP-OAP”. Next, we describe the main steps of the hierarchical partitioning version using the UP-OAP method which we call “HUP-OAP”.

3.1. Unsupervised Partitioning Method by Optimized Affinity Propagation (UP-OAP)

In contrast to the standard AP, in the UP-OAP method, all the parameters and criteria are calculated in an adaptive way, taking into account the presence of identical objects in the dataset to be partitioned. Finally, the criterion for identifying the exemplars of each class is reformulated. The flowchart of the UP-OAP method is shown in Figure 1.

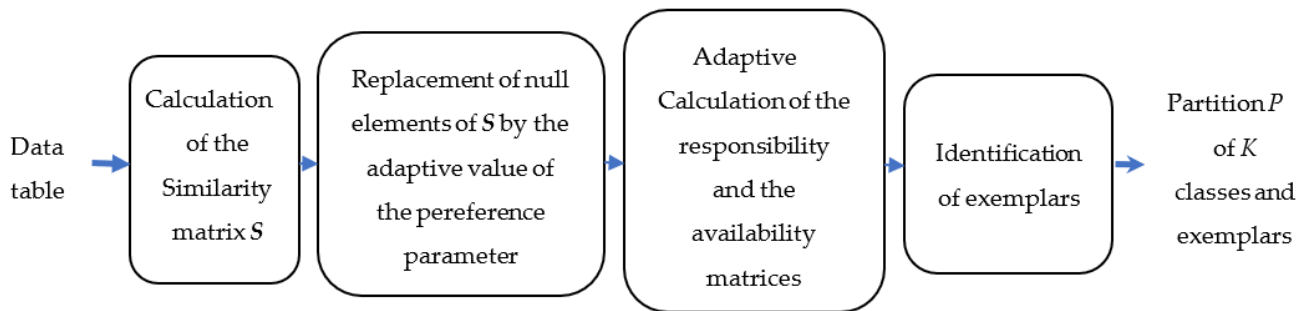


Figure 1. Flowchart of the proposed UP-OAP method.

3.1.1. Preference Parameter and Responsibility-Availability Criteria

In the standard version of AP, the preference parameter, p , is set to the minimum or median value of the similarity matrix and does not take into account the variations of the similarity values between the objects of the dataset, X . To associate to each object, x_i , its preference parameter, p_i , in the similarity matrix S (row i of S) calculated on X , when $s(x_i, x_k) = 0$, for $i = k$ or $i \neq k$, this parameter is calculated as follows:

$$\bar{p}_i = \frac{1}{N} \sum_{k=1}^N s(x_i, x_k) \quad (8)$$

where $s(x_i, x_k)$ denotes the elements of matrix S and N is the size of X .

To take into account the presence of identical objects in the dataset to be partitioned, that is, $s(x_i, x_k) = 0$, for $i \neq k$, two modifications are made: (i) assigning the value of the preference parameter, \bar{p}_k , to all the null elements of the matrix S , in the same way as those on its diagonal; and (ii) calculating the elements of R and A , as follows:

$$r(x_i, x_k)_{i \neq k} = r(x_k, x_k) = \bar{p}_k - \max_{k', k' \neq k} [s(x_k, x_{k'}) + a(x_k, x_{k'})] \quad (9)$$

$$a(x_i, x_k)_{i \neq k} = a(x_k, x_k) = \sum_{k', k' \neq k} \max[0, r(x_{k'}, x_k)] \quad (10)$$

To update R and A in Equations (5) and (6) without a choice of the damping factor, λ , a smoothing operation is introduced as follows:

$$\hat{r}(x_i, x_k)_m = [\hat{r}(x_i, x_k)_{m-3} + \hat{r}(x_i, x_k)_{m-2} + \hat{r}(x_i, x_k)_{m-1} + r(x_i, x_k)_m] / 4 \quad (11)$$

$$\hat{a}(x_i, x_k)_m = [\hat{a}(x_i, x_k)_{m-3} + \hat{a}(x_i, x_k)_{m-2} + \hat{a}(x_i, x_k)_{m-1} + a(x_i, x_k)_m] / 4 \quad (12)$$

This operation provides less biased estimates of R and A matrices, in contrast to those obtained in [42,63].

3.1.2. Identification Criterion of Exemplars

The decision criterion E^* (see Equation (7)) maximizing $(R + A)$ to identify exemplars is inappropriate because availability, A , is in some cases incoherent with responsibility, R . Consequently, some exemplars are not detected, which leads to the aggregation of a truly existing class to another. To overcome this drawback, only the criterion using responsibility is used, as justified in the following.

Proposition 1. Let $x_i, x_j \in X$, be two highly similar objects ($s(x_j, x_i) \cong 0$) and $\forall x_q \in X, s(x_j, x_i) \gg s(x_j, x_q)$, where x_j and x_q are dissimilar ($s(x_j, x_q) \ll 0$), that is, x_j can only aggregate with x_i . Assume that x_i and x_q are two candidate exemplars for x_j , where x_j

will be better represented by x_i than x_q : $r(x_j, x_i) > r(x_j, x_q)$.

Moreover, assume that x_i was not chosen as an exemplar for any object, i.e., $r(x_i, x_i) + a(x_i, x_i) < 0$, therefore $r(x_j, x_i) + a(x_j, x_i) < 0$. Then, x_j will be aggregated with another exemplar, $x_v, v \neq \{i, j\}$.

This proposition shows that when the responsibility between two objects, x_j and x_i , is positive and larger than all other responsibilities, but the candidate exemplar, x_i , is not chosen as an exemplar for any object, the availability in absolute value exceeds the responsibility value and assigns object x_j to another class, even if these objects cannot be aggregated.

Proof of Proposition 1. Let x_i, x_j , and $x_q \in X$. Assume that $r(x_j, x_i) > 0$ and $r(x_j, x_q) > 0$, with $r(x_j, x_i) > r(x_j, x_q)$ (i.e., it is better for x_j to be represented by x_i than x_q). Moreover, assume that x_i was not chosen for any object as an exemplar. \square

$$\begin{aligned} a(x_j, x_i) &= \min\{0, r(x_i, x_i) + \sum_{i', i' \neq \{j, i\}} \max[0, r(x_{i'}, x_i)]\} \\ \Rightarrow a(x_j, x_i) &= \min\{0, r(x_i, x_i) + \sum_{i', i' \neq i} \max[0, r(x_{i'}, x_i)] - \max[0, r(x_j, x_i)]\} \\ \Rightarrow a(x_j, x_i) &= \min\{0, r(x_i, x_i) + a(x_i, x_i) - \max[0, r(x_j, x_i)]\} \\ \text{Because } r(x_j, x_i) > 0, \text{ then : } \max[0, r(x_j, x_i)] &= r(x_j, x_i), \\ \Rightarrow a(x_j, x_i) &= \min\{0, r(x_i, x_i) + a(x_i, x_i) - r(x_j, x_i)\} \\ \Rightarrow a(x_j, x_i) &\leq r(x_i, x_i) + a(x_i, x_i) - r(x_j, x_i) \\ \Rightarrow a(x_j, x_i) + r(x_j, x_i) &\leq r(x_i, x_i) + a(x_i, x_i). \end{aligned}$$

Because x_i is not chosen as an exemplar, then:

$$r(x_i, x_i) + a(x_i, x_i) < 0 \Rightarrow a(x_j, x_i) + r(x_j, x_i) < 0.$$

Because $a(x_j, x_i) < 0, r(x_j, x_i) > 0$ and under the assumptions of Proposition 1, we have shown that when $a(x_j, x_i) + r(x_j, x_i) < 0$, then $|a(x_j, x_i)| > r(x_j, x_i)$.

