Shallow water bathymetry from WorldView-2 stereo imagery using two-media photogrammetry

Bincai Cao\textsuperscript{a,b}, Yong Fang\textsuperscript{a,b}, Zhengzhi Jiang\textsuperscript{a,b}, Li Gao\textsuperscript{a,b} and Haiyan Hu\textsuperscript{a,b}

\textsuperscript{a}State Key Laboratory of Geo-information Engineering, Xi’an, China; \textsuperscript{b}Xi’an Research Institute of Surveying and Mapping, Xi’an, China

\textbf{ABSTRACT}

Two-media photogrammetry appears to be a useful technology for mapping shallow water depth. The objective of this paper is to explore the feasibility of shallow water bathymetry using satellite two-media photogrammetry. This paper presents a two-media photogrammetry technique for WorldView-2 multispectral imagery. In this method, near-infrared band is used for sun glint elimination. A combination of Scale Invariant Feature Transform and Semi-Global Matching algorithm is adopted for dense stereo matching. Rational Function Model is applied for raw Digital Elevation Model generation and the water surface elevation is determined by interpolating water edge elevation. Finally, the approximate refraction correction model is adopted in the water region to correct the vertical coordinate offsets. Our experimental results of Ganquand Island and Zhaozhou Island show that a reasonable depth estimate with relative error less than 17% and 14% can be obtained for the water depth range between 5 and 20 m, and the root-mean-square error are 2.09 m and 1.76 m, respectively. However, less accurate estimate is found for the very shallow water depth range between 0 and 5 m. The overall accuracy of satellite two-media photogrammetry is at the same level with the traditional multispectral-based model for bottom depth retrieval.

\section*{Introduction}

Shallow seawater depth data is important basic geospatial information for shallow sea environmental management, shallow sea resources development and utilization, shallow sea navigation and other applications. Accurate, efficient and economical access to shallow seawater depth data is the goal pursued by marine surveying and mapping. At present, shallow bathymetry mainly relies on shipborne sonar and airborne lidar system. Satellite remote sensing has the characteristics of large range, periodicity, and can quickly acquire earth surface images, making Satellite-Derived Bathymetry (SDB) superior to sonar or lidar system in terms of efficiency and cost (Su et al., 2014). The currently used SDB methods include physics-based approach and empirical models. The physics-based approach relies on the radiative transfer equation of light in water to invert the water depth (Chen, Yang, Xu, & Huang, 2019; Lee, Carder, Mobley, Steward, & Patch, 1999), which requires accurate atmospheric correction. The empirical model inverts the water depth by calculating the relationship between the radiance of the image and the in-situ depth data, depending on existing depth data, good water quality and uniform bottom substrate (Lyzenga, Malinas, & Tanis, 2006; Stumpf, Holderied, & Sinclair, 2003).

Two-media photogrammetry provides a third option for shallow bathymetry (Hodul et al., 2018). This technique does not require special atmospheric corrections or in-situ depth data, so it can avoid the shortcomings of the first two methods to some extent. Two-media photogrammetric bathymetry is not a novel technique and researchers have applied it to study bathymetry since the 1980s. The basic principle is that the geometrical characteristic can be derived by stereo-pairs in the case when the camera is above the water surface and the object is submerged (Wang, 1990; Zhou et al., 2015). Wang, Han, and Wang (1988), Chang (1991), and Shan (1993) have discussed the general geometric relation, basic equations, and relative orientation procedure. Fryer and Kniest (1985) discussed errors in-depth determination caused by waves. Murase et al. (2008) analysed the horizontal differences when the incident angles of light rays from an underwater point to two cameras are different. Skarlatos and Agrafiotis (2018) carried out two-media bathymetric experiments using UAV images and discussed the method of embedding refraction correction in the existing photogrammetric process. Although there are some successful applications of two-media photogrammetric bathymetry technique, such as tunnel movement measurement (Yang, 1987), gravel-bed rivers mapping (Westaway, Lane, & Hicks, 2003), underwater archaeology (Georgopoulos & Agrafiotis, 2012) and marine biological conservation (dolphin body measure) (Chong & Schneider, 2001), some...
practical issues have prevented its broad use: (1) the elevation of water surface must be known or modelled prior to apply; (2) rich underwater textures (depend on water depth, turbidity, and bottom sediment types) on the images are usually required in order to apply automated image matching techniques; (3) the presence of “white water” and sun glint affects image matching results, and (4) in most cases, the water surface is not planar, thus complicating the geometry of refraction through the surface (Fryer & Kniest, 1985), and in addition, water surface may change during the time lag between image capturing exposures (Westaway, Lane, & Hicks, 2000).

The recent available high-resolution satellites such as GeoEye-1 and WorldView-2 (WV2) have provided new data sources for shallow water bathymetry using photogrammetry method. Taking WV2 as an example, the stereoscopic collection on a single pass ensures image continuity and minimizes water surface change (Digital Globe, 2010). As the target’s 3D coordinates can be calculated from stereo-pairs without ground control data, the two-media photogrammetry method appears more attractive as compared with the traditional multispectral retrieval method (Lyzenga et al., 2006; Stumpf et al., 2003) in those remote shallow water area where lack of in-situ water depth. Moreover, the enhanced satellites’ agility and short revisit period make it possible to choose proper imaging conditions such as good weather, low wind speed, and relatively calm water surface to acquire better stereo imageries. Hodul, Bird, Knudby, and Chenier (2018) and Chenier, Faucher, Ahola, Shelat, and Sagram (2018) selected the WV2 stereo data of bay regions with clear underwater texture and calm water surface, and carried out the satellite photogrammetric bathymetry experiment. Results show that the vertical accuracy meets the IHO (International Hydrographic Organization) level C requirement for hydrographic surveys. The aim of this paper is to provide a more detailed process and to further assess the capabilities of extracting island bathymetry information using satellite photogrammetric technique from WV2 stereo imagery. In the following sections, we first review the fundamentals of two-media photogrammetry, and then deduce the approximate water refraction correction model. Error sources of our two-media photogrammetric bathymetry technique are analysed. Next, we introduce the processing workflow for two-media photogrammetry using satellite stereo pairs, and the WV2 bathymetry results are presented later. By applying the proposed processing workflow, test results using WV2 multispectral imagery demonstrate that 17% depth relative error is possible to be achieved. In the final section, we discuss our application results and draw conclusions.

