

AUTOMATIC REGISTRATION OF OPTICAL IMAGERY WITH 3D LIDAR DATA USING LOCAL COMBINED MUTUAL INFORMATION

Ebadat G. Parmehr^{a,b,*}, Clive S. Fraser^{a,b}, Chunsun Zhang^{a,c}, Joseph Leach^b

^a Cooperative Research Centre for Spatial Information, VIC 3053, Australia

^b Department of Infrastructure Engineering, University of Melbourne, VIC 3010, Australia –
ebadatg@student.unimelb.edu.au, (c.fraser, leach)@unimelb.edu.au

^c School of Mathematical and Spatial Sciences, RMIT University, VIC 3000, Australia - chunsun.zhang@rmit.edu.au

KEY WORDS: Registration, Optical Imagery, LiDAR Point Cloud

ABSTRACT:

Automatic registration of multi-sensor data is a basic step in data fusion for photogrammetric and remote sensing applications. The effectiveness of intensity-based methods such as Mutual Information (MI) for automated registration of multi-sensor image has been previously reported for medical and remote sensing applications. In this paper, a new multivariable MI approach that exploits complementary information of inherently registered LiDAR DSM and intensity data to improve the robustness of registering optical imagery and LiDAR point cloud, is presented. LiDAR DSM and intensity information has been utilised in measuring the similarity of LiDAR and optical imagery via the Combined MI. An effective histogramming technique is adopted to facilitate estimation of a 3D probability density function (pdf). In addition, a local similarity measure is introduced to decrease the complexity of optimisation at higher dimensions and computation cost. Therefore, the reliability of registration is improved due to the use of redundant observations of similarity. The performance of the proposed method for registration of satellite and aerial images with LiDAR data in urban and rural areas is experimentally evaluated and the results obtained are discussed.

1. INTRODUCTION

The automatic registration of multi-sensor data has generated much research interest in remote sensing and digital photogrammetry because of both the increasing availability of large volumes of Earth observation data and the need for automatic integration of multi-sensor data to generate redundant and complementary spatial information products. In fact, integration of complementary multi-sensor data such as LiDAR (Light Detection And Ranging) and optical imagery is required for many applications such as feature extraction, building reconstruction, 3D city modelling and change detection.

LiDAR systems record 3D coordinates of ground objects, with the resulting point cloud generally giving poor definition of break lines. The photometric data acquired with the 3D point cloud includes only the intensity of each reflected laser strike. Optical images, on the other hand, provide comprehensive photometric data, but it can be difficult to fully extract 3D point clouds from imagery alone due to shortcomings in both imaging geometry and the robustness and reliability of image matching procedures (Shin et al., 2007). The advantages of using complementary information from multi-sensor datasets can be fully exploited only when the datasets are co-registered in a common reference system. Otherwise, the presence of small errors in registration can have a large impact on the accuracy of subsequent analysis. Data fusion applications that utilise both 3D point clouds and optical imagery generally require accurate and automatic registration of the two datasets. Although most Earth observation datasets are georeferenced, misregistration errors often exist between complementary datasets, generally as a result of perturbations in direct georeferencing and/or instabilities in sensor system calibration.

This paper is organized as follows: Section 2 provides a literature review related to the automated registration of optical imagery to airborne LiDAR point clouds. Section 3 defines Mutual Information (MI) and introduces the proposed approach. Then, in Section 4, the implementation of the method is explained, and its performance is experimentally evaluated in Section 5. Conclusions follow in Section 6.

2. CURRENT METHODS

In general, automatic methods for registration of 2D images and 3D point clouds fall into two main categories, namely *feature-based* and *intensity-based* (Goshtasby, 2012).

2.1 Feature-based Methods

Feature-based methods of registration rely on the extraction of features separately from different datasets and their subsequent matching. Different types of features, such as points, linear features and patches, are used to establish the correspondences between both datasets (Habib et al. 2005; Rönnholm 2011; Toth et al. 2011). Success in the matching process for datasets with different modalities is dependent upon the presence of a sufficient number and distribution of corresponding features. Particularly, in the case of 2D/3D data registration, a feature common to both datasets may nevertheless have different meaning due to differences in the data character. In addition, feature extraction from 3D point cloud data, along with the searching for corresponding features, requires both point cloud processing and feature matching procedures that are less straightforward. More importantly, in the absence of sufficient corresponding features, the feature-based methods are often inappropriate and prone to failure.

