Compact multispectral multi-camera imaging system for small UAVs

Hans Erling Torkildsen, Trym Haavardsholm, Thomas Opsahl, Urmila Datta, Atle Skaugen, Torbjørn Skauli*
Norwegian defence research establishment, P. O. Box 25, 2027 Kjeller, Norway

ABSTRACT

Cameras with filters in the focal plane provide the most compact solution for multispectral imaging. A small UAV can carry multiple such cameras, providing large area coverage rate at high spatial resolution. We investigate a camera concept where a patterned bandpass filter with six bands provides multiple interspersed recordings of all bands, enabling consistency checks for improved spectral integrity. A compact sensor payload has been built with multiple cameras and a data acquisition computer. Recorded imagery demonstrates the potential for large area coverage with good spectral integrity.

Keywords: Multispectral imaging, remote sensing, target detection, drone, filter array, UAV

1. INTRODUCTION

Multi- and hyperspectral imaging is an attractive sensing modality for small unmanned aerial vehicles (UAV) in numerous civilian and military applications, including target detection and environmental mapping. In most such applications, a large area coverage rate is desirable, to minimize flight time and cost.

Most remote sensing applications of spectral imaging can accept a scanning imaging modality, where one dimension of the image is built by the platform movement. For daylight-based imaging, the diffraction limit allows even a small camera to have a resolution such that the hemisphere underneath a UAV would be several tens of thousands of pixels across. Thus to make good use of the view from an elevated platform, it is necessary to have multiple spectral cameras with adjoining fields of view (FOV). On a small UAV, each individual camera then needs to be small. Hyperspectral "imaging spectrometer" cameras are relatively bulky compared to a conventional camera, and have track widths of at most 2000 pixels, usually significantly less. Multispectral imaging offers more compact technology options and may therefore be preferred over hyperspectral imaging, as a compromise to achieve a large area coverage rate from a small UAV. This paper explores a novel technology option for a multispectral multi-camera sensor package.

Cameras employing patterned filters on a 2D photodetector array, such as in conventional color cameras, is the most compact practical architecture for spectral imaging. (Arrays with stacked multiband detector elements, or other wavelength-separation methods [1,2,3], are in principle better, but difficult to realize in practice for more than 3 bands.) On a moving platform, multiple filters can be scanned over the scene by the platform movement. By arranging rows of detector elements with different filters across the flight direction, the spectral bands can be sampled successively along the scan. This scan geometry is widely used in satellite remote sensing, where a spectral image can be readily reconstructed from the known orbital motion. This form of filtered multispectral line imaging is more difficult on a small UAV where the flight is somewhat irregular, and where resolved 3D structures in the scene cause parallax effects that lead to inconsistencies in the reconstructed image. As discussed in Ref. [4], a multispectral image with improved spectral integrity can be recorded using the camera concept shown in Figure 1 c). Here we present initial results from a UAV sensor package employing this camera concept.

*torbjorn.skauli@ffi.no

Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXII
edited by Miguel Velez-Reyes, David W. Messinger, Proc. of SPIE Vol. 9840, 98401U
 · doi: 10.1117/12.2224495

2. IMAGING CONCEPT

In spectral imaging, coregistration of bands is crucial for the integrity of the spectral signal [5,6]. Figure 1 illustrates different concepts for spectral imaging from an airborne platform using filters in the focal plane of a conventional camera. In a), a linearly variable filter (LVF) is placed in the focal plane, as indicated in the upper part of the figure, enabling hyperspectral imagery to be recorded [7][8]. A spectral image is assembled from multiple raw image frames recorded at different positions in the flight path. However, as illustrated in the lower parts of Figure 1 a), 3D structures in the scene may cause parallax effects and coregistration errors in the reconstructed spectra: The building B obstructs the view to point P for only the blue band, and the reconstructed spectrum becomes an unphysical mixture of the spectra of ground and roof. If the 3D structure and imaging geometry is known, pixel spectra with such artifacts can be flagged with a veto mask, but the spectral information can not be recovered. Figure 1 b), illustrates a scheme for multispectral imaging where the focal plane contains strips of discrete spectral filters, one for each band. The filters can be arranged to span a narrower range of viewing angles in the flight direction than in case a). Parallax effects are then reduced, but not eliminated. A further refinement, used in this paper, is shown in Figure 1 c). Here the set of spectral filters are repeated multiple times. This provides multiple viewing angles for each band. If the 3D geometry is known, it is then possible to reconstruct valid spectra for locations in the scene that are visible through only a subset of filters for a given band[10]. Furthermore, it is possible to check spectral data for consistency and integrity by comparing different measurements of the same band in a given location, even if the 3D geometry is not accurately known. An advantage of the concepts in Figure 1 b) and c) is that a shorter length of the filter in the flight direction frees up area in the focal plane which can be used for conventional imaging. This imagery can in turn be used to estimate the imaging geometry and support reconstruction of the spectral image [9].