Theory of two-media photogrammetry

Photogrammetry is the technique to find the exact position of a 3D point based on its left and right image centres and its corresponding points on the left and right images, respectively. Two-media photogrammetry has the same basic working principle, and the main difference is the light refraction at air–water interface. Digital Elevation Models (DEMs) generated without ray refraction correction introduce systematic errors and usually, the elevation values are higher than they should be. As shown in Figure 1, underwater object P generates image rays and the image rays are refracted at the two-media interface in P1 and P2, recorded by two exposure stations S1 and S2. Two-media photogrammetry is the process of detecting the intersection point (i.e. the observed point A) of corresponding image rays and recovering the true coordinate of the target (point P).

Image rays passing from one medium to another are refracted according to Snell’s law,

$$n = \frac{\sin i_1}{\sin i_2}$$

(1)

where $i_1$ and $i_2$ are incident angles from the underwater point P to exposure stations S1 and S2; $r_1$ and $r_2$ are the corresponding refraction angles; $n$ is water refractive index and usually takes a value of 1.34. The variation of $n$ across the visible is less than 1% within a wide range of temperature and salinity conditions (Butler, Lane, Chandler, & Porfiri, 2002).

Refraction correction model

Considering refraction correction, three factors should be taken into account: water refractive index, water surface elevation and surface normal direction. The ideal condition has three assumptions: the water surface is planar, the local water quality is uniform and the refractive index is constant. The basic two-media geometric relationship of satellite push-broom sensor is shown as Figure 1. X-axis is along the flight direction forming a right-handed coordinate system. Under general conditions, corresponding light rays $S_P$ and $S_P$ will not intersect at a certain point if refraction angles $r_1$ and $r_2$ are different, and then the photogrammetrically observed point A is a least-square solution instead of a strict solution.

As shown in Figure 1, the observed point A and the true point P are different both in horizontal and vertical coordinates. We assume the following special along-track case (see Figure 2) to assess the X value differences: the line of refraction points $P_1P_2$ is parallel to the flight direction, and point A is not in the
perpendicular bisector plane of $S_1$ and $S_2$ ($D_1 \neq D_2$). Refraction angles $r_1$ and $r_2$ are defined by air–water interface combined with lines $S_1A$ and $S_2A$, then incident angles $i_1$ and $i_2$ can be calculated by Equation (1). It is assumed that image rays intersect with a virtual underwater horizontal plane in-depth $h_c$ at $P_{C1}$ and $P_{C2}$, respectively. The intersection point of $P_1P_{C1}$ and $P_2P_{C2}$, namely $P$, is the true position corresponding to the observed point $A$.

In Figure 2, $y_P = Y_A$, $k$ is the distance between $P_1$ and $P_2$,

\[ k = (\tan r_1 + \tan r_2) \cdot h_A = (\tan i_1 + \tan i_2) \cdot h \] (2)

\[ h = \frac{(\tan r_1 + \tan r_2) \cdot h_A}{(\tan i_1 + \tan i_2)} \] (3)

According to the geometric relations and Snell’s law, we get

\[ \tan r_1 = D_1/(H + h_A) \] (4)

\[ \tan r_2 = D_2/(H + h_A) \] (5)
\[
\tan i_1 = \tan(\sin^{-1}(\sin r_1/n)) = \tan(\sin^{-1}(n \cdot D_1/S_1A)) \quad (6)
\]
\[
\tan i_2 = \tan(\sin^{-1}(\sin r_2/n)) = \tan(\sin^{-1}(n \cdot D_2/S_2A)) \quad (7)
\]
Then, the X relation between the true point P and the observed point A is expressed as
\[
X_P = X_P + h \tan i_1 = X_A - h \tan r_1 + h \tan i_1 \quad (8)
\]
By taking Equation (3) into Equation (8), we get
\[
\Delta X = X_A - X_P = \frac{\tan i_2 \cdot tan r - tan i_1 \cdot tan r_2}{\tan i_1 + \tan i_2} \cdot h_A \quad (9)
\]
Equation (9) demonstrates the X value differences between the true point P and the observed point A. As shown in Section IV, calculations demonstrate that \(\Delta X\) can be ignored both in aerial or space cases.

