Proceedings of Australasian Conference on Robotics and Automation, 3-5 Dec 2012, Victoria University of Wellington, New Zealand.

# A Mutual Information Approach to Automatic Calibration of Camera and Lidar in Natural Environments

**Zachary Taylor and Juan Nieto**
Australian Centre for Field Robotics
University of Sydney, Australia
{z.taylor, j.nieto}@acfr.usyd.edu.au

## Abstract

This paper presents a method for calibrating the extrinsic and intrinsic parameters of a camera and a lidar scanner. The approach uses normalised mutual information to compare an image with a lidar scan. A camera model that takes into account orientation, location and focal length is used to create a 2D lidar image, with the intensity of the pixels defined by the angle of the normals in the lidar scan. Particle swarm optimisation is used to find the optimal model parameters. The method is successfully validated in a natural environment with images collected by a hyperspectral and a 3D lidar scanner.

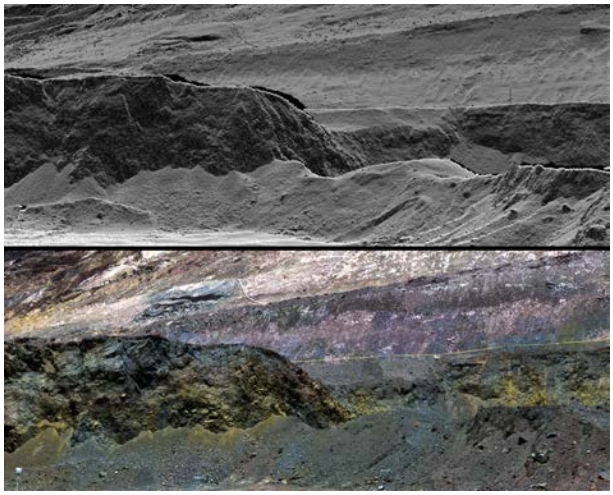

Figure 1: Generated Lidar image (top) and Hyperspectral image (bottom) of area 3 of the mine dataset

## 1 Introduction

Accurate calibration between lidar scanners and cameras is important as it allows each point in the cloud produced by the scanner to have a colour associated with it. These coloured points can then be used to build up richer models of the area. This calibration is quite challenging due to the very different modalities of the sensors and the nature of the output information. Due to the difficulty of aligning the sensors the majority of these systems are calibrated by hand. This is currently done using reflective markers, chequerboards or by painstakingly hand labelling large numbers of points. These methods are slow, labour intensive and often produce results with significant errors.

Our application for building a camera-lidar system is in mining, where hyperspectral cameras are used to detect ore quality. Hyperspectral cameras detect hundreds or even thousands of different light frequencies and their relative intensities, which is unlike regular cameras that are typically sensitive to only three different frequencies (red, green and blue). From all of the hyperspectral camera readings an estimate of the reflectance of a point as function of its frequency can be generated. These reflectance plots can then be used to detect what material is present at the location.

These cameras can be used to scan cliff faces in mines and generate an image that shows the quality of ore present in that face. If a depth map of the same face is also available, the geometry of the face can be combined with this map of the ore quality to give a 3D visualisation of the ore in the face. This information can be used to estimate the amount of ore and greatly improve the efficiency with which an area can be mined. It can also prevent a large amount of wasted effort in the excavation and processing of areas that only contain low quality ore. This process however relies on knowing which pixel in the hyperspectral image is assigned to each point detected by the lidar and thus requires calibration of the sensors.

To make the scanning of these mining areas easy and practical it is desirable that the hyperspectral and lidar scans of an area can be made independently without consideration for the exact location and orientation of the scanners. The system also needs to operate using a

single scan without any calibration objects such as chequerboards placed in the scene. With these constraints in mind this paper looks at a method for automatically aligning a camera and a lidar scanner using scans of an arbitrary environment. The method estimates the extrinsic and some intrinsic parameters of the camera in an automated process with minimal human interaction. The process makes minimal assumptions about the sensors and the environment making it applicable for a wide range of lidar-camera calibration problems. The approach is demonstrated by solving the calibration problem for a mining site that has been outlined. An example of the data used is shown in Fig. 1.

The proposed method operates by creating a camera model and projecting the lidar point cloud through it. When the camera model has the same extrinsic and intrinsic parameters as the hyperspectral camera, a point in the image generated by this model will be in the same location as a point in the hyperspectral image. Then, each point in the lidar point cloud can be assigned to a hyperspectral pixel. The parameters for this camera model are found in an automated fashion using the normalised mutual information (NMI) between the generated and hyperspectral image as a measure of the alignment. The approach uses surface normals for the generated images pixel intensities. While normals have been used for aligning 3D models using mutual information as presented in [Corsini *et al.*, 2009], it appears to never have been applied to lidar image alignment. Parameters are optimised using particle swarm [Kennedy and Eberhart, 1995]. Particle swarm optimisation was used as it can find the global optima of a function with a large number of local optima as our problem has.

