ADVANCES IN FOREST FIRE RESEARCH

2022

Edited by
DOMINGOS XAVIER VIEGAS
LUÍS MÁRIO RIBEIRO

A Multilayer Approach to Wildfire Aerial Thermal Image Segmentation Using Unsupervised Methods

Tiago Garcia*[1]; Alexandre Bernardino[2]; Ricardo Ribeiro[2]

[1]*IST. Lisbon, Portugal, {tiagojosegarcia@tecnico.ulisboa.pt}*
[2] *ISR. Lisbon, Portugal, {alex, ribeiro}@isr.tecnico.ulisboa.pt*

**Corresponding author*

Keywords

Thermal Images, Image Segmentation, Airborne Sensors, Wildfire Monitoring, Level Set Segmentation

Abstract

The infrared (IR) thermal images of a propagating wildfire taken by manned or unmanned aerial vehicles can help the firefighting authorities on the ground. The segmentation of such images in regions of fire vs non-fire is a necessary step to measure the fire perimeter and determine the location of the fire front. This work proposes a segmentation method based on level sets, which have the property of handling topology, which makes them suitable to segment wildfire images since the fire areas may be spread out across the image. The results were compared against other common unsupervised segmentation methods, including Otsu, K-means and Mean Shift. The Level Set method was optimized to ensure contour smoothness and reliability, as well as reduce the computation time. Although falling out of use in relation to Deep Learning methods, unsupervised segmentation can still be very useful when annotated datasets are unavailable. The experimental results were compared using hand-drawn labels over a set of images provided by the Portuguese Air Force (FAP) as a ground truth. These labels were carefully drawn by the author to ensure they complied with the requirements indicated by the Portuguese National Authority for Emergency and Civil Protection (ANEPC).

1. Motivation and Objectives

The development of UAVs equipped with cameras and reliable communication and navigation systems enabled a powerful tool in the observation of forest areas, both in the early detection of fires, during the active wildfire combat and the aftermath stage. For early detection, it is important to accurately identify images containing wildfires, while avoiding false positives, for example, with similar color patterns, such as sunsets of dry foliage. During the active stage, it is important to have a tool that allows the authorities to geo-localize the fire front, measure the fire extension, and realize any external effects that can affect the wildfire evolution, such as strong winds. Understanding the evolution of the wildfire and its potential spread areas can help minimize damages and contribute to a more effective combat. When the wildfire is eventually declared as controlled, it is still important to detect if and where reburns occur to quickly extinguish them, preventing a new critical situation.

This work aims at finding the best method for segmenting aerial wildfire thermal images. The thermal cameras can prove extremely useful in situations where smoke and vegetation obscure the field of vision of the aircraft. The algorithms implemented should be able to systematically receive the thermal images and perform multiple segmentations over the same image. Specifically, they should be capable of dividing the image in different areas, sorted according to the fire intensities in each area. The contour lines will then resemble level curves, to eventually construct a thermal map of the image. The contours will also be used to obtain the perimeter of the wildfire and the extension of the fire front in future work.

The layout of this paper is as follows: Section 2 presents the state of the art in thermal image segmentation, Section 3 presents an overview of the theory behind the models used, Section 4 refers some aspects of the implementation, Section 5 presents the datasets used and results obtained. Finally, Section 6 concludes this paper and presents suggestions for future work.

2. Related work

Segmentation of IR images has been done in the past mainly applied to medical images, using watershed (Grau *et al.* 2004), K-means (Ng *et al.* 2006) and level set methods (Banerjee *et al.* 2010). The segmentation of aerial

IR wildfire images for tracking of the active front was attempted using, for example, a combination of Otsu and optical flow (Yuan *et al.* 2017) and Canny edge detectors (Valero *et al.* 2018). However, these methods only present a division of the image in 2 classes.

Multilayer level set methods, like the ones used in Huang and Wu (2010), Huang *et al.* (2013) and Wen *et al.* (2020) were applied to construct thermal mapping of materials. This work applies a similar algorithm but instead to thermal images of wildfires, using only one level set function, which makes the algorithm faster, and adding an edge stopping term.

