

Robust Automatic Generation of True Orthoimages From Very High-Resolution Panchromatic Satellite Imagery Based on Image Incidence Angle for Occlusion Detection

Aleš Marsetič 

Abstract—Orthoimages have proven to provide useful reference data for a variety of geospatial applications. Their improved versions, true orthoimages, are even more valuable as they have been adjusted also for high objects. Even though true orthoimages have been—over the last decades—generated mainly from aerial images, we believe there is a need to produce true orthoimages also from satellite images. The aim of this article is to present the already developed automatic procedure for generating true orthoimages from very high-resolution optical satellite images. The automatic workflow consists of five modules, starting with the extraction of ground control points, through the geometric processing of image blocks, occlusion detection, orthorectification, and finally the generation of true orthoimages. The orthorectification module uses a high-resolution digital surface model derived from laser scanning data. Occlusion detection is performed using a ray-tracing method based on the incidence angle, which is suitable for images acquired by satellites and processed with rational function models. The usability of the workflow was demonstrated with the test of six Pléiades images that were acquired on different dates and with different viewing angles. The research presents the accuracy of the geometric correction and the quality of the true orthoimage generated with different numbers of images. The results show that satellite-based true orthoimages are suitable for the extraction of spatial information that can be used directly by the end-users.

Index Terms—Automatic processing, digital surface model, orthorectification, true orthoimages, very high-resolution satellite imagery.

I. INTRODUCTION

THE PRODUCTION of orthoimages has a long tradition. The need for images in an orthogonal projection within a designated coordinate system emerged simultaneously with the ability of imaging cameras to acquire images in higher spatial resolutions. Satellite orthoimages have been generated from 1970s onward, however, orthoimages have been used in aerial photography already prior to this. In many countries, orthoimages currently present a critical component of the national

Manuscript received September 29, 2020; revised December 10, 2020, January 26, 2021, and February 25, 2021; accepted March 16, 2021. Date of publication March 19, 2021; date of current version April 14, 2021. The work was entirely funded by the Slovenian Research Agency under Code ID Z2-8193.

The author is with the Institute of Anthropological and Spatial Studies, Research Center of the Slovenian Academy of Sciences and Arts, 1000 Ljubljana, Slovenia (e-mail: ales.marsetic@zrc-sazu.si).

Digital Object Identifier 10.1109/JSTARS.2021.3067457

spatial data infrastructure, however, their production from aerial imagery is usually periodical.

Orthoimages have a high usability level as they contain both the image characteristics of a photograph and the geometric properties of a map [1]. Orthoimages are especially indispensable as references in a variety of geospatial applications. They can replace vector data when this is not available, or complement it. They are used as raster base maps in geolocated spatial databases to generate geodetic data for planning purposes, and can also be used to realistically overlay three-dimensional (3-D) models with image data [2]. Orthorectified images are of crucial importance when one wants to use satellite images in combination with other georeferenced data obtained from other sources to conduct various spatial analyses.

Alongside high usability, traditional orthoimages also have a few drawbacks (especially in areas with high buildings) [1], due to which their usefulness in industry, government, and elsewhere is significantly reduced. The drawbacks are mostly a consequence of using a digital terrain model (DTM) during orthorectification. A DTM provides only the elevation of the terrain but does not account for the objects above ground. For this reason, the interest in generating so-called true orthoimages has been on the increase over the past several years. True orthoimage is an orthoimage in which not only geometric distortions, but also distortions that result from high objects (e.g., high buildings), are minimized. Generation of true orthoimages is a more complex procedure than the generation of traditional orthoimages, since this requires more images of the same area, acquired from different angles (the so-called image block). A block of images is required as each generated orthoimage contains occlusions that are a consequence of the high objects in the image combined with the use of the digital surface model (DSM). With the use of multiple orthoimages these occlusions can be compensated for, resulting in the creation of the final true orthoimage, which is basically a mosaic created from several images. Optical satellite images have a high range of spatial resolutions, from low to very high. Basically, only those with very high-resolution (VHR; resolution of 2 m or less) are used in the creation of true orthoimages.

Some areas of spatial data application (e.g., natural disaster management or urban planning) require spatial information of

the highest quality, which is not provided by usual orthoimages, however, it can be obtained from true orthoimages. Therefore, true orthoimage generation techniques are essential in assuring reliable interpretability and maintaining the high quality of the available data. True orthoimages have already been used advantageously in aerial photogrammetry (in urban centers) as well as in terrestrial photogrammetry (in architecture and archaeology).

The generation of true orthoimages started with aerial images in the 1990s [3], usually relying upon digital building models (DBMs). Most of the existing true orthoimage generation techniques are based on the Z-buffer algorithm and its improvements (e.g., [1], [4], [5]). The drawbacks of this method were later addressed with the angle method in [6], and its optimization, which has been proposed in [7].

