REVIEW ARTICLE

Review of geometric fusion of remote sensing imagery and laser scanning data

Bo Wu* and Shengjun Tang

Department of Land Surveying and Geo-Informatics, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

(Received 9 December 2014; accepted 13 February 2015)

Imagery and laser scanning data are two major sources of 3D information. Each dataset has distinct characteristics that render it preferable for certain applications. The fusion of imagery and laser scanning data is a prerequisite to utilising the complementary characteristics of both datasets. In the past decade, a number of methods have been developed for the geometrical fusion of the two types of datasets for better 3D mapping in various applications. This article presents a systematic review of these methods. First, comparative analysis of the derivation of 3D information from imagery through photogrammetry and laser scanning is presented. Then, three categories of methods for the geometric fusion of imagery and laser scanning data are detailed, namely, laser scanning data used as controls for imagery, imagery used as controls for laser scanning data and the combined adjustment of imagery and laser scanning data. The advantages and limitations of the three categories of methods are analysed. Finally, suggestions for future research in this area are discussed, and concluding remarks are given.

Keywords: imagery; laser scanning; photogrammetry; geometric fusion

1. Introduction

The methods used to derive three-dimensional (3D) information from remote sensing technology generally fall into two distinct types. The first is based on imagery collected by cameras through the photogrammetric process, and the second is laser scanning using laser light to measure the distance from a sensor to the object directly. Both types can be applied to terrestrial, airborne, or space-borne platforms. The advantages of imagery use include a high degree of image clarity, highly detailed texture information and a relatively low cost. Furthermore, imagery-derived 3D information has a long history in the fields of photogrammetry and computer vision (Helava 1991, Schenk 1999, Fraser et al. 2002). Laser scanning, in contrast, has been one of the most important technological advances of the past decade. It offers delivery times for 3D information, which are much faster than those of traditional photogrammetric techniques, and provides extensive information containing a mass of detail due to the extraordinary number of points involved (Vosselman 2000, Shan and Toth 2008). Imagery and laser scanning data have distinct characteristics that render them preferable in certain applications. The respective advantages and disadvantages of the two types of datasets suggest that their fusion would achieve better performance in various applications than that can be achieved using a single type of data (Baltsavias 1999, Beraldin 2004, Wu et al. 2011, Di et al. 2012, 2014).

*Corresponding author. Email: bo.wu@polyu.edu.hk

© 2015 Taylor & Francis
The fusion of imagery and laser scanning data has been pursued in a range of areas, including topographic mapping (Anderson and Parker 2002, Schenk and Csathó 2002, Lee 2003, Wu et al. 2011, Di et al. 2012, 2014), object extraction (Baltsavias et al. 2001, Fraser et al. 2002, Forkuo and King 2004, Sohn and Dowman 2007), image classification (Syed et al. 2005, Li et al. 2007a) and environmental monitoring (Lennon et al. 2006, Erdody and Moskal 2010, Valbuena et al. 2011). This article focuses on the geometric fusion of imagery and laser scanning data for topographic mapping or 3D surface reconstruction. The investigation is based on the following two main reasons. First, advances in surveying technology have made imagery and laser scanning datasets more readily available in recent years, and it is not unusual to see inconsistencies between the 3D information generated from the two types of data collected by different sensors in the same area. Second, 3D information generated from imagery through the photogrammetric process generally provides better accuracy in the horizontal direction than in the vertical direction (Li et al. 2007b, Wu et al. 2011, 2014), whereas laser scanning data are known to produce better vertical than horizontal accuracy (Baltsavias 1999, May and Toth 2007, Qiao et al. 2010). Therefore, the geometric fusion of the two types of datasets has the potential to produce 3D information that is both more precise and consistent.

This article presents a systematic review of the geometric fusion of remote sensing imagery and laser scanning data, which can be classified into three categories: (1) laser scanning data used as controls for imagery in the photogrammetric process, (2) imagery used as controls for the adjustment of laser scanning data and (3) the combined adjustment of imagery and laser scanning data. Figure 1 illustrates the three categories and presents a timeline of the most representative methods developed in each category. In the following discussion, the performance and accuracy of the geometric fusion of the two datasets are indicated by comparing the results with certain ground truth data or some known accurate data source.

The remainder of the article is organised as follows. The next section presents a comparison of the two major methods for deriving 3D information: photogrammetric processing of imagery and laser scanning. The subsequent section then presents the three categories of the geometric fusion of imagery and laser scanning data in detail and analyses the advantages and disadvantages of the methods in each category. Finally, suggestions for future research are made and concluding remarks given.

2. Comparison of 3D information derivation from imagery via photogrammetry and from laser scanning

Before discussing 3D information derivation from imagery via photogrammetry and from laser scanning, it is necessary to compare the two types of data sources or techniques. The following presents such a comparison in terms of data acquisition, data characteristics and data accuracy. It should be noted that the comparison is based on imagery and laser scanning data acquired through airborne platforms, although the general observations are also applicable to datasets collected from space-borne or terrestrial platforms.

