Detecting and tracking moving targets through foliage is difficult, and for many cases even impossible in regular aerial images and videos. We present an initial light-weight and drone-operated 1D camera array that supports parallel synthetic aperture aerial imaging. Our main finding is that color anomaly detection benefits significantly from image integration when compared to conventional raw images or video frames (on average 97% vs. 42% in precision in our field experiments). We demonstrate that these two contributions can lead to the detection and tracking of moving people through densely occluding forest.

1. Introduction

With Airborne Optical Sectioning (AOS, [1–10]), we have introduced a wide synthetic aperture imaging technique that employs conventional drones to sample images above the forest. These images are computationally combined (registered to the ground and averaged) to integral images which suppress strong occlusion and make hidden targets visible. AOS relies on a statistical chance that a point on the forest ground is unoccluded by vegetation from multiple perspectives, as explained by the statistical probability model in [2]. The integral images can be further analyzed to support, for instance, automated person classification with advanced deep neural networks. In [9], we have shown that integrating raw images before classification rather than combining classification results of raw images is significantly more effective when classifying partially occluded persons in aerial thermal images (92% vs. 25% average precision). In [10], we demonstrate a first fully autonomous drone for search and rescue based on AOS. The main advantages of AOS over alternatives such as airborne LiDAR [11–14] or Synthetic Aperture Radar [15–18] are its real-time computational performance with high spatial resolution when deployed for occlusion removal on low cost system-on-chip computers (SoCC); its applicability to other wavelengths, such as far infrared for wildlife observations and search and rescue, or near infrared for agriculture and forestry applications. AOS is a passive system with limited reliability in providing depth information, as depth maps derived from AOS focal stacks by depth-from-defocus techniques [19] are still imprecise while LiDAR and SAR still remain the preferred choice.

Thus far, the sequential sampling nature of AOS limits its applications to static targets only. Moving targets, such as walking people or running animals, lead to motion blur in integral images that are nearly impossible to detect and track (cf. Figure 1, for example). Applying AOS to far infrared (thermal imaging), as in [9, 10], restricts it to cold environment temperatures, while using it in the visible range (RGB imaging) as in [1] often suffers from too little sunlight penetrating through dense vegetation.

In this article, we make two main contributions: First, we present an initial light-weight (<1 kg) and drone-operated 1D camera array that supports parallel AOS sampling. While 1D and 2D camera arrays have already been used for implementing various visual effects (e.g., [20–26]), they have not been applied for aerial imaging (in particular with drones) because of their size and weight. Second, we show that color anomaly detection (e.g., [27, 28]) benefits significantly from AOS integral images when compared to conventional raw images (on average 97% vs. 42% in precision). Color anomaly detection is often used for automated aerial image analysis in search and rescue applications (e.g., [29–31]) because of its independence to environment temperature (in contrast to thermal imaging). It finds pixels or clusters of pixels in an
image with significant color differences in comparison to their neighbours. However, color anomaly detection fails in presence of occlusion. We demonstrate that these two contributions can lead to the detection and tracking of moving people through dense forest.

2. Materials and Methods

2.1. Integral Imaging with Airborne Camera Arrays. As illustrated in Figure 2, our new payload captures multiple aerial images with a drone-operated 1D camera array. It samples the forest in parallel at flying altitude (h) within the range of a synthetic aperture (SA) that equals the size of the camera array. This results in a structured 4D light-field formed by image pixels which are represented as light rays in a 3D volume, as discussed in [32, 33]. With known camera intrinsic, camera poses, and a representation of the terrain (either a digital elevation model representing only the height information without any further definition about the surface as utilized in [10], or a focal plane approximation if not, as in [4]), each ray’s origin on the ground can be reconstructed. An occlusion-reduced integral image can be obtained by averaging all the rays that have the same origin. Depending on occlusion density, more or fewer rays of a surface point contain information of random occluders, while others contain the signal information of the target, as shown in Figure 2. Therefore, integrating multiple rays (i.e., averaging their corresponding pixels) results in focus of the target and defocus of the occluders. This increases the probability of detecting the target reliably under strong occlusion conditions [9].

