SUPERPIXEL-BASED STATISTICAL ANOMALY DETECTION FOR SENSE AND AVOID

Odysseas A. Pappas, Alin M. Achim, David R. Bull

Visual Information Laboratory, University of Bristol, Bristol BS8 1UB, UK

ABSTRACT

This paper presents a novel preprocessing method for detecting small objects of interest within a high-resolution image, applied to the problem of visually detecting possible aircraft collisions (Sense and Avoid) for UAV platforms. The method is based on superpixel image segmentation combined with subsequent statistical analysis and anomaly detection. The existence of a possible target within a superpixel is described in terms of how it affects the local superpixel statistics and this signature statistical profile is consequently used to identify regions of interest throughout the image. The approach eliminates upwards of 90% of the total image area, significantly reducing the workload of further processing stages.

Index Terms— Superpixels, Anomaly Detection, Sense and Avoid, Statistical Detector

1. INTRODUCTION

A commonly encountered and challenging scenario in image analysis is the detection of a small target present in a comparatively much larger and cluttered image. A lack of information about the exact visual properties of the target of interest can make this task significantly more challenging. Such a problem can be viewed as an anomaly detection task, where an anomaly is defined as an otherwise unspecified object that stands out from its background, i.e. differs from the immediate context in which it finds itself [1].

Anomaly detection typically refers to the task of detecting events and objects within a dataset that do not conform to expected patterns and behaviours. In the field of image processing, such methods have previously been used to tackle the aforementioned problem of detecting small targets in high-resolution remote sensing data. Notable examples of anomaly detection applied to hyperspectral images include Stein et al. [2] and Theiler and Prasad's context-dependent detector [1].

This paper presents a general approach to tackling such detection problems using superpixels instead of previously used segmentation methods like HierArchitect [1], with an example case study drawn from the problem of visual Sense and Avoid for Unmanned Aerial Vehicles. Sense and Avoid refers to the process of detecting and avoiding a possible collision with another object (such as another aircraft). In the field of UAVs, this problem is of prime importance since there is no human pilot present. In order to enable fully autonomous UAVs to be employed in a variety of civilian applications in both controlled and uncontrolled airspace, an automated system for detecting and avoiding collisions is required [3].

Sense and Avoid can generally be performed using a plethora of modalities, including cooperative systems such as TCAS as well as non-cooperative methods such as radar and passive imaging sensors. The comparatively low cost, weight and power consumption of visual sensors makes them particularly suited for usage on small, civilian UAVs. For an in-depth discussion of these modalities, along with a general literature review on Sense and Avoid systems, the reader is directed to [4].

Previous approaches to visual Sense and Avoid have employed tiered processing systems, starting with a preprocessing step to identify possible targets/regions of interest through the image, followed by more elaborate object recognition and tracking algorithms. The two best known systems proposed in the literature [5] [6] employ simple morphological filtering for preprocessing. Operations such as the top-hat, bottom-hat and close-minus-open filter can be used to detect small regions of contrast difference, i.e. potential targets that differ in contrast from the local background. The usage of such filters requires assumptions about the target size, shape and contrast relative to background. In complex natural images, these methods are likely to detect, besides the targets of interest, a very large number of spurious artefacts.

Viewing the Sense and Avoid preprocessing task in terms of image anomaly detection, our approach employs homogeneous local neighbourhoods derived from SLIC superpixel segmentation. Anomalies of interest are now superpixels containing a possible target, the presence of which affects their local statistics and differentiates them from the remaining majority of homogeneous superpixels which can be discarded from any further processing stages.

The paper is organised as follows: Section 2 provides a discussion on superpixel segmentation and Section 3 describes the statistical detector approach. Section 4 presents results of the aforementioned method along with a relevant discussion with the conclusion found in Section 5.

The authors would like to acknowledge EPSRC for funding this project via the Centre for Doctoral Training in Communications as well as THALES UK for the provision of test data.