## 2 Related work

Several methods exist for aligning a single image with a 3D view of the same location. A recently proposed method for registering aerial images and lidar scans based on edges and corners is presented in [Li *et al.*, 2012]. Their method works by constructing closed polygons from edges detected in both the lidar scan and images. Once the polygons have been extracted they are used as features and matched to align the sensors. The method was only intended for and thus tested using aerial photos of urban environments.

A method similar to our approach, registration by creating an image from the lidar is presented in [Mastin *et al.*, 2009]. The intensity of the pixels in the image generated from the lidar scan were either the intensity of the return the laser had or the height from the ground. The images were compared using mutual information and optimisation was done via downhill simplex. This method operates quickly and produced accurate results although its search space was rather limited requiring an initial guess of the orientation of the camera that was correct to within 0.5 degrees for roll and pitch. The method was only tested in an urban environment where buildings provided a strong relationship between height and image colour.

A large number of methods exist that exploit the detection of straight edges in a scene [Lee *et al.*, 2002; Lyngbaek and Zakhor, 2008; Liu and Stamos, 2007]. While these methods work well in cities and with images of buildings they are unable to correctly register natural environments due to the lack of strong straight edges.

In a similar problem [Levinson and Thrun, 2012] calibrated the extrinsic parameters of a camera and velodyne system using a series of 100 image scan pairs. Their method involves finding edge images for both the laser and images and using an element wise multiplication of these images, assuming that when this is maximised the two sensors are correctly aligned. There is also some extra processing done to improve the robustness and convergence of the method.

Two methods have been presented for aligning hyperspectral images with lidar scans of cliff faces. In [Nieto *et al.*, 2010] the authors used a pre-calibrated camera mounted on top of the lidar to render colour information for the lidar scans. This colour information was used to create an image with approximately the same parameters as the hyperspectral image. Both images were then converted to greyscale and roughly aligned using SIFT features. Fine registration was performed by a piecewise linear transform where sum of squares differences (SSD) was used as the measure of match strength. While this method produces good results it requires an additional camera that has been pre-calibrated with the lidar. A second approach is presented in [Kurz *et al.*, 2011]. Their method of registration is based on the use of retro-reflective targets whose positions, once detected by both sensors can be used to calibrate them.

In [Corsini *et al.*, 2009] the authors looked into different techniques for generating an image from a 3D model so that mutual information would successfully register the image with a real photo of the object. They used NEWUOA optimisation in their registration and looked at using the silhouette, normals, specular map, ambient occlusion and combinations of these to create an image that would robustly be registered with the real image. They found surface normals and a combination of normal and ambient occlusion to be the most effective.

A fairly in depth look at many of the different methods for aligning images with lidar scans can be found in [Mishra and Zhang, 2012].

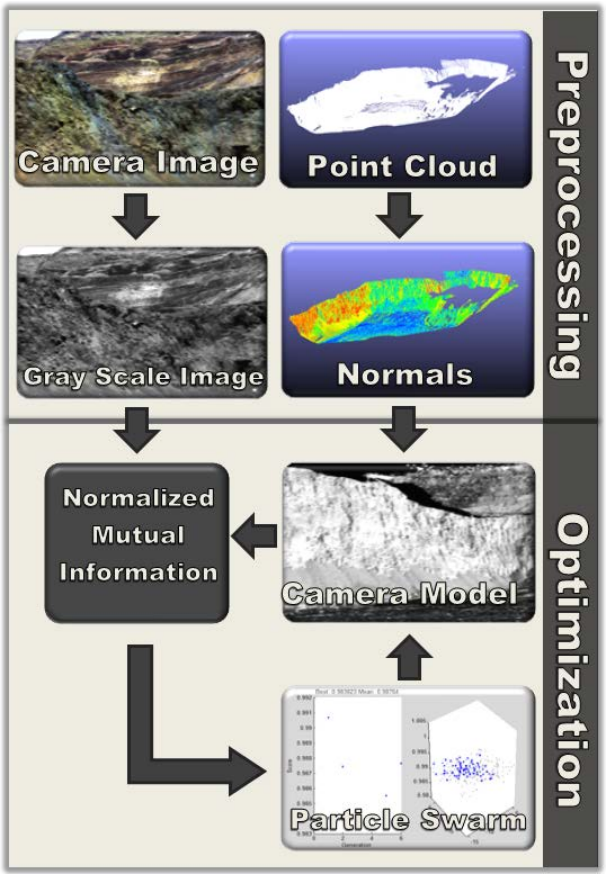


Figure 2: Overview of alignment method

## 3 Methodology

Fig. 2 shows a block diagram of the whole process. An RGB image is obtained from the hyperspectral one, which is then converted to grey scale. For the data given by the laser scanner, first an approximation of the normals of each point is estimated. After that, a camera model is used to generate a 2D image from the point cloud, where the intensity of a pixel is defined as directly proportional to the angle of its normal. Normalised mutual information is then used as a measure to compare the hyperspectral image with the laser projection. This process is repeated for different camera model parameters during the optimisation until convergence occurs.

