# A MULTI-SENSOR SUBSPACE-BASED CLUSTERING ALGORITHM USING RGB AND HYPERSPECTRAL DATA

*Kasra Rafiezadeh Shahi[1,2], Pedram Ghamisi[1,2], Robert Jackisch[1], Behnood Rasti[1], Paul Scheunders[2], Richard Gloaguen[1]*

[1] Helmholtz-Zentrum Dresden-Rossendorf (HZDR),
Helmholtz Institute Freiberg for Resource Technology (HIF), Germany
[2] Imec-Visionlab, department of Physics, University of Antwerp, Belgium

## ABSTRACT

In this work, we introduce a multi-sensor subspace-based clustering algorithm that benefits from fine spectral-resolution hyperspectral images (HSIs) and fine spatial-resolution RGB images. In order to extract spatial information, a hidden Markov random field (HMRF) is employed on the fine spatial-resolution RGB image, whereas, spectral information is derived from an HSI using an advanced sparse subspace clustering algorithm. The proposed algorithm is validated on two real geological data sets. The experimental results in this study show that the proposed algorithm outperforms the state-of-the-art clustering algorithms in terms of clustering accuracy.

***Index Terms***— Hyperspectral images, RGB images, UAV data, hidden Markov random field, spectral-spatial clustering, sparse representation, data fusion

## 1. INTRODUCTION

The recent advances in imaging technologies enable to acquire high spectral-spatial resolution data using different sensors. Among the different sensors, hyperspectral sensors become the main source to acquire fine spectral-resolution data in hundreds of narrow spectral bands in a broad spectrum range covering VNIR, SWIR, and LWIR [1]. As a result, a hyperspectral image (HSI) provides valuable information to identify, and discriminate different materials and objects. Nevertheless, due to (1) its high dimensionality and (2) the highly mixed nature of an HSI, the analysis of such data can be complicated [2]. Therefore, recently, (supervised and unsupervised) machine learning algorithms have been proposed to address the aforementioned problems. Among those algorithms, sparse representation-based algorithms have received much attention due to their empirical success in many fields (e.g., feature fusion [3] and image segmentation [4, 5]). Sparse subspace clustering (SSC) is a well-known clustering algorithm that benefits from the so-called *self-expressiveness* property, in which, data points are written as a linear combination of other data points from the same subspace [5]. However, SSC has some drawbacks: (1) it requires the entire data set to calculate sparse coefficients, and (2) it uses only spectral information [4]. In [4], the authors proposed a fast exemplar sparse subspace-based clustering algorithm (ESC). ESC utilizes a search function to select a subset of representative samples (atoms) and uses this subset to cluster the HSI. Although ESC provides fast clustering results, similar to SSC, the algorithm only uses spectral information.

In different studies [6, 7], the influence of spatial information by fusing multi-sensor based data sets, is investigated. A Hidden Markov Random Field (HMRF) is a stochastic process generated by a Markov Random Field (MRF), that can be used as a powerful tool to extract spatial information from images [7]. The general objective of HMRF is to encourage a local neighborhood of pixels to belong to the same class type [8]. However originally, the HMRF was proposed for single-sensor based data sets.

In this work, we propose a multi-sensor HMRF subspace-based clustering algorithm to accurately analyze multi-sensor data sets. Our proposed algorithm applies the ESC algorithm to extract spectral information from an HSI. By employing HMRF using a fine spatial-resolution RGB image, spatial information is incorporated in the clustering procedure.

The rest of the paper is structured as follows: Section 2 contains the applied methodology. In section 3, the data description, the experimental results, and discussions are presented. The conclusions are drawn in section 4.

## 2. METHODOLOGY

In the following section, the proposed algorithm is described in detail. The workflow of the proposed algorithm is presented in Fig. 1.