* Corresponding author.

2.2 Intensity-based Methods

Intensity-based methods for registration operate under the assumption that strong statistical relationships exist between the datasets (Hajnal et al., 2001). They do not rely on extracted common features; but instead upon matching based on similarities in intensity characteristics, e.g. on grey values of the images or on elevation values (mapped as intensities) within 3D point clouds. This can make these intensity-based methods more general, robust and efficient for registration of multi-sensor data. The performance of MI as a similarity measure for intensity-based registration methods has been reported in medical imaging for some two decades (Hajnal et al., 2001; Pluim et al., 2003). Promising results achieved in medical imaging have encouraged the photogrammetry and remote sensing communities to apply this approach for multi-sensor data registration (LeMoigne et al., 2011).

The amount of MI between two datasets is defined via the joint probability density function (pdf) of the datasets, instead of through a similarity determination via intensity matching. This particular property has made MI suitable for measuring the similarity of data with either linear or non-linear relationships in uni- or multi-modal datasets.

The maximum value of MI is expected to be achieved when the datasets are geometrically aligned (Maes et al., 1997; Viola and Wells III, 1997). Traditional registration of optical imagery with 3D point clouds has been carried out under the assumption that MI operates well when the datasets have a functional relationship between them. MI of optical imagery and surface normals has been used as a similarity measure (Viola and Wells III, 1997). This method can suffer from occlusions and different lighting conditions, so illumination-related geometric properties including the surface normals, ambient occlusions and reflection directions have been incorporated into MI to improve the performance of the 2D to 3D data registration procedure (Corsini et al., 2009). It is noteworthy that due to its statistical basis, MI can find the correspondence between 3D LiDAR point clouds and optical images without any consideration of functional relationships. The exterior orientation of the camera is computed by implementing the sum of joint entropies of optical images with LiDAR elevation and LiDAR intensity in urban areas, without using surface normals (Mastin et al., 2009). Aerial orthoimages have been registered to LiDAR DSM and intensity data by defining a general weighting function for the joint pdf to overcome the problems of a non-monotonic convergence surface of MI (Parmehr et al., 2012).

Contrary to the method proposed in this paper, the approaches referred to above are not generally able to fully utilise both the LiDAR intensity and 3D point cloud data in the registration process. Moreover, the methods referred to above suffer from unreliable results due to a lack of observational redundancy in determination of the similarity measure. In order to efficiently accommodate the registration of multi-sensor data including satellite and aerial images and airborne LiDAR data, a novel multi-resolution local combined MI has been developed.

3. METHODOLOGY

3.1 Mutual Information

MI is based on information theory and has been defined in the literature in various forms such as entropy, conditional entropy and probability divergence (Cover and Thomas, 1991). In this

paper, MI of two variables A and B with a joint pdf of $P(A,B)$ is defined using Shannon entropy (Shannon et al., 1949) as

$$MI(A;B) = H(A) + H(B) - H(A,B). \quad (1)$$

Here, $H(A)$ and $H(B)$ are the marginal entropy of A and B , respectively, and $H(A,B)$ is their joint entropy. $H(A)$, $H(B)$ and $H(A,B)$ are defined as

$$\begin{aligned} H(A) &= -\sum P(a)\log(P(a)) \\ H(B) &= -\sum P(b)\log(P(b)) \\ H(A,B) &= -\sum P(a,b)\log(P(a,b)) \end{aligned} \quad (2)$$

MI measures the amount of predictable information in one variable from that in the other. In intensity-based registration methods, images are assumed as random variables. More importantly, it is assumed that a maximum value of MI can be achieved if images are geometrically aligned. In Eq.1, the maximum value of MI is achieved when the joint entropy is minimized. That is, in the case of misregistration, corresponding elements of images such as regions are duplicated in the joint pdf, which increases the value of the joint entropy. In registered data, on the other hand, corresponding elements appear once, which yields smaller joint entropy.