### 3.1 Surface normals

The normals of the points are estimated by first taking the difference between consecutive points in the scan. Once this difference is obtained the angle between this point and the horizontal $x-z$ plane is obtained and stored for use in the image generation process. This process is shown in Eq. 1.

$$n_m = \arctan\left(\frac{y_m - y_{m+1}}{\sqrt{(x_m - x_{m+1})^2 + (z_m - z_{m+1})^2}}\right) \quad (1)$$

Where $n$ is the normal value, $m$ is the point index and $x$,$y$,$z$ are the points location.

Surface normals were chosen for colouring the pixels as the method results in a fairly strong correspondence between the generated laser image and the image from the hyperspectral camera. It also works well with mutual information as was demonstrated by [Corsini *et al.*, 2009]. This relationship exists as the angle between a surface, the camera and the light sources plays a large role in determining the amount of light a surface reflects into the camera and thus its intensity in the generated image. Normals were used rather than the angle between the surface and the camera as the normals are independent of the camera location and so can be pre-calculated. This is important as the calculation of the image from the lidar scan is the most expensive computational step in the optimisation. The angle between the normals and a horizontal plane is used as it is assumed that most of the light is coming from above and thus this angle has the largest influence on the intensity.

### 3.2 Mutual Information

Mutual information is a measure of mutual dependence between two signals. It was first developed in information theory using the idea of Shannon entropy [Pluim *et al.*, 2003]. Shannon entropy is a measure of how much information is contained in a signal and its discrete version is defined as [Shannon, 1948]:

$$H(X) = H(p_X) = \sum_{i=1}^{n} p_i log(\frac{1}{p_i}) \quad (2)$$

where X is a discrete random variable with n elements and the probability distribution $p_X = (p_1, ..., p_n)$. For this purpose $0log\infty = 0$. Using this idea of Shannon entropy, mutual information is defined as

$$MI(M,N) = H(M) + H(N) - H(M,N) \quad (3)$$

where $H(M,N)$ is the joint entropy which is defined as

$$H(M,N) = H(p(m,n)) = \sum_m \sum_n p(m,n) log(\frac{1}{p(m,n)}) \quad (4)$$

Mutual information when used for registration purposes suffers from an issue in that it can be influenced by the amount of total information contained in images causing it to favour images with less overlap [Studholme *et al.*, 1999]. This unwanted effect is reduced by using normalised mutual information defined as

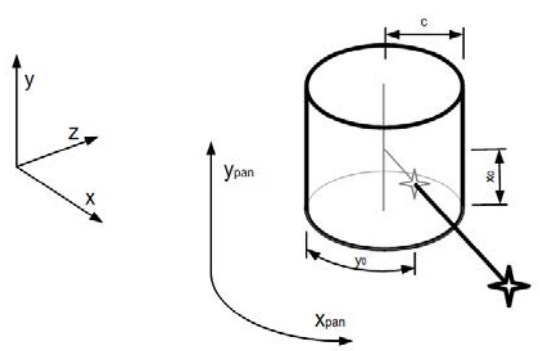


Figure 3: Cylinder model used to represent the hyperspectral imaging system

$$NMI(M,N) = \frac{H(M) + H(N)}{H(M,N)} \tag{5}$$

In practice, for images, the required probabilities $p(M)$ and $p(N)$ can be estimated using a histogram of the distribution of intensity values.

Normalised mutual information is used as the metric for evaluating the strength of the alignment between the two images as it can cope with the non-linear relationship between angle and intensity. It also accounts for issues such as how different materials can appear dissimilar in different sensor modalities. This strength means that it can be assumed that the global maximum of normalised mutual information (NMI) occurs when the images are best aligned.

### 3.3 Camera model

To convert the laser data from a list of 3D points to a 2D image the points are first passed into a transformation matrix that aligns the cameras and the world axis. After this has been performed, a basic panoramic camera model that projects the points onto a cylinder is used, a rough depiction of this is shown in Fig. 3. This model projects the points using Eqs. 6 and 7 [Schneider and Maas, 2003].

$$x_{pan} = x_0 - c\arctan(\frac{-y}{x}) + \Delta x_{pan} \tag{6}$$

$$y_{pan} = y_0 - \frac{cz}{\sqrt{x^2 + y^2}} + \Delta y_{pan} \tag{7}$$

where

$\mathbf{x}_{pan}$ **,** $\mathbf{y}_{pan}$ are the x and y position of the point in the image.