In fact, the horizontal differences (\(\Delta X\) and \(\Delta Y\)) between A and P are both negligible, which can also be deduced by the geometric relations. In order to simplify the question, in previous studies, the approximate refraction correction model, which only compensates vertical offsets, was adopted for two-media photogrammetry (Murase et al., 2008). If we consider both the left and right triangle geometrical parameters, the true water depth \(h\) can be calculated by Equations (1) and (3). Equation (1) can be transformed into
\[
\begin{cases}
i_1 = \sin^{-1}(\frac{\sin r_1}{n}) \\
i_2 = \sin^{-1}(\frac{\sin r_2}{n})
\end{cases} \quad (10)
\]
Substituting Equation (10) into Equation (3) gives
\[
h = \frac{h_A \cdot (\tan r_1 + \tan r_2)}{\tan[\sin^{-1}(\sin r_1/n)] + \tan[\sin^{-1}(\sin r_2/n)]} \quad (11)
\]
According to the results by Fryer and Kniest (1985), the accuracy of two-media photogrammetric bathymetry is closely related to the triangle geometry. The depth calculated from the better-conditioned triangle gives a better estimate for the depth, and large errors could occur if the point was almost beneath one of the camera stations (Fryer & Kniest, 1985). Therefore, one side model with the better-conditioned triangle is usually adopted. Taking the left part in Figure 2 as an example, the true water depth \(h\) is expressed as
\[
h = \frac{h_A \cdot \tan r_1}{\tan[\sin^{-1}(\sin r_1/n)]} \quad (12)
\]
For the airborne frame sensor, the refraction angle \(r_i\) of pixel \(i\) can be calculated as
\[
r_i = \cos^{-1}\left(-\frac{Z_{Ai} - Z_{S_i}}{S_i A_i}\right) \quad (13)
\]
where the length of \(S_i A_i\) can be calculated as
\[
S_i A_i = \sqrt{(X_{Ai} - X_{S_i})^2 + (Y_{Ai} - Y_{S_i})^2 + (Z_{Ai} - Z_{S_i})^2} \quad (14)
\]
For WV2 satellite push-broom imagery, the camera’s exterior orientation elements (position and orientation) are unavailable. Instead, the off-nadir angle (including the maximum, minimum and average values) which is defined as the spacecraft elevation angle measured from nadir to the product centreline as seen from the spacecraft is provided for the user. We take the average off-nadir angle for each pixel as the refraction angle and the resulting bathymetric error are analysed in the following Section.

**Error analysis of two-media bathymetry using WV2 multispectral imagery**

**Horizontal error caused by approximate refraction correction model**

We take Equation (9) for horizontal error quantitative analysis and results are shown in Figure 3. In this case, flight height \(H\) is 1500 m; photographic baseline length \(S_1, S_2\) is 500 m; water depth is, respectively, set at 1, 15 and 25 m. Green, Mumby, Edwards, and Clark (2000) pointed out that 25 m is the maximum penetration depth under most conditions, and the corresponding \(X\) difference is only 0.035 m. Moreover, as satellite imagery has the characteristics of high orbit and narrow view field, calculations got by bringing these parameters into Equation (9) indicate that the \(X\) difference is much smaller than the air-borne case. In a word, the horizontal error caused by an approximate refraction correction model can be ignored.

**Vertical error caused by average refraction angle**

In WV2’s push-broom imagery, refraction angles are equal for the same row at different lines if the water surface is planar. Also, for pixels in the same row and different columns in the image, the angle difference is quite modest. Figure 4 shows the geometric relations in the case when the sensor is scanning forward or backward along the flight direction. \(H\) is flight height, \(W\) is the ground swath and \(a_0\) is the refraction angle of the middle point. Then, the angle \(a_i\) at the edge area is
\[
a_i = \arctan\frac{H}{\sqrt{H^2 + \tan^2 a_0 + W^2/4}} \quad (15)
\]
Substituting WV2’s parameters, \(H = 770\) km and \(W = 16.4\) km into Equation (19), if \(a_0 = 30^\circ\), then \(a_1 \approx 29.99^\circ\). If observed water depth \(h_A = 25\) m, the elevation offset caused by refraction angle difference can be calculated by Equation (12), \(\Delta h = h_1 - h_0 \approx 0.02\) m. When the sensor scans along or crosses the track within a certain angle,
the refraction angles corresponding to each pixel in the same line are nearly the same for WV2 sensor. As a result, an average refraction angle can be used to estimate water depth for all the imagery pixels.

**Other errors**

From the previous analysis, we know that adopting an approximate refraction correction model and averaging refraction angles will not cause big errors. In fact, the following four aspects are major error sources of two-media photogrammetric bathymetry.

The geometric model error. WV2 has been regarded as one of the most precise direct positioning commercial satellites with the planimetric accuracy of 4.1 m CE90% without ground control points (GCPs) (Digital Globe, 2010). While this direct positioning accuracy is quite high compared to moderate resolution satellites, it is still not high enough for two-media photogrammetric bathymetry. In general, several GCPs with high accuracy are still needed to achieve sub-pixel positioning accuracy. The GCPs are usually collected by differential global positioning system (DGPS) receivers in real-time kinematic mode (RTK) and the 3D accuracy could reach centimetre level. Under the condition of GCPs, the positioning accuracy of WV2 stereo-pair can be greatly improved. For instance, two evenly distributed GCPs can improve planimetric accuracy from 4.1 m CE90% to 1.5 m CE90% (Aguilar, Saldana, Aguilar, & Lorca, 2014).

Underwater mismatching. Point matching is the critical issue in photogrammetry as the output 3D coordinates are directly related to its input image coordinates. Many factors can affect underwater points matching. Firstly, aerosols and sun glint will
gravely blur the image. Secondly, the visibility of underwater texture is closely related to water depth, water turbidity and bottom sediment types. Finally, different matching algorithms and gross error detection strategies could induce different matching results. Underwater points mismatching leads to incorrect calculations of the observed location (i.e. point A in Figure 1), and worse still, water refraction will magnify this depth error. For example, in one of our tests, one-pixel mismatching has caused 3.8-m vertical offset error at the water depth of 17 m.