With the popularization of aerial laser scanning, lidar data started to be used for the generation of true orthoimages. In [8], orthoimages were generated with a lidar digital surface model (DSM) and a building vector map. The results without the vector data were not entirely satisfactory. In [9], the lidar surface model was used to detect the occluded areas. Good results with lidar and the angle method were obtained in [10].

In recent years, true orthoimages have also been generated from unmanned aerial vehicle (UAV) data. In [11], a hybrid TIN model was used for the orthorectification and generation of a 3-D building model. In [12], orthophotos were generated from Ultra-Cam imagery. They used the elevation buffer technique which works with plane elevations and is sensor independent. Despite the promising results there is still a myriad of possibilities for further improvements, especially in the detection and correction of moving objects, object destretching and radiometric corrections of shadow areas (deshadowing). In [13], a structure from a motion (SfM) workflow was used to generate high-precision true digital orthophoto maps.

Very few studies have been conducted on the subject of generating true orthoimages from satellite imagery. One of the first was the work in [14], in which true orthoimages were generated from QuickBird images. In this case, the production of true orthoimages comprised of three steps: generating orthoimages with a DTM, adjusting the images with a DBM, and detecting and filling the occluded areas. Orthorectification was performed with a rational function model (RFM). In [15], a sorted digital elevation model method for generating true orthoimages from the IKONOS satellite was proposed. In this method, the orthoimage pixels are filled in the order of their elevation, and each pixel in the source image is accessed only once. Therefore, this method requires less memory than the traditional Z-buffer method. The work in [16] shows a procedure to improve the satellite orthoimage accuracy with the iterative ray-tracing method. About 94% of all the occluded areas were correctly determined with the use of the lidar elevation model and the proposed method.

Many studies have focused on radiometric normalization and mosaicing of images, which are essential procedures in the true orthoimage generation process. More recent research on satellite imagery has been carried out in a few studies, most of which focused on multispectral high-resolution images. A mixed radiometric normalization method was presented in [17], whereas

in [18] a radiometric normalization method for multitemporal images during mosaicing using an iteratively reweighted radiometric adjustment without master images was proposed. A robust procedure for generating high-quality and seamless urban imagery with a spatially adaptive method and stepwise local moment matching are proposed in [19] to remove all kinds of clouds. A multitemporal mosaicing algorithm specially designed for operational S3 biophysical or other products within the context of the S3/FLEX tandem missions [20] also showed good results for derived products. Most of the methods mentioned are computationally expensive and depend on a successful atmospheric correction. A comprehensive and quantitative summary of recent methods for remote sensing mosaicing in relation of the aspects of radiometric normalization, seamline detection, and image blending is presented in [21].

Due to the increase in the quantity of aerial and satellite data, the researchers in the fields of remote sensing and photogrammetry are developing highly automated or completely automated methods for image orthorectification. The relatively recent introduction of powerful instruments, processing stations and tools led to an increase in the research in this field (e.g., [22]–[29]). Although the numbers of new methods are on the rise, the emphasis is placed on the complete automation of already existing techniques.

Only a few studies have focused on fully automatic generation of true orthoimages and they revolve exclusively around aerial images. The first attempts took place at the end of the last millennium. In [30], an automated method for images from a full-frame aerial camera was developed. In order to generate radiometrically harmonized orthoimages they focused on the image mosaicing method. In [31], Biasion *et al.* worked on a method for densifying terrain models and orthorectification. A semi-automatic method was developed in [32]. A lidar surface model was used to generate the true orthoimages which were then corrected with digitized building outlines in a vector format. Only the first part of the processing was automatic. In [33], an automatic method for the production of a lidar surface model suitable for orthorectification was proposed. In the event of irregularly shaped buildings, a semi-automatic procedure for shape correction is included in the software. In [34], an improved method for the automatic generation of urban true orthoimages using a terrain model, building model, and aerial images was suggested. The method achieves an absolute planar accuracy of about a pixel and can automatically fill the occluded areas. The method is applicable only for images acquired with a single projection center (most aerial images). Therefore, it cannot be used for images acquired by linear scanners that are utilized by very high-resolution satellites.

Although the topic of automated generation of true orthoimages was mentioned on a few occasions, it has never been thoroughly discussed. In addition, automated processes have never been considered when dealing with satellite images that acquire images in pushbroom mode and are processed with an RFM.