In order to adapt MI for registration of optical images and LiDAR data, the 3D LiDAR dataset is treated as a 2D image, such that the elevation and intensity values in the LiDAR data are both treated as image grey values.

3.2 Combined Mutual Information

The statistical dependence of multiple variables can be measured by multivariable generation of MI via a higher order joint entropy (McGill, 1954; Watanabe, 1960). The MI of three variables A , B and C is defined as

$$\begin{aligned} I(A;B;C) &= H(A) + H(B) + H(C) - H(A,B) \\ &\quad - H(B,C) - H(A,C) + H(A,B,C). \end{aligned} \quad (4)$$

While multivariable MI supports high dimensional observations, it does not take into account the known relative geometry between the variables.

In order to consider existing geometric dependence between a pair of variables, Combined MI has been proposed (McGill, 1954). Combined MI of a variable C and two already aligned variables A and B is defined as

$$I((A,B);C) = H(A,B) + H(C) - H(A,B,C). \quad (5)$$

Eq. 5 can be rewritten using the chain rule of multivariable MI (Cover and Thomas, 1991) as

$$I((A,B);C) = I(A,C) + I(B,C) - I(A;B;C). \quad (6)$$

The schematic presentation of multivariable and combined MI of three variables is given in Figure 1. The darker area presents information among variables.

$I((A,B);C)$ measures the amount of information of C that would be gained from the pair of (A,B) simultaneously. Since LiDAR intensity data is inherently registered with the DSM (3D point

cloud), it provides complementary information. In this paper, the CMI is applied to exploit both the geometric dependence and complementary character of intensity data.

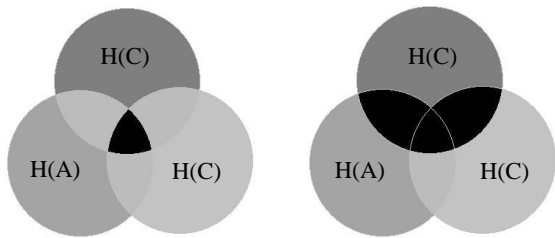


Figure 1. Venn diagram of multivariable (left) and Combined MI (right).

3.3 PDF Estimation

According to Eqs. 1 and 2, the statistical dependence between the image and LiDAR data is calculated using the probability density function. It can be estimated by either parametric or non-parametric approaches. Histogramming as a non-parametric approach of pdf estimation has been adopted in this work because of its simplicity and efficiency (Bishop, 1995). In histogramming, each entry of the joint histogram of the datasets denotes the number of elements of data with coinciding intensities. Dividing the joint histogram by the total number of members in the data yields the joint pdf, and the marginal pdf of each dataset can be found by summing over the rows and columns, respectively.

The number of bins plays a crucial role as a smoothing parameter in the process of pdf estimation. As an example, a high number of bins (small bin) causes a spiky pdf, while the use of a small number of bins (large bin) provides a very smooth pdf. In addition, the number of bins selected affects both the MI value and the results of registration. In other words, an optimum number of bins can preserve the necessary information for similarity measurement within the datasets, as well as reducing the cost of computation, particularly in the case of estimating a higher dimensional pdf (multivariable pdf).

Conventional statistical methods for finding the optimal number of bins (Scott, 1979; Freedman and Diaconis, 1981), namely that which preserves information with the fewest bins, cannot be utilised in estimation of the pdf in MI-based methods due to ignoring the effect of number of bins on the statistical similarity between the datasets. Therefore, the optimum number of bins is defined based on the number yielding the maximum value of MI, as well as the sharpest convergence surface that is measured by the kurtosis value. While the higher number of bin might highlight the role of noise, a lower number of bins can reduce the MI value through a loss of fine detail information.

