PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Estimating the pixel footprint distribution for image fusion by ray tracing lines of sight in a Monte Carlo scheme

T. Opsahl, T. V. Haavardsholm

T. Opsahl, T. V. Haavardsholm, "Estimating the pixel footprint distribution for image fusion by ray tracing lines of sight in a Monte Carlo scheme," Proc. SPIE 8743, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIX, 87431U (18 May 2013); doi: 10.1117/12.2015746

Event: SPIE Defense, Security, and Sensing, 2013, Baltimore, Maryland, United States

Estimating the pixel footprint distribution for image fusion by ray tracing lines of sight in a Monte Carlo scheme

T. Opsahl*[a], T. V. Haavardsholm[a]
[a]Norwegian Defence Research Establishment (FFI), Instituttveien 20, 2027 Kjeller, Norway

ABSTRACT

Images from airborne cameras can be a valuable resource for data fusion, but this typically requires them to be georeferenced. This usually implies that the information of every pixel should be accompanied by a single geographical position describing where the center of the pixel is located in the scene. This geospatial information is well suited for tasks like target positioning and orthorectification. But when it comes to fusion, a detailed description of the area on the ground contributing to the pixel signal would be preferable over a single position. In this paper we present a method for estimating these regions. Simple Monte Carlo simulations are used to combine the influences of the main geometrical aspects of the imaging process, such as the point spread function, the camera's motion and the topography in the scene. Since estimates of the camera motion are uncertain to some degree, this is incorporated in the simulations as well. For every simulation, a pixel's sampling point in the scene is estimated by intersecting a randomly sampled line of sight with a 3D-model of the scene. Based on the results of numerous simulations, the pixel's sampling region can be represented by a suitable probability distribution. This will be referred to as the pixel's footprint distribution (PFD). We present results for high resolution hyperspectral pushbroom images of an urban scene.

Keywords: Georeferencing, hyperspectral, ray tracing, fusion

1. INTRODUCTION

Georeferenced images captured from an airborne platform can be a great resource in aerial surveillance. The pixel signal is well suited for making detections, while the geo-spatial information allows the detections to be positioned and opens up for fusion with other georeferenced data. But attempting to fuse an image with other data immediately raises the question: Which pixels contain signal from a given point in the scene? Although rough estimates can be done purely based on proximity, a well-founded answer requires more information than the typical pixel position from the georeference.

Instead of just knowing the center position of a pixel, it would be useful to have a more detailed description of how the scene has contributed to the pixel signal. This is tightly coupled to one of the imaging system's fundamental properties, namely the Point Spread Function (PSF), which describes the imaging system's spatial response to a point source. The PSF can also be viewed as an angular distribution of the response of a single detector element (see Figure 1).

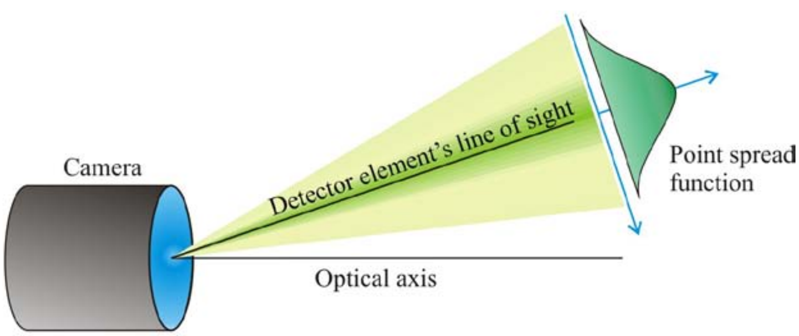

Figure 1. The point spread function (PSF) of an imaging system, can be viewed as an angular distribution of the response of a single detector element.