The presence of availability in criterion $E^*(x_i)$ for the search for exemplars can assign an exemplar to an object even if they are highly dissimilar. This disturbs the final decision and does not correctly detect the real present classes.

Under these conditions, the identification of exemplars by maximizing only R contributes to the formation of homogeneous classes representative of the observed data.

This leads to modify the decision criterion, E^* , identifying exemplars by using only the responsibility, R :

$$E^*(x_i) = \operatorname{argmax}_k [\hat{r}(x_i, x_k)] \quad (13)$$

The steps of UP-OAP method are shown in Algorithm 2.

Algorithm 2 UP-OAP**Input:** *Data table* (N objects \times B features)

1. Calculation of the similarity matrix S of size $N \times N$
 $s(x_i, x_k) = -d_1(x_i, x_k)$, where d_1 is the distance associated with L_1 -norm
2. Initialization: $r(x_i, x_k) = 0, a(x_i, x_k) = 0$
3. Replacement of the null elements of S by the value of \bar{p}_i
4. Calculation of all responsibilities given the availabilities according to Equations (1), (9), and (11)
5. Calculation of all availabilities given the responsibilities according to Equations (3), (10), and (12)
6. Identification of exemplars, x_k , that maximize $E^* = \operatorname{argmax}_k [\hat{p}(x_i, x_k)]$ (Equation (13))
7. **If** exemplars do not change, proceed to the next step, (8)
else repeat steps (4) to (6) until convergence
end if
8. Merging each object to its nearest exemplar and break

Output: Partition P of K classes and exemplar I_j of each class C_j **3.2. Hierarchical Partitioning of Large Size Hyperspectral Images (HUP-OAP)**

In this section, we detail the main steps of the hierarchical unsupervised partitioning method of large-size images based on the UP-OAP algorithm. The two main steps of this hierarchical HUP-OAP method are respectively the formation of the first and the other partitions by identifying the most relevant one according to an optimization criterion. These steps are described below.

3.2.1. First Partition of the Original Image

The application of the UP-OAP method to large-size images, such as hyperspectral aerial images, requires block partitioning and merging the block partitioning results. This subsection details the main steps of the formation of the first partition by partitioning the blocks of the original image by UP-OAP and fusion of the classes of all blocks, as shown in the flowchart of Figure 2.

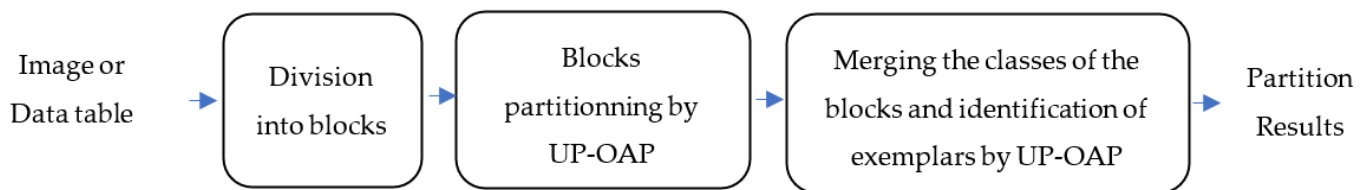


Figure 2. Flowchart of unsupervised partitioning of a large size image and identification of exemplar of each class.

To be able to partition all the pixels of an image, in a simple way, the image is divided into regular blocks of the same size, without overlapping between the blocks.

Definition 2. The number of pixels exemplars, N_0 , considered for partitioning is defined by:

$$N_0 = \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} N_{ij} \quad (14)$$

where N_{ij} is the number of classes obtained by UP-OAP on block B_{ij} , M_1 is the number of blocks in a row, and M_2 is the number of blocks in a column.

Proposition 2. If S_0 is the similarity matrix of size $N_0 \times N_0$ calculated on the new dataset X_0 , of size N_0 formed by the exemplars of classes of blocks, then we have $N_0 \ll N$.

This proposition shows that the size of the similarity matrix using the exemplars is smaller than the one calculated on the whole original image, which makes it possible to apply UP-OAP on very large-size images.

Algorithm 3 details the steps to obtain this first partition.

3.2.2. Hierarchical Partitioning

In order to give the user the ability to conduct a finer analysis of datasets, a hierarchical link between the classes of the partitions is created. These partitions are generated by using the UP-OAP method from exemplars of classes of the first partition, P_1 . Thus, to obtain several partitions in a hierarchical way, the UP-OAP method is applied iteratively at each level, i ($i \geq 2$), to exemplars of classes of partition P_{i-1} corresponding to level $i - 1$. This “exemplars-partitioning” operation is repeated as long as the partitioning results are not stable. At each level, the number of classes of the partition is automatically estimated. The optimal partition and the estimated number of classes are given by the partition that maximizes the Levine and Nazif criterion [64], denoted as LN . This procedure is detailed in Algorithm 4.

The method developed here for partitioning large size images, named Hierarchical Unsupervised Partitioning by Optimized Affinity Propagation (HUP-OAP), is composed of the three algorithms presented before: the generation of the First partition (Algorithm 3) by using the UP-OAP method (Algorithm 2) and the hierarchical partitioning (Algorithm 4). The main steps of this method are summarized in Algorithm 5.

Algorithm 3 First partition merging block classes

Input:

- Original image, I_m , or Data table to be partitioned
- Maximum block size ($Y_1 \times Y_2$) allowing application of the UP-OAP method

Procedure:

1. Dividing image I_m into N_B blocks B_{ij} , with $i \in \{1, 2, \dots, M_1\}$ and $j \in \{1, 2, \dots, M_2\}$, where M_1 is the number of blocks in a row and M_2 is the number of blocks in a column
2. Application of the UP-OAP method on each block
 - for** $i = 1$ to M_1 **do**
 - for** $j = 1$ to M_2 **do**
 - Partitioning of block B_{ij} by the UP-OAP method**
 - Let P_{ij} be the obtained partition of block B_{ij} :
 - $I_{ij} = \{I_{ij}^1, I_{ij}^2, \dots, I_{ij}^{N_{ij}}\}$, is the set of exemplars and N_{ij} is the number of classes of P_{ij}
 - $P_{ij} = \{C_{ij}^1, C_{ij}^2, \dots, C_{ij}^{N_{ij}}\}$
 - end for**
 - end for**
3. Merging classes of blocks B_{ij} by application of the UP-OAP method on the exemplars of blocks.
4. Formation of the partition P_1 : merging each object to its exemplar.

Output: Partition P of K classes and exemplar of each class

Algorithm 4 Hierarchical partitioning

Input: *Data table*, X_1 (N_1 exemplars \times B features), composed of the exemplars of the first partition P_1 of N_1 classes C_j

1. Application of UP-OAP on the dataset, X_1
2. **Repeat** UP-OAP on the new dataset X_i ($i \geq 2$)
 X_i is composed of the exemplar I_{i-1}^j of each class C_j of the partition P_{i-1} , with:

$$X_i = \{I_{i-1}^1, I_{i-1}^2, \dots, I_{i-1}^{N_{i-1}}\}$$
 Formation of the partition P_i : merging each object to its exemplar
Until the stability of the partition P_i
3. Choice of the optimal partition that maximizes the LN criterion:

$$P_{final} = \max_i [LN_i]$$

Output: The hierarchical partitions of the original dataset, the optimal partition, and a set of exemplars of its classes

Algorithm 5 HUP-OAP

Input: *Image or Data table to be partitioned*

1. Application of Algorithm 3 to obtain the first partition P_1 and its exemplars
2. Formation of the dataset X_1 composed of the exemplars of P_1
3. Application of Algorithm 4 on the dataset X_1

Output: The hierarchical partitions of the image, the optimal partition, and a set of exemplars of its classes

4. Numerical Assessment

In this section, we present the assessment of the proposed method on two hyperspectral images. The first one is a small synthetic image and the second one is a real aerial large size image. Through these assessments, the application addressed is the localization and the identification of marine algae classes.