3.4 Transformation Model

In order to register the overlapping images, a search for the maximum value of MI is required. Conventionally, one of the images is assumed as ‘fixed’ and the other as ‘moving’. The moving image is transformed under different values of transformation model parameters to find the maximum value of MI. This approach may not be feasible in the registration of high resolution satellite and aerial imagery with LiDAR point clouds due to the high cost of computation, particularly in the case of transformation models with a high number of parameters. Therefore, a new approach that determines the

similarity of the fixed and moving image in well-distributed small patches, called templates, is adopted. Then, the centres of the templates are used in the computation of transformation parameters. Only a 2D search for maximum MI is required and this decreases the risk of failure in the registration. Furthermore, an overdetermined solution for the parameters of registration supports the detection of blunders in template transformation, which improves the robustness and reliability of registration.

4. IMPLEMENTATION

In the reported implementation, a multi-resolution strategy based on a Gaussian pyramid is adopted to increase the speed of computation and provide more reliable results. The flow chart for registration is shown in Figure 2.

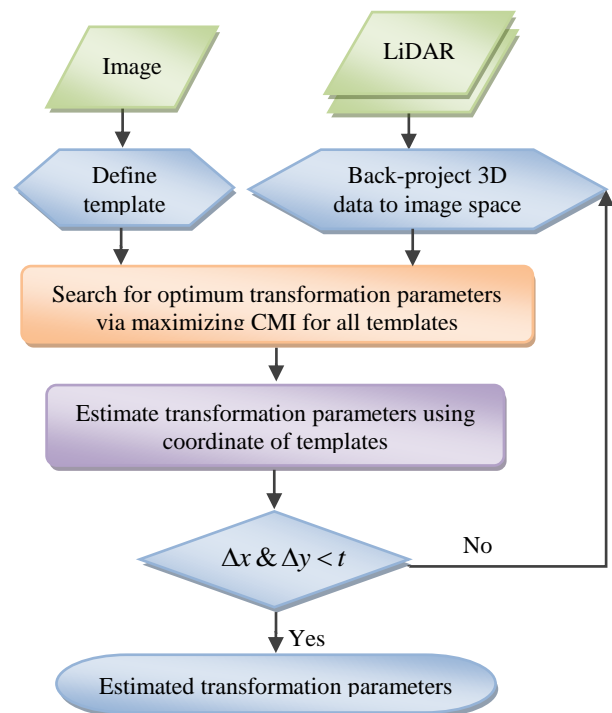


Figure 2. Flowchart of registration process.

In the estimation of the pdf, the use of 32 bins achieved the maximum value of MI and the sharpest peak. Also, the 3D affine transformation (Yamakawa and Fraser, 2004) and collinearity equations were adopted as projection models for the satellite and aerial imagery, respectively. LiDAR point clouds were then back-projected to image space using the projection model to bring the data into the same coordinate system.

For every template, the search for the maximum value of combined MI starts with the use of Powell’s optimisation method in the x and y directions, because of its performance and accuracy in MI-based registration (Maes et al., 1999). The obtained coordinates of templates are used for the estimation of the parameters of 3D affine transformation for satellite imagery, and of the exterior orientation parameters (EOPs) of the aerial imagery. Templates with detected blunders are removed from computation of parameters. Finally, in order to achieve sub-pixel accuracy for registration, the process of template matching and computation of parameters is iterated and the parameters with the minimum improvement in position of templates are considered as optimal values. This yields the final 3D affine parameters or EOPs for the imagery at the lowest level (highest resolution) of the pyramid. The use of this strategy greatly

improves the speed of computation and robustness of registration. Since the LiDAR point cloud is considered as the reference data, registration of satellite and aerial imagery also implies geo-referencing.

5. EXPERIMENTAL RESULTS AND DISCUSSION

This section presents selected experimental results obtained with the proposed local Combined MI-based registration method. Two experimental registrations, namely the registration of, first, a satellite image and second, an aerial image to LiDAR are reported.