*Thomas-Olsvik.Opsahl@ffi.no

Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIX,
edited by Sylvia S. Shen, Paul E. Lewis, Proc. of SPIE Vol. 8743, 87431U · © 2013 SPIE
CCC code: 0277-786X/13/$18 · doi: 10.1117/12.2015746

Projecting the PSF of a detector element into the scene would reveal its spatial distribution of responsivity at a given point in time, essentially describing the actual blurring of a pixel. The pixel signal is recorded over a finite integration time during which the camera typically will be moving. Hence, the PSF, the camera's scan motion during integration time and the scene's topography together determine the spatial distribution of responsivity for a pixel. Any of these can dominate over the others under the right circumstances. The topography will typically be dominant for pixels recorded by a slant viewing detector element scanning a significant vertical transition in the scene, like the roof ledge of a tall building. The scan motion during integration time can be dominating for pixels experiencing a rapid scan speed which can occur during challenging flying conditions especially if the camera is unstabilized. By the same reasoning, some pixels might experience a very low scan speed, thus making the PSF dominant. In the following we will use the term Pixel Footprint Distribution (PFD) to refer to the distribution of responsivity projected into the scene.

The PFD describes how the pixel signal was sampled in the scene. But the geometrical relation between the camera and the scene will always involve uncertainties. So in order for the PFD to be an attractive pixel property for fusion purposes, it should reflect these uncertainties. In this paper we present a method for estimating the PFD for individual pixels. The estimation is done by conducting many simple Monte Carlo simulations to produce a discrete sampling of the PFD. The PFD is then represented by a small set of parameters determined by fitting these points to a simpler continuous probability distribution like a bivariate Gaussian. We present example results for two high resolution hyperspectral pushbroom images covering an urban scenario described by a high resolution 3D-model.

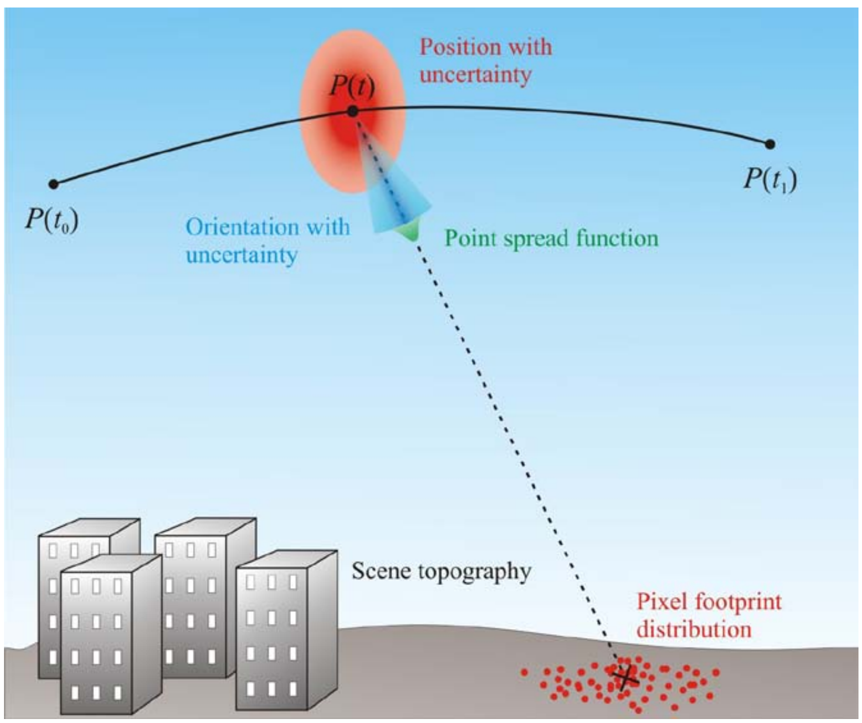

Figure 2. The pixel footprint distribution is determined by the point spread function, the camera's motion during integration time $[t_0, t_1]$ and the topography of the scene.

2. METHODOLOGY

As mentioned above, the PFD for a single pixel in the image depends on several geometrical aspects of the imaging process. An overview is presented in Figure 2. A pixel's signal is recorded by a detector element whose angular responsivity is described by a PSF. The camera is typically moving during the recording, but this motion can only be known up to some uncertainty. A high-end INS can provide accurate navigation with estimated uncertainties for the camera. The PSF and camera motion together determine how the pixel relates to the scene. But in order to actually describe this relation in terms of a PFD, the scene's topography must also be taken into account. By conducting simple Monte Carlo simulations, the combined effect of all these factors can be estimated.