4.1. Partitioning of a Synthetic Hyperspectral Image

The synthetic hyperspectral image presented in Figure 3 is used to assess the partitioning result performed by the proposed method. The size of this image is limited to 60×60 pixels (100 spectral bands), with wavelengths ranging from 404.2 nm to 978.5 nm. This image was generated from pixel samples of nine GT classes of a real hyperspectral image acquired by our aerial platform in 2013. The samples of each class are randomly selected from the GT data accompanying the real acquired hyperspectral image. The nine classes of this image can be aggregated into four main classes as detailed in Figure 4. The four classes of the GT are Water, Substrate, Algae, and Mixed class. The water class is composed of three sub-classes (Deep, Shallow, and Turbid), the substrate class is composed of two sub-classes (Pebble and Sand), and the algae class can be divided into three sub-classes (Ulva, Enteromorpha, and Fucus).

Figure 5 shows the optimal partitioning result (LN criterion: 0.25) obtained at level 2 by the proposed HUP-OAP method (Algorithm 5), where the estimated number of classes is ten. For this evaluation, the set of 100 features corresponding to the spectral signature of each pixel is considered and the image is divided into 16 blocks, where the size of each block is 15×15 pixels, to cover the whole image.



Figure 3. Synthetic hyperspectral image displayed in RGB mode.


GT image: 9 classes	Class label	Number of pixels
	■ Deep ■ Shallow ■ Turbid	989 158 177
	■ Pebble ■ Sand	281 786
	■ Ulva (green) ■ Enteromorpha (green) ■ Fucus (brown)	255 416 493
Mixed class	■ vegetation-substrate	45

Figure 4. GT classes of the synthetic hyperspectral image.

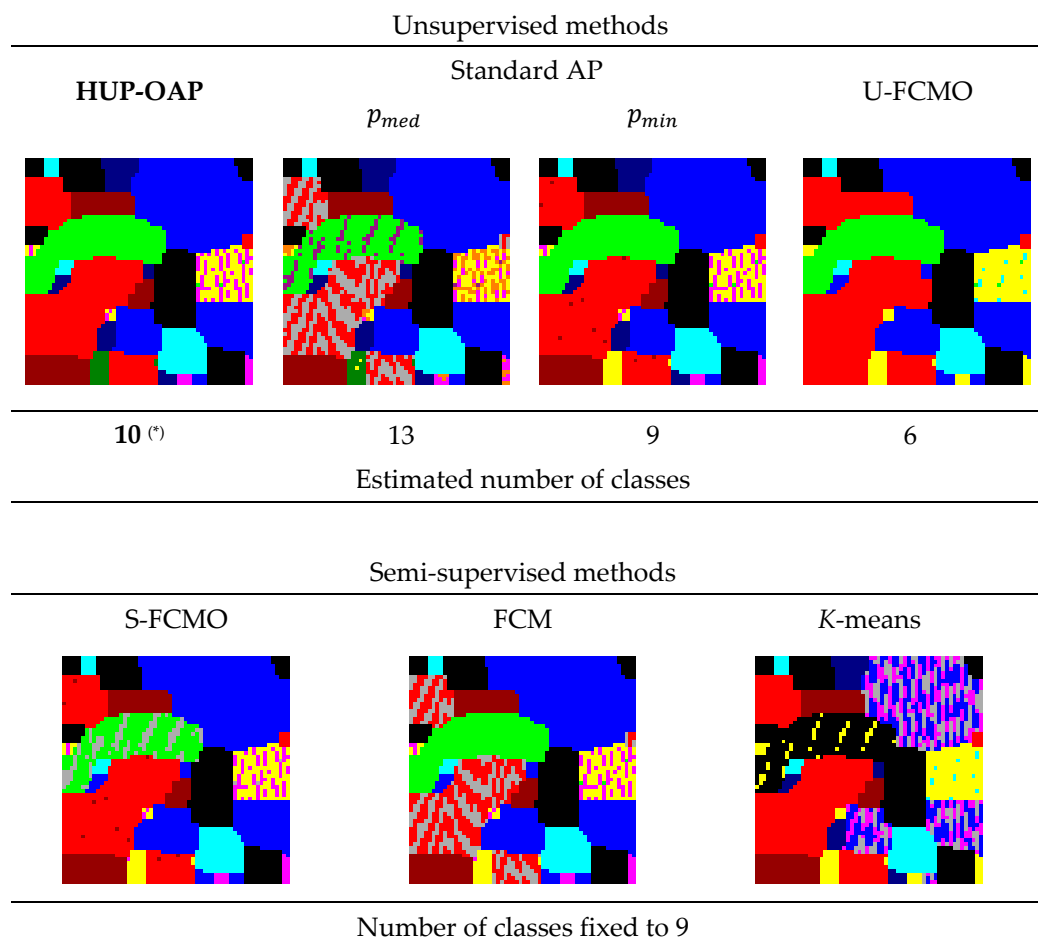


Figure 5. Partitioning results of the synthetic image by unsupervised and semi-supervised methods. (*) Optimal hierarchical partition.

The confusion matrix of Table 1 highlights the repartition of pixels of the GT classes in the classes formed by the HUP-OAP method. The quality of the partitioning result is evaluated according to the correct classification rate (CCR) criterion which is calculated from the confusion matrix as follows:

$$CCR = \frac{1}{Z} \times \left[\sum_{i=1}^{N_c} Z_i \right] \times 100 \quad (15)$$

where Z is the total number of GT points, N_c is the number of GT classes, and Z_i is the number of the GT points correctly classified in each class i of GT.

The CCR obtained by the proposed method is 96.89%. This rate can be corrected to 99.91% if we consider the homogeneity of class 10 formed only by a subset of pixels of the GT C_6 class as shown in Table 1. The example of the GT C_6 class illustrates perfectly the interest of an unsupervised method. Indeed, this class corresponds to a Green algae, but the proposed method clearly divides it into two subclasses. This is due to the fact that during the elaboration of the GT map, the two variants of this class were not specified. The classes 6 and 10 formed by the unsupervised HUP-OAP method correspond to Green algae classes, but with variations. For example, according to the depth of the water (the spectrum of Green algae through the water column), subclasses can be discriminated against. It is therefore thanks to the unsupervised nature of the proposed method that it is possible to objectively highlight the wealth of the information provided by hyperspectral imagery in the near-infrared (NIR) compared to that provided only in the visible domain. The class

10 formed may reflect, for example, the presence of algae in deeper water than the class 6 formed.

Table 1. Confusion matrix of partitioning result by HUP-OAP.




















