

GAS PLUME DETECTION AND TRACKING IN HYPERSPECTRAL VIDEO SEQUENCES USING BINARY PARTITION TREES

G. Tochon¹, J. Chanussot^{1,4}, J. Gilles², M. Dalla Mura¹, J.-M. Chang³, A. L. Bertozzi²

¹GIPSA-lab, Grenoble Institute of Technology, Saint Martin d'Hères, France

²Department of Mathematics, University of California Los Angeles, Los Angeles, USA

³Department of Mathematics and Statistics, California State University, Long Beach, USA

⁴Department of Electrical and Computer Engineering, University of Iceland, Reykjavik, Iceland

ABSTRACT

Thanks to the fast development of sensors, it is now possible to acquire sequences of hyperspectral images. Those hyperspectral video sequences are particularly suited for the detection and tracking of chemical gas plumes. However, the processing of this new type of video sequences with the additional spectral diversity, is challenging and requires the design of advanced image processing algorithms. In this paper, we present a novel method for the segmentation and tracking of a chemical gas plume diffusing in the atmosphere, recorded in a hyperspectral video sequence. In the proposed framework, the position of the plume is first estimated, using the temporal redundancy of two consecutive frames. Second, a Binary Partition Tree is built and pruned according to the previous estimate, in order to retrieve the real location and extent of the plume in the frame. The proposed method is validated on a real hyperspectral video sequence and compared with a state-of-the-art method.

Index Terms— segmentation, tracking, Binary Partition Tree, chemical gas plume, hyperspectral video sequence

1. INTRODUCTION

The detection and tracking of chemical gas plumes in the atmosphere is of great interest for several domains [1, 2]. In the defense and security area for example, such analysis could be employed in order to detect the use of chemical gas weapons. In the environmental protection field, the detection and tracking of gas plumes could be also of use to identify and repair gas leaks in order to minimize their impact on the environment and the potential harm they could cause on human populations. However, this task still remains an open research topic as most gases do not appear in the visible spectrum and hence remain invisible to human inspection or traditional color imaging systems. As a matter of fact, their spectral signature significantly responds only in a restrained portion of the infrared (IR) domain [3, 4], hence the need for a fine sampling of the electromagnetic spectrum. Additionally, the temporal dimensionality inherent to video sequences requires appropriate processings. Hyperspectral video sensors combine the ability to precisely describe spectral properties of the captured scene and to record its evolution over time, but at the cost of an important amount of data to process [5–8].

In this paper, we propose a new method to process a hyperspectral video sequence for the detection and tracking of a chemical gas plume diffusing in the atmosphere. This method relies both on spectral properties of the hyperspectral scene and on the temporal redundancy of

the video sequence. A rough estimate of the position of the plume in the frame is first computed. The actual position and extent of the plume is then retrieved using a Binary Partition Tree. The remainder of the paper is organized as follows: section 2 introduces the Binary Partition Tree (BPT) algorithm that constitutes the core of the proposed method. Section 3 further details the proposed tracking and segmentation algorithm. Section 4 displays some results obtained on a real sequence and features some comparisons. Conclusions are given in section 5.

2. BINARY PARTITION TREE (BPT)

The BPT is a hierarchical region-based representation of an image stored in a tree structure [9, 10]. Starting from an initial partition of the image, regions are iteratively merged until only one region remains, corresponding to the whole image support. The merging sequence is stored in a tree structure \mathcal{T} . In this representation, regions from the initial partition form the *leaf* nodes, the whole image represents the *root*, and each node inbetween corresponds to a region resulting from the merging of its two children. There are two notions of primary importance when building a BPT. The *region model* $\mathcal{M}_{\mathcal{R}}$ describes how regions are represented mathematically and how to model the merging of two regions. The *merging criterion* $\mathcal{O}(\mathcal{R}_i, \mathcal{R}_j)$ is a measure assessing the similarity between two neighboring regions \mathcal{R}_i and \mathcal{R}_j by measuring the distance between their region models. The merging criterion thus determines the sequence in which the regions are merged.

The *pruning* step follows the construction of the BPT. It aims at cutting off some branches in the BPT so the leaves of the pruned tree correspond to regions achieving the best segmentation with respect to the desired task. Unlike the construction of the BPT, which is generic up to the definition of the region model and merging criterion, the pruning step is application dependent, and different pruning strategies applied on the same BPT generally leads to different segmentation results [11].

