



Faculty Publications

2010-03-01

UAV Video Coverage Quality Maps and Prioritized Indexing for Wilderness Search and Rescue

Cameron Engh
cameron_engh@hotmail.com

Michael A. Goodrich
mike@cs.byu.edu

Bryan S. Morse
morse@byu.edu

Follow this and additional works at: <https://scholarsarchive.byu.edu/facpub>



Part of the [Computer Sciences Commons](#)

Original Publication Citation

B. Morse, C. Engh, and M. Goodrich, "Aerial video coverage quality maps and prioritized indexing," in Proceedings of the 5th ACM/IEEE RAS International Conference on Human-Robot Interaction, March 21.

BYU ScholarsArchive Citation

Engh, Cameron; Goodrich, Michael A.; and Morse, Bryan S., "UAV Video Coverage Quality Maps and Prioritized Indexing for Wilderness Search and Rescue" (2010). *Faculty Publications*. 107.
<https://scholarsarchive.byu.edu/facpub/107>

This Peer-Reviewed Article is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Faculty Publications by an authorized administrator of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.

UAV Video Coverage Quality Maps and Prioritized Indexing for Wilderness Search and Rescue

Bryan S. Morse, Cameron H. Engh, and Michael A. Goodrich

Department of Computer Science

Brigham Young University

Provo, Utah, United States

Email: morse@byu.edu

Abstract—Video-equipped mini unmanned aerial vehicles (mini-UAVs) are becoming increasingly popular for surveillance, remote sensing, law enforcement, and search and rescue operations, all of which rely on thorough coverage of a target observation area. However, coverage is not simply a matter of seeing the area (visibility) but of seeing it well enough to allow detection of targets of interest, a quality we here call “see-ability”. Video flashlights, mosaics, or other geospatial compositions of the video may help place the video in context and convey that an area was observed, but not necessarily how well or how often. This paper presents a method for using UAV-acquired video georegistered to terrain and aerial reference imagery to create geospatial video coverage quality maps and indices that indicate relative video quality based on detection factors such as image resolution, number of observations, and variety of viewing angles. When used for offline post-analysis of the video, or for online review, these maps also enable geospatial quality-filtered or prioritized non-sequential access to the video. We present examples of static and dynamic see-ability coverage maps in wilderness search-and-rescue scenarios, along with examples of prioritized non-sequential video access. We also present the results of a user study demonstrating the correlation between see-ability computation and human detection performance.

Keywords-unmanned aerial vehicles, wilderness search and rescue, coverage quality maps, video indexing

I. INTRODUCTION

Small lightweight mini-UAVs with 5–8 foot wingspans have seen increased use recently for aerial sensing due to their lower cost and ease of deployment. When equipped with a video camera and transmitter, these mini-UAVs can be used for surveillance, remote sensing, law enforcement, and search and rescue operations, all of which require rapid and thorough coverage of a target area. However, because of their lightweight nature, these aerial sensing platforms are highly unstable and easily buffeted by wind, and the operator’s intentions may not always correspond to the actual flight path. This makes it difficult for operators or video analysts to correctly determine what spatial areas were observed during a flight or sequence of multiple flights.

In addition to covering the target area, it is also essential to maintain sufficient resolution to allow human operators to accomplish their task. Since the altitude and orientation of the plane are highly variable due to wind or other factors, so too is the resolution of the resulting video. As the plane

banks to one side or the other, even an otherwise downward-pointing camera may end up seeing areas far away and at an oblique angle. This is compounded in varying terrain since the UAV’s height above ground may change rapidly even while maintaining constant altitude. One can try to maintain a consistent height above ground either manually or through automated means, but this is still subject to the limitations of the plane’s ability to climb or safely descend. Some flight paths, especially in difficult terrain, may make a low-altitude pass over the target area then maneuver to make another pass, providing only periodically usable video.

Our work in this area has focused on using mini-UAVs to assist in Wilderness Search and Rescue (WiSAR) operations [1]. Field trials [2] tell us that it is often difficult to tell what areas have been searched well. This assessment is an essential component of search-and-rescue applications because it is basically a prioritized search, focusing on the regions most likely to include the missing person. Also important to this task is the ability to efficiently review previously acquired video, perhaps in response to a search observation or during post hoc offline review. This can be made more efficient by providing users with the ability to intelligently access search video not only by georeferenced indexing but by coverage quality as well, allowing users to directly access usable observations of a specified target area.

Assessing the usability and coverage of aerial video is a matter not only of whether the plane’s camera *could* see a point but *how well* it saw it. Once the video is georegistered to the underlying terrain, determining whether the camera saw specific points is a simple matter of viewing geometry, what we typically think of as “visibility”. But visibility-based coverage alone isn’t enough to determine how useful the video is—one must consider the viewing resolution as well as the number of times seen, the variation of viewing angle (which can often play a role in detection), etc. We call this latter quality “see-ability”.

This paper presents a method for creating coverage quality maps based on see-ability that convey not only the video coverage of each part of a target area but also how useful that video information is for the person viewing it (Figure 1). Such coverage maps are useful for post hoc evaluation of the search, for planning either during or between flights, and for coordination with other team members.



