Aerial View Rock Obstruction Warning System (AVROWS)

Patrick V. Atienza[1], Hans L. Lim[1], Patrick T. Lo[1], Jan M. Mercadejas[1], and Macario O. Cordel[1]
[1] Center for Automation Research, College of Computer Studies, DLSU – Manila
patrick_atienza@dlsu.ph, hans_lim@dlsu.ph, patrick_t_lo@dlsu.ph, jan_mercadejas@dlsu.ph

Abstract: Being able to differentiate and quantify different objects from their surroundings is one important application of image processing that has great uses in multiple fields. One such example is in the field of helicopter automation, where distinguishing potential obstructions from their surroundings is instrumental in autonomous systems. This is because helicopters, as with all forms of aircraft, are very demanding when it comes to their landing surface, where even a few slight obstructions could prevent proper landing and even cause accidents. Moreover, because traditional sensors cannot fully analyze the environment in this manner, the use of image processing is required. This paper presents a similar concept explored by designing a system that is capable of identifying rocks on the grass from a video feed obtained through a simulated aerial view. One of the problem with processing video feed from a helicopter is that the video tends to be shaky resulting in blurred video frames. Although deblurring algorithm can be used to remove the blur in each video frame, we were able to see however that blurring the image aids in minimizing noises obtained from grass edges because the blurred image smoothens out the grass edges, while still keeping intact the rock edges, thus improving the accuracy of rock segmentation. Although this kind of object detection and segregation can prove to be quite a challenge, especially due to the noise brought to about by background. This is because if there is not sufficient contrast between the background and the object to be detected, the system will no longer be able to properly distinguish the object. If a rock is identified, a notification will be prompted by the system notifying that landing for that particular area is not safe.

Key Words: Landing; Stone Segmentation; Object Detection; Autonomous Helicopter;

1. INTRODUCTION

1.1 Overview

Object detection, and more specifically stone detection through image processing is typically hard to implement because a computer is severely limited in terms of its ability in perceiving its physical environment. As a result, it is greatly affected by the quality of stimulus it receives. This is especially true in autonomous helicopters because of how crucial obstruction detection is to ensure proper takeoff and landing. Not only does the system have to be able to identify possible obstructions, but also it should be able to analyze whether or not these potential obstructions are large enough to warrant attention.

Several difficulties can be found involving the separation of each discrete stone in an image by using its general shape and texture as parameter input for the system, contrasting it with its surroundings, as well as filtering the different kinds of noise from the image. Each stone has its own unique appearance ranging from color to texture depending on the type, which is why detecting multiple stones in a single image with completely different characteristics is fairly difficult than detecting similar stones.

The best way to segregate the stones from the rest of the image is to compress the n-bit image (usually 24-bit) to a two level image, often black and white image. However, before doing so, some segmentation and morphing methods must be applied unto the image. Since the scene is limited to a number of rocks with a grass background, multiple post-processing problems can be encountered, some examples were: (1) There are parts of the image that there was less grass so the color of that part would become the same color as the color of the stones that the program can accept. (2) Some grass blocks the stones or part of the stones hence the program might count a single stone to two or more stones or might not count the stone at all. (3) Stones are too small to count. All of problems in which can be attributed to the visibility of the stones themselves.

1.2 Review of Related Literature

Normally, Object Segmentation techniques relies on the different coherent image properties for a given shot, such as but not limited to brightness and contrast, color, texture, the depth of field and even the threshold set to be used (Malik, Belongie, Leung and Shin, 2001). There are cases wherein the process of segmenting an object based on an image are usually rendered on a monochrome image, and as such different image properties are being taken into consideration instead, that is similarity and discontinuity (Wang, 2008).

Detection or segmentation of a particular object given a still image can be easily detected using edge detection and basic morphology techniques given that the object to be identified has sufficient contrast against the background. This is because if there is sufficient change in contrast between the background and the object to be detected, one can easily identify and limit possible cases where the object to be tagged is located simply by calculating the gradient given a threshold value on a binary mask (MathWorks, 2013). However, as for the case of rock image segmentation, detecting rocks is much more complex as the characteristics of a rock can be of irregular in nature adding to the three-dimensional structure and incoherent edges and boundaries for a particular piece of rock exhibits. And since rock are usually mixed with other objects when placed on the floor such as grass and weeds, the noise that one can encounter while trying to map out the rock location can be quite overwhelming especially with just the use of simple object detection techniques (Gao and Wang, 1989). Although there were a number of studies conducted in the past on trying to use the power of image processing to identify possible rock fragments, the results were usually relying on multiple sensors, including those that are not image based sensor devices to accomplish the task.