In practice, the acquired images are pre-processed for intrinsic camera calibration, image un-distortion/rectification and are cropped to a common field of view. Furthermore, poses of all cameras at each instance in time have to be estimated. While GPS and IMU measurements enable real-time pose estimation (non-differential GPS and IMU modules available on off-the-shelf low-cost commercial drones are error prone as in [10]), computer vision techniques (e.g., multi-view stereo and structure-from-motion) applied to the recorded images (as in [9]) are more precise.

The pre-processed images are finally projected to a common ground surface representation (digital elevation model or focal plane), using the cameras’ intrinsic parameters and poses, and are averaged to the resulting integral image. Details on the concrete implementation of the prototype presented in this article are provided in Section 2.2.

2.2. Prototype. Figure 3(a) illustrates our current prototype. The drone basis system is a MikroKopter OktoXL 6S12 octocopter. The custom-built light-weight camera array is based on a truss design (Figure 3(b)) which can handle the forces and vibrations due to wind by safely distributing them over the entire structure.

The supporting tubes tilted outwards towards the outermost section at an angle α are linked to the hollow carbon fibre tubes using a 3D-printed T-shaped tube connector (Figure 3(c)) and are subjected to axial loading through an ultra-high molecular weight poly-ethylene (UHMWPE) string with a diameter of 0.9 mm and 95daN ultimate load carrying capacity. The pre-tensioning of the UHMWPE string prevents the downward bending of the structure. A labyrinth-like connection allows easy assembly of the 3D printed fixation tube (Figure 3(d)) and the UHMWPE string. The hollow detachable carbon fibre tubes with thin-walled circular cross-section are commercially available Preston response match landing rods that are manufactured through filament winding. These tubes have high bending stiffness and therefore minimize the structure deflection due to its own dead weight. The 3D-printed drone mounting bracket shown in Figure 3(e) is used to mount the structure on the drone and prevents the carbon fibre tube from spinning while the tilt fixation bracket in Figure 3(f) increases the stability of the structure. All custom-built 3D-printed connectors and brackets are made of flexible rubber-like thermoplastic polyurethane (TPU 95 Shore) to suppress crack formation and handle vibrations.

The camera array carries ten light-weight DVR pin-hole cameras (12 g each), attached equidistant (1 m) to each other on a 9 m long detachable and hollow carbon fibre tube (700 g) which is segmented into detachable sections of varying lengths and with a gradual reduction in diameter in each
section from 2.5 cm at the drone center to 1.5 cm at the outermost section. The cameras are aligned in such a way that their optical axes are parallel and pointing downwards. They record RGB images at a resolution of 1600×1200 pixels and RGB videos at a resolution of 1280×720px and 30fps to individual SD cards. All cameras receive power from two central 7.2V Ni-MH batteries and are synchronously triggered from the drone’s flight controller through a flat-band cable bus.

To ensure stable flights and to avoid resonance oscillation with our payload, the drone's PID controller had to be reconfigured. PID stands for Proportional-Integral-Derivative which is a part of the flight control software that continuously reads the information provided by the drone’s sensors and calculates how fast the motors must spin in order to retain the desired rotation speed for a stable flight. The P controller changes the motor power proportional to the angle of inclination and...
the \( I \) controller changes motor power continuously depending on the deflection angle and the time while the \( D \) controller responds to any rapid changes in the sensor data. The \( PID \) parameters were tuned to an \( I \)-dominant state (Gyro \( P \): 100, Gyro \( I \): 255, Gyro \( D \): 10, all in a 0-255 range) for increased motor power to instantly compensate imbalances caused by external forces (such as wind) and to avoid oscillation of camera array's long lever. Supplementary videos 1 and 2 show test flights with default (causing resonance oscillation) and with tuned \( PID \) parameters (ensuring stable flight).

For camera calibration, image un-distortion, and rectification, we apply OpenCV's pinhole camera model (as explained in [1, 9]). The undistorted and rectified images are cropped to a field of view of 41.10° and a resolution of 1024px×1024px. Pose estimation is carried out using the general-purpose structure-from-motion and multi-view stereo pipeline, COLMAP ([34, 35]). Image back-projection and averaging are computed for a common focal plane at the ground surface level. Integral images are always rendered from the center perspective of the camera array. OpenGL deferred rendering is utilized here with our GPU light field renderer implemented using Python, C, OpenGL, C++, and Cython integration (AOS source code, data, and publications are available at https://github.com/JKU-ICG/AOS/).