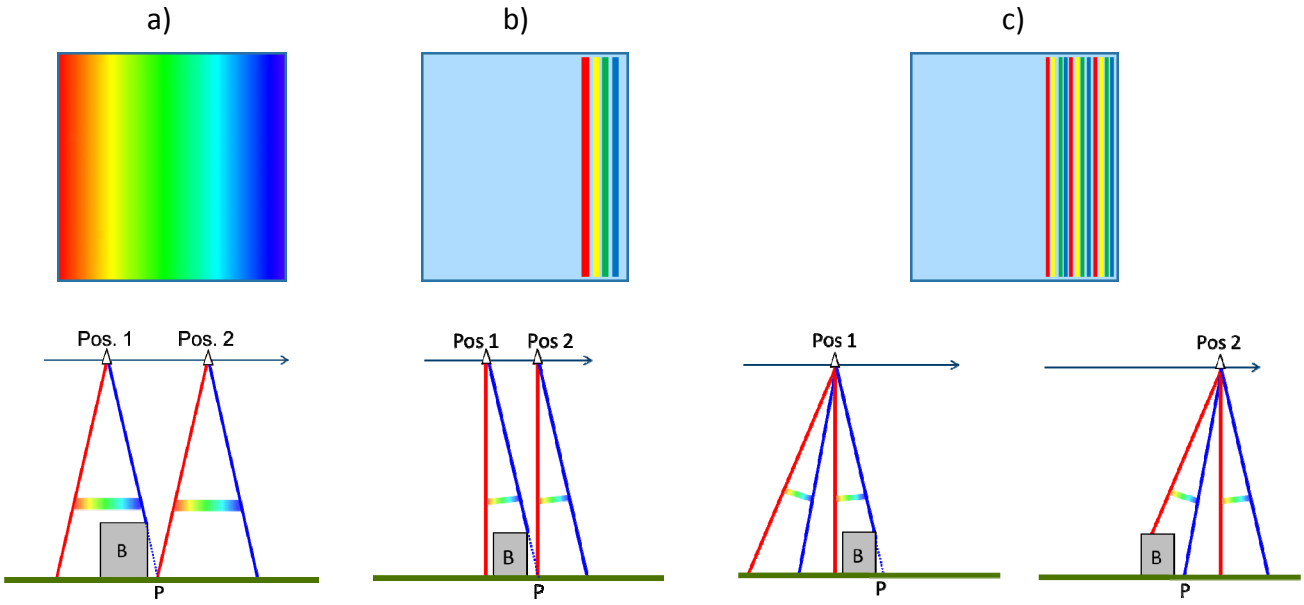

Figure 1. Different concepts for spectral imaging using a filter in the focal plane (top) in a camera viewing the ground from a UAV (bottom). Colored lines indicate the first and last band. a) Linear variable filter covering the full focal plane, enabling hyperspectral imaging, but with risk of spectral artifacts due to 3D structure. b) Multispectral imaging with strips of bandpass filters, providing reduced risk of artifacts and a part of the focal plane available for conventional imaging. c) Repeating pattern of bandpass filters, enabling consistency checking and more robust reconstruction of spectral images. Here, the second sampling intercepts point p in the blue band. (Lines of sight are drawn around nadir in all three cases for clarity, although in b) and c) all bands would tend to look forward or backward for a focal plane layout as shown.)