2. SUPERPIXEL SEGMENTATION

Superpixels are a form of image over-segmentation; they are essentially clusters of perceptually similar pixels that capture the redundancy inherent in natural images. They can provide a perceptually more meaningful alternative to the pixel grid. Superpixels can act as a convenient primitive for further image processing tasks, having the potential to significantly reduce overall algorithm complexity.

Various superpixel methods have been proposed in the literature and consequently employed in a number of image processing and computer vision applications, including image and video segmentation [7] [8], object localisation [9] and tracking [10].

As there is no rigorous definition of what constitutes a superpixel, the results of these algorithms can vary in segmentation quality, superpixel uniformity, size and number. One of the most widely used algorithms, proposed by Achanta et al. is *Simple Linear Iterative Clustering* (SLIC); a description of SLIC superpixels and comparison to other state-of-the-art methods can be found in [7].

SLIC, chosen for the work described here, performs iterative clustering in a fashion similar to k-means clustering [7]. The image is segmented into superpixels, whose total number (k) is defined by the user. The segmentation process is governed by a 5-dimensional distance metric combining spatial (*x,y* coordinates) and colour information (*L,a,b*, of the CIELAB colorspace). The distance between cluster centre (C_k) and a pixel i is shown in the equations below.

$$d_{lab} = \sqrt{(l_k - l_i)^2 + (a_k - a_i)^2 + (b_k - b_i)^2} \quad (1)$$

$$d_{xy} = \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2} \quad (2)$$

$$D_s = d_{lab} + \frac{m}{S} d_{xy} \quad (3)$$

The spatial component of the metric can dominate the end result when dealing with large superpixels (due to high image resolution or low k) and hence a scaling factor is introduced as $\frac{m}{S}$, where *S* is the initial cluster seed grid interval (dependent on image size and desired k) and *m* allows the user to control superpixel compactness and shape regularity.

Advantages of SLIC superpixels include high perceptual homogeneity within each superpixel, relative uniformity in size and shape, computational simplicity and efficiency as well as a degree of user control over the process. They also perform well in terms of standard boundary recall and under-segmentation error measures and can cope with both colour and grayscale imagery [7].

The approach described in this paper utilises superpixels in a semantically different context, where they are viewed not so much as primitives composing an image but more as an alternative paradigm to classic rectangular neighbourhoods. Neighbourhood operations are incredibly common in image processing and are typically performed over a rectangular window of arbitrary size, centred on the pixel or coefficient of interest. We here consider a pixel's local neighbourhood to be the superpixel it belongs to; neighbourhood operations are performed according to the values of the pixels within a superpixel instead of the pixels in a strictly defined grid.

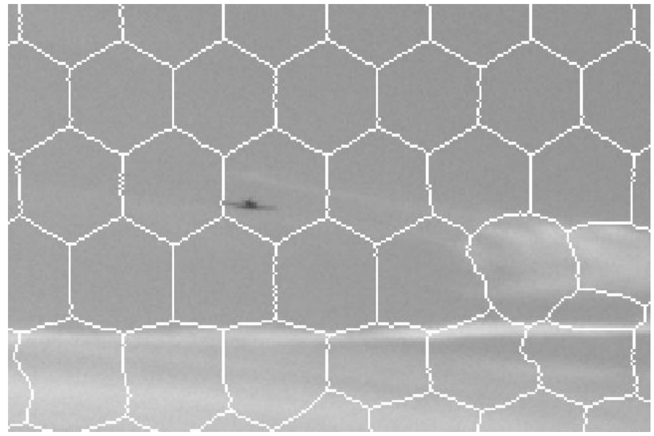


Fig. 1. Example of an aircraft target contained within a larger superpixel.

In the context of our problem, superpixel neighbourhoods can be beneficial since a homogeneous local background will cause possible targets to be more readily distinguished.