For target classification, we apply the Reed-Xiaoli (RX) anomaly detector ([27]) to the integral images. All computations are carried out offline (after landing), and require (without pose estimation but including RX detection) 896 ms per integral image on an Intel(R) Core (TM) i5-6400 CPU @2.70 GHz (64GB RAM) with NVIDIA GeForce GTX 1070 GPU. Computer vision-based pose estimation with COLMAP is slow and requires approximately 42 s on our hardware.

However, it becomes obsolete when being replaced by fast and precise online sensor measurements (such as Real Time Kinematic GPS). Furthermore, the CPU implementation of our RX detector is not performance optimized (requires 436 ms/896 ms). A GPU implementation of it will lead to an additional speed-up.

3. Results

3.1. AOS Enhanced Color Anomaly Detection. Many color anomaly detection techniques are applied to hyper-spectral or multi-spectral images ([36]), with the Reed-Xiaoli (RX) anomaly detector being the standard and most widely used ([37]). The RX anomaly detector ([27]) characterizes the image background in terms of a covariance matrix and calculates the RX score based on the Mahalanobis distance between a test pixel \( r \) and the background as follows:

\[
a_{RXD}(r) = (r - \delta)^T K^{-1} (r - \delta),
\]

where \( r \) is the spectral 3-vector of the pixel under test, \( \delta \) is the spectral mean 3-vector of the background, and \( K \) is the 3×3-covariance matrix.

The RX scores are thresholded based on its cumulative probability distribution function. The threshold represents the RX score confidence beyond which the pixel under test is considered an anomaly. The RX score confidence is set manually to maximize the detection of true positive pixels (i.e., maximum recall of the targets with a minimum one
Table 1: Comparison of color anomaly detection result between raw images and integral images before (SS-) and after (SS+) sunset.

| Flight | Set No. | C0  | C1  | C2  | C3  | C4  | C5  | C6  | C7  | C8  | C9  | PAs  | Pi   |
|--------|---------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|------|
| SS+    | 1       | 9.3 | 18.1| 8.0 | 9.7 | 14.1| 18.5| 5.7 | 10.9| 5.7 | 5.7 | 10.6 | 98.3 |
| SS-    | 2       | 0   | 37.2| 58.6| 59.2| 68.1| 45.9| 85.2| 86.7| 33.6| 0   | 47.4 | 79.9 |
| SS-    | 3       | 0   | 35.7| 64.8| 47.4| 49.2| 68.7| 29.0| 44.0| 0   | 0   | 33.9 | 100  |
| SS+    | 4       | 84.5| 81.8| 67.2| 83.5| 94.9| 89.3| 94.1| 84.9| 72.2| 0   | 75.2 | 98.8 |
| SS+    | 5       | 21.5| 16.4| 25.9| 39.5| 36.0| 34.8| 32.3| 22.7| 21.4| 0   | 25.1 | 100  |
| SS+    | 6       | 34.9| 98.0| 96.5| 95.9| 64.3| 84.7| 64.7| 96.6| 100 | 92.2| 82.8 | 100  |
| SS+    | 7       | 30.8| 66.8| 92.9| 100 | 91.3| 99.6| 100 | 100 | 72.2| 0   | 75.3 | 85.2 |
| SS+    | 8       | 75.4| 100 | 78.0| 94.1| 99.0| 100 | 79.1| 100 | 95.1| 35.4| 85.6 | 100  |
| SS+    | 9       | 47.1| 41.7| 26.6| 40.1| 50.8| 45.4| 48.1| 48.7| 6.7 | 0   | 35.5 | 95.2 |
| SS-    | 10      | 1.5 | 24.7| 39.3| 48.6| 67.2| 0.5 | 0   | 0   | 0   | 0   | 18.2 | 100  |
| SS-    | 11      | 0   | 2.9 | 16.5| 37.0| 56.4| 31.2| 100 | 57.4| 35.5| 0   | 33.7 | 100  |
| SS-    | 12      | 15.0| 38.5| 18.0| 52.0| 44.7| 43.5| 78.9| 16.5| 10.1| 0   | 31.7 | 100  |
| SS-    | 13      | 27.3| 27.6| 33.8| 56.5| 65.3| 36.4| 10.8| 2.9 | 5.9 | 0   | 26.6 | 99.2 |
| SS-    | 14      | 13.6| 43.1| 40.0| 48.8| 71.4| 35.1| 6.5 | 0   | 0   | 0   | 25.9 | 98.5 |
| SS-    | 15      | 7.9 | 72.6| 76.8| 91.3| 97.6| 60.5| 11.5| 0   | 35.1| 22.8| 47.6 | 93.1 |
| SS-    | 16      | 57.7| 42.3| 59.0| 92.8| 99.6| 68.6| 32.4| 15.4| 18.7| 0   | 48.7 | 100  |
| SS-    | 17      | 0   | 60.5| 90.4| 97.2| 61.6| 19.4| 85.0| 42.7| 1.8 | 0   | 45.9 | 100  |
| SS+    | 18      | 20.4| 25.6| 24.9| 25.4| 15.0| 16.2| 23.7| 14.2| 2.0 | 14.4| 18.2 | 91.4 |
| SS+    | 19      | 0   | 64.3| 38.1| 50.5| 30.5| 50.5| 67.7| 73.0| 30.6| 0   | 40.5 | 100  |
| SS+    | 20      | 9.5 | 38.8| 17.8| 6.0 | 24.4| 68.0| 81.4| 26.9| 0   | 0   | 27.3 | 100  |