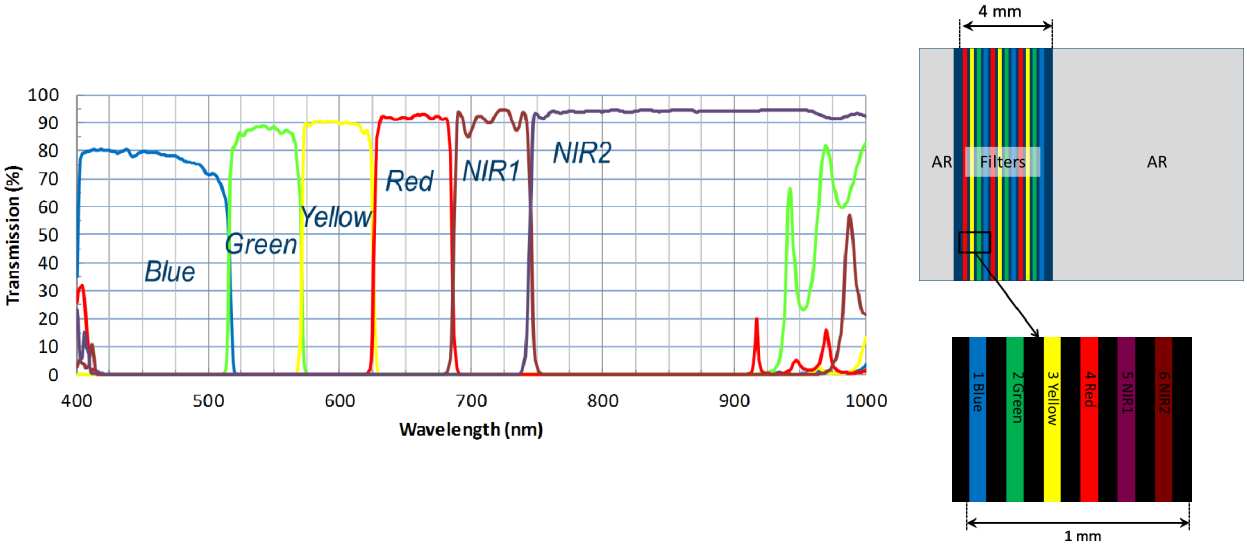

Figure 2. Left: Measured transmission spectra for each of the six different filter bands. Spectral leakage at the ends of the spectral range is eliminated with an external 450-875 nm bandpass filter. Right: Layout of the filter array.

Figure 3. Raw image from a camera, with the actual filter array seen on the right. (The brightness has been adjusted in the filter region for visual clarity.)