Figure 1. Coverage quality map overlaid in red on terrain and reference imagery. The UAV was launched from the uphill side of the terrain (orange marker on the left) and followed the flight path indicated in green.

By linking coverage maps to source video, these maps may also be used to filter or prioritize offline analysis or online review of the video. Observers can use the coverage map to specify spatial areas to show and/or exclude. They can also choose to automatically skip or fast-forward through portions of the video that are of insufficient quality, filtering the presentation and making time-critical processes more efficient. One can also compute coverage maps for only those portions of the video already seen by a user or set of users, i.e. not what or how well the plane saw certain areas but how well the user has already seen them. This can further assist a user or automated presentation system in prioritizing the presentation of the video in a non-sequential fashion by avoiding excessive repetition of areas the user has already observed well.

The system we present combines video and (approximate) telemetry from the mini-UAV with terrain models and previously acquired reference imagery of the area to georegister the video. It then uses this information to compute our see-ability quality metric for every point in the target area based on the quality of each viewing and the cumulative effect of multiple viewings. We then demonstrate how the video playback can be augmented by providing coverage-map linking, filtering, and prioritization. We also present the results of a user study demonstrating the correlation between our computed see-ability metric and user detection in the context of this task.

II. RELATED WORK

When conducting remote operations with camera-equipped robots, both ground or aerial (UAVs), it has been observed that remote operators often have difficulty knowing where the robot/camera is or where it is looking due to the limited sensory information coming back from the camera alone. This has been likened to looking at the world through a “soda straw” [3] and is an example the keyhole effect

in which humans have only a limited view of a system or problem [4], [5]. This keyhole effect has been shown to cause gaps in the space explored by such robots [6]. The user’s situational awareness can be improved by augmenting the display of the camera’s information with other available external information in the form of ecological interfaces [7], [8] though even this often still leaves the operator or video searcher uncertain of the area coverage once completed. Building a representation of the covered area is related to the mapping problem in robotics [9], though there the aim is to create a spatial model of an explored physical environment, not to represent the quality of coverage through an already (partially) known search environment. Other researchers have noted the role that sensor resolution plays in the effective coverage area for robots [10].

A key requirement for computing geospatial see-ability-based coverage maps is to create correspondence between pixels in the UAV-acquired video and points on the terrain, a process known as *georegistration*. High-precision telemetry data makes this process much simpler but still may not provide sufficient accuracy due to the error-amplifying effect of reverse projection. Mini-UAVs usually have sensors with extremely limited precision due to weight, power, or cost limitations. A common approach is to use telemetry data (if available) to provide an initial estimate of the camera’s pose, then refine this estimate using either semi-automated or automatic visual alignment to reference imagery [11]–[18]. Once the pose estimate is refined, the video may be projected back onto the terrain, replacing or being overlaid onto the reference imagery there. While this “video flashlight” [19] is highly effective for displaying video content in spatial context, it does not necessarily give the video analyst a cumulative sense of the quality of the spatial coverage. These projected frames may be composited onto the terrain in the fashion of a mosaic, but this alone likewise does not easily convey video-quality information. One could manually survey the video to visually inspect the quality, but this can take time and effort that automatic computation of see-ability seeks to avoid.

As discussed in the introduction and described further in Section V, coverage maps can be used to filter and prioritize nonsequential presentation and analysis of the acquired video. Mosaics have been similarly used to provide both spatial summarization and indexing of video [20]. By linking spatial position in the mosaic to video frames, one can use the resulting large-scale mosaic to allow indexing of the video and presentation by geospatial content rather than time. But while mosaics can summarize a sequence and assist a user in spatially identifying points of interest, they don’t directly represent the quality of the information. The proposed coverage quality maps can be used not only to provide interactive geospatial access but to help determine which portions of the video are 1) of sufficient quality to view and 2) have not already been seen well.

III. VIDEO GEOREGISTRATION

Prior to a flight or sequence of flights, we obtain terrain models and aerial imagery of the target area. Should such data not be available before the aerial search, it is gathered prior to offline (post-hoc) analysis. For terrain data, we use the National Elevation Data (NED) from the Seamless Data Distribution System (SDDS) [21]. For imagery, we use data from Google Maps, itself a composite of multiple sources and resolutions provided by third parties. These web-based APIs for terrain and image data allow on-demand online access to the desired data. The choice of terrain model and reference imagery are subject to preference or availability and is not specific to the see-ability methods proposed here.

To render the terrain, we triangulate the NED points and project the reference imagery onto the model. Although many techniques have been developed for interactive rendering of terrain models [22], [23] we have found that even simple structures and rendering are sufficient given the relatively modest search areas coverable by mini-UAVs.

Using methods similar to those in [18], we use terrain models and reference imagery to refine the pose estimate of the UAV’s camera as illustrated in Figure 2. We use telemetry from the plane as an initial estimate of the pose and render the reference-image-textured terrain model as seen from that pose. Comparison of these images allows us to refine our estimate of the pose, after which the process is repeated. This process is made more difficult, though, because of the differences between the reference imagery and the UAV-acquired video due to sensor differences, lighting, changes of season, etc., which we handle in a manner similar to that described in [18].