As these topics were never properly addressed, this article proposes and presents a new automatic workflow for the generation of satellite true orthoimages from very-high resolution

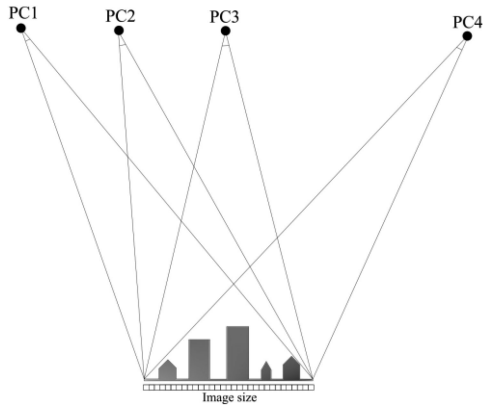


Fig. 1. Example of the acquisition of a block of four satellite images over a city area. In reality the image rays that derive from the perspective center of each image are almost parallel.

optical satellite images. The workflow is composed from three major steps: the first extracts ground control points (GCPs) and performs geometric modeling, the second generates orthoimages from an arbitrary number of input images (see Fig. 1), whereas the third generates the final true orthoimage (when at least two orthoimages are produced in the first step). All processes used to generate the final results are interconnected, automated, and based on stable and thoroughly tested automatic orthorectification procedures that were previously developed within the research team.

In order to develop an automatic processing method, a number of critical points had to be taken into account during development. The automatic workflow consists of many steps that had to be combined in a robust and smoothly functioning automatic system. The automatic procedure is very complex and contains many iterations and controls that guarantee the stability and correctness of the results. Since there is no human interaction, it has to deal with errors that may occur and different types of data. In the automatic workflow for true orthoimage generation, many iterative procedures are involved in the extraction of GCPs and removal of gross errors in the geometric modeling. Before generating the true orthoimage, the procedure must subset each orthoimage to the same AOI. Rapid quality checks are also performed after each step, but the main validation is usually performed after the generation of the final result.

In addition to the robust automation of the true orthoimage workflow, the article proposes a new approach for occlusion detection that is suited to VHR satellite images. This new method utilizes the incidence angles to generate an occlusion mask that is used in image orthorectification. This helps us avoid the double mapping effect and allows us to use the three orthoimages directly in the occlusion compensation process.

The rest of the article is organized in three sections. Section II introduces the steps within the implemented automatic workflow with the emphasis on the generation of true orthoimages. Section III describes the performed experiments, assesses the procedure performance, and discusses the usability and limitations of the results. Section IV presents the main conclusions and ideas for future work.

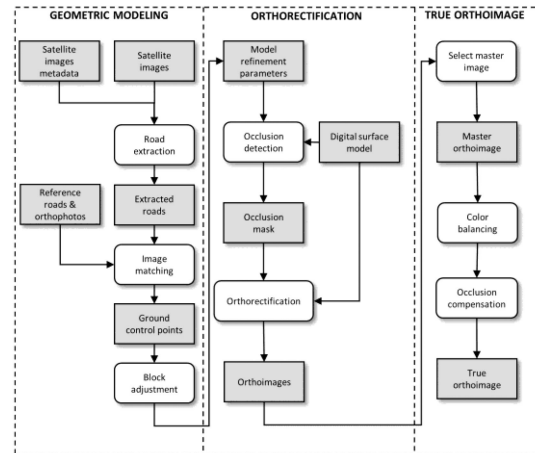


Fig. 2. Overview of the three-step true orthoimage generation workflow.

II. AUTOMATIC TRUE ORTHOIMAGE GENERATION WORKFLOW

The proposed true orthoimage generation workflow is comprised of three major steps (see Fig. 2). The first step was already implemented in the STORM processing chain [28]. STORM is an automatic processing chain that performs all processing steps from raw satellite imagery to map-ready products. Its geometric correction module can automatically generate orthoimages from various high resolution (HR) and very-high resolution (VHR) images. The module is composed of three submodules: automatic extraction of GCPs that is based on matching the roads detected in the satellite image onto either reference rasterized roads (two-step algorithm) or reference rasterized roads and orthophotos (three-step algorithm), followed by automatic sensor modeling, during which either the proprietary physical sensor model or RFM are utilized, and ending with orthorectification. The orthorectification procedure for HR images is presented in [27]; the upgraded algorithms that support image resolutions of up to 2 m were tested in [28], and the current version of the improved and upgraded geometric module that enables the processing of submeter panchromatic images is described in [29]. The same geometric modeling method was used in the present workflow for the generation of true orthoimages.

The new developments address the second and third step. The second step generates the occlusion masks and orthoimages for each image in the block. The third and last step generates the true orthoimage. This part of the workflow takes individual orthorectified images, which were produced in the second step, as the input. First, it selects the master image, then it continues with color (radiometric) balancing, proceeds to occlusion compensation, and ends with the mosaicing of the images into a single true orthoimage.

