

Smoke Detection via Linear Separation

Banafsheh Rekabdar, Duncan Wilson
Department of Mathematics and Statistics
University of Nevada, Reno

May 6, 2015

Abstract

This is an implementation of two of the methods presented in Tian, Li et al's paper Smoke Detection in Video: An Image Separation Approach[2] on a data set of images from the Lake Tahoe and southern California region. Here we present a matting technique to determine salient regions in large images with small pixel regions corresponding to smoke. This allows for a significant amount the unimportant regions to be filtered out when passing along images to the classifier. We hope for this work to have positive utility in furthering the preprocessing stage of the algorithmic smoke detection pipeline.

1 Introduction

In Tian, Li et al, they propose three novel methods to detect regions of smoke in an image. This directly aids in the preprocessing phase of algorithmic smoke detection and we have implemented two of these solutions on data sets from the Lake Tahoe and the southern California region. Here we will disclose the structure for both of these algorithms that were implemented, followed by some analysis of our results and possible uses of this work.

2 Structuring the Problem

2.1 Background

The problem of building a system for smoke detection has many sub-problems outlined the smoke detection pipeline. This work is focused on the matting problem, which entails splitting an image into regions (possibly overlapping) and giving a prediction of whether or not the region contains smoke.

Since wildfires spread quickly this means that there is a great need for them to be detected as quickly as possible. Due to the structure of the classification pipeline, this means the matting phase needs to perform well in the cases where smoke is first appearing and spreading quickly. We've incorporated a correlation model into our algorithms to account for the rapid change which appears in the regions of quickly spreading smoke.

2.2 Smoke Structure

Since the problem of matting simplifies to detecting whether or not a region of pixels is smoke or not, the first thing we must examine is the structure of smoke. In Narasimhan and Nayar (2002)[1] they lay out multiple atmospheric scattering models of which two are assumed for this work.

The first of these models is that of attenuation. In physics attenuation refers to the phenomena of a loss of flux when a field passes through a medium. In this context the change in the flux of light through the medium of smoke is of the greatest importance. While Narasimhan and Nayar explicate an exponential law model, for this application the assumption of attenuation is all that matters.

The second model which is assumed is that of airlight, which is a phenomena by which light provokes the atmosphere to act like a light source. This is caused by particles in the atmosphere scattering the illumination of the light source. As with the attenuation model, there are analytical solutions relating to the scattering coefficient presented in Narasimhan and Nayar, but again for matting the assumption that smoke will create airlight is what is most important.

2.3 Linear Modeling

Given the two models above, we now will examine the problem's formulation and possible solutions: We treat each image as a collection of subsections of pixels $\mathbf{f}_t \in \mathbb{R}^N$. If, as is reasonable, we assume that the scattering coefficient only varies slightly within regions of

smoke we can treat regions of the image (represented by \mathbf{f}_t) as a linear combination of the background vector of the image without smoke in it (\mathbf{b}_t) and a smoke vector (\mathbf{s}_t):

$$\mathbf{f}_t = \alpha_t \mathbf{s}_t + (1 - \alpha) \mathbf{b}_t + n_t \quad (1)$$

With $n_t \in \mathbb{R}^N$ representing the noise in the model. With the assumption of the airlight model \mathbf{s}_t can be thought of as the scattering component with an infinite thickness. This seemingly strange assumption allows for the airlight model to be incorporated to be used in this linear fashion.

In (1), α_t can be thought of as a blending parameter that is $\in [0, 1]$. (1) depends on both the thickness of the smoke as well as the scattering (smoke) coefficient at time t . So while the the blending will vary region from region, we will further assume that α won't vary with in a specific region of smoke, due to the likely uniform thickness in a specific small region.

As the team in the seismology lab have informed us, \mathbf{f}_t is obtained from a camera which is not moving and due to there being a plethora of non-smoke containing images from the camera we need not doing any background modeling in order to find \mathbf{b}_t .