### 2.1. Notation

Let us denote an HSI as $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N]^T \in \mathbb{R}^{\mathcal{F} \times N}$, where $\mathcal{F}$ is the number of spectral channels and $N$ is the number of samples in $\mathbf{X}$. $\mathbf{x}_i = [x_{i1}, x_{i2}, ..., x_{i\mathcal{F}}]^T$ represents

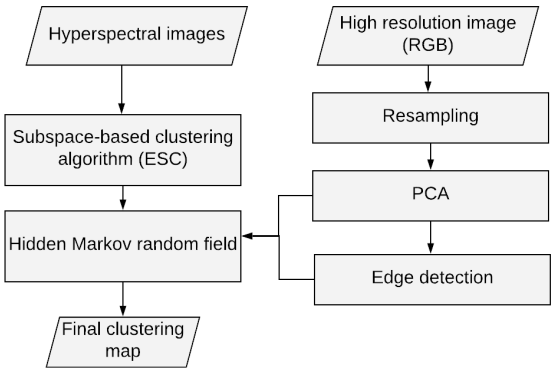


**Fig. 1**: The flowchart of the proposed algorithm

a pixel as a spectral vector and $i \in \{1, 2, ..., N\}$ is the pixel index. Let us denote an RGB image as $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_{N1}]^T \in \mathbb{R}^{3\times N1}$, where $N1$ is the number of samples in the RGB image, $\mathbf{RE} = [\mathbf{re}_1, \mathbf{re}_2, ..., \mathbf{re}_N]^T \in \mathbb{R}^{3\times N}$ represents the resampled $\mathbf{Y}$ with similar pixel size as $\mathbf{X}$. When applying PCA on $\mathbf{RE}$, the first principal component can be written as $\mathbf{Z} = [z_1, z_2, ..., z_N]$. $\mathcal{X}_0 = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_P]^T \in \mathbb{R}^{\mathcal{F}\times P}$ is a subset of a predefined number $P$ ($P << N$) of representative samples from $\mathbf{X}$. The clustering result can be expressed as $\mathbf{M} = [m_1, m_2, ..., m_N]$, where $m_i \in \mathcal{L}$, and $\mathcal{L} = \{1, 2, ..., l\}$ are the class label indices. $\mathbf{G} = [g_1, g_2, ..., g_N]$, where $g_i \in \mathcal{L}$ are ground truth labels for each pixel.

### 2.2. Exemplar-based subspace clustering (ESC)

In ESC [4], a subset of $P$ representative samples, also known as "*exemplars*", are selected to shape the sparse coefficient matrix. The optimization problem used in ESC can be formulated as:

$$\min_{\mathbf{C},\mathcal{X}_0} \quad ||\mathbf{C}||_1 + \frac{\lambda}{2}||\mathbf{X} - \mathcal{X}_0\mathbf{C}||_F^2 \tag{1}$$

where $\mathbf{C} = [\mathbf{c}_1, \mathbf{c}_2, ..., \mathbf{c}_N] \in \mathbb{R}^{P\times N}$ is the sparse coefficient matrix and $\lambda$ is a trade-off parameter between the penalty and the fidelity term. In order to initiate the clustering procedure, a random sample is selected. Subsequently the rest of the representative samples are selected based on a farthest first search method, that uses Eq.(1) as a cost function.

The obtained $\mathbf{C}$ from Eq.(1) is used to compute a similarity graph, which can be represented as $\mathbf{W} \in \mathbb{R}^{N\times N}$. For this, a $t$-nearest neighbor graph is constructed on $\mathbf{C}$ using the k-d tree algorithm [4]. From this graph, $W_{ij}$ is 1 if $\mathbf{c}_j$ is a $t$-nearest neighbor of $\mathbf{c}_i$ and 0 otherwise. To ensure that all data points are connected to each other, a symmetrical similarity graph $\mathbf{A} = |\mathbf{W}| + |\mathbf{W}|^T$ is computed. In order to obtain the final clustering map ($\mathbf{M}$), spectral clustering is performed on the similarity graph, i.e. standard K-means clustering is applied on the relevant eigenvectors of the normalized Laplacian matrix of $\mathbf{A}$ [4]. Although in ESC, the demand of computational power is drastically reduced due to using $\mathcal{X}_0$ instead of $\mathbf{X}$ as the spectral dictionary in Eq.(1), ESC has no constraints to include spatial information.

### 2.3. Hidden Markov random field (HMRF)

HMRF is a probability-based model, in which the parameters $\Theta = [\theta_l, l \in \mathcal{L}]$ need to be estimated, from $\mathbf{Z}$, the first principal component of the RGB image. $\theta_l = (\mu_l, \sigma_l^2)$ are the mean and standard deviation of the probability density function from all points $z_i$ with class label $l$.

The HMRF method requires initial values for the clustering labels ($\mathbf{M^{(0)}}$) and the set of parameters ($\Theta^{(0)}$). Here, we propose to use the clustering result obtained by ESC to estimate the initial labels and parameters. Consequently, two iterative procedures are run simultaneously to update the class labels and the set of parameters using maximum a posteriori (MAP) and expectation-maximization (EM), respectively.