3. THE SUPERPIXEL STATISTICAL DETECTOR

Our preprocessing approach relies on segmenting the image into superpixels whose size is significantly larger than the expected size of the target of interest; in our case, an approaching aircraft. The minimum allowed SLIC superpixel size is set at such a threshold that any target will be too small to form a superpixel by itself and will instead appear contained, along with its local background, in a much larger superpixel (Fig. 1). Only approximate knowledge of the target size is required; we have found superpixel size selection to be rather lenient, with sizes from one to two orders of magnitude above the expected target size to work well. The only assumption made about the target is that it is actually visible, i.e. it is a collection of pixels that differ in contrast from their immediate local background.

Since the images of interest are natural images containing large features such as expanses of sky, clouds or terrain, the vast majority of the generated superpixels are clusters of largely uniform pixels, the products of successful SLIC segmentation. However, any target present inside a superpixel will manifest itself as a "blob" of pixels of differing intensity, disturbing this expected uniformity. The task of finding potential targets and regions of interest in the image is now reduced to finding those superpixels throughout the image that demonstrate this "statistical anomaly".

An attempt to formulate this "statistical anomaly" can begin by examining the statistical moments of the pixel values within a given superpixel, for example measures describing statistical dispersion such as range, standard deviation σ and variance σ^2 of the constituent intensity values. For example, a superpixel that is a cluster of perceptually uniform pixels would be expected to have a relatively low standard deviation. A visible target present inside a superpixel is effectively a small set of values that differ from their local background to some extent, as it will appear darker or lighter than the local background. Such a superpixel would exhibit higher statistical dispersion than the uniform case and a -by comparison- higher standard deviation.

Another moment of interest is the sample kurtosis (Eq. 4). A superpixel that does not contain a target is likely to be largely uniform, having a relatively flat (platykurtic) distribution and hence relatively low kurtosis. The presence of a target will disturb this uniformity resulting in a distribution among the superpixel that is much "peakier" and exhibits heavy tails and a relatively high sample kurtosis. It should be noted that the superpixel kurtosis values themselves follow a heavy-tailed distribution (Fig. 2); the majority of superpixels have scores just under the global average with a handful of extremely high-scoring outliers.

$$g_2 = \frac{m_4}{m_2^2} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2)^2} \tag{4}$$

As implied above, our interest lies not in estimating the exact standard deviation and kurtosis values within superpixels but rather in how they relate to the values of other superpixels in the image. Statistical variations can help identify superpixels that may contain a target and differentiate them from the much more uniform ones that can be safely discarded from further processing.

The problem of deciding whether a superpixel contains a possible target or not can be viewed as a composite hypothesis testing problem, where the distributions describing each hypothesis are only partially known. The previously discussed superpixel statistics can inform as to the formulation of the hypotheses describing the case of a largely uniform superpixel, unlikely to contain a target of interest (H_0) and a leptokurtic superpixel containing a possible target (H_1).

One of the most common methods employed when faced with such a composite hypothesis testing problem is that of a Generalised Likelihood Ratio Test [11] shown in Eq. 5. As the parameter spaces of the distributions describing the two hypotheses are not fully known they need to be estimated using Maximum Likelihood Estimation on the dataset of interest (i.e. the intensity values contained in the superpixel).

$$L_G(x) = \frac{p(x;\hat{\theta}_1, H_1)}{p(x;\hat{\theta}_0, H_0)} > \gamma \tag{5}$$

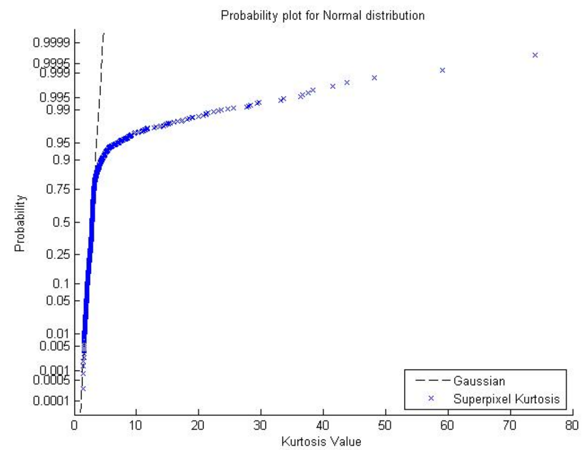


Fig. 2. Normal Probability Plot for the calculated sample kurtosis values of 2000 superpixels throughout an image, demonstrating the rarity of a much-higher-than-average result.