**x, y, z** are the coordinates of points in the environment.

**c** is the principle distance of the model

$\mathbf{x}_0$ **,** $\mathbf{y}_0$ are the location of the principle point in the image.

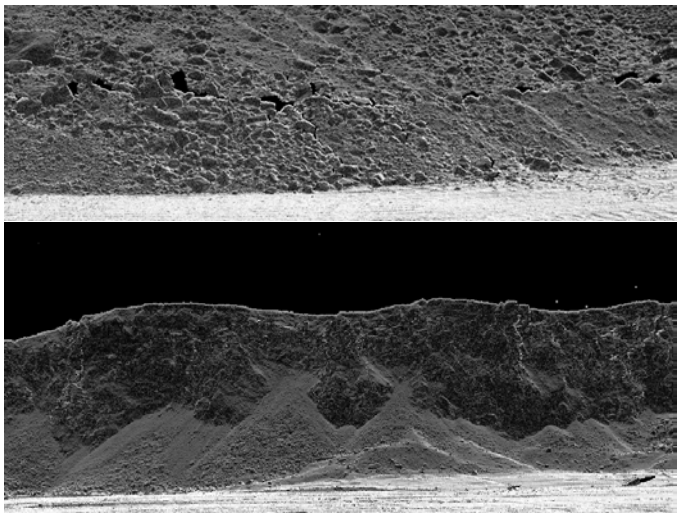

Figure 4: Examples of images obtained with the panoramic camera model

**Δx , Δy** are the correction terms used to account for several imperfections in the camera.

These models ignore the effects of several other parameters such as the x and y axis of the camera not being perfectly perpendicular and the radial distortion of the lens. However, the simplified model can be justified as for a similar camera with a resolution of 10000 by 60000 pixels, the authors in [Schneider and Maas, 2003] showed that the error caused by these parameters is less than 10 pixels. For applications in remote sensing this level of error in the images was taken to be acceptable. The resulting output of the camera model can be seen in Fig. 4.

### 3.4 Optimisation

Depending on the assumptions made by the camera model and the accuracy of the initial scans position the problem either has 4, 7 or 9 variables to solve. This search space is also highly non-convex with a large amount of local maximums. An example of the typical shape when NMI is plotted in two dimensions is shown in Fig. 5. With the simple histogram method of calculating the mutual information used in this paper there is also no information on the derivatives available. These difficulties are further compound by the relatively expensive process of generating an image from a point cloud that is required for every function evaluation.

The fairly large range that the correct values can lie in coupled with the local maximums mean that simple gradient accent type methods as used by others to solve image lidar registration [Mastin *et al.*, 2009; Mishra and Zhang, 2012] cannot be used here. To solve these problems particle swarm optimisation is used [Kennedy and Eberhart, 1995; Mikki and Kishk, 2008].

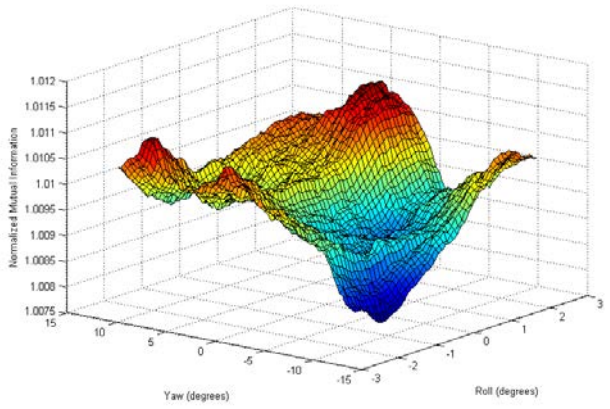


Figure 5: Example of NMI values for changing roll and yaw

Particle swarm optimisation works by placing an initial population of particles randomly in the search space. Each iteration a particle moves to a new location based on three factors: 1) it moves towards the best location found by any particle. 2) it moves towards the best location it has ever found itself. 3) it moves in a random direction. The optimiser stops once all particles have converged.

The entire algorithm for registration is summarised in section 3.5.