1) MAP can be formulated as the following optimization problem:

$$\widehat{\mathbf{M}} = \arg\min_{\mathbf{M}}[U(\mathbf{Z}|\mathbf{M}) + U(\mathbf{M})] \tag{2}$$

where $U(\mathbf{Z}|\mathbf{M}) = \sum_{i=1}^{N}[\frac{(z_i-\mu_{m_i})^2}{2\sigma_{m_i}^2} + \frac{\log\sigma_{m_i}^2}{2}]$ is the fitness term of the model and $U(\mathbf{M})$ is the penalty term to involve spatial information. This term is a sum of so-called clique potentials and penalizes situations in which neighboring pixels have different labels. Moreover, to preserve the edges and prevent oversmoothing, an edge detection technique is employed on $\mathbf{Z}$. In this work, the Canny edge detection technique is employed [9]. The penalty term is constructed in such a way that the estimation procedure is only carried out on the pixels that are not located at edges.

2) The second step in HMRF estimates the parameters $\Theta$. The initial $\Theta^{(0)}$ is calculated using the initial clustering result $\mathbf{M}^{(0)}$. To estimate $\Theta$, we used an EM algorithm, for which each iteration is given by:

$$\Theta^{(k+1)} = \arg\max_{\Theta} Q(\Theta|\Theta^k) \tag{3}$$

where $k$ is the iteration number, $Q$ is the EM functional, defined as: $Q(\Theta|\Theta^k) = E[log\ p(\mathbf{M}, \mathbf{Z}|\Theta)|\mathbf{Z}, \Theta^{(k)}]$. For more details, we refer readers to [8].

## 3. EXPERIMENTAL RESULTS

### 3.1. Data sets

The HSI and RGB images were acquired during a field campaign on sites in central Finland during September 2018. An unmanned aerial vehicle (UAV) was deployed to obtain the

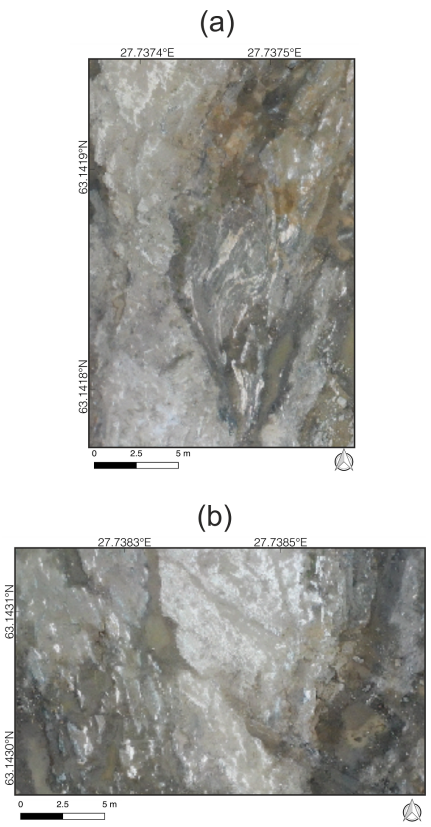


**Fig. 2**: (a) RGB image of the first scene and (b) RGB image of the second scene. Both images were captured over Siilinjrvi in Finland.

HSI and RGB images for mineral exploration. A hyperspectral frame-based camera (0.6 Mpix Rikola hyperspectral imager) and an RGB camera (20 MPix Parrot SODA) were deployed on a hexacopter UAV (Aibotix Aibot X6v2), respectively an eBee Plus fixed-wing UAV, that flew over geological areas of interest. The RGB images of both scenes are shown in Fig. 2. The acquired HSIs in both scenes contain 50 spectral bands, covering the electromagnetic spectrum between 504-900 nm. Both data sets have a spatial resolution of 3.3 cm for the HSI, and 1.5 cm for the RGB image before resampling. The HSI image of the first scene is composed of $706 \times 484$ pixels. The HSI image of the second scene is composed of $250 \times 416$ pixels. The HSIs were firstly pre-processed before further analysis. For more details on the pre-processing steps we refer the interested readers to [10]. Ground truth was acquired during an extended field campaign on the sites. Rock specimens and handheld hyperspectral scans were taken and located with precise GPS measurements. Based on those observations, we constructed the ground truth/validation data set, including the applied labels. The main geological classes in both scenes are Feldspar-Pegmatite, Glimmerite and Carbonatite, and the mixed class represents dark rocks containing a mixture of topsoil with Glimmerite.