Taking the expectation of "high kurtosis" in the presence of a target, we can model the largely homogeneous case of H_0 as a relatively flat distribution, for example the Uniform distribution. If a possible target is present in a superpixel then it will not be as homogeneous and its statistics would better match a much more leptokurtic distribution. H_1 is hence defined as a superpixel following the Generalised Extreme Value Distribution (pdf shown in Eq. 6 and 7).

$$GEV pdf = \frac{1}{\sigma} t(x)^{\xi+1} e^{-t(x)} \tag{6}$$

$$t(x) = \begin{cases} (1 + (\frac{x-\mu}{\sigma})\xi)^{-1/\xi} & \text{if } \xi \neq 0 \\ e^{-(x-\mu)/\sigma} & \text{if } \xi = 0 \end{cases} \tag{7}$$

The GEV distribution is chosen as one of many possible leptokurtic candidates, others being the Cauchy or the Generalised Laplacian distributions. Our interest again lies not so much in how well this distribution fits a target containing superpixel but more in the contrasting relationship between an extremely platykurtic and an extremely leptokurtic distribution.

The above metrics are just some examples of the many possible identifiers of an anomalous superpixel; one may wish to employ one or more of these as well as any new metrics (such as skewness, Kolmogorov-Smirnov tests, the Mahalanobis distance) according to knowledge of the problem-specific data.

4. RESULTS

Our Statistical Superpixel Detector has been tested on a number of frames from 3 sequences of field-recorded data provided by THALES UK, a sample frame of which can be seen in Fig. 3. Our primary aim here is not so much to accurately detect the target aircraft but instead to eliminate parts of the image from unnecessary further processing. Table 1 shows the results of the detector over a series of 30 frames from each of the 3 datasets. The images have been segmented using the SLIC algorithm ($k = 2000, m = 10$) and each frame contains a single target. Superpixels are characterised according to one or more the aforementioned metrics and those scoring above threshold are selected for further processing (those discarded have been masked out for ease of representation). To illustrate the robustness of the approach in terms of superpixel size, we include dataset 3*, a version of dataset 3 segmented into ca. 3500 superpixels per frame.

Dataset	1	2	3	3*
Total Superpixels	58500	58600	58496	104491
Detected Superpixels	4153	5371	7129	14624
Detected Targets	30/30	30/30	30/30	30/30

Table 1. Results of detector over series of frames. Range and kurtosis threshold set at global average.

Our approach delivers excellent results, with the reduction in image area of interest often being in excess of 90%. The target superpixel is retained on all frames together with a few false positives (usually caused by small cloud formations). This results in a significantly lower number of detections per frame when compared to morphological filtering. Our detector selects about 10% of the created superpixels per frame (consequently, 10% of the total image area given the relative size uniformity of the superpixels), even when thresholds are set at a conservative global average. By comparison, a close-minus-open operation with an appropriate structuring element applied to samples from all 3 datasets marked pixels corresponding to (on average) 47% of the image area.

The significant reduction in image area of interest greatly reduces the load of further processing stages. A variety of methods (such as deformable template matching) can then be employed to easily distinguish between the actual target and any false positives.

Algorithm performance is of course dependent on threshold selection. Statistical dispersion and kurtosis thresholds set at the global average can be considered conservative as the superpixels of interest produce values that lie deep into the right-hand tails of the global distribution, as discussed previously. A GLRT threshold of 1 can also be considered conservative as a non-uniform superpixel is certain to be closer to H_1 than to the uniform distribution of H_0.

