AUTONOMOUS IMAGE EXTRACTION AND SEGMENTATION OF IMAGE USING UAV'S

Radha Krishna Rambola, Associate Professor, NMIMS University, India
Akash Agrawal , Student at NMIMS University, India

ABSTRACT

Due to the large demand of Unmanned Aerial Vehicle (UAV) for various surveillance and search operations it is important to develop a process for autonomous segmentation and extraction of objects from aerial images. This paper presents a novel image processing based approach that image processing using low cost camera for a target within related domains. Real time autonomous extraction of visually significant targets in the image thus becomes imperative and an important problem whose solution has been provided in this paper. Our proposed relies on heuristically-based and deep learning methods to detect and identify target sightings in images retrieved from the aircraft through the ground control system (GCS).

Keywords: Image Processing, Drone, Navigation and Surveillance, Image.

INTRODUCTION

In recent years, there has been increased use of Unmanned Aerial Vehicles. Complex applications like navigation, surveillance, search and rescue operations typically involve processing video or imagery feed to identify significant objects from a wide open, uncluttered and uniform environment. Search and surveillance operations is a complex and highly expensive involving many personnel and multiple airborne search platforms. These are often employed to areas such as mountainous, ocean, forest or desert areas. Autonomously scanning the video feed to locate and extract visually significant objects is an interesting computer vision problem to solve and when applied to real world applications. It significantly improves the autonomy of such operations and reduces the cost and need for personnel for continuously monitoring the feed.

To locate and identify a unique unknown object in a uniform, uncluttered environment often involves identifying regions in the image where an unknown object might be present. Previous works in the field of computer vision have tackled this problem using visual saliency based detection. Term "saliency" was first used on work of rapid scene analysis (Ernst Niebur, 1998). Unlike images considered in previous work, aerial images captured by a point and shoot camera may contain multiple objects of interest which individually occupy a very small percentage of area within an image.

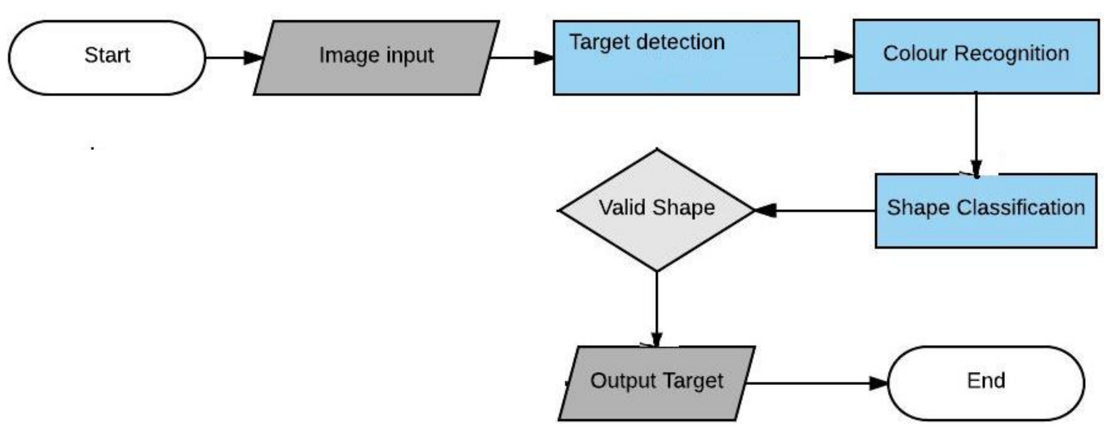


Figure 1: Flowchart of Image extraction and segmentation

PROPOSED WORK

Preprocessing

Aerial images captured often have a unique characteristic that a large portion in images is uniform and uncluttered. There can be multiple object of interest scattered throughout the image, which may not have a fixed shape and size. These objects occupy a very minute portion of the image which varies according to the altitude of the UAV. The system should be able to identify the locations of all such objects of interest in the image before they can be extracted. The algorithm thus starts from basic pre-processing steps followed by construction of saliency maps. These steps are described in detail in the following sections:

A. Smoothening and Mean Shift Filtering

The first step is pre-processing the image to remove noise and texture from the image. The small variations caused by the texture of the background like grass can add to the difficulty of finding contrast and hence this step is important to make the image analysis easier. The details of the process are as follows:

Step 1: Gaussian Smoothing is performed on the original image shown in Figure 2(a) to remove unwanted noise. The process of saliency map generation will involve performing Canny Edge Detection and thus requires a smoothing operation. Gaussian Blur is known to perform better than other smoothing filters in reducing noise before edge detection algorithms. Figure 2(b) is the image obtained after smoothing.