		Classes Formed by HUP-OAP									
		1 	2 	3 	4 	5 	6 	7 	8 	9 	10 
GT Classes	C ₁ 	989	0	0	0	0	0	0	0	0	0
	C ₂ 	0	158	0	0	0	0	0	0	0	0
	C ₃ 	0	0	177	0	0	0	0	0	0	0
	C ₄ 	0	0	0	281	0	0	0	0	0	0
	C ₅ 	0	0	0	0	786	0	0	0	0	0
	C ₆ 	0	0	0	0	0	143	3	0	0	109
	C ₇ 	0	0	0	0	0	0	416	0	0	0
	C ₈ 	0	0	0	0	0	0	0	493	0	0
	C ₉ 	0	0	0	0	0	0	0	0	45	0
	CCR	96.89% (99.91% with homogenous class 10)									

Table 2 shows that for the image in Figure 3, the partitioning results are the same regardless of the block size. We can notice that the partitioning result of the image without division into blocks is identical to that obtained by division into blocks. We can also see that the smaller the block size, the lower the CPU time and memory space.

Table 2. HUP-OAP results according to the block size.

Number of Blocks	(Level, K)	CCR (%)	CPU Time (s)	Memory Space (Mb)
Full image (60 × 60 pixels)	(3, 10)	96.89	49.76	302.86
2 (30 × 60 pixels)	(2, 10)		20.2	153.54
4 (30 × 30 pixels)			10.77	79.23
8 (15 × 30 pixels)			7.08	41.93
16 (15 × 15 pixels)			3.73	23.19

In order to evaluate the performances of the developed method compared to others, we have chosen unsupervised methods (Standard AP and U-FCMO) and methods that require a minimum of a priori knowledge, i.e., the knowledge of the number of classes without any training sample (Stable FCMO, FCM, and K-means).

For the semi-supervised methods, the number of classes was set to 9 in order to match the number of classes of GT image, and for FCM, U-FCMO, and S-FCMO methods, the fuzzification parameter was fixed at 2. We specify that for the K-means and the FCM methods, the rates given are the average of five fluctuating results due to their non-stability with respect to initialization. For methods of the state of the art, the metric used to calculate the similarity matrix is the Euclidean distance (d_2).

Table 3 gives the performances of the three unsupervised and three semi-supervised methods by computing four criteria: CCR (%), CCR with homogenous output classes (%), CPU time (s), and Memory space (Mb).

Table 3. Performances of the developed method and the five other compared methods.

Methods	Unsupervised (Estimated Number of Classes)			Semi-Supervised (Fixed Number of Classes)			
	HUP-OAP	Standard AP		U-FCMO	S-FCMO	FCM (*)	K-Means (*)
		P_{med}	P_{min}				
Number of classes	10	13	9	6	9	9	9
CCR (%)	96.89	83.17	94.94	86.14	86.55	83.07	72.03
CCR with homogenous output classes (%)	99.91	97.83	98.38	86.14	93.71	84.80	86.85
CPU time (s)	3.73	61.88	72.01	35.63	4.46	64.73	52.82
Memory space (MB)	23.19	296.93	296.93	3.82	3.17	2.99	2.75

(*) Average rate of 5 CCR.

These results show that the developed method gives the best results according to three criteria (CCR, CCR with homogenous output classes, and CPU time), in addition to its unsupervised advantage. On the other hand, it requires more memory space than the U-FCMO, S-FCMO [26], FCM, and K-means methods because of the responsibility and availability matrices, but considerably less than the standard AP method. We can also note that the semi-supervised FCM and K-means methods give overall the least interesting results according to the CCR criterion. In addition, their results are not stable from one run to another, despite the introduction of the number of classes.

4.2. Partitioning of a Real Large Size Hyperspectral Aerial Image

The objective of this experimental database is to identify two main algae species (Green and Brown) and to provide an accurate mapping of their coverage rate.

We specify that the HUP-OAP method in its design is developed for partitioning large-size images which can be larger than the example treated here.

The large-size image (630×1800 pixels) of Figure 6a used for this assessment was acquired on 27 May 2013 (part of the French seashore) using the AISA Eagle sensor integrated in the aerial acquisition platform available at the TSI2M Laboratory. The ground spatial resolution of this image is 0.6 m, and the number of spectral bands is 100, covering the V-NIR spectral range from 404.2 nm to 978.5 nm.

To allow the evaluation and validation of the results of the proposed unsupervised HUP-OAP method, we used a field campaign performed at the same time as the aerial survey. The field spectra measurements were acquired with a spectroradiometer coupled with a GPS. After this step, the GT points were validated [21,22], where only ground points with similar spectral signature to their corresponding pixels in the original aerial hyperspectral image were selected.

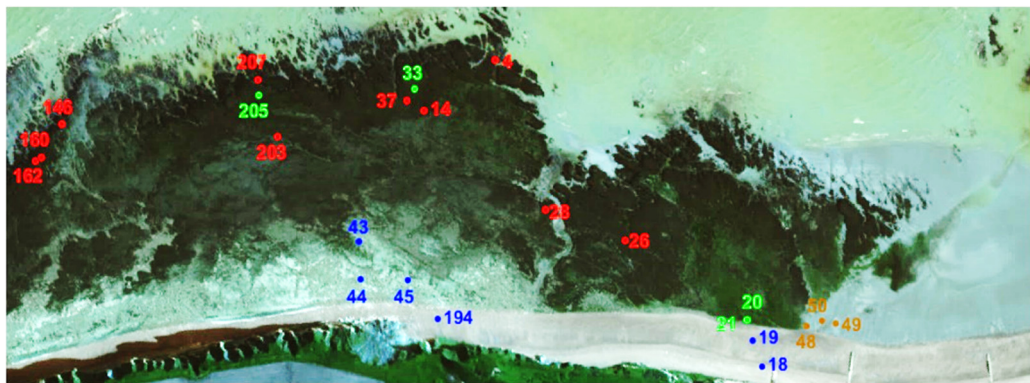
This example proves that it is impossible in practice to elaborate a GT for all the pixels of a large-size image. For this reason, we limited ourselves to a few survey points in the field to assess and validate the unsupervised partitioning method developed in this paper.

Figure 6b shows the location of the field measurements of four classes over the original hyperspectral image. Figure 7 highlights the validated GT spectral signature and average \pm standard deviation (in luminance) of these four main classes: Brown algae, Green algae, Rocks and Pebbles, and Sand. These last two classes can be aggregated into a single substrate class.

To partition this large size hyperspectral image (630×1800 pixels \times 100 spectral bands: the size of the data table is 1,134,000 pixels \times 100 features) by the HUP-OAP method (Algorithm 5), the chosen size of each block is 63×90 pixels (200 blocks). Figure 8 shows the optimal partitioning result of the hyperspectral image in Figure 6a maximizing the LN criterion which is obtained at level 4. The estimated number of classes for this partition is 5. Table 4 gives for each partitioning level the number of classes estimated by the HUP-OAP method and the value of the LN optimization criterion.



(a)



(b)

- Green algae (4 points: 20, 21, 33, 205)
- Brown algae (10 points: 4, 14, 26, 28, 37, 146, 160, 162, 203, 207)
- Substrate—Sand (3 points: 48, 49, 50)
- Substrate—Rocks and Pebbles (6 points: 18, 19, 43, 44, 45, 194)

Figure 6. Hyperspectral image 630×1800 pixels (100 bands) displayed in RGB mode and GT points: (a) Original hyperspectral image; (b) Locations of field survey points over the original hyperspectral image.