3. PROPOSED METHOD

3.1. Data set and pre-processing

The data set used in this study was acquired and provided by the John Hopkins Applied Physics Laboratory. The spectral radiance of the scene was recorded by a long wave IR spectrometer, about 2 kilometers away from the gas release, producing a hyperspectral video sequence at a frame rate of 0.2 Hz. Each frame of the sequence is therefore a hyperspectral image of size 128×320 pixels and

This material is based upon work supported by the National Science Foundation under grant no. DMS-1118971 and no. DMS-0914856.

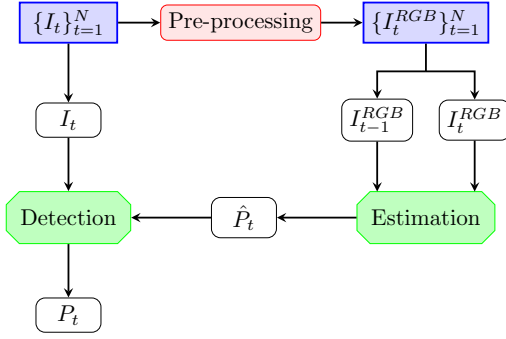


Fig. 1: General framework of the proposed method.

comprising 129 spectral bands corresponding to wavelengths evenly distributed between 7830 nm and 11700 nm. Let $\{I_t\}_{t=1}^N$ denote the hyperspectral video sequence, N being the total number of frames. In this study, $N = 23$. An initial pre-processing is applied to the whole sequence. It comprises a Principal Component Analysis (PCA) done on each frame, where the three first Principal Components (PCs) are retained, followed by a Midway equalization [12] to ease the visualization of the data, as detailed in [5]. The output of the pre-processing step is the false color representation sequence labeled $\{I_t^{RGB}\}_{t=1}^N$. Figures 3a and 3b display two consecutive frames of the false color representation sequence.

The proposed method is organized in two steps:

- The estimation of the position of the plume in the current frame by taking advantage of the temporal redundancy inherent to the video sequence.
- The validation and refinement of the previous estimate using the BPT.

Figure 1 illustrates the proposed workflow. Please note that the pre-processing is only used to provide a rough estimate of the position of the plume. The actual segmentation is performed using the initial full hyperspectral frame. For each input frame I_t , the output of the proposed algorithm is the binary mask P_t featuring the position of the plume in the current frame.

3.2. Estimation step

The goal of the estimation step is to produce a reliable estimate of the position of the plume and use this estimate as *a priori* information when pruning the BPT. Note that only the false color representation video sequence is considered at this stage of the algorithm. The whole estimation process is featured by the workflow in figure 2, and is based on the temporal redundancy between consecutive frames. More specifically, it is assumed that only the plume is moving between two consecutive frames, and the background does not change. Consequently, the image difference between two consecutive frames is expected to contain low values in areas that do not feature any change between the two frames, and higher values when significant change occurs. The first step of the estimation stage is to identify those areas. This operation is illustrated in figure 2. More precisely:

1. Input frames are initially very noisy as it can be seen in figures 3a and 3b. Therefore, a preliminary denoising is applied on both images. This is achieved using the Block Matching 3D algorithm (BM3D) [13].

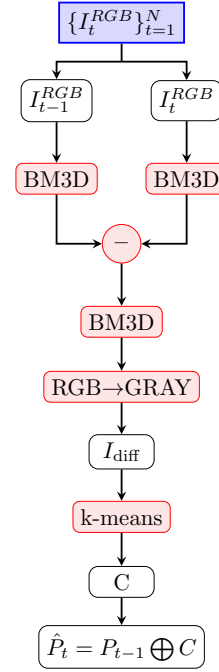


Fig. 2: Workflow of the estimation step.

2. The image difference is then computed, denoised again to remove residual noise, and converted in grayscale to produce the image labeled I_{diff} (see figure 3c). The underlying idea is to consider I_{diff} as a topographical image where peaks (bright regions) correspond to areas that are significantly changing between the two consecutive frames. Those areas correspond to regions that are either invaded or left by the plume.
3. Peaks are extracted by thresholding the topographical image with an automatically set threshold value defined by a two-class K-means algorithm, producing the binary image C displayed in figure 3d.