Once each frame of the UAV-acquired video is georegistered, it can be projected back onto the terrain in a fashion similar to [19]. For this, we project the frames back onto the terrain using projective texturing [24] and shadow mapping [25]. Shadow mapping comes with computational overhead, even when hardware assisted, so we use it only in cases where the camera views the terrain at a highly oblique angle such as when the plane is banking. In our experience, shadow mapping is not necessary very often since the plane is most commonly looking directly down or at slightly oblique angles (typically slightly forward-looking) from a height well above the variations in the local terrain.

IV. COMPUTING PERCEPTUAL COVERAGE MAPS

The basic rendering framework used to project video frames onto the terrain also provides the geometry necessary for assessing the usability of the acquired video—how see-able the content is. Simple visibility calculations can be made using the camera’s pose and the terrain model (Figure 2), but to determine the quality (see-ability) of the content in the video we also consider the viewing distance, angle of viewing incidence on the terrain, number of frames that see that target point, and multiplicity and uniqueness of

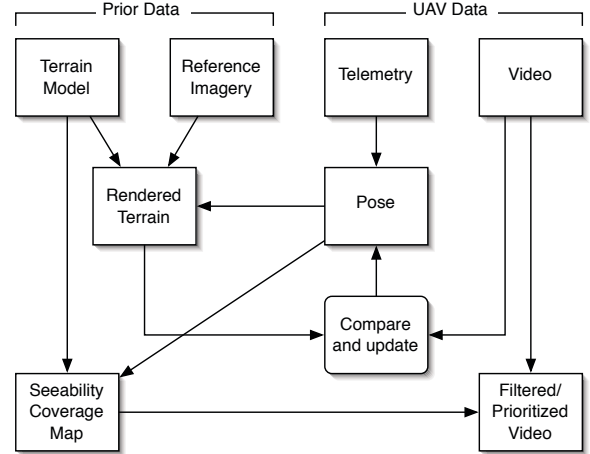


Figure 2. See-ability coverage map computation and filtered/prioritized video presentation. Previously acquired terrain and reference imagery are combined with (approximate) telemetry from the UAV to estimate the pose of the camera and to georegister the video. The pose and terrain model are used to compute the immediate see-ability for each frame, which is composited into a cumulative coverage quality map. This map can then be used to index, filter, and/or prioritize offline analysis of the video.

viewing directions. Examples of maps based on each of these criteria individually may be seen in Figure 3. Of particular interest is 3c, which shows that consideration of only the number of times each point was seen mistakenly suggests that the area was thoroughly covered, though obviously this only tells part of the story (3a,b,d).

The first two of these criteria may be used on a per-frame basis to determine the quality of the view at each pixel—what we call “immediate see-ability”. The latter two criteria are cumulative over multiple frames (or the entire video sequence), which we call “cumulative see-ability”.

To build a see-ability coverage map, we subsample the terrain model at twice the sampling of the original NED points, then compute and store see-ability data for each of these points. The resulting see-ability map may be overlaid on the original terrain (with or without the reference or video imagery) to provide a coverage map. These may be viewed in either a nadir 2D view (Figure 3) or in 3D (Figure 1).

A. Immediate Seeability

To compute the immediate (per frame) see-ability for each grid point in the coverage map, we first use view-frustum culling to select only those points potentially viewable from the camera’s pose for that frame. We then project each point back to the camera and use ray casting to test whether that point is visible to the camera or is occluded by other terrain. If it is visible, we proceed to compute the see-ability; if not, we mark the point as occluded.

Our measure for immediate see-ability is based on the effective (on-the-ground) resolution at which the video is able to image the target area. In addition to a number of constant factors such as the camera’s sampling array, the point spread function caused by the optics and the sensors, etc., the

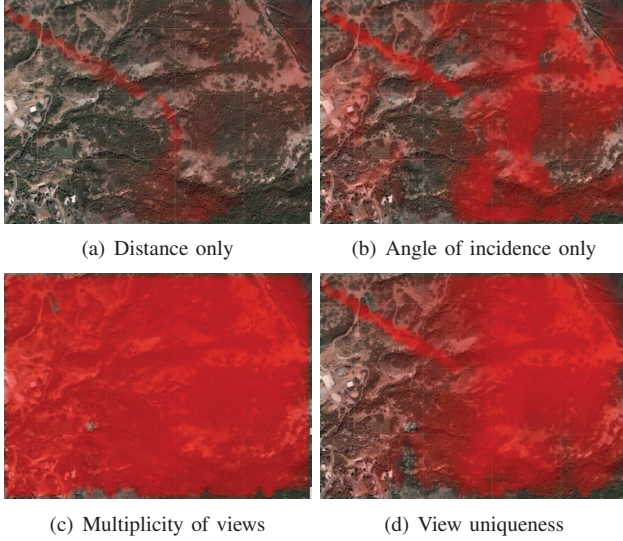


Figure 3. Factors contributing to see-ability computation (again overlaid in red on reference imagery). Considering only viewing distance (a) shows the high-altitude search to the right and the lower single pass towards the upper left. Considering only the angle of incidence (b) shows that the areas searched on the left were seen from good (downward) angles while those to the far left were seen only at high angles of incidence such as when the plane was banking in other areas. Comparing these first two maps (a,b) to the map showing the multiplicity of views (c) shows that all areas were seen multiple times but with unequal quality. The map of view uniqueness (d) shows that while areas on the right were seen from multiple compass angles, the areas on the far left were seen from few directions (mainly from the right).