Another rock image segmentation problem is the image itself if it was blurred or vague. The very nature that is the composition of the rock themselves add the level of complexity in terms of rock image segmentation due to the loose definition of a rock such as having different texture, rock albedo, and shape for each different piece of rock. The difference in size itself also adds the level of difficulty in classifying if an object is indeed a rock since anything from a small pebble to a large boulder can be considered as a rock (Dunlop, 2006). Excluding degradation, noise and distortion are the main aspects of the sharpness of the image.

Common methods of identifying classes of rocks focus around the texture of the rock at hand because texture varies with each natural rock type and the texture is slightly more unique for each type than color. In rock texture classification, most recent systems emphasize more on the textural and specular components of rock images to extract data (Kachanubal and Udomhunsakul, 2008). Spatial frequency measurement is usually implemented in mining a rock image's textural features using a statistical method of classification. Color can also be used in segmenting natural rock images in conjunction with texture. By having specular information from the image, one can identify which rock type it is and use feature extraction in a more straightforward manner. Machine learning can also be implemented in order to identify and classify rocks by combining the different features that a typical rock has, such as its color, shape, size and texture (Dunlop, 2006).

A system which tries to identify each object in an image can produce several outcomes in its output. For example, among the objects it identifies can be a false positive or an object which should not have been identified and the rest of the image may have an object which should have been identified or a false negative. These outcomes must be included during the process of quantifying the results to concretely establish how successful the system is. (Sokolova, Japkowicz, and Szpakowicz, 2006). The F-score is a weighted measure which considers these outcomes in a distribution called precision and recall. Precision dictates how many real elements were identified divided by the total number of elements which were identified while recall is the number of real elements identified divided by the total number of real elements identified and real elements that should have been identified(Sokolova, et. al, 2006) (Ye, Chai, Lee, and Chieu, 2012).

2. METHODOLOGY

2.1 Background

Every type of green color shares the same RGB pattern, as seen in Table 1, that it has a very high green channel value compared to its red and blue counterparts. This assumption is very important to determine which is a grass or which is a rock.

Table 1. Variations of green color and RGB values

Color Name	RGB Value	Sample
Dark Green	0-100-0	
Sea Green	46-139-87	
Pale Green	152-251-152	
Spring Green	0-255-127	
Lawn Green	124-252-0	
Green-Yellow	173-255-47	
Yellow-Green	154-205-50	
Forest Green	34-139-34	

Since this study concentrates on detecting multiple rocks of different sizes given a single image, the image samples which are used as the primary inputs for the stone segmentation system are taken from a bird's eye view perspective and contains rocks of varying sizes to show how effective the system is when an image contains a set of rocks with arbitrary sizes. As shown in Figure 1 below, the images are fed into the pre-processing module of the system, which is detailed in Figure 2. The images then sent to succeeding stages of the system successively where a specific image processing technique tries to manipulate the image until all features excluding rocks such as surrounding backgrounds and noises which are not supposed to remain on the image are removed. From there, the post processing will then merge the original image to the filtered image to highlight the particular rocks based from the original unmodified image. Background subtraction, dilation process, and erosion process are part of the whole rock segmentation scheme. Minimal functional methods are used to slightly improve the performance of the system. These functions are removing small pixels in a binary image and filling hollowed objects.