5.1 Registration of a satellite image to LiDAR data

The imagery for first experimental dataset comprised a WorldView2 level 1A satellite image, acquired in 2010 and shown in Figure 3, covering the South Wharf area (2.5 km by 6 km) of Melbourne, Australia. The corresponding LiDAR point cloud, with intensity data, was acquired in 2007 at a density of 4 pts/m². It is noteworthy that the WorldView2 image and LiDAR data were acquired three years apart, and changes had occurred within the scene. In addition, the satellite imagery was affected by perspective distortion, occlusions and some shadowing

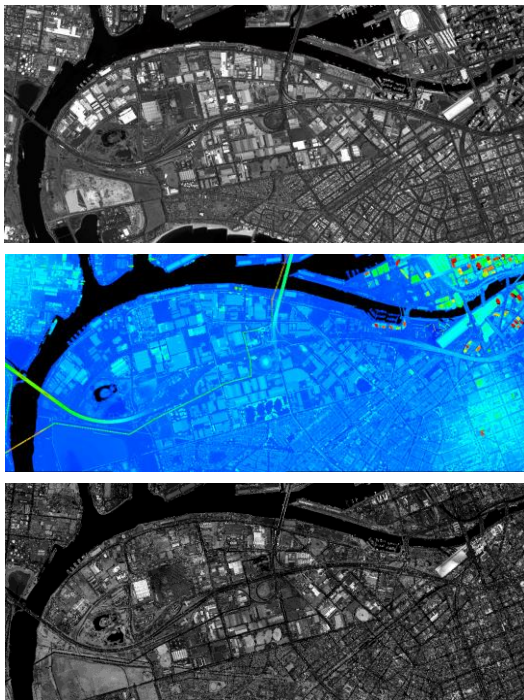


Figure 3. Satellite image (top), colour-coded LiDAR DSM (middle) and LiDAR intensity image (bottom).

An array of 2 by 6 well-distributed templates with 250 by 250 pixels was utilised to estimate the parameters of 3D affine transformation. An example of a favourable template and the corresponding peaks for the MI and CMI of the WorldView2 imagery with LiDAR data are shown in Figure 4. Also, the results of registration of imagery and LiDAR data of the template are shown in Figure 5. Since the LiDAR intensity image provides more common regions with the satellite image, its MI peak is sharper than the MI peak for the image and LiDAR DSM. The CMI, which uses both the LiDAR point cloud and intensity data, yielded a higher value of MI as well as

a sharper peak. A colourisation of the LiDAR point cloud with registered pansharpened satellite image is shown in Figure 6 to illustrate the quality of registration.

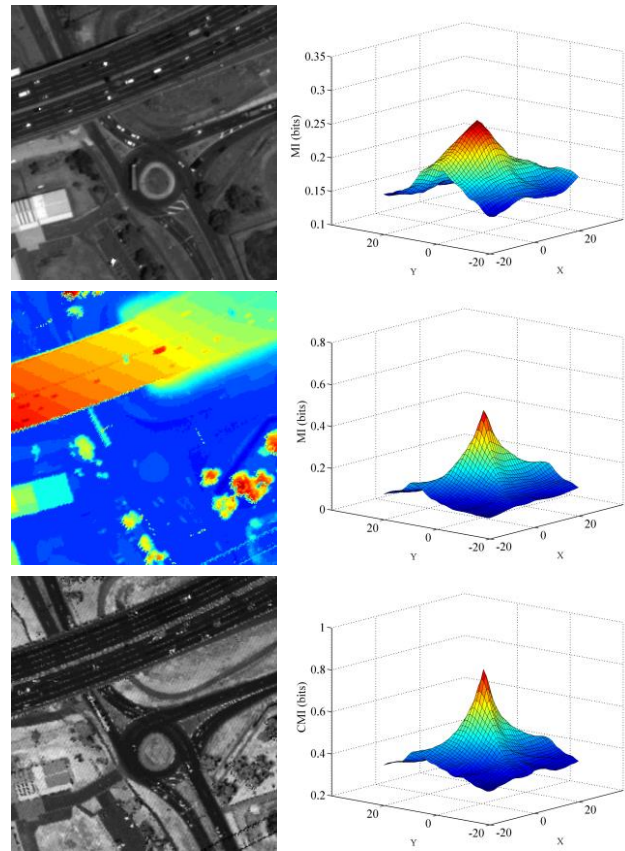


Figure 4. Satellite image, LiDAR DSM and intensity image (left column). Also, MI convergence surface of LiDAR DSM and intensity image and CMI convergence surface of image and LiDAR data for translation in x and y directions (right column).