### 3.2. Qualitative and quantitative assessment of clustering results

The performance of our proposed algorithm is compared to ESC and an entropy-based consensus clustering (ECC) which is based on K-means [11]. To ease the explanation, we use the following abbreviation for the applied clustering algorithms: $AL_{sensor}$, where $sensor$ is either the HSI or a fusion of HSI and RGB (HSI+RGB). The concatenation of HSI and RGB images are used in the multi-sensor versions of the ECC and ESC algorithms. $AL$ refers the applied clustering algorithm. Furthermore, HMRF subspace-based clustering will be referred to as HMRFsub in the rest of the text. In the experiments, HMRFsub$_{HSI}$ refers to a single-sensor version of the HMRF algorithm, which uses the first PC, obtained from the HSI to extract the spatial information. In all applied clustering methods, the default parameters suggested by their developers in [4, 8] are used. Additionally, bestMap($\mathbf{G}, \mathbf{M}$) is applied as a matching function to match the clustering results and ground-truth. The bestMap(.) function is based on a Hungarian algorithm to solve the assignment problem, further details on bestMap(.) can be found in [4].

Quantitative assessment is carried out by reporting the overall accuracy (OA), average accuracy (AA), and Kappa coefficient in Tables 1-2. The clustering maps of the first and second data set are presented in Figs. 3 and 4 respectively. As can be observed from the tables, overall, the fusion of the RGB and HSI images improves the clustering results from ECC and our proposed method, while this is not the case for ESC. Compared to HMRFsub$_{HSI}$, which applies the proposed method on the HSI alone, without using the RGB image, the proposed multi-sensor algorithm HMRFsub$_{HSI+RGB}$ improves the clustering results by 5% and 14% in the first and the second data set, respectively. In both data sets, HMRFsub$_{HSI+RGB}$ performs well in capturing geological features compared to its competitive algorithms (first data set: OA = 68.15% and second data set: OA = 77.67%). In the first data set, although HMRFsub$_{HSI+RGB}$ achieves the most accurate overall clustering result, it had a weak performance in capturing Carbonatite, while ECC$_{HSI}$ could capture Carbonatite with an accuracy of 71.00%. Both HMRFsub$_{HSI+RGB}$ and HMRFsub$_{HSI}$ could distinguish Feldspar-Pegmatite better than the other applied clustering algorithms, with an accuracy of 89.50% and 93.00% respectively. In the second data set, HMRFsub$_{HSI+RGB}$ performed the best in capturing Feldspar-Pegmatite with an accuracy of 89.22%. In addition, Carbonatite is best mapped using ESC$_{HSI+RGB}$ with an accuracy of 75.00%.

**Table 1**: Quantitative assessment of the performances of the clustering algorithms applied on the first data set. The clustering performance is evaluated using overall accuracy (OA), average accuracy (AA), and Kappa coefficient.

| Clusters | No. ground truth samples | ECC$_{HSI}$ | ECC$_{HSI+RGB}$ | ESC$_{HSI}$ | ESC$_{HSI+RGB}$ | HMRFsub$_{HSI}$ | HMRFsub$_{HSI+RGB}$ |
|---|---|---|---|---|---|---|---|
| Dust-soil | 431 | 76.27 | 74.82 | 47.70 | 38.98 | 40.68 | **80.87** |
| Feldspar-Pegmatite | 413 | 41.58 | 56.02 | 78.34 | 33.04 | **93.00** | 89.50 |
| Glimmerite | 315 | 80.32 | 45.71 | 61.27 | **97.78** | 86.35 | 73.65 |
| Carbonatite | 457 | **71.00** | 58.24 | 28.31 | 21.11 | 11.14 | 9.50 |
| Mixed | 274 | 0.00 | 82.12 | **100** | 82.85 | **100** | **100** |
| OA (%) | | 56.30 | 62.70 | 60.53 | 49.63 | 62.80 | **68.15** |
| AA (%) | | 53.83 | 63.38 | 63.12 | 54.75 | 66.23 | **70.61** |
| Kappa | | 0.45 | 0.53 | 0.51 | 0.37 | 0.54 | **0.60** |

**Table 2**: Quantitative assessment of the performances of the clustering algorithms applied on the second data set. The clustering performance is evaluated using overall accuracy (OA), average accuracy (AA), and Kappa coefficient.