3. Active Contours

3.1. Level Set Segmentation

The segmentation methods that are based on level sets make use of the idea of implicitly representing a propagating front (or a contour) C through the evolution of a higher dimensional, continuous differentiable function ϕ (dubbed the level set function). In the case of image segmentation, the 2D contour is represented by the level curve of a 3D surface represented by function $\phi: \mathbb{R}^2 \rightarrow \mathbb{R}$. To ensure curve smoothness, the evolution of this function depends on the contour curvature k along its normal direction N which are given by (see Osher and Sethian 1988 for proof)

$$N = \frac{\nabla\phi}{|\nabla\phi|}$$

$$k = \mathrm{div}\left(\frac{\nabla\phi}{|\nabla\phi|}\right)$$

where $\nabla\phi$ represents the gradient of ϕ and $|\nabla\phi|$ the respective norm, while $\mathrm{div}()$ is the divergence operator. The evolution of the level set function is given, according to Osher and Sethian (1988) and Malladi and Sethian (1996), by

$$\frac{\partial\phi}{\partial t} = |\nabla\phi|\mathrm{div}\left(\frac{\nabla\phi}{|\nabla\phi|}\right)$$

For a given open region $\Omega(x, y)$ with smooth boundary, ϕ satisfies

$$\begin{cases} \phi(x,y) > 0, (x,y) \in \Omega \\ \phi(x,y) = 0, (x,y) \in \partial\Omega \\ \phi(x,y) < 0, (x,y) \in \bar{\Omega} \end{cases}$$

where $\partial\Omega$ represents the contour of Ω, which is the contour that is to be obtained. In image processing, all the curve derivatives are taken in a discrete rectangular grid instead of a continuous curve (Malladi and Sethian 1996), which simplifies their calculation.

3.2. Edge-based ACM

Edge-based active contour models use gradient information to stop the contour evolution near strong edges. This can be achieved through the usage of an edge stopping function g that has the property of being positive and decreasing with gradient magnitude, becoming near zero at the edges and neutral when there are no differences in gradient. An example chosen in Caselles *et al.* (1993) is

$$g(\nabla I) = \frac{1}{1 + |\nabla[G_\sigma(x,y) * I(x,y)]|^2} \quad \textbf{(1)}$$

where the image I is convolved with a gaussian kernel G_σ with standard deviation σ.

Edge based methods, such as the Geodesic Active Contour (GAC) model (Caselles *et al.* 1997) are sensible to the placement of the initial contour. If it is initialized away from the object, it may never reach it and instead detect another strong edge, since it has no information about the whole image. This is especially critical if the

image is very noisy, since then the smoothing Gaussian will have to be strong and thus will smooth the edges too.

3.3. Area-based ACM

Area based active contour models use global image information, attempting to divide an image in regions by associating the pixels that share common properties, such as similar intensities. One of the most famous examples is the method introduced by Chan and Vese (Chan and Vese 2001), which is based on the Mumford-Shah functional (Mumford and Shah 1989). The Mumford-Shah formulation for image segmentation comprises the decomposition of an image in *n* regions, such that the intensities vary slowly within each region and briskly across the borders between them.

The Chan-Vese method, initially developed for a two-class segmentation which was achieved by cutting the level set function at level set zero, was generalized to multi-class segmentation (Vese and Chan 2002; Chung and Vese 2005, 2009; He and Osher 2007). For n=m+1 regions, the multi-layer Chan-Vese model determines the evolution of the level set function according to

$$\frac{\partial \phi}{\partial t} = \mu \sum_{i=1}^{m} \left[\delta_\varepsilon(\phi - l_i) div\left(\frac{\nabla \phi}{|\nabla \phi|}\right)\right] + \delta_\varepsilon(l_1 - \phi)|I - c_1|^2 - \delta_\varepsilon(\phi - l_m)|I - c_{m+1}|^2 + \sum_{i=2}^{m} [-\delta_\varepsilon(\phi - l_{i-1}) H_\varepsilon(l_1 - \phi)|I - c_i|^2 + \delta_\varepsilon(l_i - \phi) H_\varepsilon(\phi - l_{i-1})|I - c_i|^2] \quad (2)$$

where δ_ε and H_ε represent the regularized Dirac and Heaviside functions (Chung and Vese 2005), respectively, μ is a weight term, $l_1, \dots, l_m$ are the defined level set positions and $c_1, \dots, c_n$ are the intensity means in each region. The first term in (2) is a regulation term, the rest acts as a 'balloon force' that expands or shrinks the contour to minimize, for each region, the differences between the pixel intensities in the image I and the mean intensity calculated for that region.

4. Implementation

4.1. Multigrid implementation

The calculations required to apply the law in (2) may have an extremely high computational cost if the input image has a high resolution. Even at a resolution of 768x576, which is the most common resolution of the IR images supplied by the FAP, the time needed to achieve convergence on the level set function can be over one minute. To address this issue, we compute several lower resolution images, each one with half the pixels of the previous one in each direction. The level set function is initialized over the coarser grid, evolved until convergence, and extrapolated to the next finer grid. This process is repeated until convergence is reached on the grid with the original image resolution, as explained in Figure 1.