The entire procedure is automatic. The user only needs to select the paths to the input image folders and click the start button, which makes the procedure read the images and their metadata, thus starting the first step. The processing is finished when the true orthoimage and the processing information file are generated. Due to the availability of the DSM and other ancillary

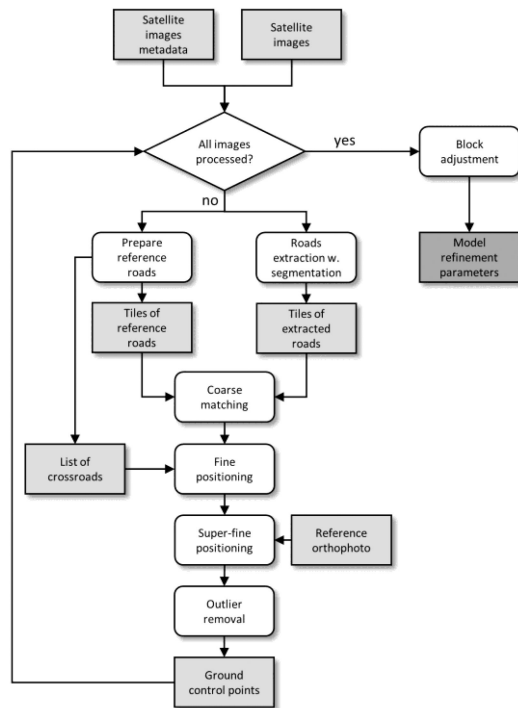


Fig. 3. Automatic extraction of GCPs and block adjustment workflow.

data needed during the procedure, the workflow currently works only for Slovenia.

The following two sections (II-A, II-B) provide an overview of the first two workflow submodules which have been extensively described in [29] and are only briefly mentioned in this article. The final three sections (II-C, II-D, II-E) present a detailed view of the new developments that produce the resulting true orthoimage.

A. Automatic Extraction of Ground Control Points

The first processing step automatically retrieves the reference roads and orthophotos based on the image map coordinates in the metadata file. The obtained data is then used in image matching. The basis of the implemented GCP extraction is to match the features extracted from the input images to reference data. The extraction follows a number of steps (see Fig. 3). First a two-step algorithm that utilizes vector roads as a reference layer (combination of the national road database and OpenStreetMap road data) and delivers GCPs for HR images is used. The matching is performed between the reference data (road map) and the roads extracted from satellite images with watershed segmentation and decision tree classification. Rather than using a single global mapping function for the entire satellite image, the algorithm matches individual image tiles. In order to support VHR data, the module was upgraded with a super-fine positioning of individual GCPs on aerial orthophoto chips. The final stage utilizes least square matching for sub-pixel positioning.

When extracting GCPs, national spatial reference data are required to process VHR images. Although this may limit the generality of the procedure in some regions, some reference data

can also be found in freely accessible web databases or national spatial databases. The reference roads can be found in national topographic databases or web databases (e.g., OpenStreetMap - OSM). OSM road layers are globally available, and although their accuracy is questionable, they can still be used as a supporting layer or when more accurate data is not available.

Difficulties may arise when there are few or no roads in the images, or when the accuracy of the reference data is not high enough. In such cases, the extracted GCP may inadequately refine the geometric model. This can lead to orthoimages that do not match and can produce unusable true orthoimages.

In order to achieve a higher number of GCPs and enable automatic selection of points on the same locations but in various images, GCP candidates were recruited solely on crossroads. This enabled us to obtain two or more images that might share the same GCPs—points with the same reference (cartographic) coordinates, but different image pixel coordinates for each image. Such GCPs connect the images during the adjustment process, which results in improved alignment once geometric modeling has been completed.

B. Geometric Modeling

The extracted GCPs serve as the input for geometric modeling. The modeling is performed with an RFM. The RFM is an empirical model and provides a good alternative to the physical model with equivalent results [35]. The list of GCPs obtained from the extraction procedure is additionally filtered with gross error removal algorithms that produce the final set of points used to refine the rational function geometric model. Because the generation of a true orthoimage demands a set of aligned orthoimages, all input images are processed together in a block adjustment. The obtained bias-corrected results are used for the orthorectification of the images.

C. Occlusion Detection

Detecting occluded pixels is a crucial step in creating true orthoimages. In rectified images the occluded areas are located on the side of high objects, where they are not visible from the camera's point of view. Many methods detect occlusions already during the rectification step (e.g., [10]) as the same method is used to project the rays from the perspective center to the terrain.

Orthorectification with rational polynomial coefficients (RPCs) has become a standard procedure when using VHR satellite images. Although RPCs are easy to use and reliable for orthorectification, they cannot be used in the most established occlusion detection methods that were developed for aerial images. In these methods we need information pertaining to internal and external orientation parameters, which is not available when we use an RFM with satellite images.