### 3.5 Algorithm

**Let**

$\boldsymbol{r^i(t)}$ be the position of particle i at time t

$\boldsymbol{v^i(t)}$ be the velocity of particle i at time t

$\boldsymbol{p_n^{i,L}}$ be the local best of the ith particle for the nth dimension

$\boldsymbol{p_n^g}$ be the global best for the nth dimension

$\boldsymbol{n \in 1, 2, ...N}$

$\boldsymbol{t}$ is the time

$\boldsymbol{\Delta t}$ is the time step

$\boldsymbol{c_1}$ and $\boldsymbol{c_2}$ are the cognitive and social factor constants

$\boldsymbol{\phi_1}$ and $\boldsymbol{\phi_2}$ are two statistically independent random variables uniformly distributed between 0 and 1

$\boldsymbol{w}$ is the inertial factor

**foreach** iteration, $l$

**if** $f(r^i(l+1)) > f(p^{i,L}(l))$ **then**

$p^{i,L}(l+1) = r^i$

**end**

**if** $f(r^i(l+1)) > f(p^g(l))$ **then**

$p^g(l+1) = r^i$

**end**

$v_n^i(t+\Delta t) = wv_n^i(t) + c_1\phi_1[p_n^{i,L} - x_n^i(t)]\Delta t + c_2\phi_2[p_n^g - x_n^i(t)]\Delta t$

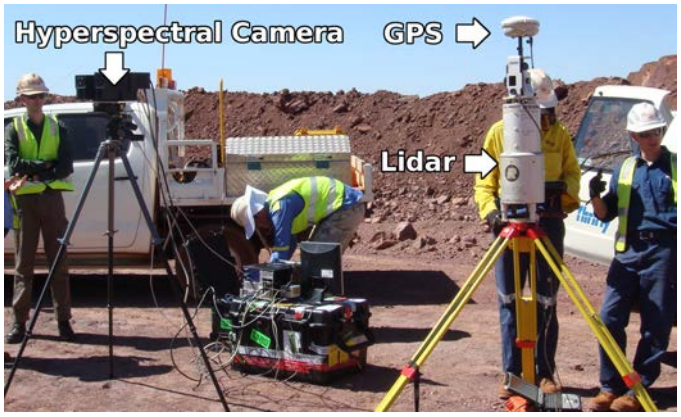


Figure 6: Hyperspectral camera and lidar setup used to collect the data

$r_n^i(t+\Delta t) = r_n^i(t) + \Delta t v_n^i(t)$

**end**

## 4 Experiments

The method presented was evaluated on a dataset collected in an open pit mine in western Australia [Nieto *et al.*, 2010]. The laser used was a Riegl LMS-Z420i and the hyperspectral camera was a Neo HySpex VNIR and SWIR, the setup can be seen in Fig. 6. RTK GPS was used to provide the exact location of the camera and laser scanner, however, due to the geography of the place, at two of the locations this signal failed and only a standard GPS location was given. The hyperspectral camera was readjusted and the focal length changed before taking each image so its intrinsics cannot be assumed to be the same between images. Scan and image pairs from four different sections of the mine were used. These images were taken over the course of two days. These areas of the mine were labelled $a_{1,...,4}$. An initial guess at the orientation of the camera was made. This guess was chosen such that a comparison of the hyperspectral and initial lidar scan could clearly show the alignment to be incorrect by a few degrees. The initial location of the camera was taken to be the GPS coordinates. The initial guess of the values are shown in Table 1

### 4.1 Optimisation

Each optimisation was run 3 times, each time with different set of parameters $\theta$. Eq. 8 shows the three configurations evaluated.

$$\begin{aligned} \theta_1 &= [roll, pitch, yaw, c] \\ \theta_2 &= [roll, pitch, yaw, c, x, y, z] \\ \theta_3 &= [roll, pitch, yaw, c, x, y, z, \Delta x_{pan}, \Delta y_{pan}] \end{aligned} \quad (8)$$

| location | $\theta$ | $\Delta$x | $\Delta$y | $\Delta$z | roll | pitch | yaw | c | $\Delta x_{pan}$ | $\Delta y_{pan}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $a_1$ | 0 | 0 | 0 | 0 | 0 | 0 | -44 | 1320 | 0 | 0 |
| | 1 | n/a | n/a | n/a | -0.1 | 0.1 | -45.1 | 1332 | n/a | n/a |
| | 2 | -0.138 | -0.109 | -0.196 | -0.82 | 0.1 | -45.2 | 1334 | n/a | n/a |
| | 3 | -0.049 | -0.03 | -0.205 | -0.75 | -0.1 | -45.9 | 1333 | 16.8 | 19.5 |
| $a_2$ | 0 | 0 | 0 | 0 | 0 | 0 | 37 | 1320 | 0 | 0 |
| | 1 | n/a | n/a | n/a | -2.6 | 5.9 | 39.1 | 1321 | n/a | n/a |
| | 2 | 0.77 | 1.3 | 0.265 | -2.5 | 5.6 | 40.4 | 1315 | n/a | n/a |
| | 3 | 0.561 | 2.195 | -0.33 | -2.4 | 6.4 | 40.74 | 1307 | 10.0 | 1.6 |
| $a_3$ | 0 | 0 | 0 | 0 | 0 | 5 | 5 | 1320 | 0 | 0 |
| | 1 | n/a | n/a | n/a | -1.1 | 6.1 | 5.1 | 1326 | n/a | n/a |
| | 2 | 1.998 | 1.97 | 0.459 | -1.1 | 6.1 | 5.1 | 1326 | n/a | n/a |
| | 3 | 1.998 | 2 | 0.484 | -1.2 | 6.3 | 5.3 | 1325 | -3.8 | -4.1 |
| $a_4$ | 0 | 0 | 0 | 0 | 0 | 4 | 50 | 1320 | 0 | 0 |
| | 1 | n/a | n/a | n/a | -1.1 | 3.3 | 54.1 | 1337 | n/a | n/a |
| | 2 | -0.098 | -0.386 | -0.319 | -1.2 | 3.7 | 53.8 | 1338 | n/a | n/a |
| | 3 | -0.949 | -0.92 | -0.32 | -1.2 | 3.4 | 53.3 | 1332 | -9.0 | 7.0 |