Figure 2 (a) Original Image. (b) Image obtained after smoothening.

Step 2: This step performs a mean-shift filter on the smoothened image which results in a posturized image with colour and fine-grain texture flattened (Peter Meer, 1999). This will greatly enhance the saliency map constructed in the next section. Image obtained after this step is referred is shown in Figure 3

Figure 3 Image obtained after Mean-Shift Filtering

B. Constructing Saliency Map

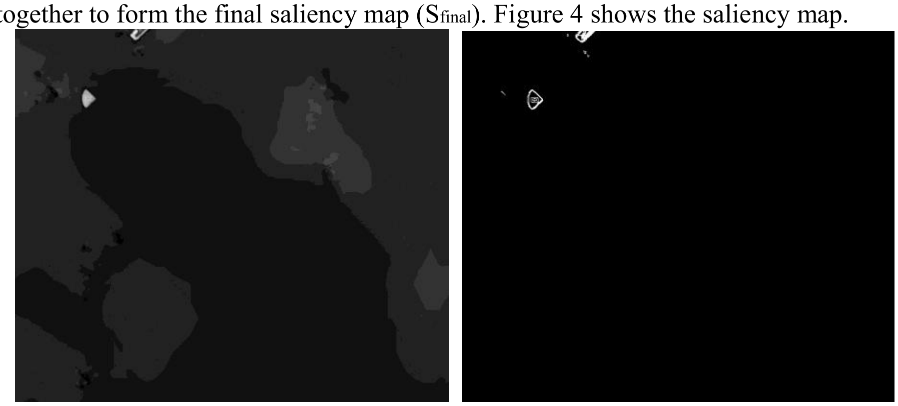
The key step in object detection in aerial images is to compute the saliency map. It represents a heap map where the white portions of the image represents area which are visually significant and may possibly contain an object of interest while the darker portions of the saliency map represents background. To compute saliency map of full resolution, the system uses an approach as described by Achanta et al. (2009) which utilizes colour and luminance information for computing a frequency tuned saliency map. The process can be briefly described as follows:

1. The image obtained after mean shift filtering (I_{msf}) and the image obtained after smoothing (I_{gs}) is converted to LAB colour space. LAB colour space has one channel (L) for luminance and other two channels (A and B) for colour. Unlike RGB colour space, the LAB colour space is perceptually more uniform. So the distances between the colour coordinates correspond to the perceived difference between two colours. The two images obtained after this step are represented as I_{msf_lab} and I_{gs_lab} for I_{msf} and I_{gs} respectively.

2. To calculate Saliency map S(x,y) , arithmetic mean of each channel for all pixels of I_{msf_lab} is calculated. Let this value be I_{mean}. The value S(x,y) is the Euclidean Distance between I_{mean} and I_{gs_lab}(x,y). It also gives a full resolution saliency maps.

C. Normalization and Canny Image

Furthermore, nine different canny edge detection images are calculated. Seven images comprise of three images obtained by red, green and blue channel of RGB channel each normalized by intensity, four broadly tuned channels of red, green, blue and yellow Itti et al. (1998). Two remaining images are canny edge detection on the original image after mean shift filtering (I_{msf}) and the Saliency Map(S). These nine images are added together to form the final saliency map (S_{final}). Figure 4 shows the saliency map.



(a) **(b)**

Figure 4 . (a) Saliency Map Image. (b) Final Saliency Map Image after adding nine different images.