Figure 5: Color anomaly detection under dark light conditions (after sunset). Ten raw images recorded by camera array (top three rows): raw RGB images, RX anomaly mask, mask overlaid over raw image. Red and green circles indicate false and true positives. Numbers are precision values in %. The corresponding integral image (raw, mask, and overlaid mask) is shown at the bottom row. Visible are three targets: two moving people and one static light source (largest cluster) for spatial reference.
and minimize the detection of false positive pixels to attain the maximum pixel-based precision value. Note that this procedure is applied to single and integral images equally, thereby obtaining the best results the RX detector can extract from the input images when the optimal confidence threshold is known. It allows comparing results independently of a more or less well-chosen threshold, as they depend only on the suitability of the input image for color anomaly detection.

Figures 4 and 5 present visual results of the RX anomaly detector applied to data sets captured by performing free flights (waypoint missions are also supported) with our drone prototype over dense mixed forest before and after sunset [performed in compliance with the legal European union Aviation Safety Agency (EASA) flight regulations, before sunset +30 min]. We either apply the RX detector to the ten single images captured by the camera array individually or to the corresponding integral image that results from registering and averaging the same ten single images. For both cases, optimal thresholds (i.e., one average threshold (CCs) for all single images and one threshold (CCI) for the integral image) are found as explained above (i.e., by maximizing true positive pixels and recall of the targets while minimizing false positive pixels). For all results, we use precision (ratio of true positive and true + false positive pixels in percent) as a quality metric. Here, $P_A$ is the precision average over all ten single images, and $P_i$ is the precision value of the corresponding integral image. Note that the number of ground truth pixels (i.e., the number of target pixels under occlusion free conditions) is unknown. Also, note that the precision value is 0% in the case of no true positives (i.e., target is completely occluded or not within the field of view of the camera) but in the presence of false positives (i.e., wrong detections). For the case of no false positives but in the presence of true positives, the precision value is 100%.

Figure 6: Motion-based single-person tracking. Subset of single video frames (two top rows) and corresponding subset of integral frames (two bottom rows): raw RGB frames, RX anomaly mask, and tracking labels overlaid. Visible are two targets: one moving person (label 2) and one static light source (label 1) for spatial reference.

Figure 7: Motion-based multi-person tracking. Subset of single video frames (two top rows) and corresponding subset of integral frames (two bottom rows): raw RGB frames, RX anomaly mask, and tracking labels overlaid. Visible are three targets: two moving people (labels 1 and 3) and one static light source (label 2) for spatial reference.
As demonstrated in Figures 4 and 5, color anomaly detection benefits significantly from image integration. While strong occlusion causes many false positives but little true positives in raw images, false positives are almost entirely removed and true positives are significantly increased in the integral images.

Table 1 presents quantitative results of twenty data sets captured before (SS-) and after (SS+) sunset. Here, $P$s are the precision values of raw images ($C_0$-$C_9$), $PAs$ the precision average over all ten images in each set, and $Pi$ the precision value of the corresponding integral image. On average, we achieve an improvement from 42% to 97% in precision when AOS is used in combination with color anomaly detection. We did not find a significant difference of this improvement when separating the low light SS+ sets (48% vs. 97%) and bright light SS- sets (37% vs. 97%).