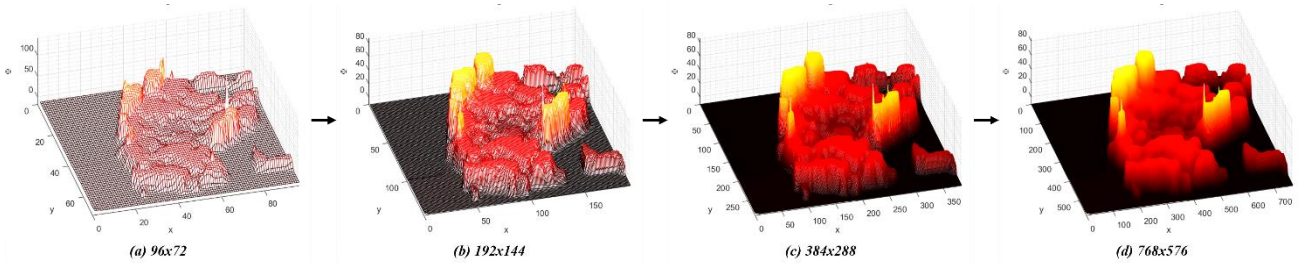

Figure 1 - State of ϕ at the end of the evolution for each resolution. The evolution is processed from the coarser grid (a) to the finer grid (d). Parameters: 6 contours, μ=0.008, α=10000, ε=1, σ=1

4.2. Inclusion of an edge-stopping term

It was verified that with the inclusion of more contours, the outer contours were detecting non-relevant edges or no edges at all, and inner contours were not "sticking" with the borders of the inner-most areas. To correct

this issue, it was decided to add an edge stopping term, like the one in (1), directly to the evolution rule (2), like what was done in Li *et al.* (2010). This term is given by

$$\text{EST} = \alpha g(\nabla I)[\delta_\varepsilon(\phi - l_1) + \delta_\varepsilon(\phi - l_m)] \quad (3)$$

where α is a weight term, $g(\nabla I)$ is given by (1) and $\delta_\varepsilon(\phi - l_i)$ allows for selective application of this term to the i^{th} level set, which is performed since the inclusion of this term has a destabilizing effect on ϕ. The term in (3) approximates the inner contour, given by the level set l_m, to the boundary of the brightest areas and keeps the outermost contour (corresponding to l_1) from propagating outside of the wildfire area.

5. Results

5.1. Datasets

The datasets used for segmentation were provided by the Portuguese Air Force. They consisted of images and videos taken both during real occurrences and during contained tests.

The main video is dubbed FOGO_1. It consists of different points of view of a wildfire, with nearly a minute of IR footage in total between RGB segments. To evaluate the segmentation quantitively, 55 images were selected, 1 second apart from each other, from the total 1342 frames. The ground truth masks were created by hand using MATLAB® Image Labeller App, following the indications given by the ANEPC. These requirements demanded a division of the images in three classes:

- Class 1 (outer): outside of the fire
- Class 2 (middle): region inside the fire that does not belong to the propagating front
- Class 3 (inner): the hottest (brightest) regions, which represent the propagating front

The contours should be smooth to later calculate the perimeter of the burning areas. Additionally, small high intensity 'dots' within the fire are to be classified as class 2, given that they most likely correspond to trees or other objects that are still burning actively but do not represent the fire front that is desired to track.

All the frames of FOGO_1 dataset were subjected to a 6-class segmentation (5 contours) using the level set method in (2) with the addition of the term in (3), with later selection and grouping of the 6 classes in the desired 3 classes (2 contours). The division in 6 classes is also useful to construct thermal maps of the fire. Performing a segmentation with more contours with further reduction, as was done in this case, allows for more reliable results without the need for heavy tuning of parameters. Figure 2 present some of the results obtained.