effective resolution depends on the viewing distance d , the (potentially variable) focal length f of the camera, and the cosine of the angle between the viewing direction \mathbf{v} (from the point to the camera) and the terrain surface normal \mathbf{n} . Ignoring the constant factors and including only the dynamic ones, the effective resolution is proportional to

$$\frac{\mathbf{n} \cdot \mathbf{v}}{d/f} \quad (1)$$

We calculate the immediate see-ability S_{ij} for point i from camera frame j with focal length f_j as follows:

$$S_{ij} = \begin{cases} \frac{\mathbf{n}_i \cdot \mathbf{v}_{ij}}{d_{ij}/f_j} & \text{if point } i \text{ is visible in frame } j \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

Because the distance d may be large relative to the focal length, S_{ij} is often quite small. This scaling presents no problems qualitatively for creating see-ability coverage maps (which may be re-scaled for display), but for practical purposes we scale the distance component of the computation at this time. For cameras with fixed focal length lenses such as those we use, the focal length may also be omitted, since it likewise simply introduces a constant scaling factor. Folding all of the proportionality constants into a single scaling factor α and modifying the computation slightly to put S_{ij} in the range $[0, 1]$ we use

$$S_{ij} = \begin{cases} \frac{\mathbf{n}_i \cdot \mathbf{v}_{ij}}{1+d_{ij}/\alpha} & \text{if point } i \text{ is visible in frame } j \\ 0 & \text{otherwise} \end{cases} \quad (3)$$



Figure 4. Cumulative see-ability map for a test flight over a park and surrounding residential area.

For our computations, we measure the distance in meters and scale it by tens of meters, though again this scaling does not qualitatively affect the results.

B. Cumulative Seeability

We combine observations of a single spatial position from multiple frames by treating each S_{ij} as a subjective probability of detecting or correctly recognizing the target or other information-carrying items. Obviously, many other factors can and do affect this task such as the contrast of the target against the background, the lighting conditions, etc. However, we limit ourselves here to factors affected by the UAV's video acquisition and not other external factors.

If we treat each observation of each point as independent, we can compute the cumulative see-ability S_i of point i from all frames $\{j\}$ as follows:

$$S_i = 1 - \prod_j (1 - S_{ij}) \quad (4)$$

Note that although we do not factor in explicitly the number of times a point is observed in different frames, Eq. 4 includes this implicitly, monotonically increasing the cumulative see-ability with each unobstructed observation.

In some cases it may not be enough to simply observe a point repeatedly (and well) but to view it from different directions. Consider a wilderness search and rescue situation in which the missing person or a clue to their presence is positioned next to a tree, rock, minor terrain variation, or other obstruction that is not reflected in the terrain model. This may cause the person or clue to be visible from certain viewing directions but occluded from others. To include this possibility, we introduce an additional directional coverage factor for each observation based on the uniqueness U_{ij} of the compass viewing direction from camera j to point i compared to other observations of the same point. If we let θ_{ij} denote the compass direction from which frame j

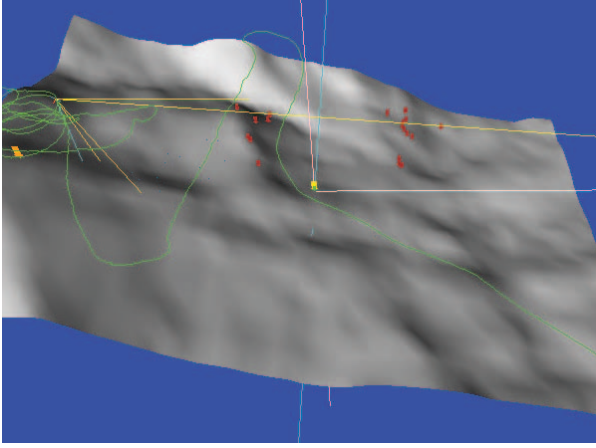


Figure 5. Occluded points may be detected and displayed (red arrows), suggesting need for further or alternative observation.

views point i , we write U_{ij} as proportional to the angular difference from the most similar view:

$$U_{ij} = \frac{\min_{k \neq j} (\theta_{ij} - \theta_{ik})}{\pi} \quad (5)$$

where the difference between compass angles is calculated as the minimum difference modulo 2π , then scaled to be in the range $[0, 1]$. Instead of using of S_{ij} as a subjective probability of detection or recognition in a single frame, we can instead now think of the product $S_{ij} U_{ij}$ as the subjective probability of this observation *adding information* to the detection process. We can thus calculate the cumulative see-ability S_i in revised form as follows:

$$S_i = 1 - \prod_j (1 - S_{ij} U_{ij}) \quad (6)$$

The cumulative see-ability calculated from the factors shown in Figure 3 can be seen in Figure 1, with an example from a flight over a different target area in Figure 4.