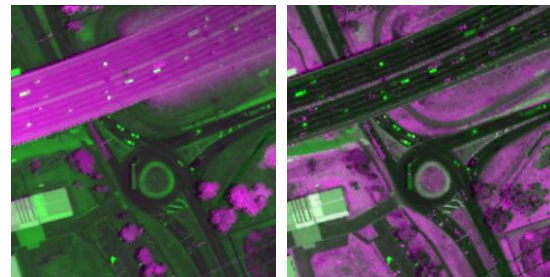


Figure 5. Registered WorldView2 image (green) with LiDAR (magenta) DSM and intensity image (left and right).

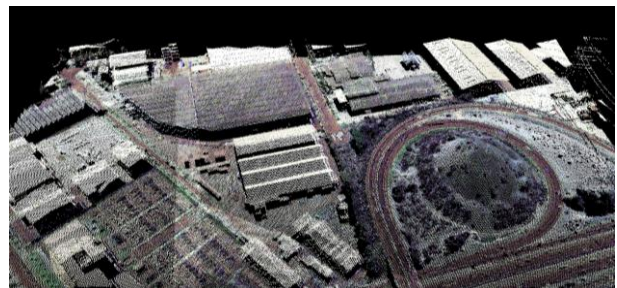


Figure 6. LiDAR point cloud colourised by registered image.

5.2 Registration of an aerial image to LiDAR data

The second experimental dataset covered a 2 km x 2 km suburban area of Bathurst, New South Wales, Australia. It comprised 20 cm ground sample distance (GSD) aerial imagery and a LiDAR point cloud with a density of 4 pts/m² acquired in 2011, again with intensity data. The aerial image and its corresponding LiDAR DSM and intensity image are shown in Figure 7.

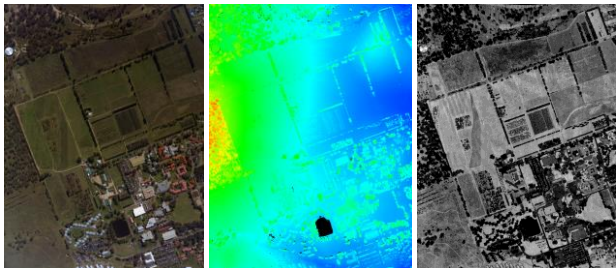


Figure 7. Aerial image (left), colour-coded LiDAR DSM (middle) and LiDAR intensity image (right).

It is noteworthy that the resolution of the dataset is sufficiently high to clearly demonstrate natural and artificial (man-made) objects in the aerial image and LiDAR data. The LiDAR points were initially back-projected to the image coordinate system, and in this case the LiDAR data was in the form of a raster DSM.

As shown in Figure 7, the LiDAR DSM highlights areas of significant elevation difference, whereas it provides limited information for matching to imagery in flat areas. LiDAR intensity data, on the other hand, offers more information for similarity determination with the aerial image, e.g. roads. The complementary nature of the LiDAR DSM and intensity image is favourably exploited in the determination of similarity between the aerial imagery and LiDAR data in the proposed CMI method.

An array of 6 by 4 well-distributed templates with 500 by 500 pixels was used to estimate the EOPs of the aerial image. An example template, demonstrating a natural scene, and the corresponding peaks for the MI and CMI of the aerial imagery with LiDAR data are shown in Figure 8. The convergence surface of MI of the image and LiDAR DSM does not provide a distinct peak and fails in finding sufficient similarity between the image and LiDAR DSM. The LiDAR intensity image, which contains more common region with the optical image, provides a distinct and sharp peak. The CMI, which exploits both LiDAR DSM and intensity data, offers the sharpest peak with the highest value of MI. It is clear that a sharper peak of MI speeds up the optimization procedure and increases the reliability and robustness of registration via avoidance of local maxima. MI of image with LiDAR DSM and intensity data succeeded for 50% and 90% of the templates, respectively. The CMI approach, on the other hand, achieved true registration in 98% of the templates. Nevertheless the CMI increases the success rate of MI-based registration method; it may still fail for individual templates that do not contain sufficient information. These must be removed in the process of estimating transformation parameters. The results of registration of the optical image to LiDAR DSM and intensity data, shown in Figure 9, indicate a precise match between corresponding features in the dataset.

