ENHANCEMENT OF UNUSUAL COLOR IN AERIAL VIDEO SEQUENCES FOR ASSISTING WILDERNESS SEARCH AND RESCUE

Nathan D. Rasmussen, Daniel R. Thornton, Bryan S. Morse

Brigham Young University
Department of Computer Science
Provo, UT 84602

ABSTRACT

The use of aerial video for search and surveillance has been popularized by the increased use of camera-equipped unmanned aerial vehicles. For many search applications, objects may also be missed by observers due to their small size, brief visibility, or the inherent monotony of the scene. This paper presents a novel method for automatically emphasizing unusually colored objects to improve their detectability. We use a hue histogram and a local saliency measure to find unusually colored objects, then boost the saturation of these objects while desaturating more common colors, thus drawing the observer's attention and facilitating video search.

Index Terms— Display human factors, video signal processing, surveillance, image color analysis, image enhancement

1. INTRODUCTION AND BACKGROUND

In recent years, aerial video has seen increased use as a tool for search, surveillance, and reconnaissance, in large part due to increased deployment of unmanned aerial vehicles (UAVs). These UAVs have many advantages over conventional aircraft: they can be launched in rugged terrain, they cost far less to operate, and they do not put human lives in danger when used. These advantages make the use of camera-equipped UAVs feasible in ways that may not have been practical previously. One such application is in Wilderness Search and Rescue (WiSAR), where low-cost UAVs can replace the use of conventional aircraft [1]. The increased availability of UAV-based aerial video can greatly augment the information obtained by searchers, reducing the time spent searching for lost individuals.

In search situations, the timeline is critical. Besides obvious concerns about the health and safety of the victim, the victim's own mobility often works against the search. For every hour that passes, the search radius must typically increase by approximately 3 km, and the probability of finding and successfully aiding the victim decreases. This creates a fundamental tradeoff for acquiring aerial video for search and rescue: flying higher and using wide-angle lenses increases the field of view and hence the search area coverage, but it also obviously reduces the resolution of the video. Flying lower or using zoom lenses increases the resolution of the video, providing greater detail and increasing the size of objects of interest, but it also reduces the area seen and hence the search coverage. These, as well as increased flight speed, can also reduce the duration of visibility for objects of interest in the video, making them less likely to be seen by human observers.
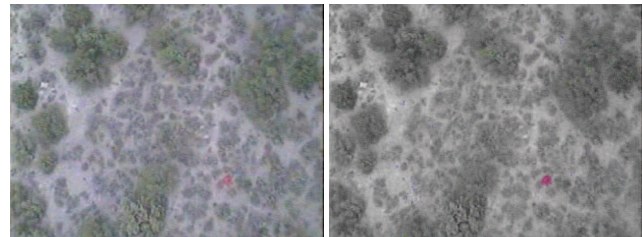
(a) Original video frame   (b) Enhanced frame

Fig. 1. Example of an unusual-color enhanced video frame. This frame includes a red object in the lower right corner against a largely brown and green background. The object stands out against its background in the enhanced frame (b) much more than in the original (a) due to the processing done to emphasize its unusual color and deemphasizing the more common surrounding colors.[1]

Another problem that exists for video observers in WiSAR tasks (and in many similar applications) comes from the nature of the scene itself. The content of the video is constantly changing and must be attended to, but most of the information is not useful or interesting, making it hard for the searcher to focus their attention for long periods of time. When items of interest are visible in the video, they are often not seen because they are small and easily overlooked due to the monotony of the task.

One way to approach this problem is to automatically detect objects of interest in the video. Caron et al. [2] classify man-made objects in images using Zipf's law. However, they require that the object be at least 2.5% of the image to be detected. In the aerial search video we are using, objects are much smaller than this. Merino et al. [3] use ratios of the RGB components to segment fire from color video in automatic detection. This technique, as well as other target-detection methods, will not work in wilderness search and rescue tasks, because searchers are looking for anything that can help them locate the lost individual and may not know ahead of time what those items will look like.

Another approach is to enhance the display of the video so as to make objects of interest easier for human observers to detect. Cheung and Milgram [4] use hyperstereoscopic video to increase the perceived depth between objects in a scene. This would only help if the differences in height between items of interest and their surroundings were significant. Gibbins et al. [5] use super-resolution to provide the viewer with enhanced views of tracked objects. Their

---

[1]Because of the subtle use of color in this paper, please refer to an electronic copy of this document when examining the images presented here.

system creates a super-resolved image after an object of interest is found. For live search we would need to have constantly super-resolved video, which adds too much computation overhead to be a viable solution. Kumar et al. [6] separate the video into its background and moving objects to show motion to the user. This method works well for helping users detect moving objects, but if objects of interest are stationary or are moving much slower than the UAV, this will do nothing to help our searchers.