This results in the salient region selection can formulated as a two variable estimation problem of α_t and \mathbf{s}_t by minimizing the following:

$$\min \|\mathbf{f}_t - \alpha_t \mathbf{s}_t - (1 - \alpha_t) \mathbf{b}_t\|^2 \text{ s.t. : } \alpha_t \in [0, 1] \quad (2)$$

But this equation has two unknowns and one equation and thus results in an infinite amount of solutions. So our best hope to estimate both variables is by constraining either \mathbf{b}_t or \mathbf{s}_t and as our background from different cameras will be completely different but in each image the smoke should retain a similar structure, we will constrain \mathbf{s}_t for the following two methods.

3 Local Smoothness

3.1 Model Formulation

The premise that this model rests on is that a given a small region of smoke there appears to be a great deal of local smoothness. Which in turn means that this region must have roughly similar pixel values. This observation allows us to modify (2) to include a measure of local smoothness:

$$\min \|\mathbf{f}_t - \alpha_t \mathbf{s}_t - (1 - \alpha_t) \mathbf{b}_t\|^2 + \lambda \sum_{i=1}^N \sum_{j \in \varphi_i} (s_{t_i} - s_{t_j})^2 \text{ s.t. : } \alpha_t \in [0, 1] \quad (3)$$

Now in (3) we see that the minimization also considers the amount of local smoothness present. λ can thought of as a tuning parameter by which more or less smoothness of subregions are considered and φ_i is the local region in which smoothness is assumed and is centered at the i' th pixel.

To make this math more compact and concise we can formulate the second term in the equation with a matrix representation. Let \mathbf{T} be an $N \times M$ matrix with elements $\in \{-1, 0, 1\}$.

M is decided when picking your size of region φ . Each the $N \times i$ element of \mathbf{T} equals 1 and all others $\in \varphi_i$ equals -1 . This means the second term in (3) becomes:

$$\sum_{i=1}^N \sum_{j \in \varphi_i} (s_{t_i} - s_{t_j})^2 = \|\mathbf{T}\mathbf{s}_t\|^2 = \mathbf{s}_t^T \mathbf{T}^T \mathbf{T} \mathbf{s}_t = \mathbf{s}_t^T \mathbf{A} \mathbf{s}_t \quad (4)$$

For simplicity's sake we've set $\mathbf{A} = \mathbf{T}^T \mathbf{T}$. Substituting the result of (4) back into (3) we get:

$$\min \|\mathbf{f}_t - \alpha_t \mathbf{s}_t - (1 - \alpha_t) \mathbf{b}_t\|^2 + \lambda \mathbf{s}_t^T \mathbf{A} \mathbf{s}_t \quad s.t. : \alpha_t \in [0, 1] \quad (5)$$

3.2 Analytical Solutions

Due to the under-determined nature of (5) to optimize you must trade off between fixing either s_t or α_t . To solve for s_t , hold α_t constant and set the partial derivative to 0. This results in the following optimal solution to s_t :

$$\hat{s}_t = (\alpha_t^2 \mathbf{I} + \lambda \mathbf{A}_t)^{-1} \alpha_t (\mathbf{f}_t - \mathbf{b}_t + \alpha_t \mathbf{b}_t) \quad (6)$$

with \mathbf{I} being an $N \times N$ identity matrix.

We also can derive an optimal solution to α by holding s_t constant:

$$\check{\alpha} = \frac{(\mathbf{b}_t - \mathbf{s}_t)^T (\mathbf{f}_t - \mathbf{b}_t)}{(\mathbf{b}_t - \mathbf{s}_t)^T (\mathbf{s}_t - \mathbf{b}_t)} \quad (7)$$

Due to the constraint that $\alpha \in [0, 1]$ in order to obtain the optimal α we constrain $\check{\alpha}$ by:

$$\hat{\alpha} = \begin{cases} 0 & \text{if } \check{\alpha} \text{ is } \leq 0 \\ \check{\alpha}, & \text{if } 0 < \check{\alpha} < 1 \\ 1 & \text{if } \check{\alpha} \text{ is } \geq 1 \end{cases} \quad (8)$$

So to optimize, oscillate between these two equations until convergence or an upper bound of maximum iterations is reached.