Ocean waves. Ocean waves or surface disturbance will change the basic geometric relationships of two-media photogrammetry and thus introduce bathymetry error, and this error caused by surface normal deviation could not be compensated by classic methods such as camera calibration (Georgopoulos & Agrafiotis, 2012). In ocean science, the waves in surf zone (high-frequency component) are difficult to be modelled purely using satellite imagery information. Although sun glint data can be used for extracting the sea surface geometry and dynamics (Cureton, 2015), most of the related applications have, so far, neglected the ocean wave influences and assumed that water surface is planar. Ideally, a calmer sea condition is preferable to minimize the error caused by waves.

Sea surface elevation. In order to correct the observed point A to the real point P, we need to calculate the water surface elevation. For imagery with both land and water inside, the surface elevation can be determined by an automatic water-land separation (use the near-infrared band) and elevation interpolation (use the raw DEM). The problem is that the raw DEM is not ensured to be so accurate because of mismatching especially in the shallow area (white-caps and low texture). For imagery with no parts exposed above the water, the surface elevation is

---

**Figure 5.** High-resolution satellite imagery-based two-media photogrammetric bathymetry workflow.
rather difficult to be determined. In order to get a more reliable water surface elevation, some active remote sensing methods should be considered. For instance, a radar altimeter can measure the water level at an accuracy of 10 cm, which could greatly help two-media photogrammetry to determine the sea surface elevation.

Two-media photogrammetric bathymetry processing workflow

The procedure of shallow water depth estimation by a two-media photogrammetric technique using WV2 multispectral imagery is shown in Figure 5. It mainly includes radiometric processing, geometric processing, water-land separation, surface elevation calculation and refraction correction. It is noteworthy that all calculations depend on the geometry information of the imagery (i.e. Rational Function Model), then the bathymetry results are relative values (corresponding to water surface elevation) at the imaging moment. Data that meet the requirement of maps and charts can be obtained after depth datum transformation.

Radiometric processing

Because photogrammetry is a geometric operation that does not rely on accurate spectral information, no special atmospheric correction is necessary (Hodul et al., 2018), but if atmospheric effects are sufficient to affect image matching, a certain atmospheric correction algorithm is recommended to make image texture clearer. In addition to atmospheric correction and image enhancement, sun glint elimination is a key issue in satellite two-media photogrammetry. Influenced by water surface condition, sun position and view of the sensor, sunlight may be directly reflected to CCD. For imagery with ground sample distance (GSD) higher than 10 m, wave characteristics can be obviously recorded. Then, sun glint is often shown as scattered bright spots or white stripes along the wave edge, which presents a significant obstacle of underwater point matching.

Kay, Hedley, and Lavender (2009) have discussed some common algorithms for sun glint elimination. We adopted Hedley algorithm (Hedley, Harborne, & Mumby, 2005) to eliminate sun glint in term of WV2’s bands setting and algorithm robustness. This method is designed based on the following two assumptions. Firstly, water exhibits strong absorption at near-infrared (NIR) wavelength and geometrically shallow water (tens of centimetres) is optically deep for NIR data. Secondly, the real refractive index of water is nearly the same at visible (VIS) and NIR wavelengths. Therefore, the water radiance is solely a function of geometry and the relation between VIS and NIR can be expressed by regression straight line. In Hedley algorithm, a subset deep-water area is used to calculate the VIS/NIR regression parameters. The corrected band becomes

\[ L_i'(\text{VIS}) = L_i(\text{VIS}) - b_i[L_i(\text{NIR}) - L_{\text{min}}(\text{NIR})] \] (16)

where \( b_i \) is the slope of regression line; \( L_i(\text{NIR}) \) is the NIR band pixel value; \( L_{\text{min}}(\text{NIR}) \) is the minimum NIR value in the subset area; \( L_i(\text{VIS}) \) and \( L_i'(\text{VIS}) \) are VIS values before and after sun glint correction.

Stereo model establishment

In order to hide sensor design parameters, satellite image data providers usually adopt the geometric approximation using Rational Function Model (RFM) for 3D mapping processing (Tao & Hu, 2001). In the RFM, image coordinates \((r_n, c_n)\) are expressed as the ratios of polynomials of object coordinates \((X_n, Y_n, Z_n)\):

\[
\begin{align*}
[r_n] = \frac{\sum_{k=0}^{m} \sum_{j=0}^{m} \sum_{i=0}^{m} a_{ijk} X_n^i Y_n^j Z_n^k}{\sum_{k=0}^{n} \sum_{j=0}^{n} \sum_{i=0}^{n} c_{ijk} X_n^i Y_n^j Z_n^k} \\
[c_n] = \frac{\sum_{k=0}^{m} \sum_{j=0}^{m} \sum_{i=0}^{m} d_{ijk} X_n^i Y_n^j Z_n^k}{\sum_{k=0}^{n} \sum_{j=0}^{n} \sum_{i=0}^{n} b_{ijk} X_n^i Y_n^j Z_n^k}
\end{align*}
\] (17)

where \( a, b, c \) and \( d \) are the rational function coefficients (RPCs); \( n \) represents the sequence point number; \( i + j + k \) defines the model degree which is no more than 3, \( i + j + k \leq 3 \). In order to avoid large differences in the magnitude, raw image and object coordinates are translated and scaled to \((-1.0, +1.0)\) range.