We propose using automatic accentuation of unusual color to help searchers more easily detect items of interest in video. This method uses a combination of selective saturation boosting and de-saturation [7] to enhance real-time video, emphasizing relatively less common colors while de-emphasizing the colors that occur more frequently in the scene. This simplifies the search task since the colors can attract the searcher's attention. It also helps break the monotony of watching the video, as the searcher can look from object to object rather than trying to constantly search over the whole area.

Because color attracts our preattentive vision [8], it can be a powerful way to draw attention to objects. This use of saturated color to accentuate an object against a colorless background is a common artistic technique. Methods for doing this digitally are commonplace (e.g., Photoshop) but require manual segmentation of the object of interest or specification of relevant object color.

Object size also affects our perception of color—due to lower color acuity in our visual systems, small objects appear achromatic, while larger objects retain their coloration. Boosting the relative saturation of small objects of interest helps them retain their apparent chromaticity and appear more well-defined.

Figure 1 shows an example of our proposed method. The small red object in Figure 1a is enhanced by boosting its saturation while simultaneously reducing the saturation of the more commonly occurring brown and green colors as shown in Figure 1b. The enhanced image appears to have almost no color other than the red dot in the bottom right corner, thus drawing the users attention to this item so it can be easily spotted in the frame. The system does not know a priori the colors of the normal background content or of the objects of interests—it automatically analyzes the video frames themselves to determine what is common and what is unusual.

2. METHODS

Our method focuses primarily on the hues of objects, not on variations in their intensity or saturation, in order to avoid sensitivity to factors such as variations in object intensity or lighting. We thus first convert each frame into a color space where these chromaticity attributes are explicit. For this we use the HSV color space, though similar color spaces should work as well. We denote the hue, saturation and value channels of the image as $H(x,y)$, $S(x,y)$ and $V(x,y)$ respectively.

After moving the image into the HSV space, we need to classify the colors, or hues, as either common or uncommon. We compile the occurrences of each hue H into a histogram $h(H)$ using $8°$ partitioning of the hues so as to combine similar colors. We then define the set of unusual hues U to be those that occur with frequency less than some specified threshold T_h:

$$U = \{H : h(H) < T_h\} \qquad (1)$$

The examples presented here use a threshold T_h of 1% of the size of the video frame, though this may be adjusted by the user based on the relative size of objects of interest in the video.

In order to avoid enhancing unusually colored but less salient features such as color noise, we also compute a saliency measure based on a center-surround difference model [9, 10, 11, 12]:

$$D(x,y) = \frac{\sum_{i=1}^{-1} \sum_{j=1}^{-1} |H(x,y) - H(x+i, y+j)|}{8} \qquad (2)$$

and treat as salient only those pixels for which $D(x,y) \leq T_d$. (We use $T_d = 1$, equating to an average neighborhood absolute difference of one $8°$ hue bin.) Thus, we consider potential objects of interest to be important both if their coloring is unusual and their saliency is high. Indeed, some authors have proposed the use of color rarity as a component of saliency [12].

Next, we enhance objects with unusual hues by adjusting the saturation of the respective pixels. We boost the saturation for the pixels with uncommon hues by a preset amount b and desaturate the pixels with common hues by a similar amount d. The new saturation $S'(x,y)$ for each pixel is thus defined as follows:

$$S'(x,y) = \begin{cases} \min(S(x,y) + b, 1.0) & \text{if } H(x,y) \in U \\ & \text{and } D(x,y) \leq T_d \\ \max(S(x,y) - d, 0.0) & \text{if } H(x,y) \notin U \\ S(x,y) & \text{otherwise} \end{cases} \qquad (3)$$

By suppressing the more common hues, we reduce the amount of color information in each frame, leaving the color information only in the unusual hues. In our implementation both b and d are user-adjustable, so that the user can determine the amount of change given the conditions in the video.