(1) Data acquisition: flying height and time, swath, flight planning and dependence on the environment are discussed for imagery and laser scanning data acquisition through airborne platforms. Owing to the restrictions of the platforms, laser power, detector sensitivity and high pulse rates in the maximum unambiguous range, the flying height of laser scanning systems is generally restricted to no more than 1 km, whereas that of airborne image acquisition platforms is typically in the range of 3–4 km (Baltsavias 1999). To ensure data accuracy and overlapping

Downloaded by [Hong Kong Polytechnic University] at 19:27 04 October 2015
strips, the flying speed of airborne laser scanning platforms may also be restricted. The lower flying height and speed of laser scanning systems result in a narrower swath and longer flying time relative to airborne image acquisition systems covering the same region (Schenk 1999, May and Toth 2007). Flight planning is also relatively simple and well established for airborne image acquisition systems, whereas that for airborne laser scanning systems can be quite complicated because of the relatively narrow strip width, low flying platform and the laser’s range of influence. With regard to dependence on the environment, as active sensors, laser scanning systems can theoretically operate 24 hours a day, with the detection of terrain objects completely dependent on the laser wavelength, whereas the land cover conditions play an important role in the quality or accuracy of the final data derived from image acquisition systems. For example, better digital terrain models (DTMs) can be generated from images obtained via photogrammetric techniques in areas in which the trees have no leaves.

(2) Data characteristics: it is obvious that imagery offers full coverage of the imaged scene. Images provide detailed texture information that can be used to interpret objects, and image contrast and boundary information benefit photogrammetric
surface reconstruction. Airborne remote sensing imagery normally has a spatial resolution of about 0.15 m/pixel. Through dense image matching and the photogrammetric process, DTMs with a resolution of 0.5 m (three times the image resolution) can be expected. However, in areas with poor textural conditions (e.g. homogeneous textures), fewer image points are likely to be matched, resulting in DTMs with sparse resolution. Instead of full coverage, laser scanning data offer point-wise sampling, meaning that no images, or sometimes monocular images, are acquired. In airborne laser scanning, the generated 3D point cloud normally has a point spacing of 0.5 m (May and Toth 2007), which can be used for high-resolution surface reconstruction. Although laser scanning data feature a high point density, the derived DTMs often tend to be smooth and leave out important geomorphologic features (e.g. road segments, building boundaries) (Schenk and Csathó 2002). DTMs derived from images through the photogrammetric process, in contrast, preserve distinct geomorphologic features that provide distinct image textures on which image matching is relatively easy and more reliable. In terms of data processing to obtain 3D information, laser scanning data inherently contain 3D coordinates. Only minor manual editing is needed for error correction and gap filling. In addition, terrain objects can be filtered out from laser scanning data automatically. Deriving 3D information from imagery through the photogrammetric process, in contrast, requires a greater amount of manual intervention. Although interior orientation is completely automated in photogrammetry, sensor orientation is difficult to automate fully. Image matching is also a challenging task in areas with poor textural conditions, rendering considerable human interaction necessary. Cost is also an important factor in data selection. The initial equipment cost and shorter lifetime result in substantially higher costs for laser scanning data relative to imagery, which benefits from its ease of availability in recent years.

(3) Data accuracy: data accuracy is a critical factor in evaluating the quality of the two types of datasets. According to the principle of photogrammetry, the height accuracy of 3D information derived from imagery through photogrammetry is determined by the base-to-height ratio of the overlapping imagery (i.e. stereo pairs) collected along and between the sensor flight paths. The horizontal accuracy (accuracy of object coordinates X and Y) is directly proportional to the image scale and is constant in relation to the image. Photogrammetric image processing generally provides better accuracy in the horizontal direction than in the vertical direction (Li et al. 2007b, Wu et al. 2014). The accuracy of laser scanning data depends on the accuracy of the navigation solution, boresight misalignment angles, range and scan angle accuracy and laser beam divergence (May and Toth 2007). The vertical accuracy of these data is reportedly two to five times better than their horizontal accuracy (Baltsavias 1999, Xhardé et al. 2006, Qiao et al. 2010). For example, Xhardé et al. (2006) analysed airborne laser scanning data collected in the south coast of Canada’s Gaspé Peninsula in 2003 and 2004, and reported that the root mean square error (RMSE) values were about 54 cm in the horizontal direction and about 16.5 cm in the vertical direction. In measurements from imagery, the degree of accuracy is relatively homogeneous within the image format, whereas that of laser scanning data is generally less uniform across the survey area because attitude errors may lead to a rapid decrease in height accuracy with an increase in scan angle, particularly at a greater flying height.

Table 1 summarises the foregoing comparative analysis of the two types of datasets. From the foregoing discussion, it is clear that imagery and laser scanning data have distinct
Table 1. Comparison of 3D information derivation from imagery via photogrammetry and laser scanning data collected from airborne platforms.