Figure 8: RGB values of target ($r_t$, red circles) and background ($r_b$, blue points) pixels for all raw images and resulting integral image of the two example sets shown in Figures 4 (top two rows) and 5 (bottom two rows). The spectral mean background vectors ($\delta$, green crosses) and 2\sigma ellipses (enclosing 95% of the pixel data) generated from the eigenvectors and eigenvalues of the covariance matrices (blue and red ellipses for background and target data, respectively) are illustrated. RGB space is projected to two principle axes in (a)–(c), and is 3D rendered, zoomed, and cropped to show details in (d).
data sets are available ([38]). We discuss the underlying cause for this enhancement in Section 4.

3.2. Tracking Occluded Targets. By making use of the camera array’s video recording functionality and by extending the AOS image integration process to compute integral videos allow the application of a simple multi-object tracker to the RX anomaly mask time series at recording speed (30fps in our case).

For tracking, we apply a blob detector to the binary RX anomaly masks which finds groups of connected pixels that differ in properties like color or brightness compared to the background. The association of the detections to the same targets depends on their motions, which is estimated using a Kalman filter.

As illustrated in Figures 6 and 7, object tracking in integral videos performs significantly better than in regular video recordings. While many false positive targets are detected, the true positive targets are not detected or are not consistently tracked throughout regular (single) video frames. However, few (none, in the examples shown in Figures 6 and 7) false positive targets are detected and all true positive targets are mostly consistently tracked throughout the integral video frames. Supplementary videos 3 and 4 show these examples in motion.

The reason for the improved tracking quality is clearly the enhanced color anomaly detection in the integral video frames.

4. Discussion

As shown in Eqn. (1), the RX anomaly detector ([27]) characterizes the image background in terms of a covariance matrix \( K \) and calculates the RX score based on the Mahalanobis distance between a test pixel \( r \) and the spectral mean vector of the background \( \delta \).

Figure 8 visualizes scatter plots of RGB pixel values for all raw images and resulting integral image of the two example sets shown in Figures 4 and 5. Clearly, image integration decreases variance and covariance, which leads to better clustering and consequently to enhanced separation of target and background pixels. This is evident by the integral images’ shrunk 2σ ellipses that results from the covariance matrices’ decreased eigenvectors and coefficients. The coefficients of \( K \) are decreased significantly (by a factor of 3-4 in our examples) for integral images when compared to the coefficient of \( K \) in raw images (as shown in Figure 9, where we plot the average of coefficients for all 20 data sets as in Table 1). As a result, the distance between the two spectrally different clusters (background and targets) increases in integral images, which can then be well segmented using the Mahalanobis distance.

Since the RX detector multiplies with the inverse of \( K \) (Eqn. (1)), lower coefficients lead to higher RX scores.

AOS is an effective wide synthetic aperture aerial imaging technique that can also be considered a variation of the standard signal averaging theorem (as explained in [2]). In contrast to signal averaging, noise (occluders, such as vegetation in our case) is highly correlated (i.e., correlated background pixels of raw images shown in Figure 8, raw). By shifting focus computationally towards the targets and by averaging multiple raw images accordingly lead to a strong and more uniform point spread of out-of-focus occluders. This improves the signal-to-noise ratio between background and targets (as can be seen in Figure 8, integral).

While detecting and tracking moving targets through foliage is difficult (and often even impossible) in regular aerial images or videos, it becomes practically feasible with image integration—which is the core principle of Airborne Optical Sectioning. We have already shown that classification significantly benefits from image integration ([9]). In this work, we demonstrate that the same holds true for color anomaly detection. This finding together with the implementation of an initial drone-operated camera array for
parallel synthetic aperture aerial imaging allows presenting first results on tracking moving people through dense forest. Besides people, other targets (e.g., vehicles or animals) can be detected and tracked in the same way. This might impact many application domains, such as search and rescue, surveillance, border control, and wildlife observation.

The utilized RX color anomaly detector and the applied combination of blob detection and Kalman filter for tracking are only implementation examples that serve a proof-of-concept. They can be replaced by more advanced techniques. However, we believe that our main finding (i.e., anomaly detection and tracking benefit significantly from image integration) will still hold true.