A Monte Carlo simulation for a single pixel is conducted as follows:

1. Sample a random time t from a uniform distribution over the pixel's integration time.
2. Sample a position and orientation for the camera at the time t. The INS typically provides position and orientation with their standard deviations at a high rate. So their values at the time t can be found by interpolation. The sampling can then be performed under the assumption that each degree of freedom has a Gaussian distribution.
3. Sample offset angles from the PSF.
4. Determine a line of sight using the sampled position, orientation and offset angles.
5. Intersect the line of sight with a 3D-model of the scene to reveal a position which is a sampling of the pixel's PFD. An efficient ray tracing method is described in [1].

After conducting enough simulations, the resulting set of points will reflect the PSF, the motion with uncertainty and the scene's topography. By fitting these points to a suitable continuous probability distribution, one can represent the PFD by a small set of parameters.

The method is straightforward in principle, but can be computationally demanding. Estimating the PFD for 10M pixels by conducting 1000 Monte Carlo simulations requires that 10^{10} lines are intersected with the 3D-model of the scene. Given a high resolution 3D-model, this requires a massive computational effort, so an efficient ray tracer should be used to keep the run times at a reasonable level.

It should be noted that the method easily can be extended by taking into account other uncertainties than just those in the navigation. One could for example also include uncertainties in the camera's calibration or in the 3D-model of the scene. The framework will be the same, but the required computational effort will increase.

3. RESULTS AND DISCUSSION

As part of the EDA project "Detection in Urban scenario using Combined Airborne imaging Sensors" (DUCAS) a large measurement campaign was conducted in Zeebrugge, Belgium, during the summer of 2011. DUCAS was a collaboration of seven nations for investigating the potential benefit of combined high spatial and spectral resolution airborne imagery for defence applications in urban areas. Participating nations were: Sweden (FOI), Germany (Fraunhofer, IOSB), Italy (CISAM), France (ONERA), Belgium (RMA), Netherlands (TNO, NLR) and Norway (FFI). Here we use a LIDAR generated digital surface model (DSM) of the scene with 10cm ground sampling distance, as illustrated in Figure 3.

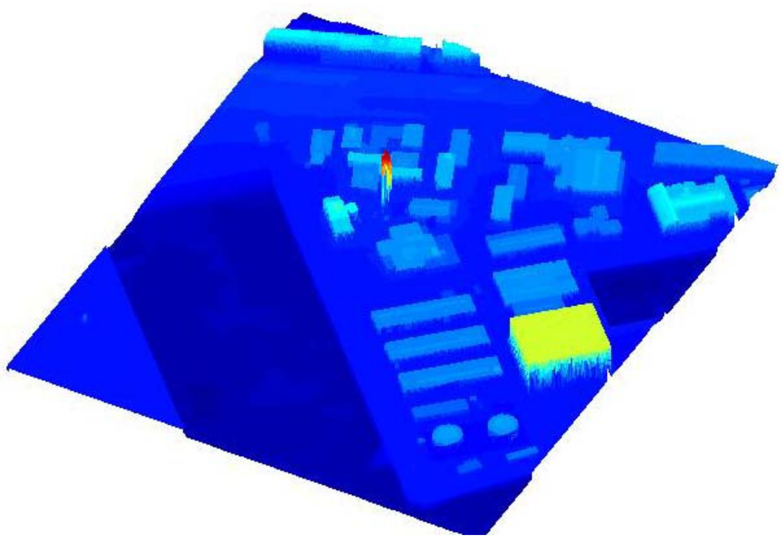

Figure 3. This 3D-model of the scene is based on the 10cm resolution LIDAR generated surface model produced by RMA for DUCAS.

The presented method was tested on two images captured by FFI's airborne hyperspectral target detection system [2] during the DUCAS campaign. They were captured only 5 minutes apart and from roughly the same altitude. To estimate the camera's position and orientation with uncertainties for each frame, the navigation software NavLab [3] was used to process data from the navigation sensors; a GPS receiver and a medium-performance IMU. The scene was represented by a mesh of triangles determined by splitting every grid element in the above mentioned DSM into two triangles. The efficient ray tracing algorithm described in [1] was used to solve the intersection problems. 1000 Monte Carlo simulations were conducted for two test images.