<table>
<thead>
<tr>
<th>Aspects</th>
<th>Imagery via photogrammetry</th>
<th>Laser scanning data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data acquisition</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flying height</td>
<td>3–4 km</td>
<td>Up to 1 km</td>
</tr>
<tr>
<td>Flying speed</td>
<td>Faster</td>
<td>Slower</td>
</tr>
<tr>
<td>Flying time</td>
<td>Less</td>
<td>More</td>
</tr>
<tr>
<td>Swath</td>
<td>Typically less than 108°</td>
<td>Typically less than 40°</td>
</tr>
<tr>
<td>Flight planning</td>
<td>Simple and well established</td>
<td>More complicated</td>
</tr>
<tr>
<td>Sensor</td>
<td>Passive</td>
<td>Active</td>
</tr>
<tr>
<td>Dependence on the environment</td>
<td>Closely correlated with land cover</td>
<td>No direct dependence on the environment</td>
</tr>
<tr>
<td>Data characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data presentation</td>
<td>Full coverage of the scene and detailed texture information</td>
<td>Point-wise sampling and no or monochromatic imaging</td>
</tr>
<tr>
<td>Data resolution</td>
<td>Typical image resolution: 0.15 m/pixel (DTMs with a resolution of 0.5 m can be generated)</td>
<td>Typical point spacing: 0.5 m</td>
</tr>
<tr>
<td>Feature preservation</td>
<td>High performance in preserving geomorphological features</td>
<td>Some geomorphological features may be missing</td>
</tr>
<tr>
<td>Automation</td>
<td>More manual intervention and editing for sensor orientation and image matching</td>
<td>Fully automated for collecting raw 3D point clouds and minor manual editing for error correction, etc.</td>
</tr>
<tr>
<td>Cost</td>
<td>Relatively low cost</td>
<td>Relatively high cost</td>
</tr>
<tr>
<td>Data accuracy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factors</td>
<td>Height accuracy is determined by the base-to-height ratio, and horizontal accuracy is proportional to the image scale</td>
<td>Related to the accuracy of the navigation solution, boresight misalignment angles, range and scan angle accuracy and laser beam divergence</td>
</tr>
<tr>
<td>Accuracy</td>
<td>Better horizontal accuracy than vertical accuracy</td>
<td>Better vertical accuracy than horizontal accuracy</td>
</tr>
<tr>
<td>Characteristics</td>
<td>More homogeneous within the image format</td>
<td>Accuracy may not be uniform over the survey area</td>
</tr>
</tbody>
</table>
characteristics that render each preferable in certain applications. Additional differences exist with respect to their technological maturity and potential for further development (Baltsavias 1999). With regard to geometric accuracy in particular, it should be noted that 3D information generated from imagery obtained through the photogrammetric process is generally more accurate in the horizontal than the vertical direction, whereas the reverse is true for laser scanning data. Moreover, due to the greater availability of both types of datasets in recent years, there are more imagery and laser scanning data covering the same area and inconsistencies at different levels between the 3D information generated from them. There is thus strong demand to integrate the two types of datasets to generate 3D information of better precision and consistency. The following sections present a number of strategies for such integration.

3. Laser scanning data used as controls for imagery

The most straightforward means of effecting the geometric fusion of laser scanning data and imagery is to use the former as a control in the photogrammetric processing of the latter. An early example was presented by Ebner et al. (1991), who used additional information (e.g. DTMs interpolated from laser scanning data) as control information for the aerial triangulation of images. The main focus of their approach was on ensuring the accuracy of medium- and small-scale photogrammetry for aerial and satellite image processing by minimising the differences between the heights derived from stereo images and using the interpolated heights from a DTM as constraints. Jaw (2000) proposed a method in which additional height information (from a DTM or laser scanning data) is integrated into the aerial triangulation workflow by hypothesising plane observations in the object space. The estimated object points derived from image measurements together with the adjusted surface points provide a better point group for describing the surface. Teo et al. (2010) investigated the block adjustment of three SPOT-5 images using a DTM as the elevation control. Common tie points were first identified from the images, with the initial ground coordinates of the tie points derived from space intersection using the image orientation parameters. The elevation coordinates of the tie points were then interpolated from the DTM based on their planimetric coordinates, and the points’ ground coordinates were further adjusted through an iteration procedure in the block adjustment. The experimental results indicated that using the DTM as an elevation control can improve geometric accuracy and reduce the geometric discrepancies between images.

Earlier work focused on multi-source data integration for topographic mapping and surface reconstruction. Toutin (1994) presented a method for integrating multi-source images and existing DTMs to generate high-precision orthoimages. Toutin’s mathematical model was based on the co-linearity condition and represented the physical realities of the full viewing geometry. He tested eight types of images from different sensors with different resolutions, with the results showing absolute accuracy of one-third of a pixel for satellite images in the visible range, one resolution cell for satellite radar images, and one to two resolution cells for airborne images. Fei et al. (2012) investigated the integration of LiDAR (light detection and ranging) data and other multi-source environmental data, including SPOT images, Shuttle Radar Topography Mission data, and 1:10000 vector geological data. They proposed a point cloud production method based on topographical constraints. A pyramid model and a linear quad tile-index method were processed to manage the data, thereby achieving multi-source data integration with multiple scales and fast browsing.