As these methods cannot be used with satellite images, we have divided the process into two parts (see Fig. 4). First, we generated a binary mask with occlusions, and then we orthorectified the image based on the occlusion mask. The generation of the mask starts simultaneously with the extraction of GCPs, whereas the orthorectification begins as soon as the mask and

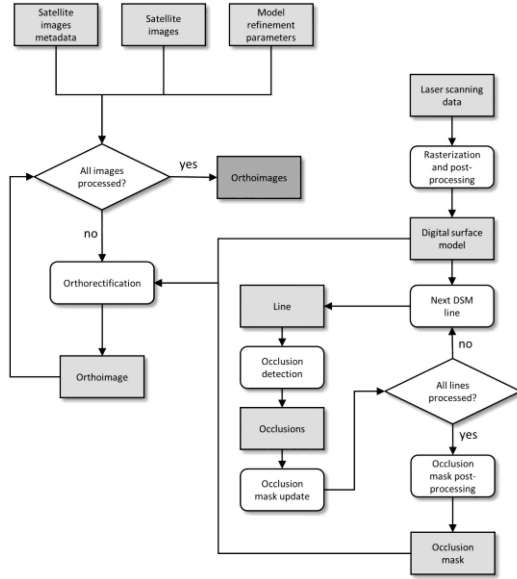


Fig. 4. Occlusion mask and orthoimage generation workflow.

model parameters are made available. Thus, we also optimized and sped up the generation of orthoimages.

In the proposed procedure, the mask of occluded areas (i.e., occlusion mask) is produced from the image incidence angle and the ray-tracing method. The ray-tracing method is a direct iterative process [36] and can be very precise when combined with a very accurate DSM. The greatest drawback of this method lies in the long processing time [27], which can only be reduced with intelligent programming and multicore processing. The mask generation procedure with ray-tracing uses precise incidence angle values that can be found in the image metadata file. In the calculations, the across and along track incidence angles are employed in order to determine the azimuth of the sensor and the incidence angle, which provides us with the elevation of the sensor. As the incidence angle is almost the same for the entire image, the implementation of the procedure is straightforward. First, a mask raster of a size that covers the input image is generated. Based on the DSM and incidence angles, each pixel within this raster is evaluated and assigned to an occluded or nonoccluded class. The procedure starts with the computation of the azimuth and elevation angle of the sensor with the following equations:

$$Az = \tan^{-1} \left(\frac{\tan(\Theta_{ac})}{\tan(\Theta_{al})} \right) \quad (1)$$

$$El = 90^\circ - \Theta \quad (2)$$

in which Az is the azimuth angle, El is the elevation angle, Θ_{ac} and Θ_{al} are the across and along track incidence angles, respectively, whereas Θ is the incidence angle in degrees. The pixel's height (Z) is determined from the DSM. In order to determine if the pixel is occluded, the height is incremented by a previously specified step and the planar coordinates on the

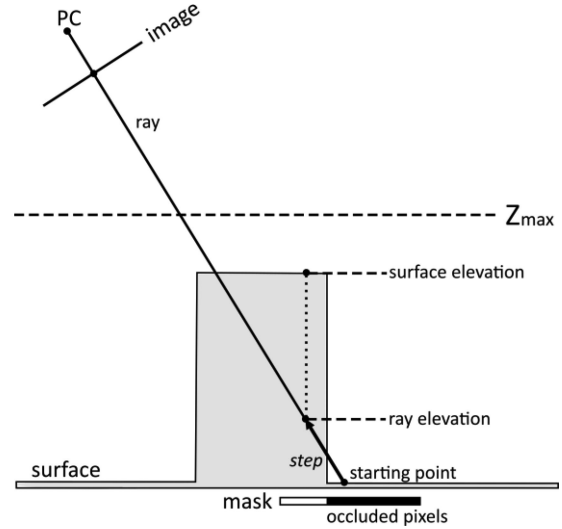


Fig. 5. Generation of the occlusion mask. The process works iteratively by gradually extending the ray until it intersects with the object (surface) or reaches the maximum DSM elevation (Z_{max}). In this example, the pixel at the starting point is classified as occluded.

viewing ray are calculated as follows:

$$X_n = \frac{(Z_n - Z) \cos(Az) \cos(El)}{\sin(El)} + X \quad (3)$$

$$Y_n = \frac{(Z_n - Z) \sin(Az) \cos(El)}{\sin(El)} + Y \quad (4)$$

where X_n , Y_n , and Z_n are the new coordinates of the viewing ray, whereas X , Y , and Z are the coordinates of the pixel.

The calculations in (3) and (4) are repeated with incremented Z coordinates until the height in the DSM is higher than that of the viewing ray (the ray intersects an object—see Fig. 5) or the ray reaches the maximum Z of the DSM. In the latter case, the pixel is considered to be nonoccluded. The iterative procedure is repeated for each pixel within the image. In the implemented algorithm, the linewise computation supports multiprocessing.

Although generally precise, the resulting mask may contain small artifacts that are usually a consequence of the DSM irregularities or errors. Artifacts can also appear if the ray-tracing step is too large. A larger step can reduce the mask generation time, so special attention must be paid to balance the accuracy of the mask with the processing time. These generated artifacts, consisting of a few pixels, are partially removed with postprocessing techniques (morphological filtering).