Table 1: Parameters found for camera model camera model. $\theta_0$ is the initial values

The search space for the optimizer was constructed assuming the following:

- The roll, pitch and yaw of the camera were within 10, 20 and 5 degrees respectively of the lasers.
- The cameras principal distance was within 20 pixels of correct (for this camera principal distance $\approx$ 1320).
- The x, y and z coordinates were either correct or within 4, 4 and 0.5 meters of correct.
- The $\Delta x_{pan}$ and $\Delta y_{pan}$ were either 0 or within 20 pixels of correct.

The particle swarm optimiser was started with 100 particles and ran until the particles all converged to within 0.1 in all dimensions of each other. This usually took between 100 and 200 iterations.

The code was written in Matlab with mex files written in C created for the generation of the lidar images and mutual information calculations. The code was run on a Dell latitude E6150 laptop with an Intel i5 M520M CPU. Each function evaluation took around 0.1 seconds and the total runtime for the code was 15 to 45 minutes depending on the image resolutions, number of points in the scan and how quickly convergence occurred.

### 4.2 Results

For all the data sets used, no ground truth as to the orientation between the lidar and the hyperspectral camera was given. This makes quantitative evidence of the accuracy of the alignment difficult. However one of the main applications of this alignment is for visualisation of multi-model information and so how accurate the alignment appears to someone viewing the two images can be used to provide some estimate on the effectiveness of the method.

The parameters estimated by our approach are listed in Table 1. For the areas $a_4$ and $a_1$ the RTK GPS was operating and so any variation in position can be taken as error. For dataset $a_1$ all results are within 0.3m of each other while in $a_4$ the error is as large as 1.4m. This error mainly occurred when all parameters were varied and may be a result of the search space becoming too large for the optimiser to locate the global maximum. In the absence of ground truth values for the other parameters few conclusions can be drawn from the Table alone. Instead, to visualise the accuracy of the alignment, four different scans are projected onto the images and shown in Fig. 7. On inspection it can clearly be seen that for all runs the approach converges to a solution that appears reasonable and is significantly better than the initial guess provided. For $a_4$ an image generated by using the calibrated camera model to colour the point cloud is shown in Fig. 8, this gives some indication to the accuracy as the only points clearly miscoloured are caused by lidar returns off dust and tape that was blowing in the wind. Note when viewing these outputs that while a tripod with the hyperspectral calibration board (for reflectance) is present in most of the images, it was often moved between the time the hyperspectral image and the lidar scan were taken meaning it cannot be used to judge the quality of the alignment.

Experiments were run with larger allowances for the roll, pitch and yaw with the aim of requiring no initial guess for these values. It was found however that for these larger tolerances 2 of the 4 images became stuck in local maximums during optimisation and converged to solutions that were clearly not optimal. These incorrect