The PFD of each pixel was represented by the mean and the covariance of the sampled points' map coordinates. In order to get a visual presentation of the results, the Circular Error Probable (CEP) was approximated from the covariance. The CEP is defined as the radius of a circle containing 50% of the sampled points when centered about the mean. The test images and CEP results are presented in Figure 4.

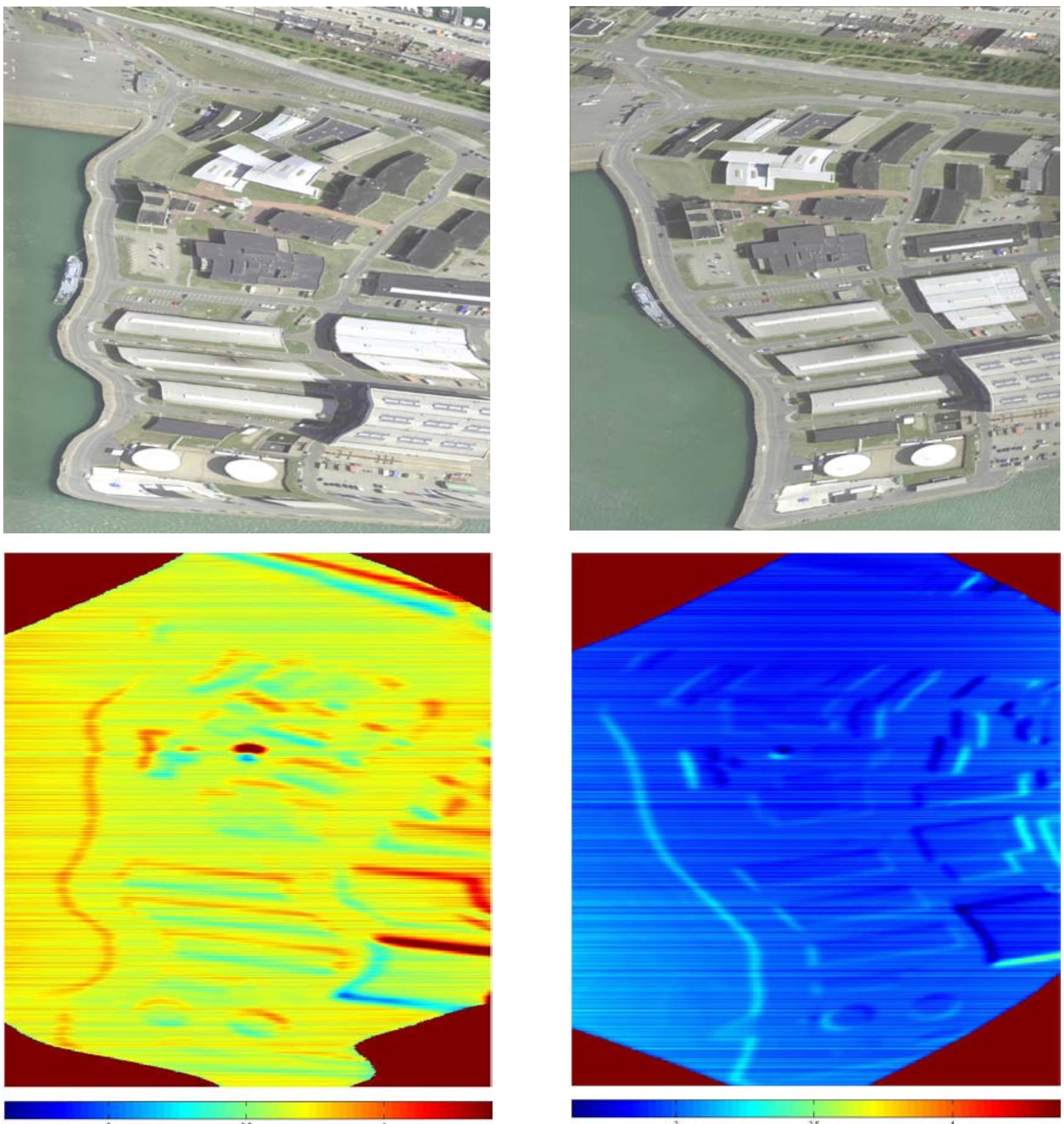

Figure 4. The top left and right images are RGB-representations of test image 1 and test image 2. Under each test image is the estimated Circular Error Probability based on 1000 simulations for each pixel. Pixels that did not intersect the DSM in at least 95% of the simulations have been given the maximal value by default.