| Clusters | No. ground truth samples | ECC$_{HSI}$ | ECC$_{HSI+RGB}$ | ESC$_{HSI}$ | ESC$_{HSI+RGB}$ | HMRFsub$_{HSI}$ | HMRFsub$_{HSI+RGB}$ |
|---|---|---|---|---|---|---|---|
| Feldspar-Pegmatite | 167 | 53.59 | 83.92 | 53.89 | 50.30 | 65.27 | **89.22** |
| Dust-soil | 153 | 35.63 | **77.06** | 41.83 | 12.42 | 71.24 | 63.40 |
| Carbonatite | 160 | **86.54** | 71.08 | 33.75 | 75.00 | 30.00 | 67.50 |
| Mixed | 161 | 36.53 | 72.33 | 80.75 | 11.18 | 94.41 | **100** |
| Water | 156 | **92.55** | 68.46 | 69.87 | 75.64 | 75.64 | 66.67 |
| OA (%) | | 60.73 | 74.24 | 56.09 | 45.04 | 63.61 | **77.67** |
| AA (%) | | 60.97 | 74.57 | 56.02 | 44.91 | 63.52 | **77.36** |
| Kappa | | 0.51 | 0.68 | 0.45 | 0.31 | 0.54 | **0.72** |

## 4. CONCLUSIONS

In this work, we investigated the effect of fusing a fine spatial-resolution RGB image and an HSI to improve the clustering performance. Our proposed algorithm consists of two phases. The first phase is used to extract subspace information from an HSI using the exemplar-based subspace clustering algorithm. The second phase incorporates spatial information in the results of the first phase by extracting information from a fine spatial-resolution RGB image, using a hidden Markov random field approach. The proposed algorithm is applied on two real geological data sets, the obtained results show that the proposed method captures geological structures well compared to other state-of-the-art clustering algorithms. As our future work, we will investigate methods to find an optimal number of "*exemplars*" in the ESC algorithm.

## 5. REFERENCES


[1] A. F. H. Goetz, G. Vane, J. E. Solomon, and B. N. Rock, "Imaging spectrometry for earth remote sensing," *Science*, vol. 228, no. 4704, pp. 1147–1153, 1985.

[2] M. D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 6, pp. 2014–2039, Jun. 2011.

[3] Q. Zhang, Y. Liu, R. S. Blum, J. Han, and D. Tao, "Sparse representation based multi-sensor image fusion for multi-focus and multi-modality images: A review," *Information Fusion*, vol. 40, pp. 57 – 75, 2018.

[4] C. You, C. Li, D. P. Robinson, and R. Vidal, "Scalable exemplar-based subspace clustering on class-imbalanced data," in *The European Conference on Computer Vision (ECCV)*, September 2018.

[5] E. Elhamifar and R. Vidal, "Sparse subspace clustering: Algorithm, theory, and applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 11, pp. 2765–2781, Nov 2013.

[6] S. Lorenz, P. Seidel, P. Ghamisi, R. Zimmermann, L. Tusa, M. Khodadadzadeh, I. C. Contreras, and R. Gloaguen, "Multi-sensor spectral imaging of geological samples: A data fusion approach using spatio-spectral feature extraction," *Sensors*, vol. 19, no. 12, 2019.

[7] P. Ghamisi, E. Maggiori, S. Li, R. Souza, Y. Tarablaka, G. Moser, A. De Giorgi, L. Fang, Y. Chen, M. Chi, S. B. Serpico, and J. A. Benediktsson, "New frontiers in spectral-spatial hyperspectral image classification: The latest advances based on mathematical morphology, markov random fields, segmentation, sparse representation, and deep learning," *IEEE Geoscience and Remote Sensing Magazine*, vol. 6, no. 3, pp. 10–43, Sep. 2018.