C. Occlusion Display

Since the computation of immediate see-ability requires visibility determination, we can incorporate this information into the visualization of the see-ability coverage maps. Rather than simply showing these occluded areas as zero see-ability (or low cumulative see-ability), we can also highlight these areas explicitly. Figure 5 shows an example indicating the occluded points seen in a single frame of the video. These can also be computed over a video sequence, giving the search team an indicator of which points may require additional search by field teams.

D. Temporal Updating and Display

Static see-ability coverage maps show what areas were seen well but not *when* they were seen. In situations where a target may be moving, this may be important to know—seeing an area well at some previous time may not mean

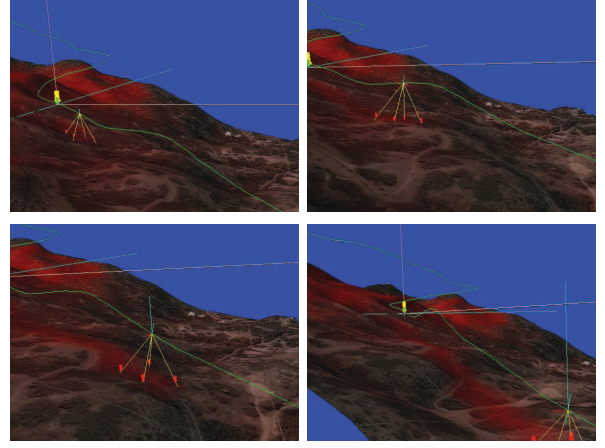


Figure 6. Coverage map updating as new video is acquired.

that the person or other target is not there now. Similarly, they do not reflect new video acquisition.

As new frames are received, coverage maps can be updated to reflect the new information gained in these frames (Figure 6). If we do not include the unique-angle factor in the computation of the cumulative see-ability (Eq. 4), the cumulative see-ability at each point in the current frame may be updated using a simple recurrence relation. If we denote the cumulative see-ability at terrain grid point i as seen in the video from the beginning through frame k as $S_i[k]$, we can compute it from the previous cumulative see-ability $S_i[k-1]$ and the immediate see-ability S_{ik} as

$$S_i[k] = 1 - (1 - S_i[k-1])(1 - S_{ik}) \quad (7)$$

By using a limited temporal window with duration n frames, one can see both new content as it arrives but also the loss of “expired” content as time passes:

$$S_i[k] = 1 - (1 - S_i[k-1]) \frac{(1 - S_{ik})}{(1 - S_{i(k-n)})} \quad (8)$$

This may also be accomplished using temporal locality weighting to cause older content to fade out gradually, though this makes the updating of the cumulative maps more expensive computationally.

V. FILTERING AND PRIORITIZATION OF OFFLINE VIDEO

Although video from the UAV is usually monitored in the field during acquisition, it is often useful to be able to access the video offline some time after it is acquired. Forms of after-the-fact searching may also be used in live acquisition situations where video searchers see something of potential interest and wish to access previously acquired video observations of that point. Georegistered video enables a number of features that rely on accurate video-terrain linking such as video indexing. Examples of these may be found both in prior literature [12]–[14], [20] and in commercial or military-use products [26], [27].

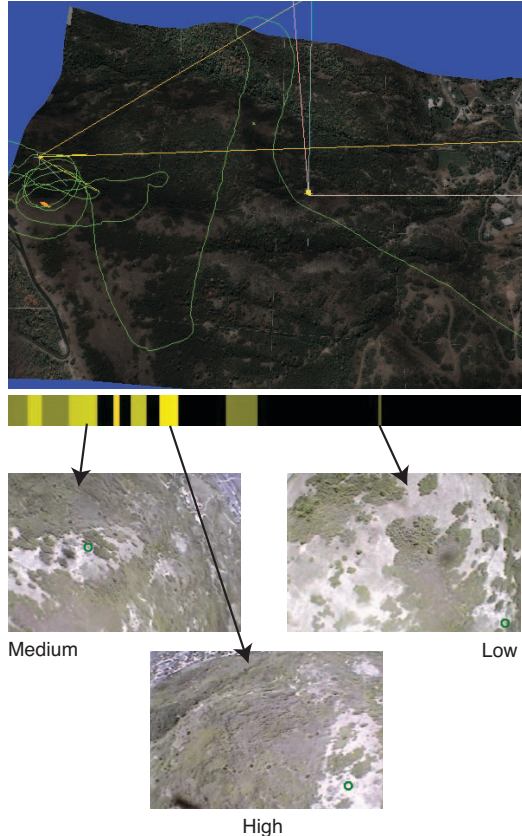


Figure 7. See-ability as an indexing cue. The visual indicator bar indicates which video segments include the target point (green marker) and the useful resolution of those segments. Shading indicates quality, from bright yellow (high quality) to black (not seen). Clicking on this indicator moves directly to those frames.