The last step of the estimation stage is to produce the estimate position of the plume. It can be done combining the position of the plume detected in the previous frame P_{t-1} and the current change map since the new position corresponds to the previous one plus the region that have been invaded, minus those that have been left. This can be mathematically formulated

$$\hat{P}_t = P_{t-1} \oplus C \quad (1)$$

where \oplus denotes the binary XOR operation, thus producing the estimate \hat{P}_t .

3.3. Detection step

The second step of the proposed method is the actual detection of the plume P_t in the current frame, using the previously computed estimate \hat{P}_t as some *a priori* knowledge. The detection is handled through the construction and pruning of a BPT, as shown by the workflow featured on figure 4.

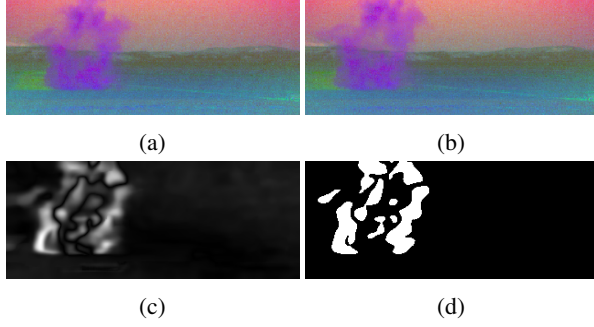


Fig. 3: Illustration of the estimation process: (a,b) two consecutive noisy frames along with (c) their image difference I_{diff} and (d) the resulting binary map.

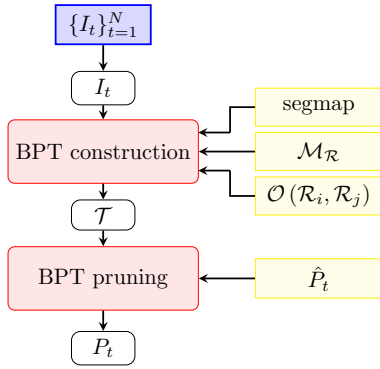


Fig. 4: Workflow of the detection step.

3.3.1. Construction of the BPT

As mentioned in section 2, the construction of the BPT asks for three input parameters:

- *The definition of the initial leaves:* starting from the pixel level generates $128 \times 320 = 40960$ leaves. This considerably impacts the computational load and leads to a high number of non-significant nodes in the BPT. On the contrary, starting with leaves corresponding to the regions of a preliminary rough segmentation significantly reduces the number of final nodes in the tree and hence the computational load, while not impacting the final segmentation results. The only required condition is to start from an over-segmentation, as initial regions will not be allowed to split in the further steps. This is easily achieved by using a watershed segmentation. The extension of the watershed to hyperspectral images presented in [14] has been selected.
- *The region model,* which describes how regions are mathematically represented. For this work, the mean spectrum region model has been implemented:

$$\mathcal{M}_{\mathcal{R}} = \bar{\mathbf{r}} = [\bar{r}^{(1)}, \dots, \bar{r}^{(q)}] \quad (2)$$

with

$$\bar{r}^{(i)} = \frac{1}{|\mathcal{R}|} \sum_{\mathbf{p} \in \mathcal{R}} p^{(i)}, \quad (3)$$

where q stands for the number of spectral band in the frame, $|\mathcal{R}|$ is the number of pixels in region \mathcal{R} , and $p^{(i)}$ is the value in the i -th band at pixel location \mathbf{p} .

- *The merging criterion* between two neighboring regions was defined as the Spectral Angle between their region models:

$$\mathcal{O}(\mathcal{R}_i, \mathcal{R}_j) = \arccos \left(\frac{\langle \bar{\mathbf{r}}_i, \bar{\mathbf{r}}_j \rangle}{\|\bar{\mathbf{r}}_i\|_2 \|\bar{\mathbf{r}}_j\|_2} \right). \quad (4)$$