4 PCA Model

4.1 Model Formulation

Although the additional term to the minimization eliminates non-smooth sub-regions, it still could fall victim to detecting other regions which are locally smooth as smoke depending on your λ parameter in (5). We'd rather have a system which incorporates the features of smoke more specifically and holistically. That leads us to the principle components:

A given pixel region of pure smoke, with high probability, lies in a lower dimensional sub-space than the totality of the image in which it is contained. Using principle component analysis (Turk and Pentland 1991 [3]) we are able to select the salient dimensions which correspond to the subspace. In order to find this subspace you first compute the covariance

matrix corresponding to N pure smoke images, which results in an $N \times N$ matrix, followed by computing the corresponding eigenvalues and vectors. Once you have the eigenvalues, sort the two sets according to magnitude of eigenvalues and select the first eigenvectors up to the upper bound of dimensionality you desire. This results in the information dense dimensions of smoke to be retained and then can be used in the detection process.

The subspace can be represented as a matrix $\mathbf{E} \in \mathbb{R}^{N \times L}$ where L is the reduced dimensionality (thus a general assumption is that $L < N$). As chosen from the aforementioned process, each column of \mathbf{E} is an eigenvector corresponding to pure smoke, an "eigensmoke". Assuming that the dimensionality reduction retained all of the salient dimensions (as we'd hope!) we can write the smoke vector at any time-step as $\mathbf{s}_t = \mathbf{E}\mathbf{y}_t$, where $\mathbf{y}_t \in \mathbb{R}^L$ and can be thought of as operational vector by which \mathbf{s}_t is projected onto \mathbf{E} . This means that we can rewrite (2) to incorporate this new formulation of \mathbf{s}_t :

$$\min \|\mathbf{f}_t - \alpha_t \mathbf{E}\mathbf{y}_t - (1 - \alpha_t)\mathbf{b}_t\|^2 \quad s.t. : \alpha_t \in [0, 1] \quad (9)$$

When holding α_t or \mathbf{y}_t constant, this equation becomes quadratic and allows for clean analytical solutions to both α_t and \mathbf{y}_t by which \mathbf{s}_t can be found via \mathbf{E} .

4.2 Analytical Solutions

When holding α fixed you can set the partial derivative, with respect to \mathbf{s}_t , to 0 find the optimal $\hat{\mathbf{s}}_t$:

$$\hat{\mathbf{s}}_t = \mathbf{E}(\alpha_t \mathbf{E}^T \mathbf{E})^{-1}(\mathbf{f}_t - \mathbf{b}_t + \alpha_t \mathbf{b}_t) \quad (10)$$

Likewise with α :

$$\check{\alpha} = \frac{(\mathbf{b}_t - \mathbf{s}_t)^T(\mathbf{f}_t - \mathbf{b}_t)}{(\mathbf{b}_t - \mathbf{s}_t)^T(\mathbf{s}_t - \mathbf{b}_t)} \quad (11)$$

To obtain $\hat{\alpha}$ a three case limiting condition as shown in (8) is needed in order for $\check{\alpha}$ to be within acceptable bounds.

4.3 Correlation

We also included an estimate of correlation in both of our models using the following equation:

$$c = \frac{(\mathbf{f}_t - \bar{\mathbf{f}}_t)(\mathbf{b}_t - \bar{\mathbf{b}}_t)}{\sqrt{(\mathbf{f}_t - \bar{\mathbf{f}}_t)^2(\mathbf{b}_t - \bar{\mathbf{b}}_t)^2}} \quad (12)$$

The only regions which are examined have low correlation, ($c < 0.001$ for PCA and $c < 0.0001$ for local smoothness). This allows us to skip selection for a large amount of regions that stay highly similar in the time from when the background photo was taken to the time which smoke is needed to be detected. Environmental conditions like wind can cause many non-important regions to still possibly be selected even with correlation.

5 Results

Our results vary by camera position and by method used. Overall this matting step in the pipeline, due to the massive size of Tahoe images, is at least minimally helpful by decreasing the number of regions to pass on to the classifier. We’ve included results from different data sets from both the Tahoe and southern California region.