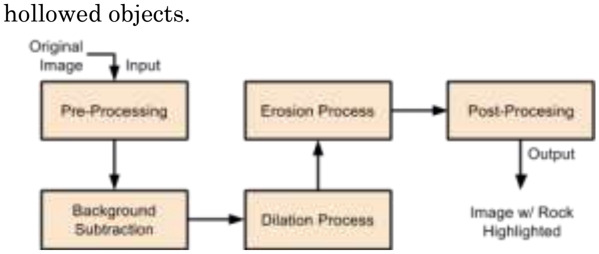

Fig. 1. Block Diagram of AVROWS Process

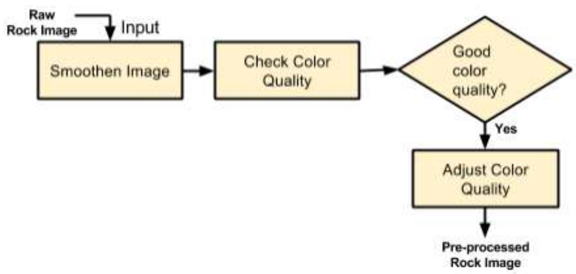

Fig. 2. AVROWS Pre-Processing Block Diagram

The output generates an image containing highlighted rocks by performing pre-processing and filtering on individual frames of the video input, as shown in Figure 1. The filtered rock image is then fed into another series of processes which can be shown in Figure 3. Initially, the filtered rock image is used as an input for the Pixel based detection algorithm. The algorithm identifies how many pixels a particular stone or rock contains which is used as a parameter for approximating the size of that particular sample while grouping similar rocks through different highlight color. Afterwards the output of the Pixel based detection algorithm will undergo differential processing wherein the final output generates an image which displays the highlighted rocks segregated based on their size. As mentioned, the image would then be overlaid on the original image.

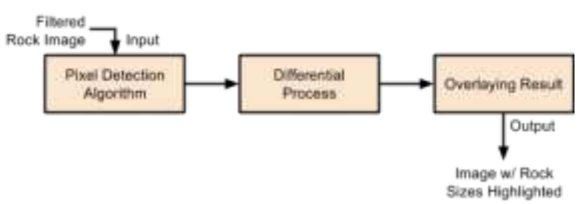

Fig. 3. AVROWS Post-Processing Block Diagram

2.2 Obstruction Detection Algorithm

The system initially performs a conversion that turns the video feed into a string of images to be process by the system. This is performed using the video reader function of MatLab. Some parameters regarding the video's properties are also read and stored into variables for future use such as the number of frames the video has, the height and width of the video frame, and also the frame rate of the actual video.

Initially, the string of images undergoes pre-processing wherein the size of image is scaled down in order to accommodate real-time image processing given the limited processing power. To further meet the speed requirements for real-time image processing, some video frames are dropped proportionally during processing of a given video feed such that the whole video is kept intact while maintaining the speed requirement. The frames were trimmed down to a resolution which has a width of 170px. The height of the frames were based on the ratio of the current width and the original width.

2.2.1 Gaussian Blurring Method

As part of the pre-processing stage, each video frame undergoes Gaussian blurring in order to remove or minimize the edges produce by the grass on the field. This is done because normally, due to the difference in size and texture between rock formations and grass, when one undergoes Gaussian burring in an image where there are rocks place on a grass area the edges of the grass becomes much easier to remove while still keeping the rock edges intact. Putting it simply, the Gaussian blurring of an image tries to prevent the grass and background elements from being incorrectly identified due to fact that it reduces the edges in the grass, which helps prevent it being misinterpreted as a rock formation during the edge detection portion of the algorithm

2.2.2 Color Quality Management

Another part of pre-processing is to determine the image's color quality. This can be done by: (1) getting the red, green and blue values of each pixel, (2) subtract the lowest value from the highest value of each pixel, (3) find the mean of all the subtracted values. The certain frame would be rejected if the results are relatively low, but if it doesn't, then the image would undergo color adjustments to improve the effectiveness of the next process which is image masking. Color adjustment is just increasing the red, green and blue values of any pixel which has a low amount of both three. This method is best for equalizing the dark areas of the image.

2.2.3 Image Masking

After the pre-processing of the video frame, the image would again split into its respective RGB channels and each of them is put into its own mask. After that, each mask undergoes binarization where a specific threshold is set that closely resembles the ideal grass area. This is done to subtract the background from the image, which consists of possible grassy areas, making it much easier to identify and emphasize possible rocks in the vicinity of a particular video frame. The connected objects which do not reach a particular size or pixel threshold are removed because there are a set of miniscule structures which are similar to noise inside the binary image which must be ignored.