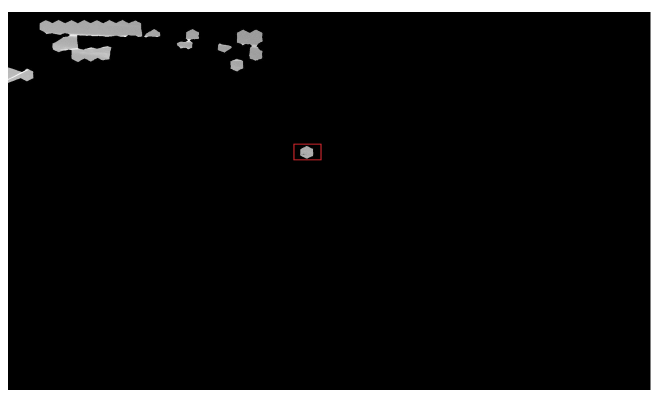
On most datasets the selection thresholds can be increased to multiples of the global average to select values lying in the upper quartile of the distribution, typically resulting in a reduction of false positives while maintaining target detection. Such thresholds are of course heavily data dependent and are most safely derived in an experimental, application specific fashion.

5. CONCLUSION

This paper has introduced a statistical anomaly detection-inspired framework for the detection of small targets in cluttered images using superpixel neighbourhoods. The approach is demonstrated as a preprocessing step in the Sense and Avoid problem, where it has been shown to greatly reduce the image search area while maintaining the target of interest.

(a) Original Image

(b) SLIC Segmentation



(c) Selected Superpixels

Fig. 3. Target detection in sample image from Dataset 2, k = 2000. Standard deviation threshold at global average, kurtosis threshold at x2 the global average.

6. REFERENCES

[1] J. Theiler and L. Prasad, "Overlapping Image Segmentation for Context-Dependent Anomaly Detection," *SPIE 8048, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVII*, 2011.

[2] D. W. J. Stein, S. G. Beaven, L. E. Hoff, E. M. Winter, A. P. Schaum, and A. D. Stocker, "Anomaly Detection from Hyperspectral Imagery," *IEEE Signal Processing Magazine*, vol. 19, pp. 58–69, 2002.

[3] P. Angelov (.ed), *Sense and Avoid in UAS: Research and Applications*, Wiley, 2012.

[4] S. B. Hottman, K. R. Hansen, and M. Berry, "*Literature Review on Detect, Sense and Avoid Technology for Unmanned Aircraft Systems*," Tech. Rep. DOT/FAA/AR-08/41, US Federal Aviation Administration, 2009.

[5] D. Dey, C. M. Geyer, S. Singh, and M. E. DiGioia, "Passive, Long-range Detection of Aircraft: Towards a Field Deployable Sense and Avoid System," *Proceedings Field and Service Robotics (FSR '09)*, pp. 113–123, 2009.

[6] J. Lai, J. J. Ford, L. Mejia, P. O'Shea, and R. Walker, "See and Avoid Using Onboard Computer Vision," *Sense and Avoid in UAS: Research and Applications, ed. P. Angelov, Wiley*, pp. 265–294, 2012.

[7] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "SLIC Superpixels Compared to State-of-the-Art Superpixel Methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 2274–2282, 2012.

[8] R. Trichet and R. Nevatia, "Video Segmentation with Spatio-Temporal Tubes," *IEEE AVSS International Conference on Advanced Video and Signal-Based Surveillance*, pp. 330–335, 2013.

[9] B. Fulkerson, A. Vedaldi, and S. Soatto, "Class Segmentation and Object Localization with Superpixel Neighborhoods," *IEEE 12th International Conference on Computer Vision*, pp. 670–677, 2009.

[10] F. Yang, H. Lu, and M. H. Yang, "Robust Superpixel tracking," *IEEE Transactions on Image Processing*, vol. 23, pp. 1639–1651, 2014.

[11] S. M. Kay, *"Fundamentals of Statistical Signal Processing: Detection Theory"*, Prentice Hall, 1998.