3.3.2. Pruning of the BPT

As the plume is a thin layer overlaying the background, the spectral response of pixels belonging to the plume only differs slightly from pixels “behind” the plume (be it ground or sky). However, the BPT is able to capture the plume as one single region while it has not too much diffused yet, or two different regions (one being the bottom half of the plume which overlays with the ground, and the other one being the top half superimposed on the sky). The goal of the pruning step is to identify in the tree structure which node or set of nodes corresponds to the plume. Therefore, the implemented pruning strategy is based on the estimated position \hat{P}_t and seeks the best node or set of nodes matching this estimate. The matching criterion is defined as follows: a set of nodes $\{\mathcal{N}_1, \dots, \mathcal{N}_m\}$ and its corresponding regions $\{\mathcal{R}_1, \dots, \mathcal{R}_m\}$ is said to match the estimate \hat{P}_t if the area covered by all the regions overlaps with at least $\gamma\%$ of \hat{P}_t , and if each region independently has at least half of its pixels belonging to \hat{P}_t . Mathematically, $\{\mathcal{R}_1, \dots, \mathcal{R}_m\}$ matches \hat{P}_t if:

$$\left| \left(\bigcup_{i=1}^m \mathcal{R}_i \right) \cap \hat{P}_t \right| \geq \gamma \times |\hat{P}_t| \quad (5)$$

$$|\mathcal{R}_i \cap \hat{P}_t| \geq |\mathcal{R}_i \setminus \hat{P}_t| \quad \forall i = 1, \dots, m \quad (6)$$

The parameter γ represents the confidence in the estimate and was empirically set to 70%. This value achieves a trade-off between trust and mistrust in the estimate. To reconstruct the plume with as few regions as possible (ideally only one), the retained pruning is the one leading to the tree with the smallest number of nodes. If several regions were found, they are fused together in a last step to obtain the final binary mask P_t .

4. RESULTS

Figures 5a to 5d display segmentation results obtained by the presented method for the second, sixth, tenth and fourteenth frames after the plume appearance, respectively. Figures 5e to 5h exhibit segmentation results on the same frames obtained by the state-of-art method [5], based on the Merriman-Bence-Osher (MBO) semi-supervised clustering scheme [15]. It is worth mentioning that the MBO utilizes only the first five principal components while our method uses the whole hyperspectral data. It can be seen how both methods accurately detect and segment the plume from the background, and track it along the frames. However, the MBO method produces some small false detection areas at the interface between the ground and the sky. This issue does not arise in our proposed method. Moreover, a small cloud of dust, triggered by the explosive release of the plume, can be seen near the bottom left part of the plume in the two middle frames of figure 5. Our proposed method is able to correctly differentiate it from the gas plume while it is included within the plume region for the MBO results. Note that [5] utilizes the output of a background subtraction step to initialize the MBO scheme while the estimation process of our method only starts when the plume appears in the video sequence. In both cases, the exact appearance time of the plume in the video sequence must be known in order to trigger the detection process.

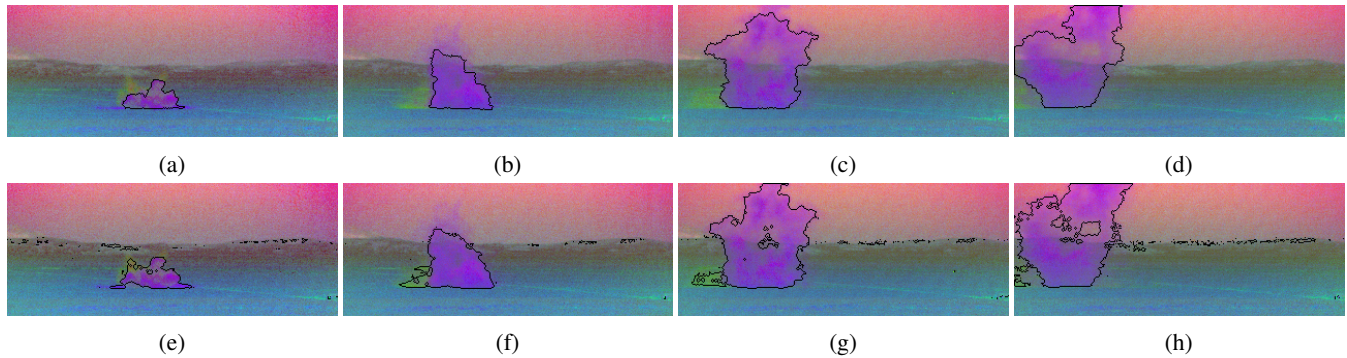


Fig. 5: Top row: segmentation results obtained by the presented method for four frames of the video sequence. Bottom row: segmentation results obtained by [5] for the same frames. Results are displayed on the false color representation video sequence.