In the 3D reconstruction process, we transform Equation (21) into

\[
\begin{cases}
[r_n] = r_{n0} \cdot F(X_n, Y_n, Z_n) + r_{n0} \\
[c_n] = c_{n0} \cdot G(X_n, Y_n, Z_n) + c_{n0}
\end{cases}
\] (18)

where \( F(X_n, Y_n, Z_n) \) and \( G(X_n, Y_n, Z_n) \) are the corresponding right side part of \( r_n \) and \( c_n \) in Equation (21). After Taylor expansion, the observation equation is expressed as:

\[
\begin{align*}
\nu_r &= \frac{\partial r_n}{\partial X_n} \Delta X_n + \frac{\partial r_n}{\partial Y_n} \Delta Y_n + \frac{\partial r_n}{\partial Z_n} \Delta Z_n - (r_n - \hat{r}_n) \\
\nu_c &= \frac{\partial c_n}{\partial X_n} \Delta X_n + \frac{\partial c_n}{\partial Y_n} \Delta Y_n + \frac{\partial c_n}{\partial Z_n} \Delta Z_n - (c_n - \hat{c}_n)
\end{align*}
\] (19)

where the terms \((\nu_r, \nu_c)\) and \((\hat{r}_n, \hat{c}_n)\) represent the corresponding observational residuals vector and the approximate image coordinates, respectively. Then, the 3D object coordinates can be iteratively calculated with the corresponding points in stereo-pairs. Either the mean offset value or the one-degree term can be
used as the initial value during the iteration process. The water-land integrated raw DEM can be derived by calculating ground coordinates of all the matched points.

**Water-land separation**

Water-land separation is implemented using the NIR band because of its low radiance in the water area. A variety of methods can be used, such as the threshold segmentation, the water index technique and the region growing methods. Special attention should be paid to the sun glint-contaminated imagery. The above-mentioned sun glint elimination method is used for correcting the blue, green and red band, the bright spots in NIR itself are preserved, and then scattered patches might appear in the water-land separated binary image. By setting a proper area threshold, those small patches (sun glint areas) and even larger parts (e.g. clouds area) can be recognized. After water-land separation, it is easy to find the water boundary, and we can interpolate water surface elevation together with the raw DEM as mentioned before.

**Shallow water bathymetry using real WV2 multispectral imagery**

**Study area and data set**

We selected two islands, namely Ganquan Island and Zhaoshu Island, in the South China Sea as study areas. Those two islands have a moderate tropical monsoon climate, the waterbody is clear and light penetration depth can reach 20 m in some areas. Ganquan Island is about 2 km in length long from north to south and about 1 km in width from west to east, while Zhaoshu Island is about 0.7 km in the north-south direction and 1 km in the east-west direction. The information about study area WV2 images are shown in Table 1. The multispectral stereo-pairs used in our analysis were acquired by WV2 satellite. This satellite has the ability to collect information stereoscopically on a single pass, so the time interval of stereo-pairs is relatively short (76 s for Ganquan Island and 79 s for Zhaoshu Island).

In the imaging quality aspect, the left image of Ganquan Island is severely contaminated by sun glint due to its small refraction angle (4.9°), and so the bottom texture is invisible (see Figure 6). The right image is almost sun glint free, the underwater texture is clear and rich. For Zhaoshu Island, the left and right refraction angles are both moderate, so the stereo-pairs are sun glint free and the bottom texture are both clear (see Figure 7). The multispectral images used here have three visible bands (blue 450–510 nm, green 510–580 nm, red 630–690 nm) and one NIR band (770–895 nm) with 2 m spatial resolution. In the following test, we use the NIR band for water-land separation and sun glint elimination, while only the blue band is used for underwater points matching since it penetrates water deeper than green and red in this area.

In terms of water depth verification data, Ganquan Island and Zhaoshu Island, respectively, carry lidar bathymetric points and single-beam sonar data. The lidar bathymetric points of Ganquan Island were collected using SHOALS-3000 (Scanning Hydrographic Operational Air-borne lidar system) by the North China Sea Air-borne Detachment of China Marine Surveillance on 19 January 2013. The detected maximum depth in Ganquan Island is about 21 m. The horizontal and vertical accuracies are about 2.5 m and 25 cm, respectively. Because lidar points have a small footprint and high density, it is inconvenient to directly compare with photogrammetric results. Therefore, the

| Study area | Stereo-pairs | Acquisition Time | Refraction angle | Sun glint |
|------------|--------------|-----------------|-----------------|-----------|
| Ganquan Island | Left 2014-04-02 (03:32:15) | 4.9°  | Severe |
| Zhaoshu Island | Right 2014-04-02 (03:33:31) | 31.8°  | Free |
| Zhaoshu Island | Left 2017-03-11(03:03:27) | 24.5°  | Free |
| Zhaoshu Island | Right 2017-03-11(03:04:46) | 29.2°  | Free |

Figure 6. Stereo-pairs of Ganquan Island shown on WV2 true-colour composite of red, green, and blue bands as R, G, B.
SHOALS-3000 lidar points were used to generate the water-land integrated DSM (see Figure 8(a)) with the same spatial resolution (2.0 m) with WV-2 multispectral imagery. In addition, air-borne high-resolution imagery (see Figure 8(b)) with 0.05 m GSD was also collected during the same flight. Since the lidar point cloud data locates in the UTM coordinate system, and no tide correction has been conducted, this paper subtracts the water boundary elevation form the point cloud elevation to obtain instantaneous lidar DEM. Visual observation shows that the water boundary of aerial image and WV2 satellite image almost coincide, and the depth obtained by two-media method is also instantaneous value, so it is reasonable to use this lidar DEM for accuracy evaluation. From Figure 8(b), we can divide Ganquan Island into three parts: land area, shallow surf zone and deep calmer zone. The bottom substrates of the shallow and deep zone are sand and coral reef, respectively. Therefore, the image texture is richer in the deep zone. In addition, due to the surf wave and white caps in the shallow zone, sea condition of the deep zone is much better. Although WV2 stereo-pairs and the air-borne imagery are acquired at different time, similar history oceanic forecast indicates that the three parts classification is also reasonable for WV2 imagery.