Then, a morphological disk-shaped structuring element is superimposed unto the objects in the image and if that object's radius is less than the radius of disk-shaped structuring element then the object is removed. Image dilation, followed by erosion, is then performed on the resulting image in order to help remove any noise elements that might be interpreted as rock formations at an arbitrary threshold. Since the texture of the rock has an effect on the result of a binarization of the image, imperfections represented as holes can be seen within the structures in the image and will be filled.

2.2.4 Post-Processing and Resulting Detection

Post-processing will try to identify rocks given a particular frame through overlaying the identified rock mask with the original video frame. The system also provides information concurrently during the process whether a particular field is indeed mark as safe for landing or not.

Figure 4 shows a sample screenshot of the user interface of the system. The user must open a video file in the program to be processed by the system. The canvas displays the frames of the video file after processing in which all visible stones are covered in a specific color. If the program identifies the vicinity of the image as safe for landing, then "Land" module beside the canvas will turn green.

Fig. 4. AVROWS GUI Program (MATLAB)

After the results are obtained, a quantitative test is then performed that aims to assess the capabilities of the system to properly identify the rock formations. It involves obtaining the F-1 score of the system to measure the system's capability to detect all of the rocks from the video feed as well as the system's resistance to producing false positives.

3. RESULTS AND DISCUSSION

Based from a simulated helicopter landing situation, the two frames or images, in Figure 5, were based on two different cameras and both of them were obtained at the Cory Aquino Democratic Space. The images in Figure 5 were already blurred by the system. The altitude is one the most important aspect in terms of a helicopter landing, thus we note that aspect and observe the rock detection in three types of altitude levels.

These images were gathered at the same spot with different cameras, and these were their results:

Fig. 5. The resulting frame of the videos from two different cameras and their color quality values

The first image to the left had a color quality of 13.8612 compared to the color quality of the other image which was 57.9289. By that means, the left image would be always rejected and the right image would proceed to the next steps.

Fig. 6. The sample rocks that were used to be detect by the system

Six rocks have been classified manually and these were their properties:

Rock A
Diameter: 12.5cm Roughness: Rough
Color: Gray

Rock B
Diameter: 3.8cm Roughness Rough
Color: Brownish gray

Rock C
Diameter: 6.7cm Roughness: Rough
Color: Brownish gray

Rock D
Diameter: 2.6cm Roughness: Rough
Color: Gray

Rock E
Diameter: 6cm Roughness: Smooth
Color: White

Rock F
Diameter: 16cm Roughness: Rough
Color: Light Brownish Gray

The videos were taken at 1500 centimeters, 1000 centimeters and 500 centimeters away from the target which are crucial landing altitudes for an RC helicopter that means all sensor must be more accurate than usual to prevent miscalculations.

The next image would show the system processes for the rock detection algorithm at 1.5 meter above the ground:

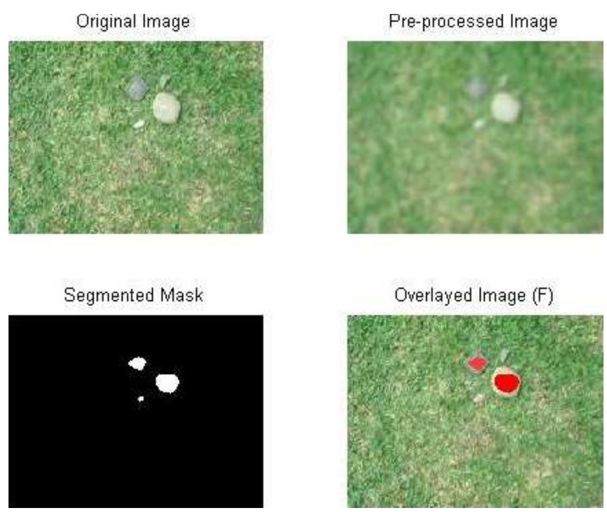

Fig. 7. Simulated landing shots taken at 1.5 meter (segmented and overlayed)

Obtaining the F-1 score based Figure 7, we were able to obtain this results:
Precision: 0.64 Recall: 0.62
F-1 Score: 62.98%

At 1500 centimeters above the ground, the readings were fairly off the mark. The main reason why inaccurate readings was happening because of ignoring small detections to make the system do fewer mistakes for very small objects that has somewhat the same properties as the rocks. As the altitude gets lower and lower, more and more rocks would be recognize and detect by the system.