Research on the integration of laser scanning data and imagery has recently become more notable in the planetary mapping community. Examples include Mars and lunar
topographic mapping. In these cases, the imagery is obtained through space-borne cameras, and the laser scanning system is a space-borne laser altimeter. The integrated processing of Mars orbital imagery and laser altimeter data has been a topic of considerable interest since the launch of NASA’s Mars Global Surveyor (MGS) in November 1996. There is a Mars orbiter camera (MOC) and Mars orbiter laser altimeter (MOLA) on board the MGS spacecraft. MOC incorporates three instruments: a narrow-angle camera that obtains greyscale (black-and-white) high-resolution images (typically 1.5–12 m/pixel) and red and blue wide-angle cameras for context (240 m/pixel) and daily global imaging (7.5 km/pixel). A large number of Mars DTMs with various resolutions have been generated using MOC imagery. In addition, MOLA mapped the majority of the Martian topography, and a DTM with a resolution of 128 pixels per degree for the planet (from −87 to +87) was released in 2003. Because of the orbital characteristics of the MGS, its resolution is latitude-dependent and highest near the poles. Kirk et al. (2003) used MOC images to produce DTMs with horizontal resolutions of 10 m to provide the topographic and slope information needed to assess the safety of candidate landing sites for the Mars Exploration Rover (MER). These DTMs are controlled by the MOLA global dataset, and are consistent with it at the limits of resolution. Dickson et al. (2007) registered MOC images with interpolated MOLA data to analyse the distribution, local topographic setting, morphology, orientation, elevation and slopes of Martian gullies. Anderson and Parker (2002) examined the precision registration between MOC imagery and MOLA data at select candidate landing sites for MER. Finally, Sakimoto et al. (1999) also used MOC imagery and MOLA data to examine the regional and local variations in the relief, slopes, vertical roughness and relative elevations of the Medusae Fossae Formation on Mars.

In addition to MOC imagery, images acquired by the high-resolution stereo camera (HRSC) on board the European Space Agency’s Mars Express spacecraft launched in June 2003 have also been widely used in Mars topographic mapping. The HRSC imaged the entire planet in full colour with a resolution of about 10 m/pixel, although select areas were imaged at a resolution of 2 m/pixel. A number of researchers have developed methods for using MOLA data as controls for the photogrammetric processing of HRSC imagery. Albertz et al. (2005) constituted the first endeavour in this respect. They improved the image exterior orientation parameters with the assistance of MOLA data to generate better quality DTMs. Spiegel (2007) developed a bundle adjustment technique for HRSC imagery in which a sparse stereo point cloud is adjusted to optimise its fit to a surface interpolated from MOLA data. This technique is now standard in the production of controlled orthorectified HRSC products. Ebner et al. (2008) combined HRSC images and MOLA data using a method in which the two datasets are processed in the same bundle adjustment and the HRSC images are adjusted to fit the MOLA data. Gwinner et al. (2009) also produced high-quality DTM products from the integration of HRSC imagery and MOLA data. Their proposed method successfully created a DTM of 50 m resolution from HRSC stereo imagery with a resolution of 12 m/pixel. The differences between the DTM produced by the proposed method and that from the MOLA data ranged from 30 to 40 m. Lin et al. (2010) developed a co-registration approach to align Mars DTMs derived from various sources, i.e. HRSC imagery, HiRISE imagery (from the high-resolution imaging science experiment, a camera on board NASA’s Mars Reconnaissance Orbiter launched in August 2005) and MOLA data. Gwinner et al. (2009) presented an assessment of the surface matching parameters for the co-registration of multi-source Mars DTMs. DTMs generated from multi-source data integration have been used in various studies.
Recent lunar exploration missions, such as China’s Chang’E-1 (launched in October 2007) and Chang’E-2 (launched in October 2010) (Ouyang et al. 2010), Japan’s selenological and engineering explorer (SELENE)/Kaguya (launched in September 2007) (Kato et al. 2008), India’s Chandrayaan-1 (launched in October 2008) (Vighnesam et al. 2010) and NASA’s lunar reconnaissance orbiter (LRO) (launched in June 2009) (Vondrak et al. 2010), have all involved the use of on board orbiter cameras and laser altimeters of different configurations. These missions have collected a vast number of lunar surface images and elevation measurements at different resolutions and levels of uncertainty. Most of the previously related research has processed the orbiter imagery and laser altimeter data for lunar topographic mapping separately (Araki et al. 2009, Mazarico et al. 2012, Di et al. 2012, Haruyama et al. 2012, Scholten et al. 2012), and there are usually inconsistencies between the derived lunar DTMs.