The single-beam sonar data of Zhaoshu Island were collected on 2 March 2014. The detected maximum depth in this region is about 18 m, and the vertical accuracy is better than 20 cm. After tide correction, 1284 sonar points were preserved for the accuracy assessment latter. In addition, the bottom substrates and sea wave conditions of Zhaoshu Island are similar to Ganquan Island.

This paper uses C++ programming language to implement the main algorithm functions based on the Microsoft Visual Studio 2010 platform, including

Figure 7. Left Image of Zhaoshu Island shown on WV2 true-colour composite of red, green, and blue bands as R, G, B.

Figure 8. Water depth validation data. (a) Land-water integrated DSM generated by LIDAR cloud points; (b) Air-borne high-resolution imagery.
sun glint elimination, RFM stereo model construction, land-water separation, refraction correction, etc. The dense stereo matching described below is developed based on OpenCV algorithm library.

**Sun glint elimination**

Figure 9 shows the sun glint contamination and elimination effects of the local white box of the left image in Figure 8. After sun glint elimination by Hedley algorithm, bright spots distributed along the wave edge disappear and the underwater texture becomes much clearer. In addition, the DN curve along the white line fluctuates wildly before sun glint elimination, after sun glint processing, the DN curve becomes more stable.

**Dense stereo matching**

Accurate image matching is one of the key issues of two-media photogrammetry. Radiometric distortion, low texture and data noise in the underwater environment are the difficulties faced by image matching techniques. Now, both local features based matching technique (Lowe, 2004) and global dense matching technique (Hirschmuller, 2008) are employed and integrated into the proposed workflow.

The local feature-based matching technique is based on keypoints extraction. SIFT (Scale Invariant Feature Transform) descriptor (Lowe, 2004) is used in the proposed processing workflow. Extracted feature points are invariant to image scale, image rotation and change in illumination. While SIFT is capable of providing enough features point numbers that cover the image region if the texture is rich and image quality is high, features point numbers are usually insufficient in the underwater case due to less texture and data noise.

The global dense matching technique considers matching issues in the entire imagery. SGM (Semi-Global Matching) (Hirschmuller, 2008) uses a pixel-wise, mutual information-based matching cost for radiometric differences compensation, and adopts global cost function as smoothness constraint of pixel-wise matching. Therefore, the disparity map produced by SGM is usually denser and more reliable than interpolation results based on feature points. In this paper, SIFT and SGM results are combined to form a reliable disparity map. The basic idea is to recognize wrong disparities of SGM results by using local SIFT feature points as shown in Figure 10.

Specific steps and results are as follows:

1. SIFT local disparity generation. Concrete steps include scale-space extrema detection, keypoint localization, orientation assignment, keypoint description, keypoint matching, and gross error detection. The distribution of corresponding points extracted by the SIFT method is shown in step 1 of Figure 11. Numbers and distribution of
SIFT points are primarily related to image texture, and depend on choices for various SIFT parameters. We can see that points are much denser in land area than water area for Ganquan Island, and the local accurate SIFT points are only used for the refinement of SGM disparity.

(2) Stereo epipolar images generation. SGM requires stereo epipolar images, where corresponding points lie on the corresponding lines on the left and right epipolar images. Gong and Wan (2014), and Cao, Qiu, Zhu, Shen, and Cao (2017) have studied the method of generating WV2 epipolar images based on the parallel characteristic, which shows that WV2 multispectral epipolar images can be produced by simple rotation and translation. We have tested this method with no less than 30 sets of data, and the obtained epipolar images have a good linear characteristic in the Island area with subpixel level of vertical-parallaxes. The rotation angles and translation values are different for each stereo-pair, and those parameters can be determined by classic projection track method (AlRousan, Cheng, Petrie, Toutin, & Zoej, 1997).

Ganquan Islands WV2 stereo epipolar pairs are shown in step 2 of Figure 10. It should be pointed out that this method does not necessarily have universal applicability to other types of terrain and satellite images, and the mechanism needs further study.

(3) Semi-global matching and disparity refinement. While SGM is suitable for occlusions and disparity discontinuous cases, it could not effectively handle the radiometric distortion and low texture issues. Our tests show that SGM may get wrong disparities in those large low texture areas, and then interpolation is necessary for filling holes in the disparity images. As shown in step 3 of Figure 10, blocks inside red lines are obviously wrong disparity. This is judged from its neighborhood SIFT keypoints disparities. By value interpolation, we can fill those holes and get an optimized dense disparity map. A circular template is adopted for the disparity refinement, and the refined disparity of the centre point $d_{ref}$ is expressed as
\[ d_{\text{ref}} = \begin{cases} d_{\text{SGM}_c} , & \text{if} \quad d_{\text{SGM}_c} \in [\tilde{d}_{\text{SIFT}} - \gamma, \tilde{d}_{\text{SIFT}} + \gamma] \\ \sum_{i=1}^{n} p_i d_{\text{SGM}_i} , & \text{if} \quad d_{\text{SGM}_c} \notin [\tilde{d}_{\text{SIFT}} - \gamma, \tilde{d}_{\text{SIFT}} + \gamma] \end{cases} \]  
\[ \sum_{i=1}^{n} p_i = 1 \quad \text{and} \quad d_i = \text{the distance between point} \ i \ \text{and the centre}. \]  

where \( d_{\text{SGM}_c} \) and \( d_{\text{SGM}_i} \) are the raw SGM disparity of the centre point and the neighbour point \( i \), respectively; \( \tilde{d}_{\text{SIFT}} \) is the local SIFT mean disparity within the circle; \( \gamma \) is a threshold and should be adjusted appropriately according to different cases; \( p_i \) is the weight of point \( i \), \( p_i = \frac{1}{d_i} \) where \( d_i \) is the distance between point \( i \) and the centre.