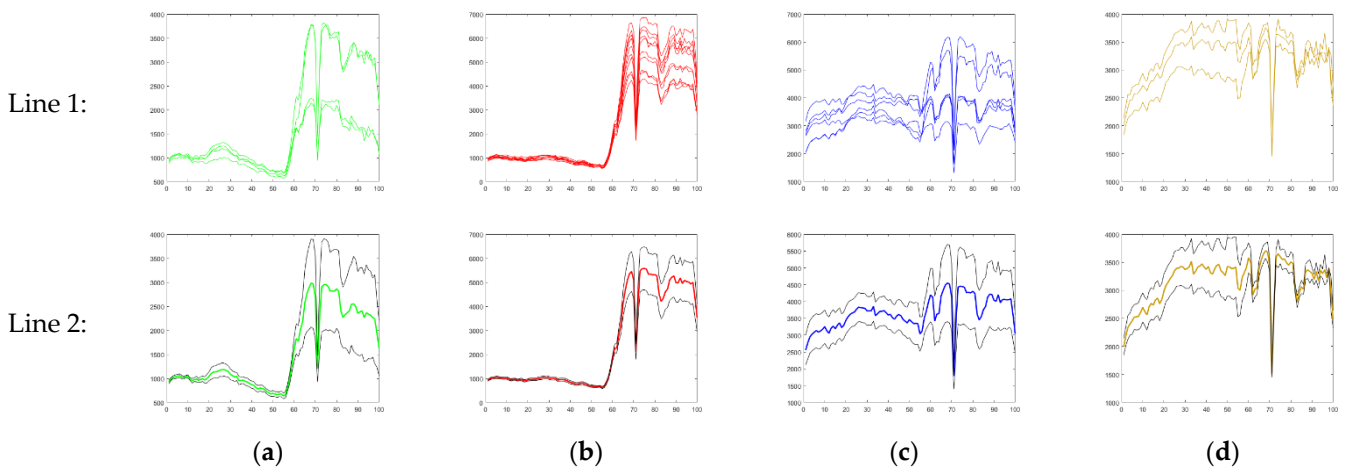
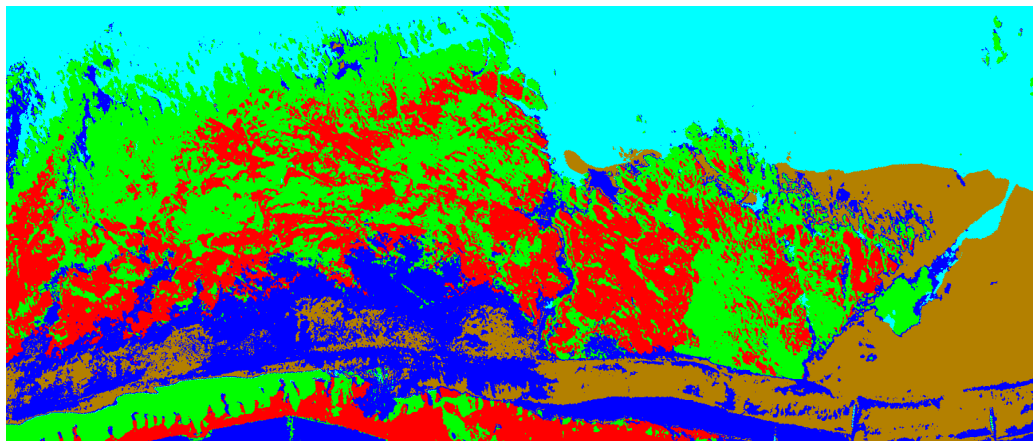


Figure 7. Spectral signatures of GT points by class. Line 1: (a) Green algae (*Ulva armoricana*); (b) Brown algae (*Fucus serratus*); Substrate ((c) Rocks and Pebbles and (d) Sand), and Line 2: corresponding average spectral signature \pm standard deviation of each class.



- Green algae (4 points: 20, 21, 33, 205)
- Brown algae (10 points: 4, 14, 26, 28, 37, 146, 160, 162, 203, 207)
- Water
- Rocks and Pebbles (6 points: 18, 19, 43, 44, 45, 194)
- Sand (3 points: 48, 49, 50)

Figure 8. Optimal partitioning result obtained by the HUP-OAP method (5 classes localized, $LN = 0.17$ at level 4).

Table 4. Estimated number of classes by HUP-OAP method per partitioning level and values of optimization criterion for each partition.

Level	Estimated Number of Classes	LN Criterion Value
1	1162	0.029
2	181	0.038
3	33	0.07
4	5	0.17
5	2	0.14

The GT points of the four classes (Green algae, Brown algae, Rocks and Pebbles, and Sand) belong to the four different formed classes. This result highlights a fifth class, whose spectral signature corresponds to that of water. We observe in Figure 9 that the average spectral signature \pm standard deviation of each formed class differs from the others and can be used as reference learning samples. The optimal partition of level 4 shows that the CCR is 100%, by checking the positions of the 23 points of the four GT classes within the formed classes.

The method thus developed gives, in addition to the optimal partitioning, other partitions which can contribute to the fine analysis and interpretation of the data according to the users' needs. It is also important to stress that the several experiments conducted show that the partitioning result is independent of the choice of block size.

Based on the reference spectral signatures, the algae coverage rate given by the optimal partition corresponds to 44.61% (19% for Brown algae and 25.61% for Green algae), as shown in Table 5.

Table 5. Coverage rate of each class obtained at level 4 by HUP-OAP.

Classes	Number of Pixels	Coverage Rate (%)
Green algae	290,452	25.61
Brown algae	215,509	19.00
Rocks and Pebbles	180,802	15.94
Sand	170,617	15.05
Water	276,620	24.39

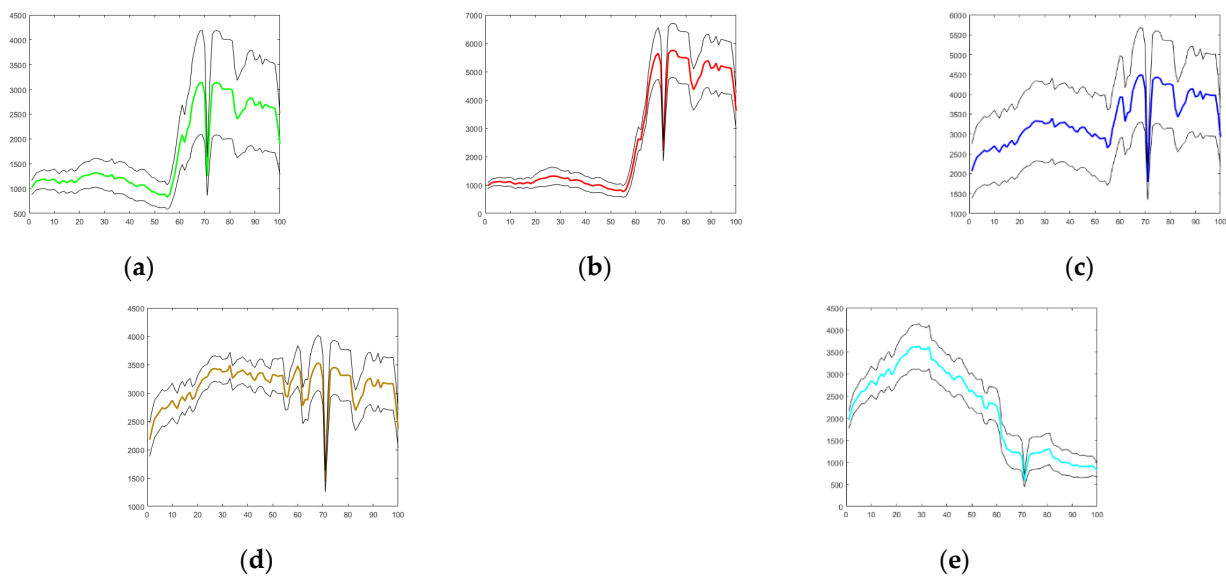


Figure 9. Average spectral signature \pm standard deviation of each class of Figure 8 obtained by HUP-OAP: (a) Green algae; (b) Brown algae; (c) Rocks and Pebbles; (d) Sand; (e) Water.