Use of Eq. 3 can sometimes create unwanted coloration of pixels that are basically achromatic (grey) but due to noise may have a slight amount of color. To avoid saturation boosting of such near-grey values, we modify Eq. 3 to boost the saturation of pixels only when they are already above a minimum level of saturation T_s:

$$S'(x,y) = \begin{cases} \min(S(x,y) + b, 1.0) & \text{if } H(x,y) \in U \\ & \text{and } D(x,y) \leq T_d \\ & \text{and } S(x,y) > T_s \\ \max(S(x,y) - d, 0.0) & \text{if } H(x,y) \notin U \\ S(x,y) & \text{otherwise} \end{cases} \qquad (4)$$

In our implementation we have used 15% as the saturation threshold. This reduces the problem of falsely coloring achromatic regions, which eases the work of the searcher.

An example of this process is illustrated in Figure 2. To enhance the image shown in Figure 1a, we first calculate its hue histogram and identify the common and unusual colors (Figure 2a) relative to our threshold. We then boost the saturation for all pixels with unusual hues while decreasing the saturation those with common hues according to Eq. 4. As can be seen in Figures 2b–c, the average saturation for common hues decreases to the point of being achromatic. (No saturation boosting of unusual colors was used in this example.) The final resulting image can be seen in Figure 1b.

Using constant boosting and desaturation values is not the only method for emphasizing the less common colors. We have also used the hue histogram to determine the respective amounts of saturation adjustment. This can be done by inverting the hue histogram and then using this as a lookup for the saturation of a color, causing colors that are more common to have lower saturations and colors that are less common to have higher saturations. This likewise emphasizes the unusual colors for the user and avoids the need for a threshold, but we have found it introduces significant overhead when used in real-time applications.

To speed up the conversions to and from the HSV color space, lookup tables are built for the transformations. We have discretized each color channel in the lookup table down to 5 bits (or 32 levels).
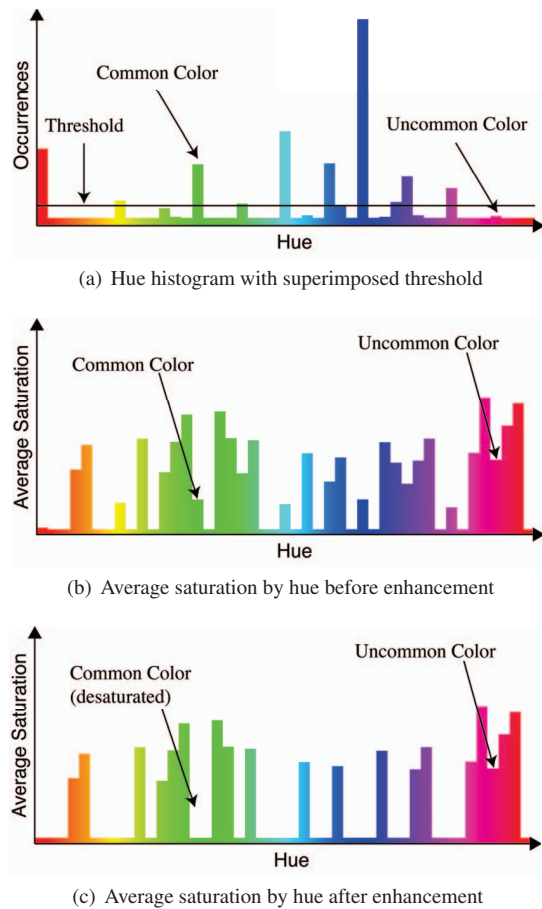
(a) Hue histogram with superimposed threshold

(b) Average saturation by hue before enhancement

(c) Average saturation by hue after enhancement

Fig. 2. Hue histogram and average saturation by hue for the images from Figure 1. The hue histogram and a predetermined threshold are used to classify hues as common or uncommon (a). For each hue that is identified as a common one, the saturation of all pixels of that hue (b) are decreased to the point of being achromatic (c). [Each of these histograms has a color bar along the bottom to show which colors are missing.]

As previously discussed, the hues are segmented into $8°$ partitions, which gives us 45 different hue groups to work with. This also allows us to use bit arithmetic when calculating the discretized values and hue groups. The resulting size of these tables allows them to fit into cache memory. While this discretization causes information loss in the image, this information is not necessary. The task being performed is centered on finding items of interest rather than exact color replication. Lookup tables can be taken one step further to convert the original RGB value into an enhanced RGB value that takes into account either the saturation or desaturation adjustments. This reduces the computation to a single lookup rather than multiple lookups when applying the enhancement to a frame.