**Two-media geometric model establishment and accuracy assessment**

At the geometric pre-processing aspect, we selected 15 evenly distributed GCPs for Ganquan Island from the existing air-borne high-resolution imagery with a spatial resolution of 0.05 m to evaluate WV2 direct positioning accuracy and provide systematic error correction parameters. Horizontal residual errors are \( dx = 0.6 \) m and \( dy = 3.9 \) m, respectively, which are within WV2’s geo-positioning accuracy statement. The aim of two-media photogrammetry is to provide relative water depth, and \( Z \) coordinates in land area will be transformed relative to water surface elevation. Therefore, it is not necessary to perform three-dimensional systematic error corrections. In the following geometric issues, we still adopt raw RFM for 3D coordinates calculation, and only horizontal coordinates \( (X_i, Y_i) \) are systematically corrected by previous 15 GCPs. For Zhaoshu Island, since there are no GCPs on the land area, the raw RFM was adopted without any corrections.

After geometric pre-processing, we performed two-media photogrammetric bathymetry following the proposed workflow. Water-land separation was conducted and the surface elevation was interpolated in the raw DEM. We calculated all coordinates of the entire water-land edge pixels and took the average as water surface position. The observed object-space coordinates were calculated by RFM forward intersection formula (see Equations (17) – (19)). Combined with water-land binary image, \( Z \) coordinates of underwater points were corrected from their observed positions to real positions using the refraction correction equations. Finally, we got the integrated DEM of Ganquan Island and Zhaoshu Island as shown in Figures 11(a) and 12(a), respectively.

We applied three error analysis criteria, namely, the relative error (RE), the root-mean-square error (RMSE) and the correlation coefficient (C), to evaluate the accuracy of estimated depth. The three equations used to calculate the error criteria are as follows:

\[ \text{RE} = \left( \frac{1}{n} \sum_{i=1}^{n} \text{abs}(x_i - y_i) / y_i \right) / n \]  
\[ \text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - y_i)^2 / n} \]  

**Figure 11.** Two-media bathymetry results and error distribution of Ganquan Island. (a) Integrated DEM with refraction corrected in water area; (b) Distribution of relative error.
\[
C = \left[ \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) \right] / \sqrt{ \sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2 }
\]

(23)

where \( x_i \) means two-media measured water depth, \( y_i \) is the corresponding true depth and \( n \) is total numbers.

To evaluate the performance of the two-media photogrammetry bathymetric technique, we compared the model calculated bottom depth values with the reference bottom depths. As shown in Table 2, our method produced bathymetric estimates down to the depth of 20 m, and the overall RMSE of Ganquan Island and Zhaoshu Island are 2.09 m and 1.76 m, respectively. The global RE for accuracy checkpoints have been mapped to show the relationship between RE and water depth (see Figures 11(b) and 12(b)). In theory, the two-media photogrammetry bathymetric technique could also obtain high accuracy in very shallow water as long as the water surface is planar. However, waves in the surf zone make this almost impossible. Usually, the wave surface in those surf zones are much more unstable than the deeper area; thus, the geometric relation of the image rays changes greatly and become unpredictable. In contrast, the water surface change in the deeper area is relatively smaller, and the geometric relationship is closer to the ideal condition. From Table 2, Figures 11(b) and 12(b), we can see that the large relative depth estimate error (red areas in Figure 11(b) and red points in Figure 12(b)) occurred in the very shallow water (less than 5 m), and the RE became much smaller in deeper water (5–20 m). For the 5–20 m depth range, the overall relative error RE of those two test area produced by our two-media method are 17% and 14%, and the correlation coefficient \( C \) reaches to 0.83 and 0.92, respectively. Due to the greater water depth, the overall RMSE obtained here is slightly larger than that of Hodul et al. (2018) and Chenier et al. (2018), but the relative error is basically the same. Skarlatos and Agrafiotos (2018) carried out UAV aerial two-media detection with higher accuracy, but without overwhelming advantages, which may be associated with image resolution, depth range, imaging conditions and other factors.

To further demonstrate the advantage of two-media photogrammetry proposed in this paper, we made a comparison with the traditional multispectral-based