The next image would show the system processes for the rock detection algorithm at 1 meter above the ground:

Fig. 8. Simulated landing shots taken at 1 meter (segmented and overlayed)

Obtaining the F-1 score based Figure 8, we were able to obtain this results:
Precision: 0.72 Recall: 0.71
F-1 Score: 71.50%

As mentioned before, if the altitude gets lower, then more rocks would notice by the system and that's the result in the second video. At the height of 1000 centimeters, the F-1 score gets higher than first one and the system was became more accurate based on counting the rocks.

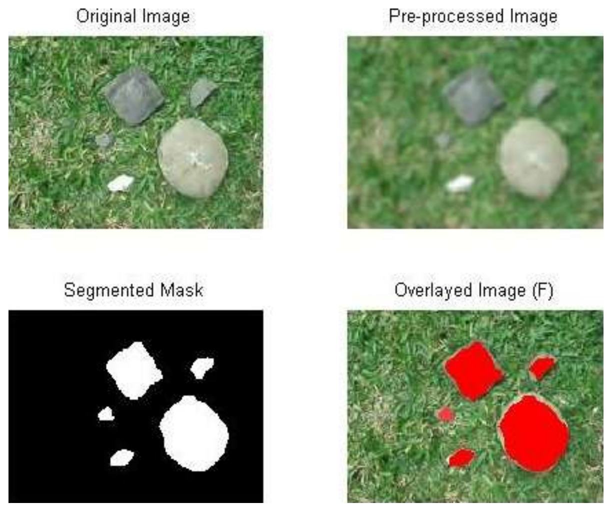

Fig. 9. Simulated landing shots taken at 0.5 meter (segmented and overlayed)

Obtaining the F-1 score based Figure 9, we were able to obtain this results:
Precision: 0.97 Recall: 0.96
F-1 Score: 96.50%

At 500 centimers above the ground level, the readings of the rocks were quite accurate because they were taken very close. For this experiment, it is good to have an accurate result for that distance.

Though it is worth noting that in some special instances, it becomes slightly more difficult to differentiate the grass from smaller rocks. In the first two localized cases, the rocks are too small and are not able to be properly segmented. Nevertheless, the algorithm proved to be more than sufficient especially because rocks of that size normally have a negligible effect on the landing capabilities of a helicraft in that area.

The last case suggests that the algorithm was largely successful in identifying rocks in the given area without significant error. Also, in this particular case it appears that there isn't a significant difference in the sources of error in this test ca se with both the Precision and Recall being roughly equal.

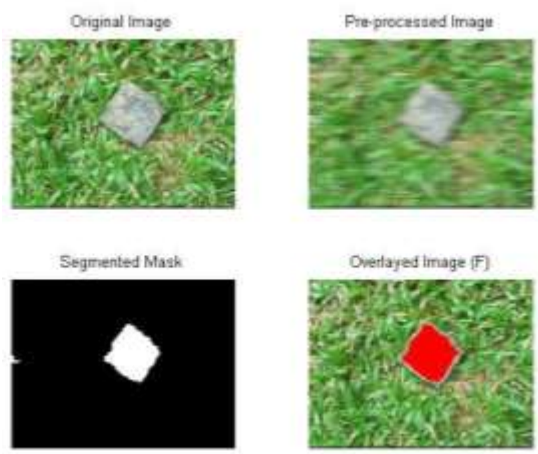

Fig. 10. Simulated Landing shots taken in Cory Aquino Democratic Square

Overall, the results are quite promising, showing considerable accuracy in the detecting and highlighting of potential obstructions. It has shown to be relatively unhindered by Gaussian blur caused by vibrations and oscillations present in the video feed, which is of great importance if this system is to be installed onto a helicopter which is prone to heavy vibrations.