Di et al. (2012) presented a method for co-registering the DTMs derived from Chang’E-1 stereo images to Chang’E-1 laser altimeter data to reduce the inconsistencies between them. An iterative closest point (ICP) algorithm was used to register the 3D points derived from the stereo Chang’E-1 images to the laser altimeter points. The key function of the ICP algorithm is to search for pairs of the closest points in the two datasets, estimating the rigid transformation and iteratively refining that transformation by repeatedly generating pairs of the closest points in the two sets by minimising an error metric. On the basis of the transformation parameters obtained, the image orientation parameters can be refined in such a way that the DTMs from the images and laser altimeter data are co-registered. Guo and Wu (2012) presented a method for co-registration between the DTMs derived from Chang’E-1 imagery and SELENE laser altimeter data. Seven transformation parameters (one scale factor, three rotations and three translations) were calculated through a least-squares surface matching method based on the tie points identified from the imagery and DTMs from which the two datasets were well aligned. Radhadevi et al. (2013) presented a geometric correction method for Chandrayaan-1 imagery that used LRO laser altimeter (LOLA) data as vertical controls.

This section has summarised research on the geometric fusion of imagery and laser scanning data in which these data are used as the absolute control for the photogrammetric processing of imagery. Corrections or adjustments are made to the image data alone to ensure that the 3D information derived from the images fits the laser scanning data. Although the methods discussed thus far are able to reduce or eliminate inconsistencies between the imagery and laser scanning data, the quality of the final topographic products is entirely dependent on the accuracy of the latter data, rendering them inapplicable when that accuracy is not assured, which constitutes an obvious limitation.

4. Imagery used as control for laser scanning data

The accuracy of laser scanning data is dependent on a number of factors. For airborne laser scanning or LiDAR data, most providers quoted their accuracy in RMSE values of 15 cm in the first several years of LiDAR’s technology development (from 2000 onwards). In fact, most laser scanning systems are highly integrated multi-sensor systems, and a variety of components and their spatial construction may introduce different types of error that can reduce accuracy. Hence, accuracy can be obtained only in ideal situations (Hodgson and Bresnahan 2004). As noted, laser scanning data normally feature better accuracy in the vertical or the range direction than in the horizontal direction (Baltsavias 1999, Xhardé et al. 2006, Qiao et al. 2010, Wu et al. 2011, 2014). Bowen and Waltermire (2002) examined the errors in LiDAR data collected in Richland County, South Carolina,
and their results show that the horizontal errors of the LiDAR points are typically large (RMSE of about 120 cm) relative to the vertical errors. Other research has demonstrated LiDAR data range accuracies for large-scale mapping applications at RMSEs ranging from 26 to 153 cm (Adams and Chandler 2002, Hodgson et al. 2003).

Csanyi and Toth (2007) investigated the use of LiDAR-specific ground control targets to improve data accuracy. They obtained test results from two flights and showed that specially designed LiDAR targets can improve the centimetre-level accuracy of the final LiDAR product. However, their improvement also relied on specially designed ground control targets.

In the area of terrestrial laser scanning, images have been used as supplemental information for the automatic registration of laser scanning data collected from multiple stations or for the quality improvement of laser scanning data. Terrestrial laser scanning is widely used for 3D modelling, particularly in the fields of heritage recording and site documentation. Multiple scans are often required to generate an occlusion-free 3D model in situations in which the object to be recorded has complex geometry. Many terrestrial laser scanning systems are now supplied with a single-lens reflex camera, meaning that the scene to be scanned can also be photographed. The provision of overlapping imagery offers an alternative, photogrammetric means of achieving point cloud registration between adjacent scans. For example, Al-Manasir and Fraser (2006) acquired images from a digital camera mounted on the top of a laser scanner and used them to first orient the network of images and then transfer that orientation to laser scanning stations to provide an exterior orientation allowing the registration of data scanned at different stations. A certain number of image points automatically measured from the images were used to orient the images, and the registration of the point clouds reached an accuracy of about 3 mm. This method offers a viable alternative to other point cloud registration approaches such as the well-known ICP algorithm.

Dold and Brenner (2006) proposed a laser scanning data registration method based on planar patches extracted from the point cloud and associated images. These planar patches were obtained through a specialised region growing algorithm and then textured automatically based on the images. The correspondences of the extracted planar patches were built by correlating the textured patches, and registration of the point cloud data was achieved by shifting the patches until they achieved the best fit. Kang et al. (2009) proposed an approach for the automatic registration of laser scanning data using panoramic reflectance images, which incorporates a two-step procedure: pair-wise registration and global registration. Their results indicated registration accuracy at the millimetre level and showed the process to be fully automatic and to converge quickly.