There are several interesting observations to be made in these results. First of all, it is clear that the CEP value, and thus the PFD, varies significantly both within the images and from one image to another. It is also clear that it can change dramatically over a few pixels. The regions with the smallest and largest CEP values seem to be near significant vertical transitions in the scene, like the roof ledges of tall buildings, when the line of sight is sufficiently non-vertical. High values are found on the far side of the building where the PFD covers an area on the ground that did not contribute directly to the pixel signal. The low values are found on the near side of the building where much of the pixel signal is from the wall, restricting the PFD's horizontal extent. The visible striping effect is caused by oscillations observed in the pitch angle. The large systematic difference in CEP values between the two results is actually caused by a difference in the level of uncertainties in camera orientation. These uncertainties are dominating contributors in the CEP calculations and they are significantly larger for the first image than for the second. They vary slowly enough to not cause any features in the results, but they blur the features caused by the other contributions.

As mentioned above, the PFD should be a convenient pixel property for fusion. In order to fuse an image with another data set collected from the same scene, a relation between the two data sets must be established. Although it is one of the basic assumptions in several simple fusion algorithms, like in hyperspectral change detection; from a general perspective, this relation will almost never be one-to-one. In general, every pixel can correspond to several data elements and vice versa. The PFD opens up for estimating a detailed one-to-many relation. The level of detail is only limited by the complexity of the continuous probability distribution used to represent the PFD. Figure 5 illustrates how the CEP of a pixel in one image can be compared with the CEP of pixels in another image to determine a set of neighbors.