D. Orthorectification

Whereas the extraction of GCPs and the geometric modeling are implemented exactly as described in [29], the orthorectification submodule was modified to accommodate the use of a DSM. The use of this model minimizes the distortions that are a result of high objects (e.g., high buildings) in the resulting orthoimages.

The accuracy of the final product depends on the precision of the utilized DSM as this must correspond precisely to the high

objects on the images that are being rectified. For this reason, two types of DSMs were tested: lidar DSM and a DSM generated by image matching of stereopairs. Although the latter DSM reflects the current state of the objects and is perfectly aligned with the orthoimages, it generally lacks the sharp edges of buildings which are important for generating precise true orthoimages. Whereas DSM generated from lidar data usually shows small misalignments—as it is obtained from a different data source and is acquired within a different time frame—it is usually more suitable for generating satellite true orthoimages as it clearly indicates the object’s shapes and edges.

Thus, lidar DSM was used to orthorectify the images. The DSM was generated from laser scanning data obtained in 2015. The country-wide scan was performed at a point density of 5 points/m². We used Lastools (v. 200101) to calculate the DSM with a spatial resolution of 0.5 m without spikes and pits [37]. The algorithm does not form needle-shaped triangles in vegetated areas and along the roofs of buildings that appear as spikes and holes in the TIN. We used a normalized point cloud and closed all the gaps that were wider than 1.5 m. Even though the time difference between the acquisition of the images and the lidar was almost 2 years, the sizes and shapes of most buildings remained almost the same, thus most of the changes were to be found in the vegetation.

Several methods were considered for projecting the image coordinates into the orthoimage coordinates. As direct methods have longer processing times and generate orthoimages with unassigned grey values [27], we decided to implement the so-called indirect method. This method starts in the object space (orthoimage) and projects the pixel onto the input image through the DSM. The grey value of the pixel is usually obtained by resampling the neighboring pixels on the input image. Indirect methods are generally better than direct ones in terms of quality and processing speed [38]. In addition, indirect methods do not produce empty pixels. However, indirectly generated orthoimages may contain double-mapped areas (ghost images) when high-resolution DSMs are utilized.

In the proposed procedure, this problem was solved with the use of a mask of occluded areas. The orthorectification was then conducted only on nonmasked areas. The result is an orthoimage, also in which the objects that are not usually defined by a traditional DTM are also represented in the orthogonal projection: the roofs are displayed on top of the foundations, without displaying the facades and without hiding the building’s surroundings. The occluded areas are pixels without an assigned value. Examples of a traditional orthoimage, an occlusion mask, and an orthoimage generated with DSM are presented in Fig. 6.

E. True Orthoimage Generation

The ultimate aim of the described automatic workflow is the generation of a true orthoimage from a set of orthoimages. A minimum of two images must be used as inputs. If more images are available, it is important to select images acquired from different directions (see Fig. 7). Only a block of images can compensate for the occlusions and generate a homogeneous result. The generation of a true orthoimage can also be referred to

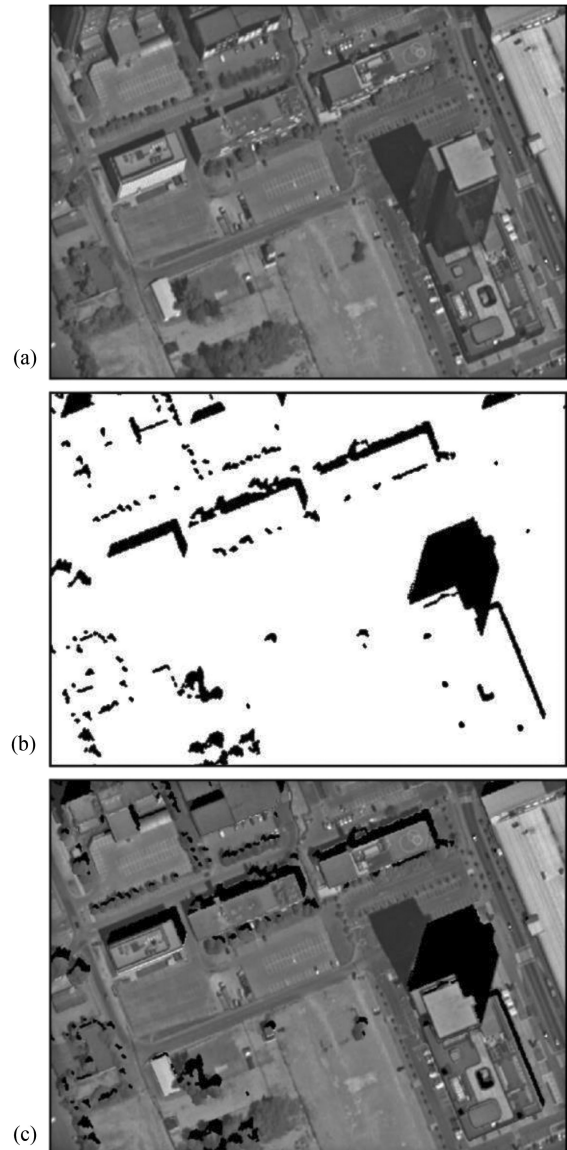


Fig. 6. Examples of orthorectification data generated during the workflow. (a) Orthoimage generated with DTM (for the purpose of comparison). (b) Occlusion mask. (c) Orthoimage generated with a DSM.

image mosaicing in which patches of different images are used to compose the final product. The procedure employed in our method consists of various steps which are presented in Fig. 8 and described in the following sections.