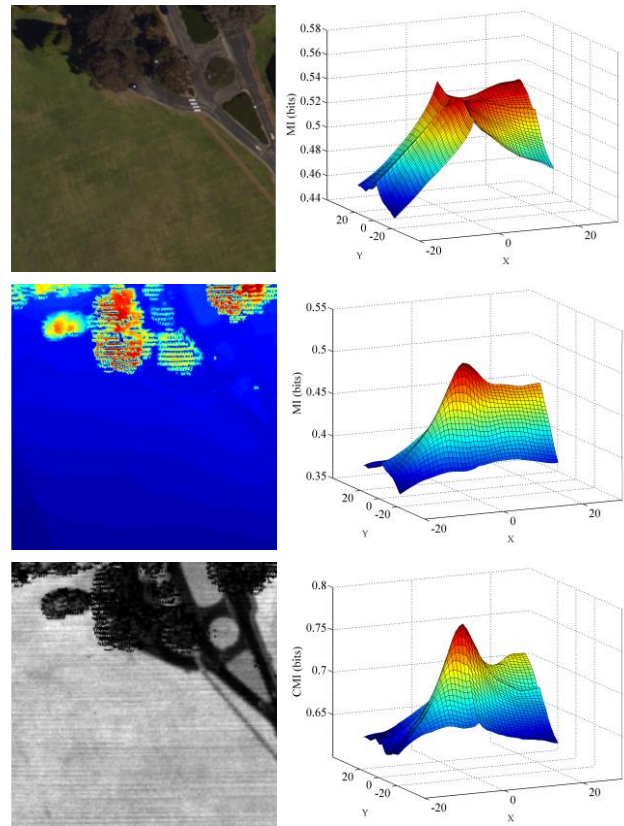


Figure 8. Aerial image, LiDAR DSM and intensity image (left column). Also, MI convergence surface of LiDAR DSM and intensity image and CMI convergence surface of image and LiDAR data for translation in x and y directions (right column).

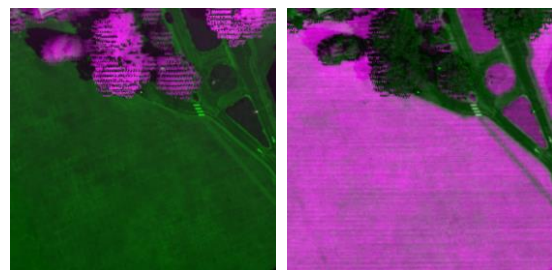


Figure 9. Registered aerial image (green) with LiDAR (magenta) DSM and intensity image (left and right).

In addition, a small cut out portion of the overlaid, back-projected LiDAR point cloud and aerial image is shown in Figure 10.

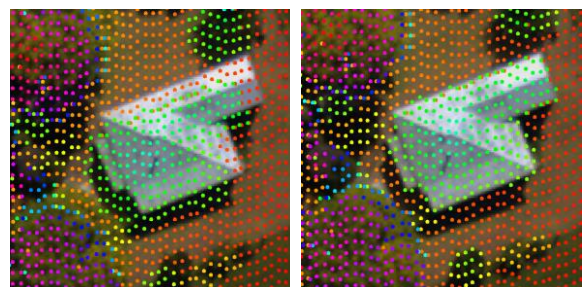


Figure 10. Overlay of back-projected LiDAR point clouds and aerial image before and after registration (left and right).

As shown in Figure 10, natural and artificial features of LiDAR and imagery data are matched very well, which indicate the quality of registration. Finally, a colourisation of LiDAR point cloud with registered colour image is shown in Figure 11.



Figure 11. Colourised LiDAR point cloud.