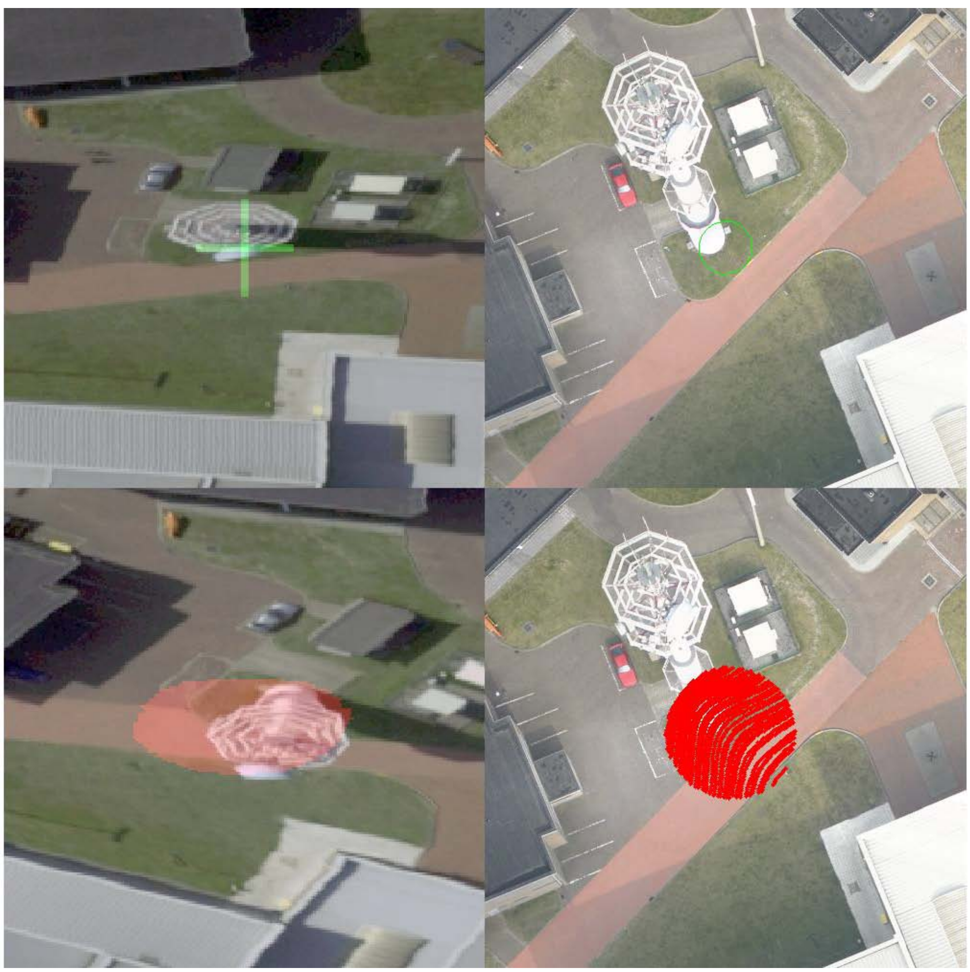

Figure 5. The top left image highlights a chosen pixel in test image 2. The top right image shows the CEP region for this pixel on top of an orthophoto. The bottom right image illustrates the center position of all pixels in test image 1 with an overlapping CEP region. The bottom left image highlights the neighboring pixels in test image 2. This illustrates the principles of how the pixel footprint distribution (PFD) can be used to establish a one-to-many relation between two images.

Figure 6 illustrates how the topography and angle of view can cause pixels from different images that are close in center position, to have completely different PFDs. Given a detailed representation of the PFD, it would be clear that the chosen pixels do not contain information from a common part of the scene.

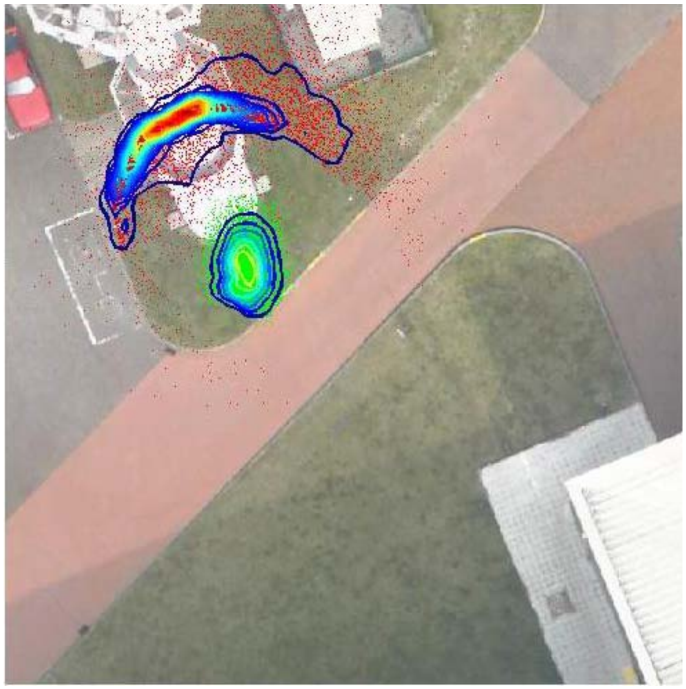

Figure 6. This image illustrates how different the pixel footprint distribution (PFD) can be for pixels from different images even though the distance between their center positions is quite small.

4. CONCLUSION

We have presented a method for estimating the PFD for individual pixels in an image. This pixel specific spatial distribution of responsivity reflects the combined geometrical limitations induced by the imaging system's point spread function, the motion of the camera during integration time with uncertainties as well as the potentially complicated topography of the scene. The PFD could be a valuable asset for data fusion, for example in hyperspectral change detection.

REFERENCES

[1] Opsahl, T., Haavardsholm, T. V., Winjum, I., "Real-time georeferencing for an airborne hyperspectral imaging system", Proc. SPIE 8048, Algorithms and Technologies for Multispectral, Hyperspectral and Ultraspectral Imagery XVII, 80480S (2011); doi: 10.1117/12.885069.

[2] Skauli, T., Haavardsholm, T. V., Kåsen, I., Arisholm G., Kavara, A., Opsahl, T., Skaugen, A., "An airborne real-time hyperspectral target detection system", Proc. SPIE 7695, Algorithms and Technologies for Multispectral, Hyperspectral and Ultraspectral Imagery XVI, 76950A (2010); doi: 10.1117/12.850443.

[3] Gade, K., 'NAVLAB, a Generic Simulation and Post-processing Tool for Navigation", European Journal of Navigation, Vol. 2, no. 4, (2004).