As a part of the automatic procedure, the true orthoimage generation starts only once all orthoimages in the block have been generated. These orthoimages represent the sole input data in the final step, which begins with the selection of the master image.

1) *Master Image Selection*: Unlike aerial images which are usually taken in blocks with predefined overlaps (frontal and side) and nadir angles, satellite images are acquired with different angles and from different sides of the target area of interest (AOI). Therefore, the selection of the cost function that determines which image is chosen to fill the true orthoimage is

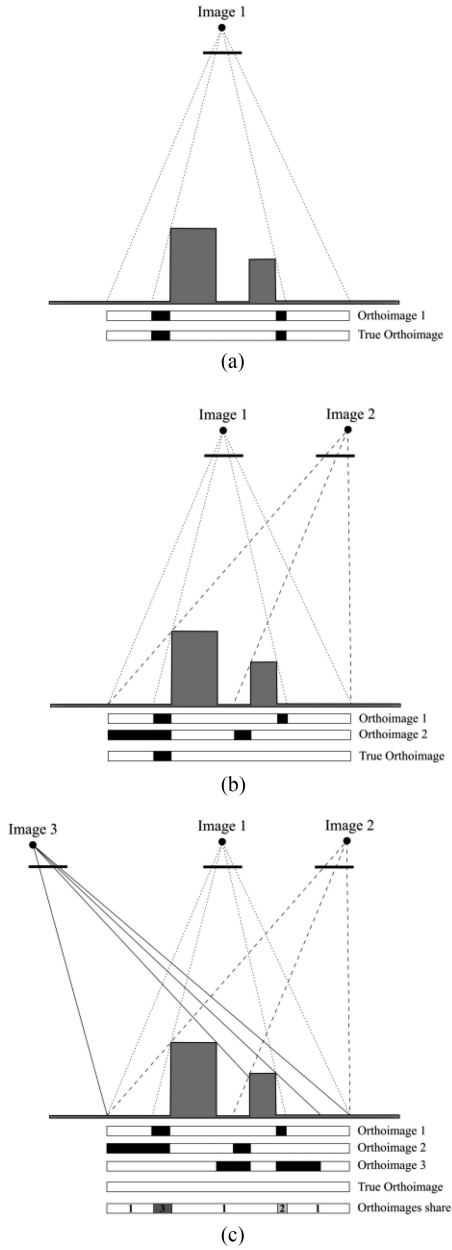


Fig. 7. Generation of a true orthoimage from satellite images. The number of images defines the amount of occlusions that remain present in the result. (a) Use of a single image—many occlusions that are found in the image are reflected in the true orthoimage. (b) Use of two images—less occlusions are present in the true orthoimage. (c) Use of three images—the images compensate for all occlusions; because of the different viewing angles, all three images contribute to the true orthoimage.

slightly different. When working with aerial images, the most common cost functions are based on the minimum angle or nearest ground nadir methods (e.g., [39]). With these methods the true orthoimage is almost evenly covered with block images—i.e., no master image is selected.

As satellite images usually cover the entire AOI, the master image can be selected and can be entirely based on the viewing angle (i.e., minimum angle in the block). The image with the

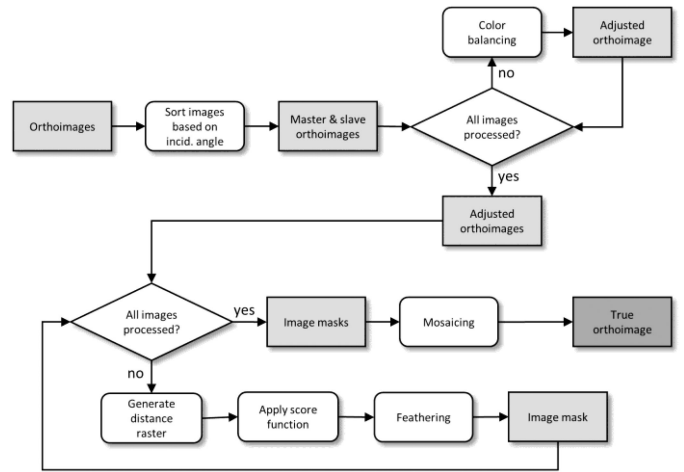


Fig. 8. True orthoimage generation workflow.

lowest angle has smaller occlusions as well as minimal distortions caused by objects (e.g., vegetation and cars) that are not accurately modeled by the DSM. The satellite block is usually composed of a few images with specific viewing angles that facilitate the choice of the master image. For the next step—radiometric balancing—the master image needs to be selected prior to mosaicing, as this step needs a reference image in order to perform the radiometric adjustment. The final selection of image patches for mosaicing is described in Section II-E3.