OBJECT DETECTION

A. Surf Detection

All salient objects are detected by the SURF which works as a blob detector that continuously thresholds the image and extracts those objects that remain stable over a certain specified number of thresholds while removing any false positives. The outputs are a set of crops with potential targets. A contour algorithm outlines the outermost edges. The target from a grayscale target crop is extracted by creating a mask. The mask is created by selecting the longest continuous outer edge present, which corresponds to the shape of the desired target

B. Graph Cut based Object Extraction

To segment the foreground (visually salient objects in our case) from the image, an interactive technique based on seminal work on graph cuts Rother et al. (2004). One of the drawbacks of this approach is the initial user interaction required to label some part of the image as background and foreground. This information is then used by the proposed algorithm to find the alpha-matte of the foreground.

A semi-autonomous approach like this is not suitable for real-time applications involving aerial images. To overcome the problem of user interaction, the information obtained after calculating bounding boxes in step 2 is used. The area inside bounding box is marked as foreground while the area outside it is marked as background. The information given by the bounding box makes the process fully autonomous and requires no human interaction.

Figure 5 (a) represents the cropped image considered for the grab cut operation. Red box in this image marks the area represented as foreground during the operation. Area outside this rectangle is marked as background. Notice how the larger image is cropped to a much smaller region for computational efficiency.

C. Masking

The last step for object extraction is to use the alpha-matte calculated in the previous step as a mask on the original image to extract the colored object from the original image.

Figure 5(b) shows the mask obtained. This image is now used to extract the colored object from the original image. The final object of interest obtained after masking is shown in figure 5 (c).

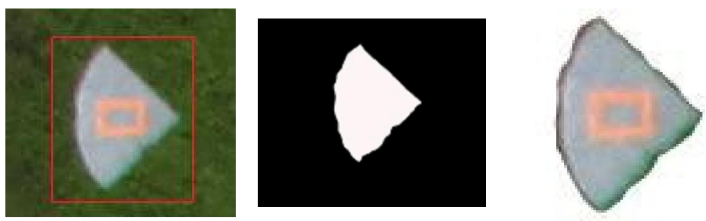


Figure: 5 **(a),** **5(b),** **5(c)**

Shape Classification

The shape classification method uses Fourier analysis on the shape contours to retrieve approximate shape descriptors. These are classified as shapes by a neural network trained on generated data. Once the segmented regions and shape are determined, the average color of each region is classified as a particular color based on its distance to known color values.

CONCLUSION

The proposed algorithm was tested on images captured in various test flights. The system successfully extracted 86% of the meaningful target autonomously out of the total targets. The number of false positives detected by the algorithm was extremely low. Out of all the targets detected approximately 8% of the targets were false positives. The algorithm achieves a high detection rate at the cost of some erroneous detection, but it is better to have a few erroneous detections than to miss the original targets. The future work will involve a feedback system from the positioning system and other sensor information from the UAV to better estimate the location. We also plan to work on estimating the GPS coordinates of the targets based on the location data so that we can plan a revisit to the object to capture a zoomed in high resolution photograph of the target again. We are also looking at the possibility of incorporating a supervised learning approach to reduce the number of false positives and increase the accuracy rate further.

REFERENCES

Itti, Laurent, Koch Christof, and Ernst Niebur (1998). A Model of Saliency- Based Visual Attention for Rapid Scene Analysis. Pattern Analysis and Machine Intelligence, *IEEE Transactions on 20.11: 1254-1259.*

Sokalski, Jan, Breckon Toby P., and Ian Cowling (2010). Automatic Salient Object Detection in Uav Imagery, *Proc. 25th International Unmanned Air Vehicle Systems: 11-1.*

Achanta, Radhakrishna, et al. (2009). Frequency-Tuned Salient Region Detection. Computer Vision and Pattern Recognition, 2009. *CVPR*. IEEE Conference.

Comaniciu, Dorin and Peter Meer (1999). Mean Shift Analysis and Applications. Computer Vision. *The Proceedings of the Seventh IEEE International Conference On. Vol. 2. IEEE*

Rother, Carsten, Kolmogorov Vladimir, and Blake Andrew (2004). Grabcut: Interactive Foreground Extraction Using Iterated Graph Cuts. *ACM Transactions on Graphics (Tog). Vol. 23. No. 3. ACM.*