5. CONCLUSION

In this article, we presented a novel algorithm for the detection and tracking of chemical gas plume in a hyperspectral video sequence. The proposed method is organized in two stages being the estimation of the position of the plume in the current frame and the detection of the real plume, which relies on the previous estimate. While the first step is based on the temporal redundancy inherent to video sequences, the second one involves the construction and pruning of a Binary Partition Tree. The proposed algorithm gives satisfactory visual results for the presented video sequence. Future work includes the design of a method to quantitatively assess the quality of the obtained segmentation and tracking despite the lack of ground-truth data. The use of anomaly detection technics to blindly detect the release instant of the plume will also be investigated in order to make the proposed method fully unsupervised.

6. REFERENCES

- [1] V. Farley, A. Vallires, A. Villemaire, M. Chamberland, P. Lagueux, and J. Giroux, "Chemical agent detection and identification with a hyperspectral imaging infrared sensor," in *SPIE Defense, Security, and Sensing*, 2007, vol. 6739, pp. 673918–673918–12.
- [2] M. Hinnrichs, "Imaging spectrometer for fugitive gas leak detection," in *SPIE Defense, Security, and Sensing*, 1999, vol. 3853, pp. 152–161.
- [3] J. B. Broadwater, T. S. Spisz, and A. K. Carr, "Detection of gas plumes in cluttered environments using long-wave infrared hyperspectral sensors," in *SPIE Defense and Security Symposium*. International Society for Optics and Photonics, 2008, pp. 69540R–69540R.
- [4] A. Vallires, A. Villemaire, M. Chamberland, L. Belhumeur, V. Farley, J. Giroux, and J-F. Legault, "Algorithms for chemical detection, identification and quantification for thermal hyperspectral imagers," in *SPIE Defense, Security, and Sensing*, 2005, vol. 5995, pp. 59950G–59950G–11.
- [5] T. Gerhart, J. Sunu, L. Lieu, E. Merkurjev, J-M. Chang, J. Gilles, and A. L. Bertozzi, "Detection and tracking of gas plumes in LWIR hyperspectral video sequence data," in *SPIE Defense, Security, and Sensing*. International Society for Optics and Photonics, 2013, pp. 87430J–87430J.
- [6] A. Banerjee, P. Burlina, and J. B. Broadwater, "Hyperspectral video for illumination-invariant tracking," in *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, 2009. WHISPERS '09. First Workshop on*, August 2009, pp. 1–4.
- [7] H. Van Nguyen, A. Banerjee, and R. Chellappa, "Tracking via object reflectance using a hyperspectral video camera," in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*. IEEE, 2010, pp. 44–51.
- [8] C. S. Grant, T. K. Moon, J. H. Gunther, M. R. Stites, and G. P. Williams, "Detection of amorphously shaped objects using spatial information detection enhancement (SIDE)," *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, vol. 5, no. 2, pp. 478–487, April 2012.
- [9] P. Salembier and L. Garrido, "Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval," *Image Processing, IEEE Transactions on*, vol. 9, no. 4, pp. 561–576, 2000.
- [10] S. Valero, P. Salembier, and J. Chanussot, "Hyperspectral image representation and processing with binary partition trees," *Image Processing, IEEE Transactions on*, vol. 22, no. 4, pp. 1430–1443, 2013.
- [11] S. Valero, P. Salembier, and J. Chanussot, "Comparison of merging orders and pruning strategies for binary partition tree in hyperspectral data," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*. IEEE, 2010, pp. 2565–2568.
- [12] J. Delon, "Midway image equalization," *Journal of Mathematical Imaging and Vision*, vol. 21, no. 2, pp. 119–134, 2004.
- [13] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, et al., "Image denoising with block-matching and 3D filtering," in *Proceedings of SPIE*, 2006, vol. 6064, pp. 354–365.
- [14] Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, "Segmentation and classification of hyperspectral images using watershed transformation," *Pattern Recognition*, vol. 43, no. 7, pp. 2367–2379, 2010.
- [15] E. Merkurjev, T. Kostic, and A. L. Bertozzi, "An MBO scheme on graphs for segmentation and image processing," *UCLA CAM Report*, pp. 12–46, 2012.