However, one potential issue is that it is still limited by its lack of depth perception, which limits it to differentiating objects based on their contrast with the background image. Though for the current test environment it is sufficient, when operating on rocky terrain for example the system may no longer work as effectively due to the similarities between the potential obstructions and the background.

The problem of determining the color quality may arise in this system. We have three assumptions if the image failed the color quality test: (1) the camera used is not that colourful enough, (2) too many rocks in the ground, (3) the altitude is very low. Since it's best to use a colored camera and the detection of rocks would end at least 1000 centimeters off the ground, the main concern of this system is the multiple rocks.

Another key issue discovered during the testing phase is the amount of time it takes to process the images. It results in a 0.2 to 0.5 second delay before the current frame is fully processed. Originally, this time was even longer however due to a few optimizations made to the algorithm, as well as downscaling the video feed, the processing time was kept manageable.

Despite that however, this means that there are some limitations to the overall efficacy of the system due to the slight time delay present. This processing weakness is compounded further when the system is mounted onto the actual helicopter, which will undoubtedly have a considerably weaker processor to work with.

4. CONCLUSIONS

Overall, the algorithm we were able to develop is still in its alpha stage, although based on the test we were able to properly segment and determine if there are potential hazard obstructing for safe landing, which the latter being the primary goal of the system. However, there are still some bugs floating around from time to time, such as not being able to properly tag the specific rock given a single video frame properly. But then, based on the test, the main goal of the overall project was actually met and that is determining if the landing space is viable enough for landing of an RC helicopter or not.

As for future work, there is still a lot of room for improvements and one of those is simply removing the reliance of the current system on contrast between the background and objects. One way that this can be performed is by using an additional camera to capture the same image from a different perspective, which can be used to map the objects against their background in 3 dimensions. This would not only improve the detection capabilities of the system, but it would also provide another criterion for object detection due to its augmented ability to perceive objects in 3 dimensions.

Another key improvement that should be considered is the optimization of the algorithm to improve the system's performance. This is essential because if the system were to be deployed onto an actual helicopter the processing capabilities of the onboard computer might be considerably weaker than the current test system. And since the system should be able to perform the processing in real time, great care has to be put in the optimization to ensure that the system can still perform in such conditions.

5. REFERENCES

Dunlop, H. (2006). Automatic Rock Detection and Classification in Natural Scenes. Robotics Institute, Carnegie Mellon University. Retrieved November 2013, from http://dunlop1.net/doc/thesis.pdf

Eddins, S. (2002). The Watershed Transform: Strategies for Image Segmentation. Retrieved October 2013, from http://www.mathworks.com/company/newsletters/articles/the-watershed-transform-strategies-for-image-segmentation.html.

Gao, Q. G., & Wang, A. K. (1989). Rock Image Segmentation. Vision Interface '89, 125-133.

Kachanubal, T., & Udomhunsakul, S. (2008). Rock Textures Classification Based on Textural and Spectral Features. International Journal of Information and Mathematical Sciences, 4:4, 240-246.

Malik, J., Belongie, S., Leung, T., & Shi, J. (2001). Contour and texture analysis for image segmentation. International Jwenal of Computer Vision, 43(1), 7-27.

MathWorks (2013). Detecting a Cell Using Image Segmentation. Retrieved November 2013, from http://www.mathworks.com/help/images/examples/detecting-a-cell-using-image-segmentation.html

Wang, W. X. (2008). Rock Particle Image Segmentation and Systems. Pattern Recognition Techniques,Technology and Applications, Peng-Yeng Yin, ISBN: 978-953-7619-24-4, InTech.

Sokolova, M., Japkowicz, N., & Szpakowicz, S. (2006). Beyond Accuracy, F-Score and ROC: a Family of Discriminant Measures for Performance Evaluation. American Association of Artificial Intelligence. Retrieved December 2013, from http://rali.iro.umontreal.ca/rali/sites/default/files/publis/EvAAAI06finMay10.pdf

Ye, N., Chai. K. M., Lee, W. S., & Chieu, H. L. (2012). Optimizing F-Measures: A Tale of Two Approaches. 29th International Conference on Machine Learning. Retrieved December 2013, from http://icml.cc/2012/papers/175.pdf