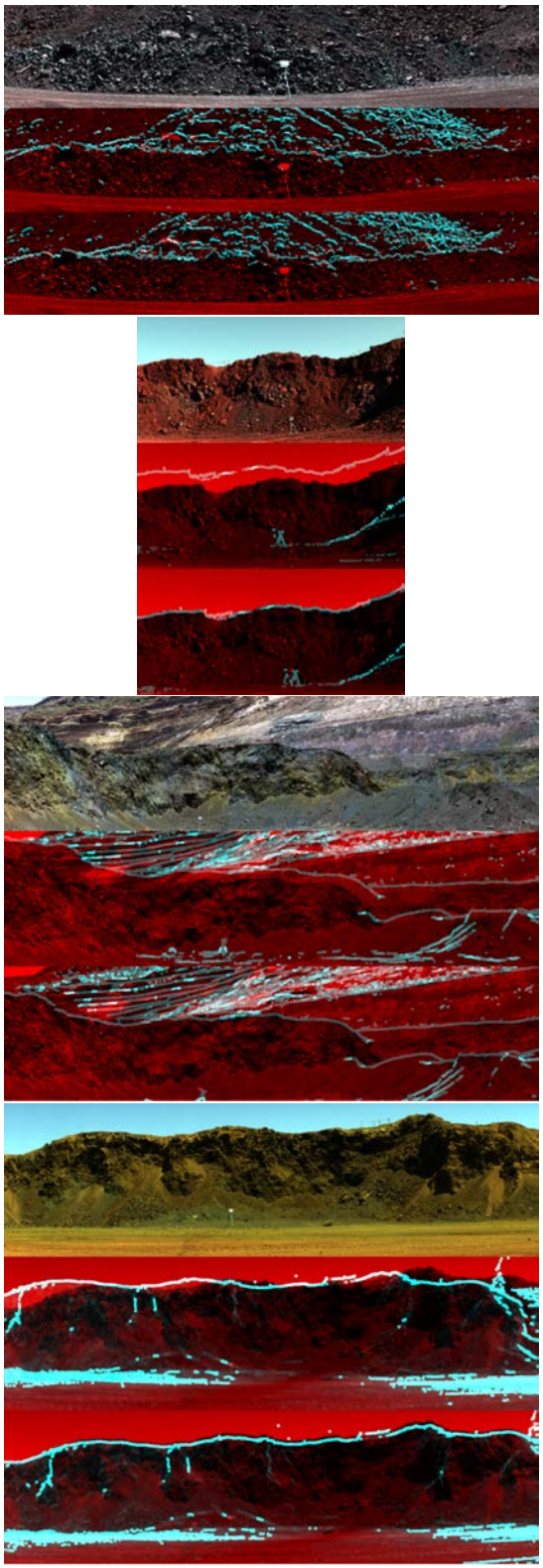

Figure 7: Results of optimisation. To visualise results strong edges in the lidar data have been found and are overlaid in blue on top of the red hyperspectral image. Each area has shown the rgb bands of the hyperspectral (top), the initial guess (middle) and the alignment found changing $\theta_2$. The sites from top to bottom are $a_1$, $a_2$, $a_3$ and $a_4$.

results were caused by a problem with the optimisation and not with the metric as the NMI value found for these incorrect solutions was less than that the one found by the more constrained optimisation.

## 5 Comparisons with other methods

For the dataset used with natural images, very few methods were appropriate to attempt comparison with. We implemented the approach presented in [Mastin *et al.*, 2009]. However no images were successfully aligned by this method. This failure was expected as the method was designed for aerial images of urban scenes and the intensity of the lasers return was not available. As our problem involves calibration of range and image data, the method presented in [Nieto *et al.*, 2010] could not be used in its standard form. We tried a modified version of the method replacing the colour image with our laser image coloured by the normals. We found that the images could be successfully registered for some of the dataset. This method however suffered from problems with outliers during SIFT matching due to the large differences in modalities between the normals image and the hyperspectral. The method also suffers from the issue that it operates by warping the images and so cannot be used to calibrate the sensors. No other methods were found that claimed to be able to register a ground based lidar scan with a photo in a natural environment.

## 6 Conclusions and Future Work

A method for aligning a camera with a lidar scan suitable for natural environments was presented. This method operates by creating an image using a camera model and the directions of the normals of points. Normalised mutual information is used to compare the images and maximised to find correct alignment of the images. The method was demonstrated to successfully work to calibrate a hyperspectral camera and a 3D laser in a dataset collected in a mine. In future work a dataset with ground truth values will be obtained so that the accuracy of the results can be quantitatively measured. Currently the implementation is slow when compared to similar methods. The reason for this is transformation of the lidar points which take over 90% of the time. An implementation that uses the GPU for this task or a method for thinning the initial point cloud would solve this issue. A pyramid implementation is also being looked into to help this process.

## Acknowledgements

This work has been supported by the Rio Tinto Centre for Mine Automation and the Australian Centre for Field Robotics, University of Sydney.