The computing time and memory space for partitioning this image with the proposed method (data table of size: $1.134.000 \text{ pixels} \times 100 \text{ features}$) on an Intel(R) Core (TM) i7-7700 CPU processor with 3.6 GHz and 16 Go memory are given in Table 6 for two different block sizes. We can see that they decrease with the size of the blocks. The indicative time given here can be greatly reduced because the block partitioning can be done in parallel, on a multiprocessor machine.

Table 6. CPU time and Memory space of partitioned image of Figure 6 by HUP-AOP.

Number of Blocks	CPU Time (s)	Memory Space (Mb)
200 ($63 \times 90 \text{ pixels}$)	4932.67	148,850
5040 ($15 \times 15 \text{ pixels}$)	982.60	837

In comparison (see Table 7), the developed unsupervised method yields better performances with respect to FCMO [26] in its unsupervised and semi-supervised versions, denoted as U-FCMO and S-FCMO respectively, and to the semi-supervised K -means and FCM methods. The number of classes for these last three methods has been set to five and the metric used to calculate the similarity matrix is the Euclidean distance (d_2). We recall that the standard AP algorithm cannot be applied to partition this large-size image.

Table 7. Performance of the developed method, HUP-OAP, U-FCMO, S-FCMO, FCM, and K -means.

Methods	Number of Classes	CCR (%)
HUP-OAP	5 (estimated)	100
U-FCMO	4 (estimated)	86.85
S-FCMO	5 (fixed)	91.30
FCM	5 (fixed)	79.12 (*)
K -means	5 (fixed)	78.25 (*)

(*) Average rate of 5 CCR.

The analysis of the U-FCMO result shows that the number of estimated classes correspond to the Green algae, Brown algae, Substrate, and Water classes. In this case, the GT points belonging to the Rocks and Pebbles and Sand are aggregated in the same class. This means that the GT points 48, 49, and 50 of the sand class are misclassified, which

gives a rate of 86.95%. If we do not take into account the discrimination between these two Substrate classes, the rate can therefore be 100%. However, this result is less accurate than that of the HUP-OAP method which gives more detail by splitting the substrate class into two subclasses. Another interesting piece of information is given by the partition of level 5, where the algae classes are merged into one class and the others (substrates and water) in another class. Furthermore, the user can use the other partitions of the hierarchy for more details. In the case of the S-FCMO method, the GT points 19 and 44 of the Rocks and Pebbles class were aggregated in the Sand class, which gives a CCR of 91.30%.

5. Conclusions

In this paper, we presented a new unsupervised hierarchical partitioning method adapted to large-size datasets, such as hyperspectral aerial imagery. This method has eight main advantages that can be objectively listed: (1) no a priori knowledge is required, especially the non-introduction of training samples, to give the user the possibility to detect and locate known classes and new classes called “discovery classes”; (2) stability of the results thanks to its deterministic character; (3) selection of the exemplar of each class and the assignment of a pixel (or an object) to a class are done in a very elaborate way according to optimization criteria; (4) very low computing time with block processing, in contrast to the compared methods; (5) applicable to data or images whatever their size with the possibility of parallelizing the block partitioning; (6) possibility of elaborating several hierarchical partitions by indicating the most relevant one according to an objective criterion; (7) possibility of objectively selecting the samples of the classes in a learning system in order to be able to detect them afterwards; finally, (8) applicable to several domains without learning constraints. However, it requires more memory space.

Evaluations of the developed method on synthetic and real hyperspectral images show that the results are relevant without any intervention of the end-user, and its application to large-size images gives the same optimal result regardless of the block size used.

The correct classification rates (CCR) obtained by our method are better than those of semi-supervised methods such as Stable FCM (S-FCMO), *K*-means, and FCM, despite its unsupervised character (estimation of the number of classes and classification of pixels without any a priori knowledge). It also outperforms compared unsupervised methods such as U-FCMO and the standard AP method.

This true unsupervised method meets the partitioning requirements of large-size images provided by modern hyperspectral sensors. Moreover, it can be applied to a wide range of applications to objectively highlight all existing classes in an image [22]. From this complete partition, the user can exploit all or some of the obtained classes. It also offers the possibility to use the samples of the classes in learning processes.

In perspective, this method will be assessed and validated on other databases extended to several application fields in order to prove the relevance and the benefits of its unsupervised operating mode. An optimization in memory space also remains to be performed.

Author Contributions: Conceptualization, K.C.; methodology, K.C. and J.A.; software, J.A.; validation, K.C.; formal analysis, J.A. and K.C.; investigation, J.A. and K.C.; resources, K.C.; data curation, J.A. and K.C.; writing—original draft preparation, K.C.; writing—review and editing, K.C.; visualization, J.A. and K.C.; supervision, K.C.; project administration, K.C.; funding acquisition, K.C.; advice during the coding and review of the manuscript, C.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the department of Côtes d’Armor in Brittany (France) under Grant INR00131, and SHINE-TSI2M team.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: This work was supported by the department of Côtes d’Armor in Brittany (France), who co-financed this study (Project PANSIADÉ). The authors thank the ‘Electricité de France’ (EDF) for providing hyperspectral images and the ‘Centre d’Étude et de Valorisation des Algues’ (CEVA) for providing the ground truth data.