In all results that follow the color of the points corresponds with different α values. Specifically green is plotted if $\alpha \geq 0.9$, blue if $0.5 \leq \alpha < 0.9$, and red if $0.3 < \alpha < 0.5$. Also for all results involving the local smoothness model, it should be assumed that $\lambda = 0.7$. (determined as at least semi-optimal via experiment)

5.1 Local Smoothness

The local smoothness model works very well on a majority of the southern California data but, less than perfect on the majority of the images from Tahoe. As the example below shows, the local smoothness model, even when incorporating correlation, still misclassifies many regions in the image as smoke. But it nearly consistently classifies the smoke very well. Due to the rarity of false negatives in the region detection, we feel that this should offer a reliable means to largely reduce the amount of regions needed to be processed by the classifier.

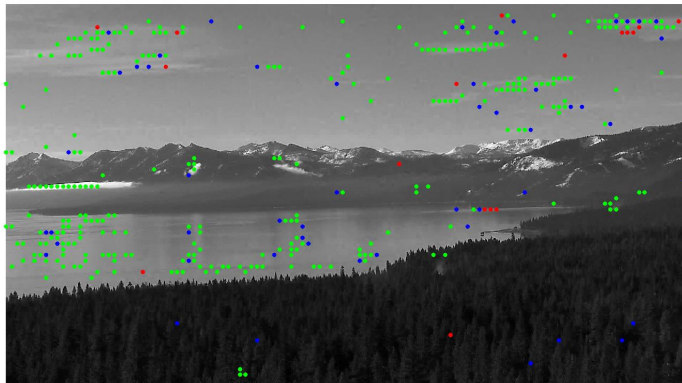


Figure 1: This photo from an early stage of a fire from the Tahoe data set

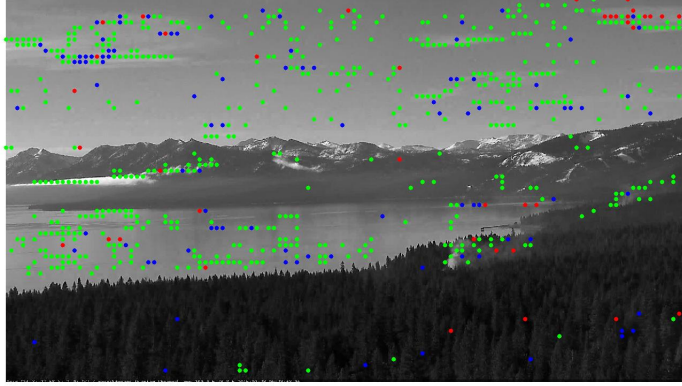


Figure 2: This photo from a later stage of the same fire as above, with regions in both fires on the right of the image well selected for.

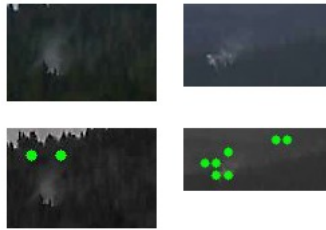


Figure 3: Close ups of the fire and it's region selection via local smoothness from Figure 1.

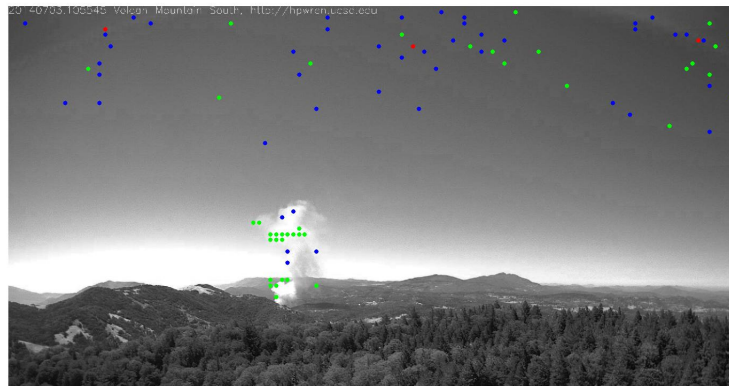


Figure 4: The best result using local smoothness. (from the southern California data set)

5.2 PCA

5.2.1 Eigensmoke

We collected the training data used in Lian, Ti et al of pure smoke in order to be able to compute the most information rich dimensions.

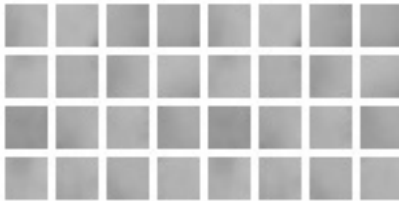


Figure 5: The subset of the pure smoke data set by which the eigensmokes were computed.