A. Spatially Constrained Playback

Using terrain-video indexing, playback can be constrained to only the sequences that correspond to a particular point or area of interest in a manner similar to [20]. This is particularly useful when one sees something of interest and wants to directly access all other video sequences that provide alternative views of that point. Users can click on a point of interest in one frame, reference the corresponding position in the terrain, and access all other video sequences that cover that point. In order to provide sequences rather than individual frames, we look for sequences that include the point of interest without gaps longer than a specified duration. To provide additional context, we also include the immediately preceding and following frames.

B. See-ability Video Filtering

See-ability coverage maps add an additional component to spatial indexing of georegistered video by allowing playback or interactive random access not only by spatial position but by usability. In order to facilitate selective playback, it is necessary not only to store which terrain points are seen from which frames but to include their respective see-ability as well. Since the full set of immediate see-ability $\{S_{ij}\}$

is extremely sparse, we store for each grid point a list of relevant frames and corresponding see-ability. During the computation of immediate see-ability for each frame we add to the list for each visible grid point a link to the current frame and the see-ability of that point from that frame. To link from the video to the corresponding point on the terrain, we store for each frame its camera pose. Selecting the corresponding terrain point from a pixel is then just a single ray cast.

When a user clicks on a point in the coverage map, we indicate not only which portions of the video include that point but the respective quality of each viewing (Figure 7). The user can choose to move to these frames manually using a shaded video “scrubber”, or the presentation system can automatically filter or prioritize the video playback.

During filtered playback, we display only the sequences that saw the point of interest with at least a minimum desired see-ability. Rather than requiring the user to specify this minimum level, we find it useful to allow the user to click in one frame to indicate a geospatial point of interest and corresponding minimum see-ability. This interaction and resulting presentation is not just “show me more” but “show me *better*”, potentially greatly reducing the time to investigate and confirm or reject items of potential interest.

This notion can also be extended to the entire video, causing the video presentation system to include only those portions that are of a desired see-ability. This filtering removes sequences when the plane is seeing distant terrain while banking or when it is flying at a higher altitude as is often done while moving to or from the launch point to the target area or from one area to another.

C. See-ability Video Prioritization

Since offline analysis of video may be motivated by the need to find a missing person or other target as quickly as possible, we have also found it useful to prioritize the presentation of the video. Rather than simply filtering out low-quality content, we can also sort the presented video sequences by their usefulness. A user may thus indicate a portion of the terrain and jump directly to the highest-quality sequence that included that spatial position, or they could be shown multiple sequences of decreasing usefulness.

D. User’s See-ability Filtering

See-ability coverage maps can also be computed using only the video already seen by a user in offline viewing. This allows us to map not what the *camera* has seen but what the *user* or a set of users has analyzed. Combining coverage maps for both the camera’s acquired video and what the user has already viewed (Figure 8) allows us to further filter or prioritize the presentation of the video. The user(s) can then choose to be shown only sequences that are of sufficient see-ability to be useful *and* include areas that they have not already seen well. Alternatively, they may

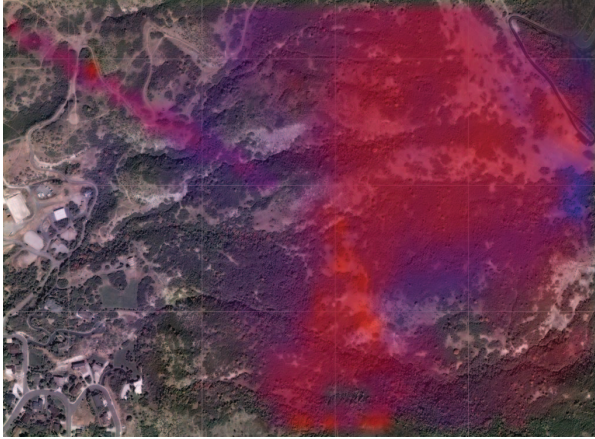


Figure 8. Coverage maps showing what the user has analyzed so far (blue) and the remaining video (red).

choose to view all of the acquired video but prioritized by both usability and lack of redundancy—i.e., by the amount of new and usable information each sequence brings.

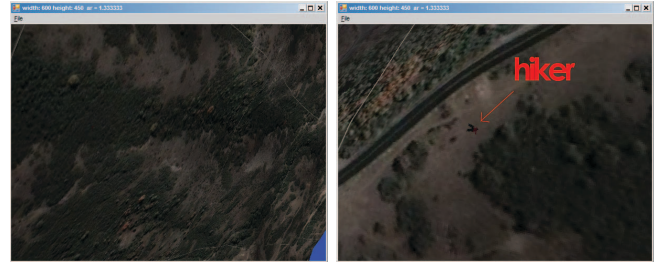
VI. PARTIAL VALIDATION OF IMMEDIATE SEE-ABILITY

Our see-ability metric relates directly to the effective resolution with which the UAV’s camera sees the terrain. Although this is a well-known (and obvious) factor in detection and recognition, it is still worth validating the metric in the context of specific tasks. To do so for immediate see-ability in a search-and-rescue context, we conducted a user study in which subjects were asked to detect a simulated missing person in actual footage from a search flight.