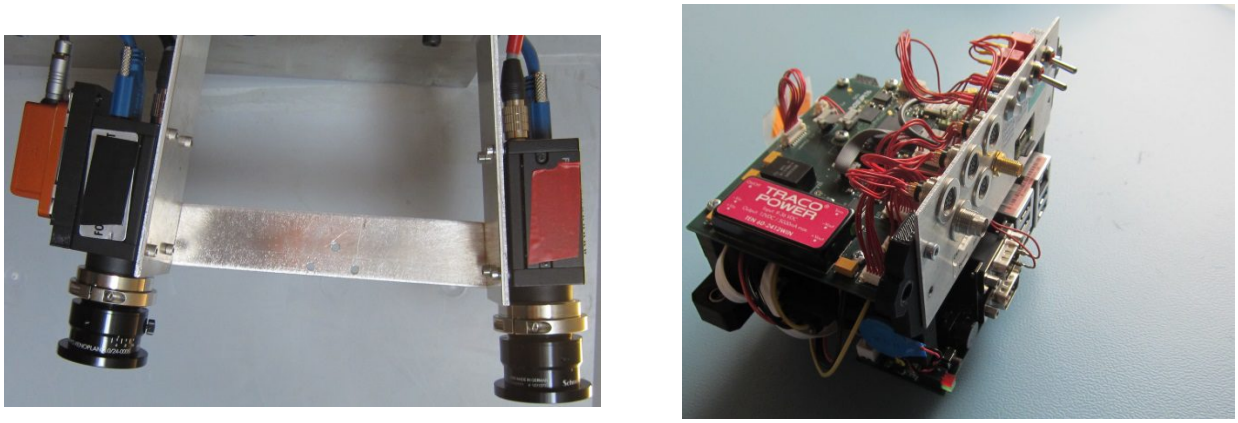

Figure 4. Main parts of the payload package. Left: Cameras mounted on a bracket to give partially overlapping fields of view across track. An IMU is mounted on the leftmost camera. Right: Control computer with custom board for logging navigation data.

3. UAV SENSOR PACKAGE

3.1 Cameras

Our UAV sensor package uses the same filter array as shown in Refs. [4] and [10]. The layout is shown in Figure 2, with an array of strip-shaped filters laid out across the FOV. The filter array has six spectral bands in the visible and near infrared, roughly matching bands used on remote sensing satellites. The set of six bands are repeated four times across the filter array, which extends a total of 4 mm in the scan direction. Each of the 24 individual filters is about 10 pixels across. The filter array is mounted in close proximity to the image sensor, with the individual filter strips oriented parallel to the short edge of the image sensor. Thus the scan direction is nominally parallel to the long edge of the image sensor. Clear areas outside the array are used to record panchromatic images.

The system employs cameras based on a Sony IMX174 monochrome CMOS image sensor with 1920x1200 pixels and 5.86 µm pixel pitch. For the tests presented here we used Schneider Apo-Xenoplan 2.0/24 lenses, giving a FOV of 26x17 degrees for a single camera. The cameras have USB 3.0 interface and can record up to 163 frames per second. Figure 3 shows a single raw image recorded with one of the cameras. The cross-track FOV for a single camera is 17 degrees. Here we have mounted the cameras with about 25% overlap to enable good image-based coregistration, so that the total FOV is 25.5 degrees.

The results presented here were obtained using a payload package with two cameras, mounted as shown in Figure 4 so that the fields of view of the two cameras point nominally to either side of nadir with a small overlap.

3.2 Control, navigation and data acquisition

A Pico 880 microcomputer from Axiomtek is connected to the cameras using the USB 3.0 interface. The computer controls the camera exposure adaptively and stores the image data using two solid state drives. The system also contains a GPS receiver and a MEMS inertial motion unit (IMU), the latter mounted directly on one of the cameras. These navigation sensors are read out by a custom FPGA-based navigation logging board which also triggers the cameras. Navigation data are time stamped and stored on an SD card together with the time stamps of the trigger pulses, to be used for navigation post-processing [11] and image georeferencing.

The cameras are operated in a high dynamic range mode, alternating between a long and a short integration time. The short integration time of each camera is continuously updated to best exploit the dynamic range of the panchromatic part of the image sensor while avoiding too many saturated pixels. The long integration is set to be 16 times longer, up to a limit set by the camera or flying speed. The regulation loop calculates a target value for exposure time from each frame so that the 98.5 percentile of the pixel values is at 80 percent of full scale. The mean of the last calculated N consecutive target values are then used as the short integration time, where N is the number of frames recorded during the time it takes to fly a distance equal to the width of the unfiltered region, projected on the ground, at nominal flying speed.

Due to shortcomings of the current version of the camera firmware, the frame rate is limited to 86 Hz when the cameras are operated in the high dynamic range mode. To ensure that every pulse from the FPGA triggers an image capture, the frame rate of the system was set to 80 Hz. Using 8 bits to represent the DN value of each pixel, the Pico 880 computer can control and store all the images from up to 4 cameras. Alternatively, if high dynamic range is not used, the system can control and store images from two cameras triggered at the maximum frame rate of 163 images per second.

The complete mass of the key components of the data acquisition and navigational system is 170 grams. The mass of each of the cameras included in the system is also 170 grams. Thus, the total mass of a two camera system, using off the shelf components, is 510 grams plus cables, connectors, power supply parts and material for structural assembly. The total mass of a payload can thus be well below 1 kg.