Color anomaly detection is clearly limited to detectable target color. In our experiments, targets were colored in white, black, blue, and red. Greenish color would have most likely not been detected. A combination of color (RGB), thermal (IR), and time (motion itself) channels for anomaly detection might result in further improvements. This has to be investigated in the future. Furthermore, the implications of parallel-sequential sampling strategies and other sampling devices, such as re-configurable drone swarms instead of camera arrays with a fixed sampling pattern, have to be explored. We would like to also explore fusion-based pose estimation approaches that offers high accuracy and precision in real-time.

Data Availability

All twenty data sets captured before (SS-) and after (SS+) are available in the Zenodo data repository (10.5281/zenodo.5680949). The AOS source code, data, and publications are available at https://github.com/JKU-ICG/AOS/.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Authors’ Contributions

R.J.A.A. Nathan assigned to the software, data curation, and writing - original draft preparation. I. Kurmi assigned to the software, data curation, and writing - original draft preparation. D.C. Schedd assigned to the software and data curation. O. Bimber assigned to the conceptualization of this study, methodology, and writing - original draft preparation.

Acknowledgments

The camera array structure was designed and constructed by JKU’s Institute of Structural Lightweight Design. This research was funded by the Austrian Science Fund (FWF) under grant number P 32185-NBL and by the State of Upper Austria and the Austrian Federal Ministry of Education, Science and Research via the LIT – Linz Institute of Technology under grant number LIT-2019-8-SEE-114.

Supplementary Materials

Video 1. Test flight with default PID parameters (resonance oscillation). Video 2. Test flight with tuned PID parameters (stable flight). Video 3. Motion-based single-person tracking. Video 4. Motion-based multi-person tracking. (Supplementary Materials)

References

[1] I. Kurmi, D. C. Schedd, and O. Bimber, "Airborne optical sectioning," Journal of Imaging, vol. 4, no. 8, p. 102, 2018.
[2] I. Kurmi, D. C. Schedd, and O. Bimber, "A statistical view on synthetic aperture imaging for occlusion removal," IEEE Sensors Journal, vol. 19, no. 20, pp. 9374–9383, 2019.
[3] I. Kurmi, D. C. Schedd, and O. Bimber, "Thermal airborne optical sectioning," Remote Sensing, vol. 11, no. 14, p. 1668, 2019.
[4] I. Kurmi, D. C. Schedd, and O. Bimber, "Fast automatic visibility optimization for thermal synthetic aperture visualization," IEEE Geoscience and Remote Sensing Letters, vol. 18, no. 5, pp. 836–840, 2021.
[5] I. Kurmi, D. C. Schedd, and O. Bimber, "Combined person classification with airborne optical sectioning," Scientific Reports, vol. 12, no. 1, pp. 1–11, 2022.
[6] I. Kurmi, D. C. Schedd, and O. Bimber, "Pose error reduction for focus enhancement in thermal synthetic aperture visualization," IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1–5, 2022.
[7] O. Bimber, I. Kurmi, D. C. Schedd, and M. Potel, "Synthetic aperture imaging with drones," IEEE Computer Graphics and Applications, vol. 39, no. 3, pp. 8–15, 2019.
[8] D. C. Schedd, I. Kurmi, and O. Bimber, "Airborne optical sectioning for nesting observation," Nature Scientific Reports, vol. 10, no. 1, pp. 1–7, 2020.
[9] D. C. Schedd, I. Kurmi, and O. Bimber, "Search and rescue with airborne optical sectioning," Nature Machine Intelligence, vol. 2, no. 12, pp. 783–790, 2020.
[10] D. C. Schedd, I. Kurmi, and O. Bimber, "An autonomous drone for search and rescue in forests using airborne optical sectioning," Science Robotics, vol. 6, no. 55, p. eabg1188, 2021.
[11] F. Rodriguez-Puerta, E. Gomez-Garcia, S. Martin-Garcia, F. Perez-Rodriguez, and E. Prada, "Uav-based lidar scanning for individual tree detection and height measurement in young forest permanent trials," Remote Sensing, vol. 14, no. 1, p. 170, 2022.
[12] J. N. Hayton, T. Barros, C. Premebida, M. J. Coombes, and U. J. Nunes, “Cnn-based human detection using a 3d lidar onboard a uav,” in 2020 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC), pp. 312–318, Ponta Delgada, Portugal, 2020.
[13] O. Risdal and L. Gustavsen, “LiDAR from drones employed for mapping archaeology – potential, benefits and challenges,” Archaeological Prospection, vol. 25, no. 4, pp. 329–338, 2018.
[14] K.-W. Chiang, G.-J. Tsai, Y.-H. Li, and N. El-Sheimy, “Development of lidar-based uav system for environment reconstruction,” IEEE Geoscience and Remote Sensing Letters, vol. 14, no. 10, pp. 1790–1794, 2017.
[15] S. Palm, R. Sommer, D. Janssen, A. Tessmann, and U. Stilla, “Airborne circular W-band SAR for multiple aspect urban site
monitoring,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 6996–7016, 2019.