By computing the covariance matrices and subsequently the eigenvalues and vectors, we were able to compute our "eigensmokes", below are twenty of the eigensmokes.



Figure 6: The top 20 eigensmokes from the training data set



Figure 7: The top 20 eigensmokes from Tian, Li, et al.

5.2.2 PCA results

Principle component analysis fared quite a bit better in not returning nearly as many false positives as than the local smoothness assumption.

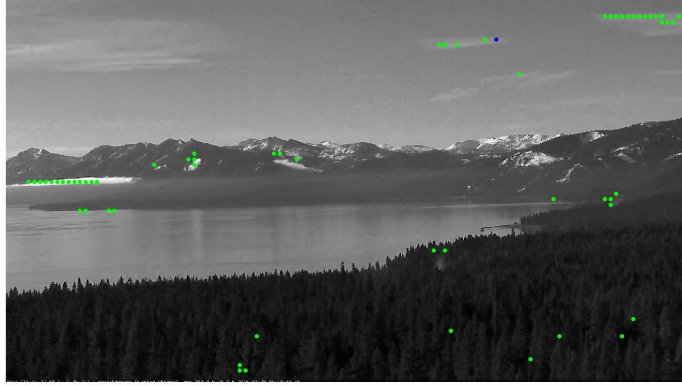


Figure 8: PCA's region selection on the same image as Figure 1.

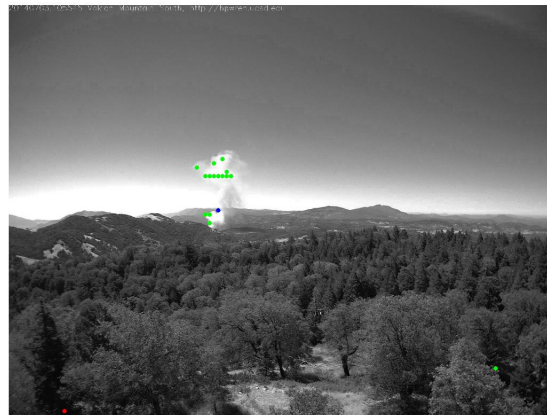


Figure 9: The best result using PCA. (from the southern California data set)

There are many more results of both good performance as well as one false positive via PCA on a very cloudy/foggy picture.

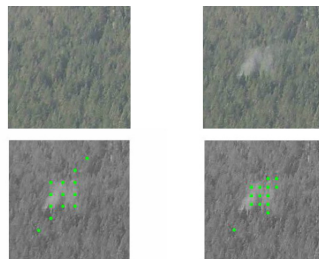


Figure 10:

Top: Background and foreground images Bottom: local smoothness' and PCA's regions

6 Conclusions

Our results serve as a good matting tool for preprocessing out the salient regions in a given image. Using the principle component analysis method there seems to be a intrinsic flaw in the structure of the eigensmokes. They look a lot like clouds! This means that on clear days there wouldn't be a high rate of cloud containing regions being passed along to the classifier.

While local smoothness has the drawback of generating lots of false positive, we never found a time that it generated a false negative. Our results imply that PCA is much more accurate but on the other hand, when there is heavy cloud cover or fog can miss the smoke. When combined with the other preprocessing methods, we see this work being reliable enough to be integrated into the final system.

Acknowledgements

We'd like to greatly thank Dr. Raul Rojas for the amazing semester and personal involvement in this work. We're also quite grateful to Tian, Li et al for outlining the methodology used here, as well as sharing their pure smoke data set so that PCA could be used.

References

- [1] Srinivasa G Narasimhan and Shree K Nayar. Vision and the atmosphere. *IJCV*, 48(3):233–254, 2002.
- [2] Hongda Tian, Wanqing Li, Lei Wang, and Philip Ogunbona. Smoke detection in video: An image separation approach. *International Journal of Computer Vision*, 106(2):192–209, 2014.
- [3] Matthew A Turk and Alex P Pentland. Face recognition using eigenfaces. In *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR'91., IEEE Computer Society Conference on*, pages 586–591. IEEE, 1991.