3.3 Test data

The camera payload has been integrated for use on a Freefly CineStar 8 octocopter. Single cameras have been flown successfully several times. Unfortunately, a string of glitches in electronics and logistics has kept us from actually flying the multi-camera setup on the octocopter in time for the deadline for this paper. We therefore resort to images recorded with the camera payload from a rooftop about 10 m above ground. The cameras were mounted at about 45-degree angle off nadir and "flown" on a trolley that was pushed along manually, creating some irregularity in the movement. Taking the distance to the scene to be about 15 m, the pixel size in the scene is about 4 mm. The speed of the "flight" was about 0.2 m/s. This slow speed allowed us to emulate different scan speeds by subsampling the stream of images.

The imaged scene is seen in Figure 4. This is FFI's lawn in March, partially covered in snow, with a couple of painted plates as test objects. We also see a road with cars, behind a fence. This scene actually illustrates well the features of the camera concept, as discussed in the following.

4. IMAGE PROCESSING

To reconstruct a spectral image from a set of raw images, we need to know the corresponding relative camera poses and the scene geometry well enough to achieve sub-pixel co-registration between the images. To accomplish this level of camera navigation accuracy, we use the panchromatic parts of the raw images to perform image-based camera pose estimation with Structure-from-Motion (SfM) using the generic 3D reconstruction software RealityCapture. This method also provides us with a sparse 3D point cloud of the scene, which we use to estimate a dominating ground plane with RANSAC-based plane estimation. We define a reference view, from a point far behind the mean camera pose, as the view in which the output pixels are defined. For the results shown here, the height of the output image was chosen at 2400 pixels so that output pixel size is close to the camera resolution.

All raw images can be projected onto the ground plane and then to the reference view with a projective transformation known as the homography induced by the plane [12]. With the correct relative camera poses, all pixels from the raw images that lie on the ground plane will be mapped to the same corresponding pixel in the reference view, while off-plane pixels will be dispersed along lines described by the pair-wise epipolar geometry between the raw images and the reference view. We will here use the ground plane as our scene model, and use the appearance of spectral inconsistencies between corresponding filter samples to detect areas that are inconsistent with this simple model.

By transforming and fusing each filter strip from each raw image into the reference view, it is possible to reconstruct four 6-band spectral images, corresponding to each group of 6 filters in the repeating filter array. This intermediate image product can be checked for consistency to try to compensate for geometric errors and produce a valid image [10], or to produce a veto mask to flag invalid pixels. Finally, valid samples from the same band are averaged together in each pixel to form the output image. Here we apply only a simplistic consistency check by finding the spectral angle between different intermediate images in each pixel and then vetoing pixels for which the spectral angle exceeds a threshold.