[16] S. Palm and U. Stilla, “3-d point cloud generation from airborne single-pass and single-channel circular SAR data,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 10, pp. 8398–8417, 2020.

[17] S. Kim, J. Yu, S.-Y. Jeon, A. Dewantari, and M.-H. Ka, “Signal processing for a multiple-input, multiple-output (MIMO) video synthetic aperture radar (SAR) with beat frequency division frequency-modulated continuous wave (FMCW),” *Remote Sensing*, vol. 9, no. 5, p. 491, 2017.

[18] J. Svedin, A. Bernland, A. Gustafsson, E. Claar, and J. Luong, “Small UAV-based SAR system using low-cost radar, position, and attitude sensors with onboard imaging capability,” *International Journal of Microwave and Wireless Technologies*, vol. 13, no. 6, pp. 602–613, 2021.

[19] A. P. Pentland, “A new sense for depth of field,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-9, no. 4, pp. 523–531, 1987.

[20] V. Vaish, B. Wilburn, N. Joshi, and M. Levoy, “Using plane + parallax for calibrating dense camera arrays,” in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2004. CVPR 2004, vol. 1, Washington, DC, USA, 2004.

[21] V. Vaish, M. Levoy, R. Szeliski, C. L. Zitnick, and S. B. Kang, “Reconstructing occluded surfaces using synthetic apertures: Stereo, focus and robust measures,” in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (CVPR’06), IEEE, vol. 2, pp. 2331–2338, New York, NY, USA, 2006.

[22] H. Zhang, X. Jin, and Q. Dai, “Synthetic aperture based on plenoptic camera for seeing through occlusions,” in *Pacific Rim Conference on Multimedia*, pp. 158–167, Springer, 2018.

[23] T. Yang, W. Ma, S. Wang, J. Li, J. Yu, and Y. Zhang, “Kinect based real-time synthetic aperture imaging through occlusion,” *Multimedia Tools and Applications*, vol. 75, no. 12, pp. 6925–6943, 2016.

[24] N. Joshi, S. Avidan, W. Matusik, and D. J. Kriegman, “Synthetic aperture tracking: tracking through occlusions,” in *2007 IEEE 11th International Conference on Computer Vision*, pp. 1–8, Rio de Janeiro, Brazil, October 2007.

[25] Z. Pei, Y. Zhang, X. Chen, and Y.-H. Yang, “Synthetic aperture imaging using pixel labeling via energy minimization,” *Pattern Recognition*, vol. 46, no. 1, pp. 174–187, 2013.

[26] T. Yang, Y. Zhang, J. Yu et al., “All-in-focus synthetic aperture imaging,” in *Computer Vision – ECCV 2014*, pp. 1–15, Springer International Publishing, Cham, 2014.

[27] I. S. Reed and X. Yu, “Adaptive multiple-band cfar detection of an optical pattern with unknown spectral distribution,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 10, pp. 1760–1770, 1990.

[28] T. Ehret, A. Davy, J.-M. Morel, and M. Dabracco, “Image anomalies: a review and synthesis of detection methods,” *Journal of Mathematical Imaging and Vision*, vol. 61, no. 5, pp. 710–743, 2019.

[29] B. S. Morse, D. Thornton, and M. A. Goodrich, “Color anomaly detection and suggestion for wilderness search and rescue,” in *2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 455–462, Boston, MA, USA, 2012.