After georegistering the video footage to the relevant terrain and reference imagery, we selected 27 points on the terrain and 82 video frames that included one or more of these points so as to create a range of viewing distances and angles of incidence. In 52 of these frames, we synthetically inserted at one of these target points a rendered figure of a person wearing a red shirt and blue jeans (Figure 9). This use of synthetic targets allowed us to control their placement and viewing while otherwise using real footage. The other 30 frames did not include targets and acted as a control set.

Each of 18 volunteer subjects observed each of the 82 frames in randomized order in a controlled setting. The subjects were asked to indicate for each frame whether the frame included the target, which they indicated by pressing a key on the keyboard. (No response was required if they did not see it, as would be the case when observing continuous search footage.) Each frame was presented for two seconds, followed by a one-second masking image [28] before presentation of the next frame.

For each of the 82 frames, we computed the immediate see-ability of the point at which the target was placed in that frame, or for frames with no targets we used the see-ability of the center of the frame, and compared this prediction to subject performance. Results showed a 0.56 correlation



(a) Frame without target

(b) Frame with target

Figure 9. Examples frames without (left) and with (right) detection targets as used in the user study.

coefficient with high statistical significance ($p < 0.0001$), indicating a moderately strong correlation [29].

As already suggested, there are obviously other factors that can affect performance in this task besides the effective resolution of the video. The results of our study show that even when omitting these other factors, resolution alone is a strong enough predictor of performance that see-ability-based coverage maps can be used to qualitatively visualize the relative coverage and usability of search video.

VII. DISCUSSION AND FUTURE WORK

For this work, we have chosen to focus solely on see-ability factors related to the position and control of the UAV. While resolution is a key element of video quality, there are obviously other factors as well. Subjects in the user study noted anecdotally that the placement of the missing person (against brown dirt, green shrub brush, etc.) was often the dominant factor given even modest resolution. If there is reason to believe that the missing person or other potential target of interest is a particular color, it should be possible to include this in see-ability models.

Similarly, this work has not attempted to define exactly what constitutes a minimum usable resolution, since this depends on various other factors of the imaging system as well as external factors just described. Because of the many internal and external factors involved in various tasks, we believe this may have to be done on a system- and application-specific basis.

The approaches presented here, while driven in part by problems associated with assessing coverage for relatively unstable mini-UAVs, could be extended to any form of aerial search, though specifics of the model may depend on task- or platform-specific factors. For wide-angle imaging systems, for example, one might need to handle the non-uniform resolution induced by the lens or catadioptric system. For manned aircraft, there might not be a camera but rather human spotters. It might be interesting to couple telemetry data from the plane with spotter head-tracking information.

VIII. CONCLUSION

We have presented methods for using see-ability, or resolution-based usability for human observers, to create

perceptual coverage quality maps for UAV-acquired search video, including both single-frame immediate see-ability and cumulative see-ability over video sequences. Assessing the spatial coverage quality for search video is essential to WiSAR, as well as to other spatial-search tasks. Both static (post hoc) and dynamic (during flight) see-ability coverage maps can be used for coverage assessment, for guidance and planning, and for coordination. Although we have presented this work in the context of UAV-based search operations, these methods can be extended to other forms of aerial search or to other applications in which it is important to know what parts of a target area have been observed well.

We have also presented ways to use these coverage maps to provide both geospatial and quality-based indexing of search video. Users can interact dynamically with the video not just in temporal order or filtered by geospatial position but filtered and prioritized by how usable that video is.

A complete validation of the relevance of see-ability to WiSAR experts is beyond the scope of this paper, but the experiment presented here indicated that even a simple see-ability metric accounted for a large portion of what is important for evaluating quality of search coverage.

ACKNOWLEDGMENT

The work was partially supported by the National Science Foundation under grant numbers 0534736 and 0812653. Any opinions, findings, and conclusions or recommendations expressed are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