Our implementation also takes advantage of being in the HSV space to provide basic user-adjustable brightness and contrast enhancements of the video frames. Due to the use of the lookup tables, this can be done without taking any additional CPU cycles when converting the frame. The contrast enhancement is built into the saturate and desaturate lookup tables, only taking CPU cycles during the pre-computation of these tables.
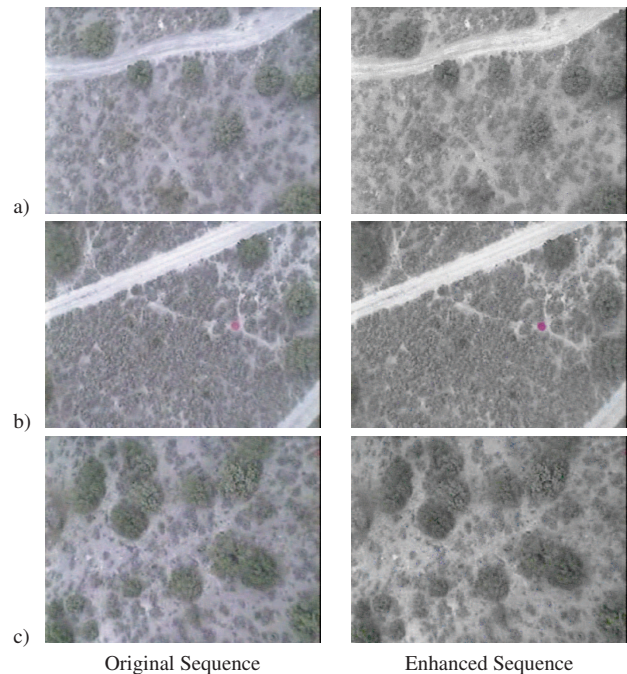

Original Sequence   Enhanced Sequence

Fig. 3. Frames from an unenhanced video sequence (left) and the corresponding frames of the enhanced video (right). Without unusually colored objects (a), the enhanced video appears grey. As a red object of interest appears (b), it appears distinctly in the enhanced video against the colorless background. After the object leaves the field of view (c), the enhanced video again returns to grey. The enhanced views are easier to search through for items of interest than their original counterparts, especially when viewed as a continually moving video sequence.

3. RESULTS AND PERFORMANCE

Figures 3 and 4 show the results of applying the proposed method to video sequences containing items of interest. These videos are 640×480 and have been compressed for storage. The enhancements were run on a 3.2 GHz Pentium 4 computer. For the enhancement, no other brightness or contrast adjustments were made. The boosting and de-boosting amounts were each 20%, though in most video sequences noticeable improvement with as little as 5%.

In Figure 3, the enhanced frames without unusually colored objects appear gray (frames a and c). When the unusually colored red object appears, its retained coloration makes it stand out sharply against the colorless background (frame b). The temporal change from colorless frames to an isolated colored objects also attracts the user's attention.

Figure 4 shows an example from another video sequence. The object of interest (an orange jacket) appears as a small area to the right of the road seen in the image. The only other color in these enhanced images comes from grass that is near the location of the jacket. This grass, just like the jacket, has a different color from its surroundings and is thus classified as an unusual color and enhanced.

The use of lookup tables allows the enhancements to be done in software at 60 full frames per second, double the speed of our NTSC video source. At this rate, the CPU is not at full load when running these methods.
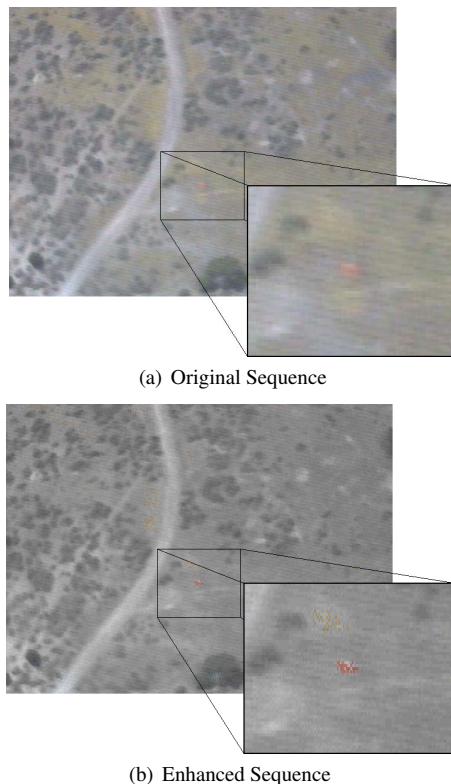
(a) Original Sequence

(b) Enhanced Sequence

Fig. 4. Original and corresponding enhanced frames from search video. This frame of the original video (a) includes an orange jacket, which is automatically enhanced in (b). Some portions of the background are also enhanced, but this does not detract from a searcher's ability to see the jacket more clearly.