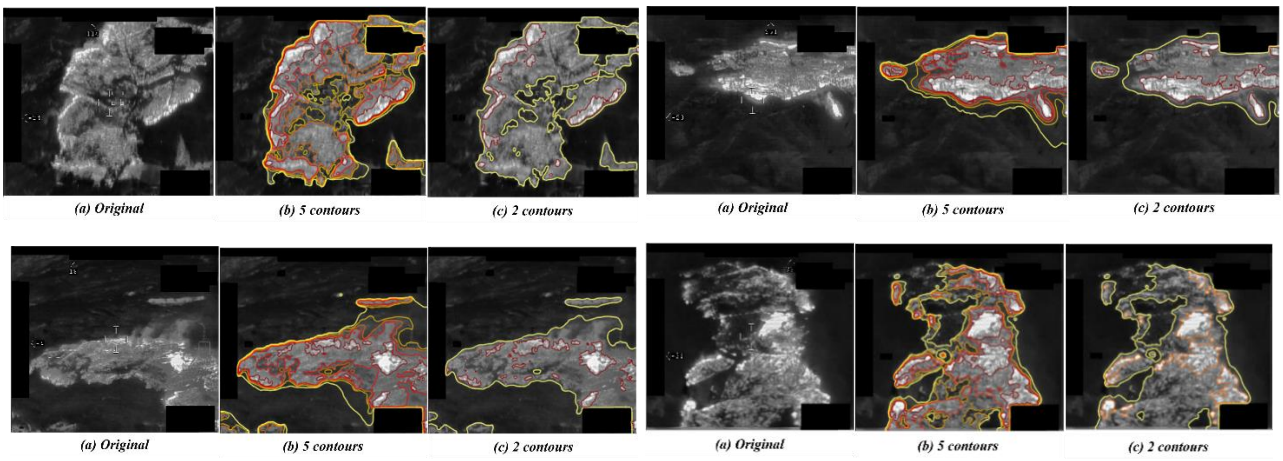

Figure 2 – Segmentation examples of FOGO_1 images using 5 contours with further reduction to 2 contours

To test the algorithm over the whole video sequence, all the frames in FOGO_1 were segmented and compiled in a video that was made available in the following link: https://youtu.be/kD9TEFOaHxw.

5.2. Comparison with other methods

With the objective of finding the best set of parameters for the level set method a grid search of parameters was conducted using the selected 55 frames, and to further compare its performance against other common unsupervised methods, they were also subjected to the application of the multi-layer Otsu (Otsu 1979; Liu and Yu 2009), K-means (Liu and Yu 2009, Sinaga and Yang 2020) and Mean Shift (Fukunaga and Hostetler 1975; Comaniciu and Meer 1999) algorithms. Class grouping was also applied when using these methods.

Experimental results, present in Table 1, were conducted in MATLAB 2020a, using a computer running in Windows 10 64-bit with an Intel® Core™ i7-6700HQ microprocessor and 8 GB of RAM. All metric values displayed correspond to the average of values for the three referred classes. For Mean Shift two types of inputs were tested, a 1-D input corresponding to the pixel intensity (i) and the 3-D input consisting of the pixel intensity (i) and position (x,y), with respective weights 0.8 and 0.2. Since Mean Shift calculates the number of classes automatically, we combine it with Otsu thresholding to aggregate classes. It is important to refer that the ground truths have compromised reliability, since they were not made by the authorities.

Table 1 – Comparison of results using FOGO_1 55 image labelled dataset

Method	Precision	Recall	Accuracy	IoU[1]	WEoS[2]	Time [s]
Proposed	**0.871**	**0.859**	**0.962**	**0.775**	**0.035**	2.86
Otsu	0.812	0.832	0.916	0.688	0.066	**0.15**
K-means	0.803	0.838	0.909	0.682	0.069	1.24
Mean Shift + Otsu (1D input)	0.862	0.741	0.926	0.646	0.056	1.65
Mean Shift + Otsu (3D input)	0.852	0.675	0.902	0.571	0.071	17.3

In terms of computational time, Otsu vastly outperforms by far all the other methods which allows for real-time segmentation. However, to obtain curve regularization, one need to increase the intensity of pre-process blurring, which leads to less reliable contours. Level sets allow for a highly customizable contour regularization therefore they present the best segmentation results.

6. Conclusions and Future Work

This work presented a method based on level sets that can detect the wildfire area and the burning fire front in wildfire thermal images which will enable the calculation and location of the fire perimeter and the most critical areas in future work. It outperforms other common unsupervised methods in terms of segmentation quality, albeit at a cost of computational time.

The lack of a labelled database dictated the non-usage of supervised learning. The hand labels created throughout the development of this work can be used to start such database. It would be interesting to apply state-of-the-art deep-learning methods to these images and compare the results with the ones obtained in this research.

7. Acknowledgements

This work was supported by FCT projects FIREFRONT (PCIF/SSI/0096/2017) and VOAMAIS (PTDC/EEI-AUT/31172/2017, 02/SAICT/2017/31172). The authors would like to thank everyone involved in these projects for the fruitful discussions and collaborations. Furthermore, we would like to express our gratitude to the

[1] Intersection over Union
[2] Weighted Error of Segmentation (same as EoS but a misclassification by just one class counts as a ‘half’ error)

Portuguese Air Force for providing the surveillance images and videos used throughout this work and to the ANEPC for the support and the many suggestions.