It is not unusual to find noise, outliers and missing measurements in the 3D surfaces obtained from terrestrial laser scanning data. Very recently, researchers employed image-based photometric stereo or shape-from-shading methods to improve the quality of these data. A representative approach is that proposed by Herbst et al. (2013), who developed a novel algorithm for the fusion of absolute laser scanning depth profiles and photometrically estimated surface normal data to produce noise-reduced and highly detailed laser scanning data. The surface normal data estimated from imagery based on the shape-from-shading method were used to help filter the outliers and noise. Eight objects with varying albedos and types of reflectance behaviour were employed for experimental evaluation, with the results showing a reduction in both outliers and noise and the preservation of large-scale shapes. At the same time, some of the surface details that were not visible in the laser scanning data were enhanced using image data.
This section has summarised research on the geometric fusion of imagery and laser scanning data in which the former is used as a control to improve the latter. Similar to the discussion in the previous section, imagery is used as the absolute control or reference in the registration or quality improvement of laser scanning data. Possible errors in the imagery itself or in the alignment of the camera and the laser scanner are ignored.

5. Combined adjustment of imagery and laser scanning data

In previous discussions in this area, laser scanning data have been used as controls for imagery or vice versa. In either case, one dataset is assumed to be fixed and is used as a reference for the adjustment of the other. Although the aforementioned integration methods have been shown to be capable of reducing the inconsistencies between the two types of datasets, the integration strategies are rather ad hoc in nature and fail to take full advantage of the coupling between them. Several researchers have emphasised the combined adjustment of imagery and laser scanning data, with both types of data adjusted simultaneously in the process based on certain strict mathematical models.

The work of Rosiek et al. (2001) constitutes an endeavour to integrate Clementine images with Clementine laser altimeter data, with the Clementine global mosaic used to establish the horizontal control and the Clementine laser altimeter points used for the vertical control. However, because of the marginal overlap between stereo models, the geometry of the photogrammetric network was weak, resulting in unaligned stereo models. Nevertheless, these researchers were the first to emphasise the different contributions of imagery and laser scanning data to the final topographic result, i.e. imagery contributes to its horizontal component and laser scanning data to its vertical component. Yoon and Shan (2005) introduced combined processing between MOC imagery and MOLA data using an adjustment method to correct the misregistration between the two datasets and accurately determine the ground positions. In their method, the MOLA ranges and ground points, MOC image orientation parameters and tie points collected from the MOC stereo images are involved in a least-squares adjustment. The outcome is the refined MOC image orientation and refined MOLA ground positions. Their results indicated that a large degree of misregistration between the two datasets can be corrected to a certain extent. Such combined adjustment is highly beneficial when the trajectory data have a large degree of inconsistency.

Wu et al. (2011) proposed an approach for integrating Chang`E-1 imagery and Chang`E-1 laser altimeter data for high-precision lunar topographic mapping through combined adjustment. In their approach, tie points are identified from the stereo Chang`E-1 images through image matching and then evenly distributed over the images. The ground coordinates of the tie points can be obtained using the tie points themselves and the initial image orientation parameters through a space intersection. With these parameters, the image coordinates of the laser ground points on the stereo images can also be derived through back projection. The image orientation parameters, tie points on the stereo images, ground coordinates of the tie points, laser ground points and image coordinates of the laser ground points are all used as components of the combined adjustment. Wu et al. (2011) constructed observation equations for each component based on the co-linearity equation. In addition, a local surface constraint was employed in the adjustment model to confine the adjustment process in such a way that the ground positions of the tie points on the images obtained through the space intersection using the image orientation parameters were consistent with the local surface determined by the nearby laser altimeter points. An empirical weight scheme was carefully designed for each
type of observation in the adjustment model, and the combined adjustment was processed iteratively based on a least-squares approach. In each iteration, the corrections were added to the unknown parameters until the precise values of those parameters were obtained. The final output was the improved orientation parameters of the Chang'E-1 imagery and improved ground positions of the Chang'E-1 laser altimeter points. This approach can reduce the inconsistencies between imagery and laser altimeter data. The researchers showed that misregistrations in the image space between the Chang'E-1 imagery and Chang'E-1 laser altimeter data can be reduced by a maximum of 18 pixels. Inconsistencies of more than 1 km between the DTMs derived from the Chang'E-1 imagery and laser altimeter data can be reduced to less than 100 m after combined adjustment, which is remarkable given the Chang'E-1 imagery’s low resolution of 120 m/pixel.