4. CONCLUSION

We have presented a method for enhancing the detectability of unusually colored objects in video sequences for wilderness search and rescue and similar applications. The method does not require prior knowledge of the scene or user interaction and relies on the color content of the video itself in order to identify unusual and common colors. For video sequences with relatively monotonous background coloration, as is the case in wilderness settings, this method is effective in assisting searchers in finding objects of interest that would otherwise be less visible in the scene. Since color helps attract attention to the objects, they are not only easier to detect in individual frames but are also attract attention to combat user inattention when viewing long sequences. In addition to search and rescue, this method could easily be generalized to many different search tasks required in other surveillance and reconnaissance situations.

Obviously, this approach is not applicable in all situations. As expected, it would not be effective when items have the same color as their background, though these items could potentially be found using other cues such as texture or shape. It also will not enhance items that are large in the field of view, but these items should already be easily seen.

Other extensions to this approach could also be pursued. Using a window of frames rather than just a single frame could improve the classification of colors. as could the inclusion of other local features.

5. REFERENCES

[1] Michael A. Goodrich, Bryan S. Morse, Damon Gerhardt, Joseph L. Cooper, Julie Adams, Curtis Humphrey, and Morgan Quigley, "Supporting wilderness search and rescue using a camera-equipped mini UAV," *Journal of Field Robotics*, 2007.

[2] Y. Caron, P. Makris, and N. Vincent, "A method for detecting artificial objects in natural environments," in *ICPR '02: Proceedings of the 16th International Conference on Pattern Recognition (ICPR'02) Volume 1*, Washington, DC, USA, 2002, p. 10600, IEEE Computer Society.

[3] L. Merino, F. Caballero, J.R. Martinez de Dios, and A. Ollero, "Cooperative fire detection using unmanned aerial vehicles," in *Proceedings of the IEEE International Conference on Robotics and Automation*, 2005, pp. 1884–1889.

[4] K. M. Cheung and P. Milgram, "Visual detection with hyperstereoscopic video for aerial search and rescue," in *Proceedings of the IEA/HFES 2000 Congress*, 2000, pp. 472–475.

[5] D. Gibbins, P Roberts, and L. Swierkowski, "A video geolocation and image enhancement tool for small unmanned air vehicles (UAVs)," in *Intelligent Sensors, Sensor Networks and Information Processing Conference*, 2004, pp. 469– 473.

[6] R. Kumar, H. Sawhney, S. Samarasekera, S. Hsu, H. Tao, Y. Guo, K. Hanna, A. Pope, R. Wildes, D. Hirvonen, M. Hansen, and P. Burt, "Aerial video surveillance and exploitation," in *Proceedings of the IEEE*, Oct 2001, vol. 89, pp. 1518–1539.

[7] L. Lucchese, S. K. Mitra, and J. Mukherjee, "A new algorithm based on saturation and desaturation in the xy chromaticity diagram for enhancement and re-rendition of color images," in *IEEE International Conference on Image Processing (ICIP)*, 2001, vol. 2, pp. 1077–1080.

[8] Haim Levkowitz, *Color Theory and Modeling for Computer Graphics, Visualization, and Multimedia Applications*, Kluwer Academic Publishers, Norwell, MA, USA, 1997.

[9] B. A. Thomas, R. N. Strickland, and J. J. Rodriguez, "Color image enhancement using spatially adaptive saturation feedback," in *IEEE International Conference on Image Processing (ICIP)*, 1997, vol. 3, pp. 30–33.

[10] Laurent Itti, Christof Koch, and Ernst Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254–1259, November 1998.

[11] Tie Liu, Jian Sun, Nan-Ning Zheng, Xiaoou Tang, and Heung-Yeung Shum, "Learning to detect a salient object," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2007, pp. 1–8.

[12] Gunhee Kim, Daniel Huber, and Martial Hebert, "Segmentation of salient regions in outdoor scenes using imagery and 3-D data," in *WACV '08: Proceedings of the Ninth IEEE Workshop on Applications of Computer Vision*. January 2008, IEEE Computer Society.