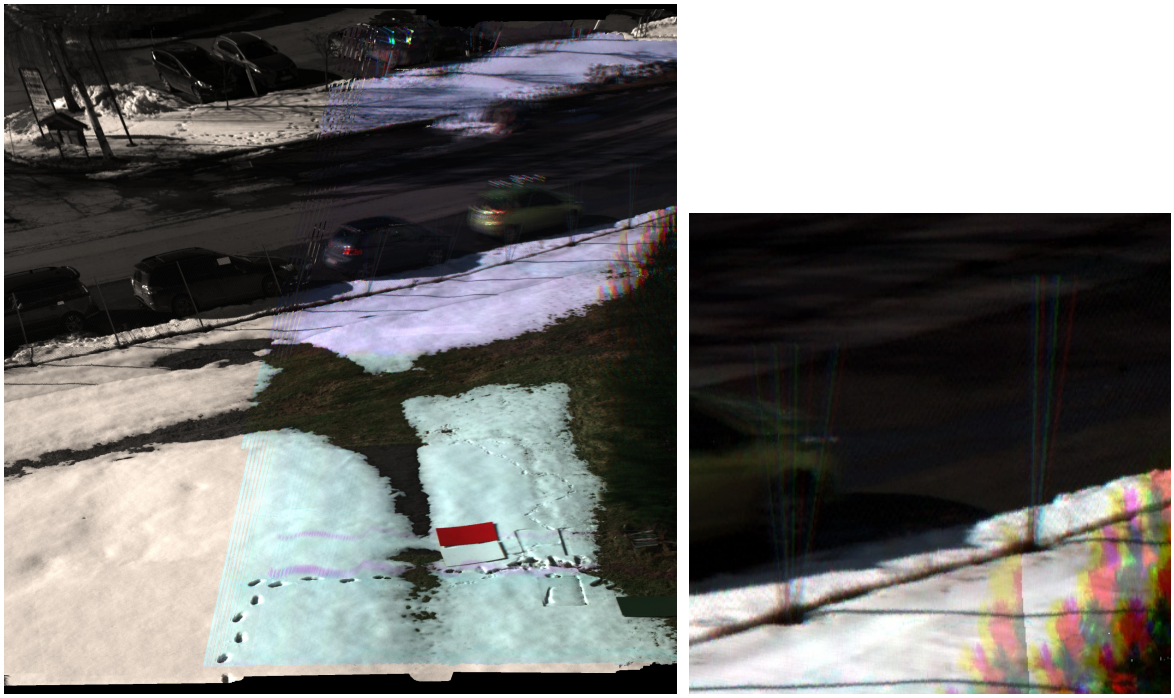

Figure 5. RGB color image reconstructed from the 6-band data, assuming the 3D geometry to be a plane at ground level. The reconstructed spectral image is overlaid on a monochrome image stitched together from multiple raw image frames. Coregistration errors are apparent in areas that deviate from the assumed 3D model, as seen in the magnified part on the right. See discussion in the text.

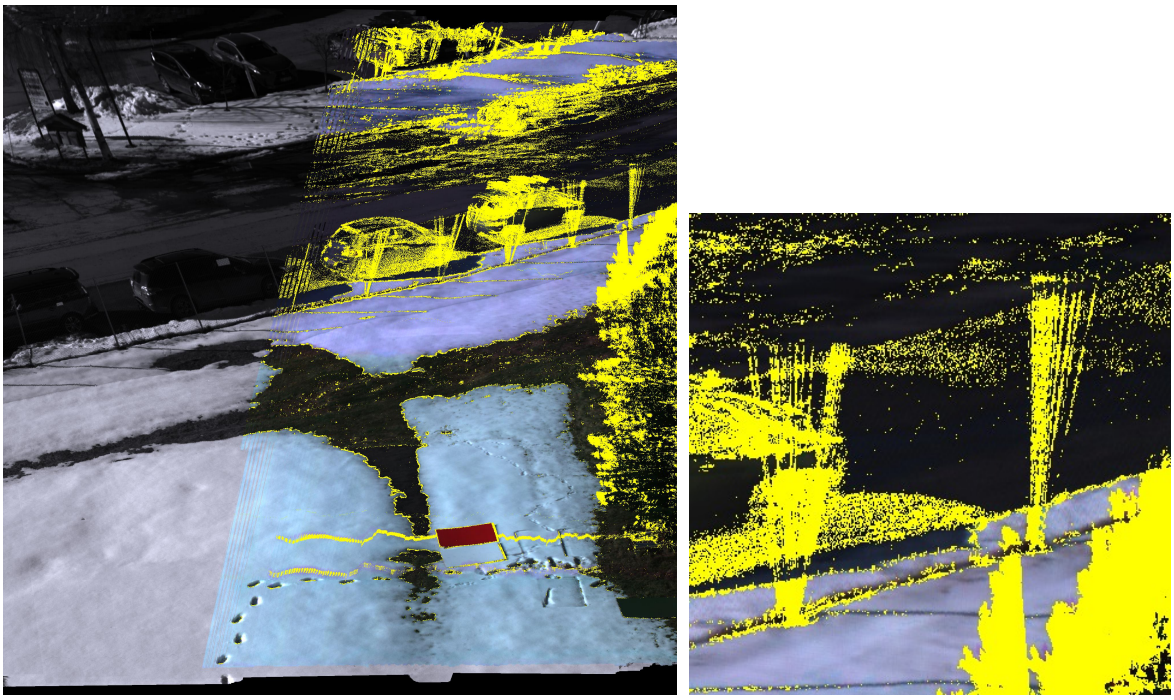

Figure 6. Images from Fig. 5 overlaid with a veto mask (yellow) derived from consistency checking on the multiple spectral samples in each band. Coregistration errors due to geometric effects are detected. See discussion in the text.