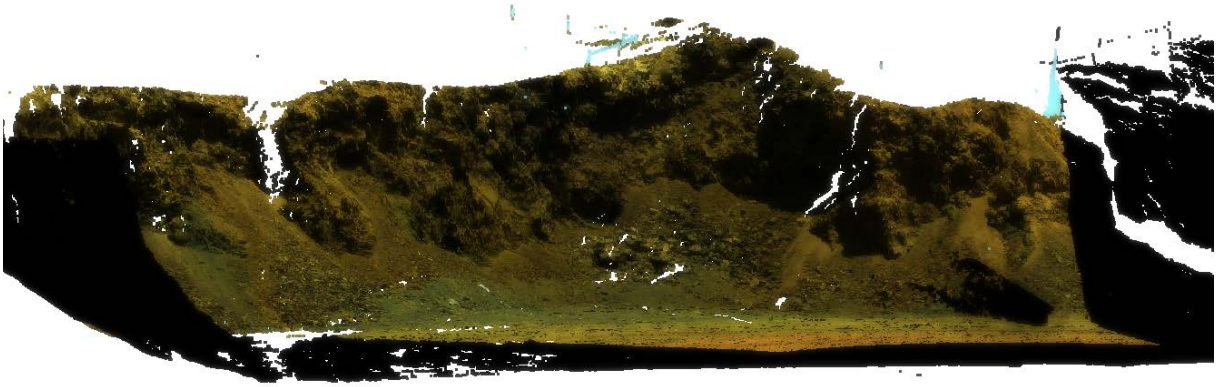

Figure 8: Point cloud of $a_4$ coloured by the visible bands of the calibrated hyperspectral image

## References


[Corsini *et al.*, 2009] Massimiliano Corsini, Matteo Dellepiane, Federico Ponchio, and Roberto Scopigno. Image to Geometry Registration: a Mutual Information Method exploiting Illumination-related Geometric Properties. *Computer Graphics Forum*, 28(7):1755–1764, 2009.

[Kennedy and Eberhart, 1995] James Kennedy and Russel Eberhart. Particle swarm optimization. *Proceedings of ICNN'95 - International Conference on Neural Networks*, 4:1942–1948, 1995.

[Kurz *et al.*, 2011] Tobias H. Kurz, Simon J. Buckley, John a. Howell, and Danilo Schneider. Integration of panoramic hyperspectral imaging with terrestrial lidar data. *The Photogrammetric Record*, 26(134):212–228, June 2011.

[Lee *et al.*, 2002] Sung Lee, Soon Jung, and Ram Nevatia. Automatic integration of facade textures into 3D building models with a projective geometry based line clustering. *Computer Graphics Forum*, 21(3), 2002.

[Levinson and Thrun, 2012] Jesse Levinson and Sebastian Thrun. Automatic Calibration of Cameras and Lasers in Arbitrary Scenes. In *International Symposium on Experimental Robotics*, pages 1–6, 2012.

[Li *et al.*, 2012] Hui Li, Cheng Zhong, and Xianfeng Huang. Reliable Registration of Lidar Data and Aerial Images without Orientation Parameters. *Sensor Review*, 32(4), 2012.

[Liu and Stamos, 2007] Lingyun Liu and Ioannis Stamos. A systematic approach for 2D-image to 3D-range registration in urban environments. *2007 IEEE 11th International Conference on Computer Vision*, pages 1–8, 2007.

[Lyngbaek and Zakhor, 2008] Kristian Lyngbaek and Avideh Zakhor. Automatic registration of aerial imagery with untextured 3D LiDAR models. *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, June 2008.

[Mastin *et al.*, 2009] Andrew Mastin, Jeremy Kepner, and John Fisher III. Automatic registration of LIDAR and optical images of urban scenes. *Computer Vision and Pattern Recognition*, pages 2639–2646, 2009.

[Mikki and Kishk, 2008] Said M. Mikki and Ahmed a. Kishk. *Particle Swarm Optimization: A Physics-Based Approach*, volume 3. January 2008.

[Mishra and Zhang, 2012] Rakesh Mishra and Yun Zhang. A Review of Optical Imagery and Airborne LiDAR Data Registration Methods. *The Open Remote Sensing Journal*, 5:54–63, 2012.

[Nieto *et al.*, 2010] Juan Nieto, Sildomar Monteiro, and Diego Viejo. 3D geological modelling using laser and hyperspectral data. *Geoscience and Remote Sensing Symposium*, pages 4568–4571, 2010.

[Pluim *et al.*, 2003] Josien P. W. Pluim, J. B. Antoine Maintz, and Max A. Viergever. Mutual-information-based registration of medical images: a survey. *Medical Imaging, IEEE*, 22(8):986–1004, 2003.

[Schneider and Maas, 2003] Danilo Schneider and Hans-Gerd Maas. Geometric modelling and calibration of a high resolution panoramic camera. *Optical 3-D Measurement Techniques VI*, 2003.

[Shannon, 1948] Claude Elwood Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3):379–423, 1948.

[Studholme *et al.*, 1999] Colin Studholme, Derek L.G. Hill, and David J Hawkes. An overlap invariant entropy measure of 3D medical image alignment. *Pattern recognition*, 32(1):71–86, January 1999.