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

References

- Lobo, A.; Garcia, E.; Barroso, G.; Martí, D.; Fernandez-Turiel, J.-L.; Ibáñez-Insa, J. Machine Learning for Mineral Identification and Ore Estimation from Hyperspectral Imagery in Tin–Tungsten Deposits: Simulation under Indoor Conditions. *Remote Sens.* **2021**, *13*, 3258. [[CrossRef](#)]
- Ortega, S.; Fabelo, H.; Iakovidis, D.K.; Koulaouzidis, A.; Callico, G.M. Use of Hyperspectral/Multispectral Imaging in Gastroenterology. Shedding Some–Different–Light into the Dark. *J. Clin. Med.* **2019**, *8*, 36. [[CrossRef](#)] [[PubMed](#)]
- Jansen-Winkeln, B.; Barberio, M.; Chalopin, C.; Schierle, K.; Diana, M.; Köhler, H.; Gockel, I.; Maktabi, M. Feedforward Artificial Neural Network-Based Colorectal Cancer Detection Using Hyperspectral Imaging: A Step towards Automatic Optical Biopsy. *Cancers* **2021**, *13*, 967. [[CrossRef](#)]
- Felli, E.; Al-Taher, M.; Collins, T.; Nkusi, R.; Felli, E.; Baiocchi, A.; Lindner, V.; Vincent, C.; Barberio, M.; Geny, B.; et al. Automatic Liver Viability Scoring with Deep Learning and Hyperspectral Imaging. *Diagnostics* **2021**, *11*, 1527. [[CrossRef](#)]
- Moulla, Y.; Buchloh, D.C.; Köhler, H.; Rademacher, S.; Denecke, T.; Meyer, H.-J.; Mehdorn, M.; Lange, U.G.; Sucher, R.; Seehofer, D.; et al. Hyperspectral Imaging (HSI)—A New Tool to Estimate the Perfusion of Upper Abdominal Organs during Pancreatoduodenectomy. *Cancers* **2021**, *13*, 2846. [[CrossRef](#)] [[PubMed](#)]
- Pallua, J.D.; Brunner, A.; Zelger, B.; W Huck, C.; Schirmer, M.; Laimer, J.; Putzer, D.; Thaler, M.; Zelger, B. New perspectives of hyperspectral imaging for clinical research. *NIR News.* **2021**, *32*, 5–13. [[CrossRef](#)]
- Räsänen, J.; Salmivuori, M.; Pölönen, I.; Grönroos, M.; Neittaanmäki, N. Hyperspectral Imaging Reveals Spectral Differences and Can Distinguish Malignant Melanoma from Pigmented Basal Cell Carcinomas: A Pilot Study. *Acta Derm. Venereol.* **2021**, *101*, 00405. [[CrossRef](#)]
- Cucuzza, P.; Serranti, S.; Bonifazi, G.; Capobianco, G. Effective Recycling Solutions for the Production of High-Quality PET Flakes Based on Hyperspectral Imaging and Variable Selection. *J. Imaging* **2021**, *7*, 181. [[CrossRef](#)] [[PubMed](#)]
- Englert, T.; Gruber, F.; Stiedl, J.; Green, S.; Jacob, T.; Rebner, K.; Grählert, W. Use of Hyperspectral Imaging for the Quantification of Organic Contaminants on Copper Surfaces for Electronic Applications. *Sensors* **2021**, *21*, 5595. [[CrossRef](#)]
- Wen, W.; Timmermans, J.; Chen, Q.; van Bodegom, P.M. A Review of Remote Sensing Challenges for Food Security with Respect to Salinity and Drought Threats. *Remote Sens.* **2021**, *13*, 6. [[CrossRef](#)]
- Bue, B.D.; Thompson, D.R.; Sellar, R.G.; Podest, E.V.; Eastwood, M.L.; Helmlinger, M.C.; McCubbin, I.B.; Mprgan, J.D. Leveraging in-scene spectra for vegetation species discrimination with MESMA-MDA. *ISPRS J. Photog. Rem. Sens.* **2015**, *108*, 33–48. [[CrossRef](#)]
- Dudley, K.L.; Dennison, P.; Roth, K.; Roberts, D.; Coates, A.R. A multi-temporal spectral library approach for mapping vegetation species across spatial and temporal phenological gradients. *Remote Sens. Environ.* **2015**, *167*, 121–134. [[CrossRef](#)]
- Kong, W.; Zhang, C.; Huang, W.; Liu, F.; He, Y. Application of hyperspectral imaging to detect sclerotinia sclerotiorum on oilseed rape stems. *Sensors* **2018**, *18*, 123. [[CrossRef](#)]
- Schmitter, P.; Steinrücken, J.; Römer, C.; Ballvora, A.; Léon, J.; Rascher, U.; Plümer, L. Unsupervised domain adaptation for early detection of drought stress in hyperspectral images. *ISPRS J. Photog. Remote Sens.* **2017**, *131*, 65–76. [[CrossRef](#)]
- Peerbhay, K.; Mutanga, O.; Lottering, R.; Bangamwabo, V.; Ismail, R. Detecting bugweed (*Solanum mauritianum*) abundance in plantation forestry using multisource remote sensing. *ISPRS J. Photog. Remote Sens.* **2016**, *121*, 167–176. [[CrossRef](#)]
- Stagakis, S.; Vanikiotis, T.; Sykioti, O. Estimating forest species abundance through linear unmixing of CHRIS/PROBA imagery. *ISPRS J. Photog. Remote Sens.* **2016**, *119*, 79–89. [[CrossRef](#)]
- Mogstad, A.A.; Johnsen, G. Spectral characteristics of coralline algae: A multi-instrumental approach, with emphasis on underwater hyperspectral imaging. *Appl. Opt.* **2017**, *56*, 9957–9975. [[CrossRef](#)]
- Mehrubeoglu, M.; Teng, M.Y.; Zimba, P.V. Resolving mixed algal species in hyperspectral images. *Sensors* **2014**, *14*, 1–21. [[CrossRef](#)] [[PubMed](#)]
- Lopatin, J.; Lopatin, J.; Fassnacht, F.E.; Kattenborn, T.; Schmidlein, K. Mapping plant species in mixed grassland communities using close range imaging spectroscopy. *Remote Sens. Environ.* **2017**, *201*, 12–23. [[CrossRef](#)]
- Walsh, S.J.; McCleary, A.L.; Mena, C.; Shao, Y.; Tuttle, J.P.; Gonzalez, A.; Atkinson, R. QuickBird and Hyperion data analysis of an invasive plant species in the Galapagos Islands of Ecuador: Implications for control and land use management. *Remote Sens. Environ.* **2008**, *112*, 1927–1941. [[CrossRef](#)]
- Chehdi, K.; Cariou, C. The true false ground truths: What interest? In Proceedings of the SPIE 10004, Image and Signal Processing for Remote Sensing XXII, Edinburgh, UK, 26–29 September 2016; pp. 1–16.
- Chehdi, K.; Cariou, C. Learning or assessment of classification algorithms relying on biased ground truth data: What interest? *J. Appl. Remote Sens.* **2019**, *13*, 1–26. [[CrossRef](#)]
- Luo, F.; Huang, H.; Ma, Z.; Liu, J. Semisupervised Sparse Manifold Discriminative Analysis for Feature Extraction of Hyperspectral Images. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 6197–6211. [[CrossRef](#)]