6. CONCLUSIONS

The ability of the proposed local combined mutual information for the registration of optical imagery and LiDAR data has been highlighted in this paper. It should be recalled that such a 3D point cloud-to-2D image registration cannot be readily handled by conventional feature-based or area-based matching methods. Based on testing involving a high number of datasets, it has been observed that MI can fail to deliver acceptable results for the registration of optical imagery to 3D LiDAR point clouds, especially of course when there is limited elevation variation within the point cloud. However, with the CMI method, which employs both LiDAR point cloud and intensity data, it has been found that more robust registration performance is achieved. Also, the proposed method of selecting the optimal number of bins for estimation of the pdf has increased the value of similarity measures and the speed of computation. Moreover, the use of local similarity measures can produce a robust registration even in the presence of blunders within transformation of templates. The proposed approach provides more robust results than conventional MI, and can be used as an efficient tool for automated registration of multi-sensor, multi-temporal and multi-resolution remote sensing data for a wide range of applications.

ACKNOWLEDGEMENTS

The Authors would like to thank the Victorian Department of Sustainability and Environment (www.dse.vic.gov.au) and Department of Land and Property Information of New South Wales (www.lpi.nsw.gov.au) for providing the LiDAR data and images for Melbourne and Bathurst datasets, respectively.

REFERENCES

- Bishop, C.M., 1995. *Neural networks for pattern recognition*. Oxford university press.
- Corsini, M., Dellepiane, M., Ponchio, F., Scopigno, R., 2009. Image-to-Geometry Registration: a Mutual Information Method exploiting Illumination-related Geometric Properties, in: *Pacific Graphics*.
- Cover, T.M., Thomas, J.A., 1991. *Elements of information theory*. John Wiley.
- Freedman, D., Diaconis, P., 1981. On the histogram as a density

estimator: L 2 theory. *Probability theory and related fields* 57, 453–476.

Goshtasby, A.A., 2012. *Image registration: principles, tools and methods*. Springer-Verlag New York Incorporated.

Habib, A., Ghanma, M., Morgan, M., Al-Ruzouq, R., 2005. Photogrammetric and LIDAR Data Registration Using Linear Features. *Photogrammetric Eng. Remote Sens* 71, 699–707.

Hajnal, J. V, Hill, D.L.G., Hawkes, D.J., 2001. *Medical image registration*. CRC.

LeMoigne, J., Netanyahu, N.S., R.D., E., 2011. *Image Registration for Remote Sensing*. Cambridge University Press.

Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P., 1997. Multimodality image registration by maximization of mutual information. *Medical Imaging, IEEE Transactions on* 16, 187–198.

Maes, F., Vandermeulen, D., Suetens, P., 1999. Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information. *Medical Image Analysis* 3, 373–386.

Mastin, A., Kepner, J., Fisher, J., 2009. Automatic registration of LIDAR and optical images of urban scenes, in: *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference On*. pp. 2639–2646.

McGill, W., 1954. Multivariate information transmission. *Information Theory, IRE Professional Group on* 4, 93–111.

Parmehr, E.G., Zhang, C., Fraser, C.S., 2012. Automatic Registration of Multi-source Data Using Mutual Information, in: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. pp. 303–308.

Pluim, J.P.W., Maintz, J.B.A., Viergever, M.A., 2003. Mutual-information-based registration of medical images: a survey. *Medical Imaging, IEEE Transactions on* 22, 986–1004.

Rönnholm, P., 2011. Registration quality - toward integration of laser scanning and photogrammetry. *EuroSDR*.

Scott, D.W., 1979. On optimal and data-based histograms. *Biometrika* 66, 605–610.

Shannon, C.E., Weaver, W., Blahut, R.E., Hajek, B., 1949. *The mathematical theory of communication*. University of Illinois press Urbana.

Shin, S.W., Habib, A.F., Ghanma, M., Kim, C., Kim, E.M., 2007. Algorithms for multi-sensor and multi-primitive photogrammetric triangulation. *ETRI journal* 29, 411–420.

Toth, C., Ju, H., Grejner-Brzezinska, D., 2011. Matching between different image domains. *Photogrammetric Image Analysis* 37–47.

Viola, P., Wells III, W.M., 1997. Alignment by maximization of mutual information. *International journal of computer vision* 24, 137–154.

Watanabe, S., 1960. Information theoretical analysis of multivariate correlation. *IBM Journal of research and development* 4, 66–82.

Yamakawa, T., Fraser, C., 2004. The affine projection model for sensor orientation: experiences with high-resolution satellite imagery. *International Archives of Photogrammetry and Remote Sensing* 8, 1–6.