8. References

Banerjee, S., & Bhattacharya, M. (2010, October). Segmentation of medical images using Selective Binary and Gaussian Filtering regularized level set (SBGFRLS) method. In *2010 3rd International Conference on Biomedical Engineering and Informatics* (Vol. 2, pp. 541-545). IEEE.

Caselles, V., Catté, F., Coll, T., & Dibos, F. (1993). A geometric model for active contours in image processing. *Numerische mathematik*, *66*(1), 1-31.

Caselles, V., Kimmel, R., & Sapiro, G. (1997). Geodesic active contours. *International journal of computer vision*, *22*(1), 61-79.

Chan, T. F., & Vese, L. A. (2001). Active contours without edges. *IEEE Transactions on Image Processing*, *10*(2), 266–277.

Chung, G., & Vese, L. A. (2005). Energy minimization based segmentation and denoising using a multilayer level set approach. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 3757 LNCS, 439–455.

Chung, G., & Vese, L. A. (2009). Image segmentation using a multilayer level-set approach. *Computing and Visualization in Science*, *12*(6), 267–285. https://doi.org/10.1007/s00791-008-0113-1.

Comaniciu, D., & Meer, P. (1999, September). Mean shift analysis and applications. In *Proceedings of the seventh IEEE international conference on computer vision* (Vol. 2, pp. 1197-1203). IEEE.

Fukunaga, K., & Hostetler, L. (1975). The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Transactions on information theory*, *21*(1), 32-40.

Grau, V., Mewes, A. U. J., Alcañiz, M., Kikinis, R., & Warfield, S. K. (2004). Improved watershed transform for medical image segmentation using prior information. *IEEE Transactions on Medical Imaging*, *23*(4), 447–458.

He, L., & Osher, S. (2007). Solving the Chan-Vese Model by a Multiphase Level Set Algorithm Based on the Topological Derivative. In *Scale Space and Variational Methods in Computer Vision* (pp. 777–788). Springer Berlin Heidelberg.

Huang, Y., & Wu, J. W. (2010). Infrared thermal image segmentations employing the multilayer level set method for non-destructive evaluation of layered structures. *NDT and E International*, *43*(1), 34–44.

Huang, Y., Lee, M. G., Lin, S. Y., & Xiaoyu, Y. I. (2013). Segmenting thermal images of pervious concrete pavement temperature with employing the multilayer level set approach. In *ICSDEC 2012: Developing the Frontier of Sustainable Design, Engineering, and Construction* (pp. 757-764).

Li, C., Xu, C., Gui, C., & Fox, M. D. (2010). Distance regularized level set evolution and its application to image segmentation. *IEEE Transactions on Image Processing*, *19*(12), 3243–3254.

Liu, D., & Yu, J. (2009, August). Otsu method and K-means. In *2009 Ninth International Conference on Hybrid Intelligent Systems* (Vol. 1, pp. 344-349). IEEE.

Malladi, R., & Sethian, J. a. (1996). Level Set and Fast Marching Methods in Image Processing and Computer Vision. 1(4), 0–3.

Mumford, D. B., & Shah, J. (1989). Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on pure and applied mathematics*.

Ng, H. P., Ong, S. H., Foong, K. W. C., Goh, P. S., & Nowinski, W. L. (2006). Medical image segmentation using k-means clustering and improved watershed algorithm. *Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation*, *2006*, 61–65.

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, *9*(1), 62-66.

Osher, S., & Sethian, J. A. (1988). Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*, *79*(1), 12–49.

Sinaga, K. P., & Yang, M. S. (2020). Unsupervised K-means clustering algorithm. *IEEE access*, *8*, 80716-80727.

Valero, M. M., Rios, O., Pastor, E., & Planas, E. (2018). Automated location of active fire perimeters in aerial infrared imaging using unsupervised edge detectors. *International Journal of Wildland Fire*, *27*(4), 241–256.

Vese, L. A., & Chan, T. F. (2002). A multiphase level set framework for image segmentation using the Mumford and Shah model. *International Journal of Computer Vision*, *50*(3), 271–293.

Wen, D., Ren, A., Ji, T., Flores-Parra, I. M., Yang, X., & Li, M. (2020). Segmentation of thermal infrared images of cucumber leaves using K-means clustering for estimating leaf wetness duration. *International Journal of Agricultural and Biological Engineering*, *13*(3), 161–167.

Yuan, C., Liu, Z., & Zhang, Y. (2017). Fire detection using infrared images for UAV-based forest fire surveillance. *2017 International Conference on Unmanned Aircraft Systems, ICUAS 2017*, 567–572.