**Table 2. Accuracy assessment and comparison with multispectral derivation method.**

| Test area       | Method   | Criteria | Water depth range (m) |
|-----------------|----------|----------|-----------------------|
|                 |          | 0–1      | 1–3      | 3–5      | 5–10     | 10–15    | >15      | 15–20    |
| Ganquan Island | Two-Media| Numbers  | 3836     | 18021    | 60405    | 144403   | 65616    | 21528    | 251547   |
|                 |          | RE       | >100     | >100     | 58       | 20       | 12       | 15       | 17       |
|                 |          | RMSE (m) | 5.82     | 4.03     | 2.58     | 2.21     | 1.93     | 2.49     | 2.09     |
|                 |          | C        | <0.01    | <0.01    | 0.36     | 0.31     | 0.49     | 0.42     | 0.83     |
|                 | Multispectral | Numbers | 59      | 58       | 68       | 210      | 164      | 41       | 415      |
|                 |          | RE       | >100     | 97       | 51       | 15       | 10       | 24       | 17       |
|                 |          | RMSE (m) | 1.77     | 1.15     | 1.09     | 1.20     | 1.58     | 2.14     | 1.36     |
|                 |          | C        | 0.60     | 0.75     | 0.73     | 0.72     | 0.66     | 0.51     | 0.92     |
| Zhaoshu Island  | Two-Media| Numbers  | 239     | 160      | 118      | 375      | 139      | 53       | 567      |
|                 |          | RE       | >100     | 90       | 41       | 18       | 6        | 11       | 14       |
|                 |          | RMSE (m) | 1.83     | 2.01     | 2.16     | 1.87     | 1.01     | 2.37     | 1.76     |
|                 |          | C        | <0.01    | 0.35     | 0.21     | 0.60     | 0.81     | 0.50     | 0.92     |
|                 | Multispectral | Numbers | 1.73   | 1.55     | 1.98     | 1.75     | 1.78     | 2.30     | 2.12     |
|                 |          | RE       | >100     | 68       | 39       | 18       | 11       | 10       | 19       |
|                 |          | RMSE (m) | 0.19     | 0.47     | 0.37     | 0.63     | 0.62     | 0.43     | 0.93     |
model for water depth retrieval as shown in Table 2. We adopted the two-band ratio model (Stumpf et al., 2003) for multispectral water depth retrieving. We randomly selected 1200 uniformly distributed points from Ganquan Island lidar data. Half of them were used for multispectral inversion model construction and the other half for accuracy checking. In Zhaoshu Island, the number of points for model construction and accuracy checking was 200 and 1048, respectively. From Table 4, we can see that the RE of the multispectral method was also very big in shallow water (0—5 m). In Ganquan Island, those three accuracy criteria of two-media method are poorer than multispectral-based method (RE 17% vs. 17%, RMSE 2.09 m vs. 1.36 m, Correlation coefficient C 0.83 vs. 0.92), while it demonstrates better results in Zhaoshu Island (RE 14% vs. 19%, RMSE 1.76 m vs. 2.12 m, Correlation coefficient C 0.92 vs. 0.93). It can be confirmed that neither two-media photogrammetry nor multispectral derivation method is suitable for very shallow water, and overall bathymetry accuracy in deep water (5—20 m) of the two methods are very close. In addition, we would like to emphasize that the multispectral method depends heavily on the in-situ water depth data and good water quality. However, many times those conditions are unavailable and thus the accuracy is relatively lower than Table 2. In this paper, our two tests are actually representative cases (low texture and sea waves). If water surface are planar, e.g. in the lagoon, the two-media technique is hopeful to acquire more accurate water depth than 14% relative error.

Conclusions and discussions

With rapid satellite imaging sensor development, now it is much easier to obtain high-resolution stereo images anywhere on the earth, it provides the opportunities to utilize the two-media photogrammetry bathymetric technique to retrieve depth information in shallow water. This paper proposes a complete two-media photogrammetric bathymetry workflow and verifies the technical feasibility with WV2 multispectral data.

Two studies using high-resolution WV2 multispectral stereo-pairs illustrate that depth estimates can be derived with a vertical accuracy of about 2.09 m and 1.76 m (RMSE) in water depths from 5 to 20 m. The proposed workflow can produce relatively more accurate depth estimates (relative error better than 17% in Ganquan Island and 14% in Zhaoshu Island) for 5–20 m range but partially poor accuracy for very shallow range (0–5 m).

Although most techniques are based on previous researches and well-developed theories (e.g. RFM), some finding is still worth to mention: the theoretical analysis suggests that although horizontal differences exist between different observation views, these horizontal differences are negligible for the air-borne case. In other words, in order to correct the observed position, only the vertical offset needs to be adjusted.

The proposed workflow tackles some challenge issues and provides several reasonable solutions, such as finding the boundaries between water and land, removing the sun glint effects, integrating sparse and dense matching results; we proposed a two-media bathymetry processing workflow and validated the accuracy using WV2 satellite’s 4 bands stereo multispectral imageries.

The overall bathymetry accuracy achieved using WV2 stereo images have reached the International Hydrographic Organization (IHO) category zone of confidence (CATZOC) C level for survey quality, which requires a plane accuracy of 500 m, a vertical accuracy of 2.5 m for 5—20 m depth range and 3.5 m for 10–30 m depth range (International Hydrographic Organization (IHO), 2017). In short, two-media photogrammetry has the ability to support the improvement of coastal mapping and has good application potential in maritime transport, coastal zone management, fisheries, tourism, environmental protection and so on.

Some further research is needed to improve and fine-tune our techniques and the proposed workflow. Extensive assessments with accurate in-situ water depths are needed to make some solid conclusions about two-media photogrammetric bathymetry accuracy. As mentioned in this paper, other error sources such as sea waves, tides, surfaces, turbidity, sun angle changes between images should be carefully modelled or retrieved. The two-media photogrammetric bathymetry technique for other satellite (such as GeoEye-1, Pleiades-1 and so on) and air-borne sensors should also be explored to get a better understanding of the potentials of two-media photogrammetric bathymetry technique, and if possible to integrate with other bathymetry techniques.

Acknowledgments

The authors are grateful for the editors and reviewers of this manuscript. In addition, we would like to thank The North China Sea Air-borne Detachment of China Marine Surveillance for providing the LIDAR water depth data.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work is supported by the major project of High-resolution Earth Observation System (GFZX01010202, GFZX04031204).
Westaway, R.M., Lane, S.N., & Hicks, D.M. (2003). Remote survey of large-scale braided, gravel-bed rivers using digital photogrammetry and image analysis. *International Journal of Remote Sensing*, 24(4), 795–815. doi:10.1080/01431160110113070

Yang, X. (1987). Two-media close range photogrammetry measurement of sandy soil movements around tunnel models. *Acta Geodetic Et Cartographica Sinica*, 16(4), 261–270.

Zhou, G., Li, Y., Ren, Y., Sheng, L., Ye, D., Fan, F., & Bai, J. (2015). Research of two-media underwater reefs depth measurement experiment based on low-altitude UAV. *Acta Geodetic Et Cartographica Sinica*, 44(5), 548–554. doi:10.11947/j.AGCS.2015.20140259