Wu et al. (2014) extended the foregoing approach for the integrated processing of the cross-platform and cross-sensor lunar imagery and laser altimeter data. The extended approach is able to process multi-strip images in an image network, and imagery and laser altimeter data collected from different platforms can be integrated for processing in a combined block adjustment model. This approach uses two types of tie points: intra-strip tie points and inter-strip tie points. Intra-strip tie points are the corresponding points identified from the stereo images (forward- and backward-looking images) of the same image strip, whereas inter-strip tie points are those identified from the overlapping area in the stereo images of two neighbouring image strips. These tie points should be distributed evenly in the image area. Based on the image orientation parameters, the 3D coordinates of the intra- and inter-strip tie points in the object space can be calculated from the photogrammetric space intersection. However, these parameters are probably not sufficiently precise, and the 3D coordinates of the tie points may thus deviate from their true locations. There may also be inconsistencies between the 3D coordinates of the same pair of inter-strip tie points identified from two neighbouring image strips. In addition, based on the image orientation parameters, the laser altimeter points can be back-projected to the stereo images. If the image orientation parameters are sufficiently precise, the back-projected points on the stereo images should then be conjugate points representing the same textural feature. However, this may not be the case if there are errors in the image orientation parameters. These errors or inconsistencies can be reduced or eliminated through the combined block adjustment of the image block and laser altimeter data. Wu et al. (2014) incorporated two additional constraints into the combined block adjustment. The first was a local surface constraint similar to the previous one, which indicated that the calculated ground coordinates of the tie points should lie on a local surface determined by its neighbouring laser altimeter ground points. The second was an orbit height constraint, which was explicitly applied to the exterior orientation parameters of the images. These two constraints proved to be critical in the combined block adjustment due to the large inconsistencies between the imagery and laser altimeter data collected from different platforms and the lack of ground control points. Without these constraints, it would be difficult to achieve convergence in the adjustment process. The weights for each type of observation were determined by their a priori standard deviations in the combined block adjustment. The final output of this adjustment was the improved image orientation parameters and improved ground coordinates of the laser altimeter points. Experimental analysis using multi-strip Chang'E-2 imagery and NASA LOLA data showed that the inconsistencies in the image space between the two datasets were drastically reduced from tens of pixels before the adjustment to the sub-pixel level after adjustment. The DTMs generated from the Chang'E-2 imagery based on the improved
orientation parameters and those from the LOLA data were closely aligned, with an average absolute elevation difference of about 15 m. Figure 2 illustrates the performance of the combined adjustment approach using Chang’E-2 imagery and LOLA data covering part of the Sinus Iridum area of the Moon.

Wu et al. (2015) further improved upon their previous developments and proposed and tested a method for the geometric integration of high-resolution satellite imagery (HRSI) and airborne laser scanning data in metropolitan areas. In the previous methods used in lunar mapping, the lunar surface is relatively smooth, which facilitates the integrated processing of the imagery and laser altimeter data. For the metropolitan areas of Earth, in contrast, the situation is much more complex. Metropolitan areas feature significant height variations due to their dense distributions of skyscrapers and other tall buildings. A sudden change in elevation or slope introduces new challenges to the integration process. For example, the surface continuity constraint that works well for relatively smooth terrains is not functional in metropolitan areas. Based on a principle of combined adjustment similar to previous methods, the new approach highlights the three following aspects in adapting to the integration of HRSI and airborne laser scanning data in metropolitan areas. First, the influence of the individual rational polynomial coefficients (RPCs) of HRSI on geo-positioning accuracy is determined, and the RPCs that dominate that accuracy are identified. The dominating RPCs (18) rather than the full RPC set (78) are employed in the combined adjustment, thereby greatly reducing the computational complexity and guaranteeing the convergence of the mathematical solution. Second, a local vertical constraint similar to the local surface constraint used in lunar mapping is involved in the combined adjustment. Although metropolitan areas feature dense populations of tall buildings, they also include plenty of open spaces such as grasslands and playgrounds. The terrain fluctuations are continuous and smooth in these regions, enabling the use of a local vertical constraint during geometric integration. A semi-automatic process is developed to identify tie points on the HRSI and laser scanning data in these open spaces for use in the local vertical constraint. This constraint then

Figure 2. Combined adjustment of Chang’E-2 imagery and LRO laser altimeter (LOLA) data, (a) the DTM generated from Chang’E-2 imagery, (b) the LOLA points covering the same area, (c) the LOLA points directly overlaid on the Chang’E-2 DTM indicating large inconsistencies and (d) the Chang’E-2 DTM well aligned with the LOLA points after combined adjustment.
ensures vertical consistency between the two types of datasets. Third, a local horizontal constraint is developed to ensure their horizontal consistency. As noted, there are numerous tall buildings in metropolitan areas. The local horizontal constraint is based on the principle that the building boundaries derived using the image RPCs should be consistent with those derived from the laser scanning data in the horizontal direction. Based on these developments, the latest approach developed by Wu et al. (2015) improves the dominating RPCs of the HRSI and the ground coordinates of the laser scanning data, thereby decreasing the inconsistencies between the two datasets and improving their geopositioning accuracy. Their approach was evaluated using ZY-3 and Pleiades-1 satellite imagery and airborne laser scanning data collected in Hong Kong. The results verify that the geometric integration model effectively decreases the inconsistencies between the 3D products generated from both datasets and improves the geo-positioning accuracy of both types of imagery and the laser scanning data. For ZY-3 imagery, significantly greater accuracy (about 1 pixel in size) is obtained after geometric integration (3.37 m/pixel for ZY-3 imagery), whereas for Pleiades-1 imagery, the improvement in accuracy is about 2.6 pixels in size (0.5 m/pixel for Pleiades-1 imagery). The horizontal accuracy of the laser scanning data can be improved to about 18 cm from the original errors of about 180 cm.