24. Barile, C.; Casavola, C.; Pappaletta, G.; Paramsamy Kannan, V. Laplacian score and K-means data clustering for damage characterization of adhesively bonded CFRP composites by means of acoustic emission technique. *Appl. Acoust.* **2021**, *185*, 108425. [[CrossRef](#)]
25. Chehdi, K.; Soltani, M.; Cariou, C. Pixel classification of large-size hyperspectral images by affinity propagation. *J. Appl. Remote Sens.* **2014**, *8*, 1–4. [[CrossRef](#)]
26. Chehdi, K.; Taher, A.; Cariou, C. Stable and unsupervised fuzzy C-means method and its validation in the context of multicomponent images. *J. Electron. Imaging* **2015**, *24*, 1–6. [[CrossRef](#)]
27. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum Likelihood from Incomplete Data via the EM Algorithm. *J. R. Stat. Soc. Ser. B Methodol.* **1977**, *39*, 1–38.
28. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
29. MacQueen, J. Some methods for classification and analysis of multivariate observations. In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability; University of California Press: Berkeley, CA, USA, 1967; pp. 281–297.
30. Peña, J.M.; Lozano, J.A.; Larrañaga, P. An empirical comparison of four initialization methods for the K-Means algorithm. *Pat. Recog. Lett.* **1999**, *20*, 1027–1040. [[CrossRef](#)]
31. Bubeck, S.; Meilă, M.; Luxburg, U.V. How the initialization affects the stability of the κ -means algorithm. *ESAIM Prob. Stat.* **2012**, *16*, 436–452. [[CrossRef](#)]
32. Xu, R.; Wunsch, D. Survey of clustering algorithms. *IEEE Trans. Neur. Netw.* **2005**, *16*, 645–678. [[CrossRef](#)]
33. Linde, Y.; Buzo, A.; Gray, R. An Algorithm for Vector Quantizer Design. *IEEE Trans. Commun.* **1980**, *28*, 84–95. [[CrossRef](#)]
34. Huang, B.; Xie, L. An improved LBG algorithm for image vector quantization. In Proceedings of the 3rd International Conference on Computer Science and Information Technology, Chengdu, China, 9–11 July 2010; pp. 467–471.
35. Fritzke, B. The LBG-U method for vector quantization—an improvement over LBG inspired from neural networks. *Neur. Process. Lett.* **1997**, *5*, 35–45. [[CrossRef](#)]
36. Patané, G.; Russo, M. The enhanced LBG algorithm. *Neur. Netw.* **2001**, *14*, 1219–1237. [[CrossRef](#)]
37. Li, X.; Lu, X.; Tian, J.; Gao, P.; Kong, H.; Xu, G. Application of fuzzy c-means clustering in data analysis of metabolomics. *Anal. Chem.* **2009**, *81*, 4468–4475. [[CrossRef](#)]
38. Bezdek, J.C.; Ehrlich, R.; Full, W. FCM: The fuzzy c-means clustering algorithm. *Comput. Geosci.* **1984**, *10*, 191–203. [[CrossRef](#)]
39. Dunn, J.C. A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters. *J. Cybern.* **1973**, *3*, 32–57. [[CrossRef](#)]
40. Bezdek, J.C. *Pattern Recognition with Fuzzy Objective Function Algorithms*; Springer: New York, NY, USA; London, UK, 1981.
41. Rosenberger, C.; Chehdi, K. Unsupervised clustering method with optimal estimation of the number of clusters: Application to image segmentation. In Proceedings of the 15th International Conference on Pattern Recognition (ICPR-2000), Barcelona, Spain, 3–7 September 2000; pp. 656–659.
42. Frey, B.J.; Dueck, D. Clustering by Passing Messages Between Data Points. *Science* **2007**, *315*, 972–976. [[CrossRef](#)] [[PubMed](#)]
43. Hang, W.; Chung, F.; Wang, S. Transfer Affinity Propagation-based Clustering. *Inf. Sci.* **2016**, *348*, 337–356. [[CrossRef](#)]
44. Lerm, S.; Saeedi, A.; Rahm, E. Extended Affinity Propagation Clustering for Multi-source Entity Resolution. *Lect. Notes Inform.* **2021**, *P-311*, 1–20.
45. Park, S.; Jo, H.-S.; Mun, C.; Yook, J.-G. RRH Clustering Using Affinity Propagation Algorithm with Adaptive Thresholding and Greedy Merging in Cloud Radio Access Network. *Sensors* **2021**, *21*, 480. [[CrossRef](#)]
46. Cardille, J.A.; White, J.C.; Wulder, M.A.; Holland, T. Representative Landscapes in the Forested Area of Canada. *Environ. Manag.* **2012**, *49*, 163–173. [[CrossRef](#)]
47. Li, P.; Ji, H.; Wang, B.; Huang, Z.; Li, H. Adjustable preference affinity propagation clustering. *Pat. Recog. Lett.* **2017**, *85*, 72–78. [[CrossRef](#)]
48. Shang, F.; Jiao, L.C.; Shi, J.; Wang, F.; Gong, M. Fast affinity propagation clustering: A multilevel approach. *Pat. Recog.* **2012**, *45*, 474–486. [[CrossRef](#)]
49. Xia, D.; Wu, F.; Zhang, X.; Zhuang, Y. Local and global approaches of affinity propagation clustering for large scale data. *J. Zhejiang Univ. Sci. A* **2008**, *9*, 1373–1381. [[CrossRef](#)]
50. Yang, Y.; Ha, H.; Fleites, F.C.; Chen, S. A Multimedia Semantic Retrieval Mobile System Based on HCFGs. *IEEE MultiMedia* **2014**, *21*, 36–46. [[CrossRef](#)]
51. Xiao, H.; Guo, P. Iris Image Analysis Based on Affinity Propagation Algorithm. In Proceedings of the Advances in Neural Networks (ISNN), Wuhan, China, 26–29 May 2009; pp. 943–949.
52. Zhou, T.; Qi, M.; Jiang, J.; Wang, X.; Hao, S.; Jin, Y. Person Re-identification based on nonlinear ranking with difference vectors. *Inf. Sci.* **2014**, *279*, 604–614. [[CrossRef](#)]
53. Wang, C.; Lai, J.; Suen, C.Y.; Zhu, J. Multi-Exemplar Affinity Propagation. *IEEE Trans. Pattern Anal. Mach. Int.* **2013**, *35*, 2223–2237. [[CrossRef](#)] [[PubMed](#)]
54. Zha, Z.; Yang, L.; Mei, T.; Wang, M.; Wang, Z. Visual query suggestion. In Proceedings of the 17th ACM International Conference on Multimedia-MM '09, Beijing, China, 19–24 October 2009; pp. 15–24.
55. Lindorff-Larsen, K.; Ferkinghoff-Borg, J. Similarity Measures for Protein Ensembles. *PLoS ONE* **2009**, *4*, e4203. [[CrossRef](#)]
56. Gan, G.; Ng, M.K.-P. Subspace clustering using affinity propagation. *Pat. Recog.* **2015**, *48*, 1455–1464. [[CrossRef](#)]

57. Zhu, Z.; Jia, S.; Ji, Z. Towards a Memetic Feature Selection Paradigm [Application Notes]. *IEEE Comput. Intell. Mag.* **2010**, *5*, 41–53. [[CrossRef](#)]
58. Guo, K.; Guo, W.; Chen, Y.; Qiu, Q.; Zhang, Q. Community discovery by propagating local and global information based on the MapReduce model. *Inf. Sci.* **2015**, *323*, 73–93. [[CrossRef](#)]
59. Taheri, S.; Bouyer, A. Community Detection in Social Networks Using Affinity Propagation with Adaptive Similarity Matrix. *Big Data* **2020**, *8*, 189–202. [[CrossRef](#)] [[PubMed](#)]
60. Bi, X.; Guo, B.; Shi, L.; Lu, Y.; Feng, L.; Lyu, Z.A. New Affinity Propagation Clustering Algorithm for V2V-Supported VANETs. *IEEE Access* **2020**, *8*, 71405–71421. [[CrossRef](#)]
61. Wang, L.; Sun, W.; Han, X.; Hao, Z.; Zhou, R.; Yu, J.; Parmar, M. An Improved Integrated Clustering Learning Strategy Based on Three-Stage Affinity Propagation Algorithm with Density Peak Optimization Theory. *Complexity* **2021**, *2021*, 6666619. [[CrossRef](#)]
62. Bandi, A.; Joshi, K.; Mulwad, V. Affinity Propagation Initialisation Based Proximity Clustering for Labeling in Natural Language Based Big Data Systems. In Proceedings of the 2020 IEEE 6th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing, (HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS), Baltimore, MD, USA, 25–27 May 2020; pp. 1–7.
63. Alameddine, J.; Chehdi, K.; Cariou, C. Optimization of unsupervised affinity propagation clustering method. In Proceedings of the Image and Signal Processing for Remote Sensing XXV, Strasbourg, France, 9–12 September 2019; pp. 115–124.
64. Levine, M.D.; Nazif, A.M. Dynamic Measurement of Computer Generated Image Segmentations. *IEEE Trans. Pattern Anal. Mach. Intell.* **1985**, *7*, 155–164. [[CrossRef](#)] [[PubMed](#)]