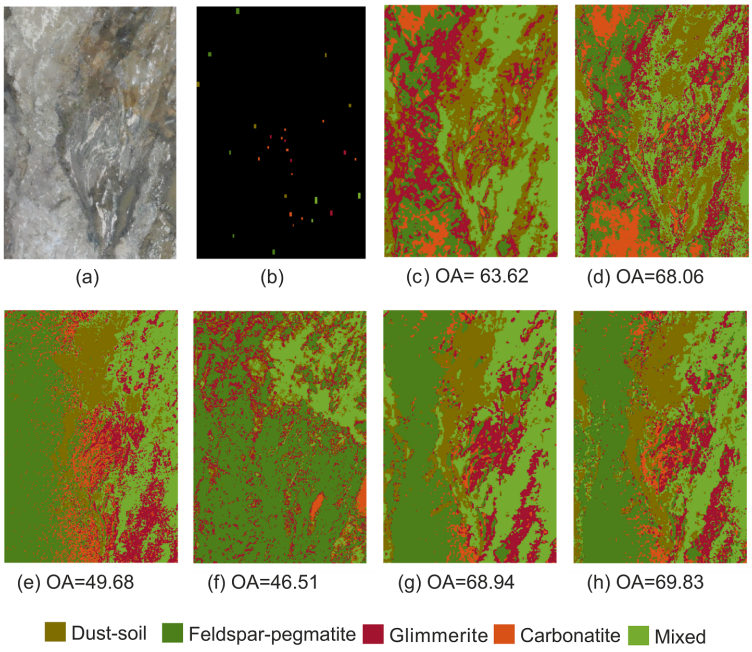


**Fig. 3**: First data set. (a) RGB, (b) ground-truth, (c) $\text{ECC}_{HSI}$, (d) $\text{ECC}_{HSI+RGB}$, (e) $\text{ESC}_{HSI}$, (f) $\text{ESC}_{HSI+RGB}$, (g) $\text{HMRFsub}_{HSI}$, (h) $\text{HMRFsub}_{HSI+RGB}$. OAs are reported in percent (%).

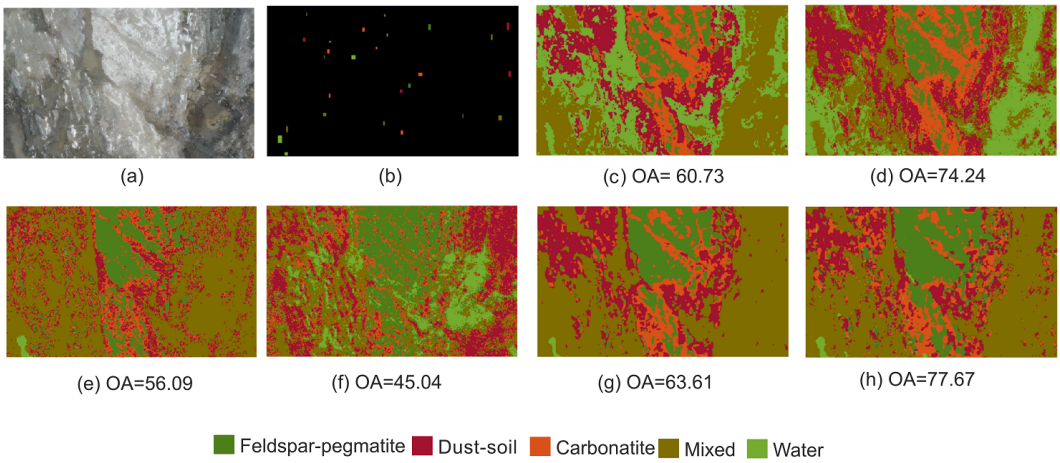


**Fig. 4**: Second data set. (a) RGB, (b) ground-truth, (c) $\text{ECC}_{HSI}$, (d) $\text{ECC}_{HSI+RGB}$, (e) $\text{ESC}_{HSI}$, (f) $\text{ESC}_{HSI+RGB}$, (g) $\text{HMRFsub}_{HSI}$, (h) $\text{HMRFsub}_{HSI+RGB}$. OAs are reported in percent (%).

[8] P. Ghamisi, J. A. Benediktsson, and M. O. Ulfarsson, "Spectralspatial classification of hyperspectral images based on hidden markov random fields," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 5, pp. 2565–2574, May 2014.

[9] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, Nov 1986.

[10] S. Jakob, R. Zimmermann, and R. Gloaguen, "The need for accurate geometric and radiometric corrections of drone-borne hyperspectral data for mineral exploration: Mephystoa toolbox for pre-processing drone-borne hyperspectral data," *Remote Sensing*, vol. 9, no. 1, 2017.

[11] H. Liu, R. Zhao, H. Fang, F. Cheng, Y. Fu, and Y. Liu, "Entropy-based consensus clustering for patient stratification," *Bioinformatics*, vol. 33, no. 17, pp. 2691–2698, 2017.