A summary of the general principles underpinning the combined adjustment approach for imagery and laser scanning data is presented in Figure 3. For combined adjustment in

![Figure 3](image_url)

Figure 3. Overview of the combined adjustment approach for imagery and laser scanning data.
relatively smooth areas such as those on the lunar surface, a local vertical constraint (or the local surface constraint used in Wu et al. 2011, 2014) may be sufficient to achieve favourable performance. In metropolitan areas and other areas in which the surface fluctuates to a considerable extent, however, a local horizontal constraint is necessary in addition to the local vertical constraint.

The major difference between the combined approach and those described in previous sections is that both the imagery and laser scanning data can be adjusted. Hence, the accuracy of both the photogrammetric processing of the imagery and that of the laser scanning data can be improved through the combined adjustment process. The specific degree of adjustment required is dependent on the \textit{a priori} accuracy of the dataset and each component thereof (e.g. the horizontal or vertical component of the dataset), which can be controlled by assigning an appropriate weight to each. A more accurate dataset (e.g. the LOLA data are more accurate than the Chang’E-2 imagery) or component (the vertical accuracy of laser scanning data is greater than their horizontal accuracy) will generally be assigned a greater weight in the combined adjustment, and thus require less adjustment. By the same token, a less accurate dataset, or the components thereof, will be assigned a smaller weight, and thus be adjusted more. In this way, 3D mapping results that are more accurate than either imagery or laser scanning data can be generated, and the inconsistencies between the 3D information generated from the two datasets can be largely eliminated.

6. Summary and discussion

The fusion of imagery and laser scanning data is a prerequisite to capitalising on the complementary characteristics of both. The effective geometric fusion of the two datasets is vital for a range of applications. In accordance with the nature of the computation principles in the fusion process, three categories of the geometric fusion of imagery and laser scanning data are detailed in this article. The discussion of each category includes a survey of classical and up-to-date methods.

Early research in the 1990s concentrated largely on the first category of integration, which uses laser scanning data as absolute controls to improve the photogrammetric processing of imagery. The resulting 3D mapping is thus completely dependent on the laser scanning data. Possible errors in the laser scanning data itself are ignored. The second category of integration uses imagery as a control or reference for laser scanning data, primarily in the area of terrestrial laser scanning. The imagery is used to assist the registration of laser scanning data from multiple stations or to improve the quality of those data based on photometric methods. Again, imagery is used as an absolute control in the process, and errors in the imagery itself or in sensor alignment are ignored. The third category of integration takes advantage of both imagery and laser scanning data. Both are processed in the same adjustment model, and adjustments are made to each dataset, allowing the generation of consistent and precise 3D mapping results.

The state-of-the-art geometric fusion of imagery and laser scanning data has significant benefits for a variety of applications, and remarkable progress has been made in this field. However, there are several difficult problems that await further investigation. They are discussed in the following paragraphs.

(1) Automation of the geometric fusion process: in the existing methods, the tie points identified from the images and common features (e.g. building boundaries) on the imagery and laser scanning data are critical components of the integration
process of the two datasets. They need to be accurate and evenly distributed to ensure favourable integration performance. In the existing methods, these components are obtained semi-automatically or manually, which is a difficult and time-consuming process. If they could be obtained automatically, then automation of the entire geometric fusion process would be possible. The development of automatic and reliable feature correspondence methods between images and between image and laser scanning data will be necessary for this task.

(2) Weight determination in combined adjustment: in the existing methods, weights are determined for each observation equation, representing their respective contributions to the integration process. Our experience suggests that it is difficult to determine an optimum weight for every observation. Inappropriate weights lead to an unstable solution for the combined adjustment process and sometimes prevent the adjustment from converging. Further research on an optimum weighting scheme is necessary to ensure the stability and performance of the combined adjustment approach.

(3) Efficiency of the approach: due to the large volume of laser scanning data, the efficiency of the existing algorithms is generally poor. Our experience indicates that it is not necessary to involve all laser scanning points in the integration process. Instead, a sampling strategy that extracts a relatively small number of laser scanning points may be preferable and could be handled by a regular computer. Such a sampling strategy should be able to extract laser scanning points that represent the overall geometry of the entire dataset. However, highly efficient algorithms with memory-saving arithmetics need to be developed to process large-scale datasets.

The geometric fusion of imagery and laser scanning data builds a bridge between the two distinct techniques of photogrammetry and laser scanning in such a way that they supplement each other for better 3D mapping. Such geometric fusion is essential for the proper calibration, registration and analysis of the two types of datasets, and also allows their full comparative and synergistic use. This approach is of significant importance, as increasing amounts of imagery and laser scanning data are now easily accessible.

Disclosure statement
No potential conflict of interest was reported by the authors.

Funding
This work was supported by the Research Grants Council of Hong Kong [project number PolyU 5330/12E] and the National Natural Science Foundation of China [project number 41471345], [project number 91338110].

References


Spiegel, M., 2007. Improvement of interior and exterior orientation of the three line camera HRSC with a simultaneous adjustment. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36 (3), W49B.