5. SAMPLE RESULTS

A sample image from the test scene is shown in Figs. 5 and 6. Consider first the reconstructed image in Figure 5 without veto mask. We see that colors and contrasts on the ground are visually well reconstructed, since these parts of the image conform to the assumption that the 3D geometry is a plane. It is also clear that images from the two cameras join smoothly together without any apparent offset and with very little radiometric discontinuity.

Away from the ground plane, the fence posts are reconstructed as four replicated color-distorted images corresponding to the four repetitions of the filters. Note that the four replicas are converging at the ground. Similarly, the tree on the right is reconstructed with progressively larger color artifacts upward from the ground. These image artifacts illustrate the risk of significant coregistration errors that is inherent in filter-based spectral imaging techniques.

Now consider the image with veto mask applied, Figure 6. Clearly, the spectra affected by geometry errors are detected by the inconsistency between repeated samplings. Note that uniform areas away from the assumed plane, such as the side of a car, are not flagged as inconsistent. These data are arguably valid, even with their large coregistration error.

A further notable effect in Figure 6 is the veto masking of many transitions on the ground, such as the edges of the snow. High-contrast edges will be sensitive to small inaccuracies in the relative positioning of images. High-contrast edges are sensitive to residual coregistration errors, this may be the main reason for veto masking along the edges of the snow.

In the lower part, two irregular lines of vetoed pixels run across the image. These are due to image defects caused by contamination of the filter on one of the cameras. Such defects can be avoided relatively easily, as evidenced by the other camera, but the result serves to illustrate that the consistency checking can be seen to provide an internal self-check of the camera. These lines of vetoed pixels also serve to illustrate the scan path across the scene.

In the left part of the reconstructed area, there are gaps in the reconstructed spectral image. This is due to a gap in the regions covered by a given filter strip from one frame to the next in a period of too fast scanning. Thus rather than generating invalid data, a too fast scan simply leads to gaps in the image.

6. DISCUSSION AND CONCLUSIONS

Lacking a proper flight test, it is of interest to estimate the area coverage rate that is feasible with this camera concept. Many configurations are possible, but consider first the actual UAV payload presented here. Assume a nominal ground sampling distance (GSD) of 10 cm, corresponding to an altitude of 400 m above ground for the optics used here. The ground track width can be on the order of 2400 pixels, or 240 m. The camera sets an upper limit on flying speed: If each individual filter is 10 pixels wide, a frame rate of 80 Hz permits a flying speed of up to 800 pixels per second, or 80 m/s, provided that the integration time is kept on the order of 1 ms to avoid motion blur. A more realistic flying speed for a fast medium-sized fixed-wing UAV is about 30 m/s (108 km/h). This leaves some margin for irregular movement of the FOV during flight. A rough estimate of area coverage rate is then 240×30 m^2/s or 26 km^2/h at 10 cm GSD for this particular payload. This coverage rate could readily be doubled by adding two more cameras. This would be well within the data bandwidth of the computer used here, and the overall system would remain compact.

It must be emphasized that the image reconstruction method applied here is simplistic. Although the assumption of a flat-earth scene serves to illustrate important points in this paper, an obvious improvement is to use the panchromatic imagery to reconstruct a full 3D model of the scene. Furthermore, many different strategies are possible for the spectral consistency check, which should preferably include physical estimates of noise [13].

It is worth noting that the resolution and sampling pattern of the reconstructed spectral image can be chosen freely. The results here are based on a reconstructed image with approximately the same resolution as the raw images. The high output resolution will inevitably make it more difficult to achieve good coregistration, and it may be reasonable to reconstruct the spectral image with a pixel size about double that of the raw images. This would tend to produce a smaller fraction of vetoed pixels than shown here. For the panchromatic image, the full resolution of the raw images can be maintained.