- [1] M. A. Goodrich, B. S. Morse, D. Gerhardt, J. L. Cooper, J. Adams, C. Humphrey, and M. Quigley, "Supporting wilderness search and rescue using a camera-equipped mini UAV," *J. Field Robotics*, vol. 25, no. 1–2, pp. 89–110, January 2008.
- [2] M. A. Goodrich, B. S. Morse, C. Engh, J. L. Cooper, and J. A. Adams, "Towards using UAVs in wilderness search and rescue: Lessons from field trials," *Interaction Studies*, vol. 10, no. 3, pp. 453–478, 2009.
- [3] J. Casper and R. R. Murphy, "Human-robot interactions during the robot-assisted urban search and rescue response at the world trade center," *IEEE Trans. Syst., Man, Cybern. B*, vol. 33, pp. 367–385, 2003.
- [4] J. Reason, *Human Error*. Cambridge University Press, 1990.
- [5] D. D. Woods and J. C. Watts, "How not to have to navigate through too many displays," in *Handbook of Human-Computer Interaction*, 2nd ed., M. G. Helander, T. K. Landauer, and P. Prabhu, Eds. Elsevier, 1997.
- [6] D. D. Woods, J. Tittle, M. Feil, and A. Roesler, "Envisioning human-robot coordination in future operations," *IEEE Trans. Syst., Man, Cybern. C*, vol. 34, pp. 210–218, 2004.
- [7] J. Cooper and M. A. Goodrich, "Towards combining UAV and sensor operator roles in UAV-enabled visual search," in *ACM/IEEE Int. Conf. on HRI*, 2008, pp. 351–358.
- [8] J. Rasmussen and K. J. Vicente, "Coping with human errors through system design: Implications for ecological interface design," *Int. J. Man-Machine Studies*, pp. 517–534, 1989.
- [9] S. Thrun, "Robotic mapping: A survey," in *Exploring Artificial Intelligence in the New Millennium*. Morgan Kaufmann, 2003, pp. 1–33.
- [10] J. Blitch, R. Murphy, and T. Durkin, "Mobile semiautonomous robots for urban search and rescue," in *Encyclopedia of Microcomputers*, A. Kent and J. G. Williams, Eds. CRC Press, 2002, vol. 28, ch. 19, pp. 211–222.
- [11] R. Cutler, C. Shekhar, B. Burns, R. Chellappa, and R. Bolles, "Monitoring human and vehicle activities using airborne video," in *28th Applied Imagery Patt. Rec. Workshop*, 1999.
- [12] K. J. Hanna, H. S. Sawhney, R. Kumar, Y. Guo, and S. Samarasekera, "Annotation of video by alignment to reference imagery," in *IEEE Int. Conf. on Multimedia Comp. & Sys.*, 1999, pp. 38–43.
- [13] R. Kumar, H. Sawhney, S. Samarasekera, S. Hsu, H. Tao, Y. Guo, K. Hanna, A. Pose, R. Wildes, D. Hirvonen, M. Hansen, and P. Burt, "Aerial video surveillance and exploitation," *Proceedings of the IEEE*, vol. 89, no. 10, pp. 1518–1539, October 2001.
- [14] R. Kumar, H. S. Sawhney, J. C. Asmuth, A. Pope, and S. Hsu, "Registration of video to geo-referenced imagery," *IEEE CVPR*, pp. 54–62, 1998.
- [15] Y. Lin and G. Medioni, "Map-enhanced UAV image sequence registration and synchronization of multiple image sequences," in *IEEE CVPR*, 2007, pp. 1–7.
- [16] H. Schultz, A. Hanson, E. Riseman, F. Stolle, Z. Zhu, C. Hayward, and D. Slaymaker, "A system for real-time generation of geo-referenced terrain models," in *SPIE Symposium on Enabling Technologies for Law Enforcement*, November 2000.
- [17] C. Shekhar and R. Chellappa, "Airborne video registration for activity monitoring," in *Video Registration*, M. Shah and R. Kumar, Eds. Springer, 2003, ch. 6.
- [18] R. P. Wildes, D. J. Hirvonen, S. C. Hsu, R. Kumar, W. B. Lehman, B. Matei, and W. Y. Zhao, "Video georegistration: algorithm and quantitative evaluation," in *IEEE ICCV*, vol. 2, 2001, pp. 343–350.
- [19] H. S. Sawhney, A. Arpa, R. Kumar, S. Samarasekera, M. Aggarwal, S. Hsu, D. Nister, and K. Hanna, "Video flashlights: real time rendering of multiple videos for immersive model visualization," in *Eurographics Workshop on Rendering*, 2002, pp. 157–168.
- [20] M. Irani and P. Anandan, "Video indexing based on mosaic representations," *Proceedings of the IEEE*, vol. 86, no. 5, pp. 905–921, 1998.
- [21] United States Geological Survey, "The national map seamless server," <http://seamless.usgs.gov/>.
- [22] D. Luebke, M. Reddy, J. Cohen, A. Varshney, B. Watson, and R. Huebner, *Level of Detail for 3D Graphics*. Morgan Kaufmann, 2002.
- [23] R. Pajarola and E. Gobbetti, "Survey of semi-regular multiresolution models for interactive terrain rendering," *The Visual Computer*, vol. 23, no. 8, pp. 583–605, August 2007.
- [24] M. Segal, C. Korobkin, R. van Widenfelt, J. Foran, and P. Haeblerli, "Fast shadows and lighting effects using texture mapping," in *ACM SIGGRAPH*, 1992, pp. 249–252.
- [25] L. Williams, "Casting curved shadows on curved surfaces," in *ACM SIGGRAPH*, 1978, pp. 270–274.
- [26] Air Force Research Laboratory, "Geo*View."
- [27] Science Applications International Corporation, "GeoViz."
- [28] H. Piéron, *The Sensations: Their Functions, Processes, and Mechanisms*. Yale University Press, 1952.
- [29] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences (2nd Edition)*. Lawrence Erlbaum Associates, 1988.