A major question for our camera concept is whether it is possible to do spectral reconstruction and target detection onboard the resource-limited platforms that are otherwise well suited to carry the system. Work on a more optimized image processing chain is ongoing and will be published elsewhere.

It must be mentioned that the current filter arrangement with 6 bands is an obvious sacrifice of capability compared to hyperspectral imaging. Compared to 3-band RGB imaging, on the other hand, a 6-band camera can provide significant capability for automated spectral detection and discrimination. Band selection studies suggest that a somewhat higher band count is likely to provide a clear benefit [14,15,16], indicating a path for future upgrades to our system.

In summary, we have demonstrated a concept for wide-area spectral remote sensing using a compact and low-cost payload suitable for a small to medium-sized UAV. Multiple sampling of bands provides for consistency checking and thus a spectral image with strongly reduced spectral artifacts compared to other filter-based imaging concepts. The results demonstrate a clear potential for use in UAV-based mapping and reconnaissance.

REFERENCES

[1] P. M. Hubel, "Foveon Technology and the Changing Landscape of Digital Camera Technology", Proc. 13th IS&T Color Imaging Conference, Scottsdale, AZ, 2005, 314 - 317

[2] G. H. Lee, K. S. Kim, Y. W. Jin, "Image sensor and electronic device including the same", US Pat. Appl. 20150041940 (2014)

[3] E. Laux, C. Genet, T. Skauli, T. W. Ebbesen, "Plasmonic photon sorters for spectral and polarimetric imaging," Nature Photonics **2**(3), 161 (2008)

[4] T. Skauli, H. E. Torkildsen, S. Nicolas, T. Opsahl, T. Haavardsholm, I. Kåsen and A. Rognmo, "Compact camera for multispectral and conventional imaging based on patterned filters," Appl. Opt. **53** (13), ISA1-ISA2 (2014)

[5] P. Mouroulis, R. O. Green, and T. G. Chrien, "Design of pushbroom imaging spectrometers for optimum recovery of spectroscopic and spatial information," Appl. Opt. 39(13), 2210–2220 (2000)

[6] T. Skauli, "An upper-bound metric for characterizing spectral and spatial coregistration errors in spectral imaging," Opt. Expr. 20, 918-933 (2012)

[7] A. M. Mika, "Linear-wedge spectrometer," Proc. SPIE 1298, 127–131 (1990)

[8] Tack, N., Lambrechts, A., Soussan, P., Haspeslagh, L., "A compact high-speed and low-cost hyperspectral imager," Proc. SPIE 8266, 82660Q-1 (2012)

[9] J. Biesemans, B. Delaure, B. Michiels, "Geometric referencing of multi-spectral data," US patent 20120257047 (2012)

[10] H. E. Torkildsen, T. Opsahl, T. V. Haavardsholm, S. Nicolas, T. Skauli, "Characterization and calibration of a compact 6-band multifunctional camera based on patterned spectral filters in the focal plane," Proc. SPIE **9088**, 908819-1 (2014)

[11] K. Gade, " NAVLAB, a Generic Simulation and Post-processing Tool for Navigation," European Journal of Navigation **2**, 1 (2004)

[12] R. Hartley, A. Zisserman, "Multiple View Geometry in Computer Vision", Cambridge University Press (2003)

[13] T. Skauli, "Sensor noise informed representation of hyperspectral data, with benefits for image storage and processing," Optics Express **19**, 13031 (2011)

[14] S. S. Shen, E. M. Bassett, "Information Theory Based Band Selection and Utility Evaluation for Reflective Spectral Systems", Proc. SPIE **4725**, 18 (2002)

[15] P. L. Carmona, A. Martínez-Usó, J. M. Sotoca, F. Pla, P. García-Sevilla, "Band selection in spectral imaging for classification and regression tasks using information theoretic measures," Proc. Euro-American Workshop on Information Optics, 2011, p.1

[16] I. Kåsen, A. Rødningsby, T. V. Haavardsholm, T. Skauli, "Band selection for hyperspectral target-detection based on a multinormal mixture anomaly detection algorithm," Proc. SPIE **6966**, 96606 (2008)