2) *Radiometric Balancing:* Radiometric balancing of satellite images is an essential step that needs to be performed before the mosaicing of the images takes place, since these images have been acquired at different times. Radiometric balancing performs a correction that compensates for the visual effects (e.g., shadows) between the two images, which occur as a result of the possible variations in the viewing angle, the position of the sun, time of day, time of year, and the atmospheric conditions. If the images were captured on the same orbit, the most important factor that influences the radiometry is the viewing angle, whereas in the case of images acquired from different orbits, the radiometric values depend on all factors.

As a result of the aforementioned factors, the radiometric variation can result in radiometric discontinuity in the final true orthoimage. In order to harmonize the patches that compose the true orthoimage, various methods can be employed when adjusting the values between two images. Satellite images available to users (level 1A and above) are radiometrically corrected throughout the image. Therefore, two images can be radiometrically harmonized with a global adjustment of one image to the other. Histogram matching [40] is a simple method for harmonizing two images. This method attempts to transform the input histogram so that it resembles the reference histogram.

The simplified histogram matching equation can be written as follows:

$$v_r = C_r^{-1} (C_i(v_i)) \quad (5)$$

where C_r and C_i are the cumulative histograms of the reference and input images, whereas v_r and v_i are the values of the respective histograms.

In our processing chain, we implemented histogram matching in order to radiometrically adjust and transform the slave images to match the master image. Before they were matched, the images were trimmed to the area that was shared by all orthoimages and which was also the size of the true orthoimage. This ensured that the orthoimages were adjusted and ready for occlusion compensation.

Although histogram matching is usually effective, it can produce undesirable results when images from different dates and/or with different atmospheric conditions are used. In such cases, matching is difficult and the algorithm may fail in matching the radiometry of the images.

3) *Occlusion Compensation*: As presented in Section II-E1 all satellite images were primarily sorted based on their viewing angle, at which the one with the lowest angle was selected as the master image. Although the viewing angle is the most important factor, it is not the only parameter that needs to be taken into account when the images are mosaiced. When stitching different images together also the position of the seamlines (the transitions between two adjacent images) is important. If the seamlines are positioned on the edge of the occlusions, the results will depend highly on the accuracy of the DSM and the (mis)alignment between the DSM and each of the images. Even with small inaccuracies, deformations will be noticeable around buildings and other high objects.

In order to reduce the impact of the DSM on the true orthoimage, the seamlines need to be placed as far away from the occlusions as possible. If we wish to achieve the desired position, the distance-from-the-occlusions parameter needs to be considered in the occlusion compensation [41]. This parameter places the seamline between the occlusions in all involved images. In this way, the images will overlap on both sides of the seamline, which will consequently enable the feathering procedure described in the next section.

When using this method, it is impossible to control the position of the seamlines in regard to the objects in the images, such as roads, hedges, buildings, etc. These seamlines will often run through these objects, splitting them into one or more parts, and will often run across roofs. The greatest anomalies occur when the seamlines run through moving objects (e.g., cars) as these are harder to correct. Although these circumstances should generally be avoided, their consequences can be minimized if appropriate radiometric balancing and feathering is performed on the images.

In order to include the distance-from-the-occlusions parameter, we generated a raster based on the binary occlusion mask, in which the pixel values resemble the distance to the nearest occlusion. This raster was generated with the distance transform function [40], [42]. The function obtains the distance raster through the use of the Euclidean distance to the occlusions. This raster is then used to determine which orthoimage is the source of the pixels in the true orthoimage.

The provenience of each pixel value is calculated with a score function that takes into account two parameters: the viewing angle and the distance to the occlusions. By combining the two parameters the true orthoimage is composed from the images with the lowest angle to the AOI and the furthest distance from

the occlusions. The distance from the occlusions is represented by the values in the distance raster, whereas for the viewing angle parameter we introduced a cost function that lowers the weights of images with higher viewing angles. Thus, images with a smaller angle are prioritized. The cost function can have different forms. In this article, two simple approaches were tested: linear and power function. The score function is obtained by multiplying the cost function with the distance raster.

The equations of the derived linear (6) and power (7) score functions can be written as follows:

$$f_i = d_i \cdot \left(-\frac{I_j}{90} + 1 \right) \quad (6)$$

$$f_i = d_i \cdot \left(-\frac{\sqrt{I_j}}{\sqrt{90}} + 1 \right) \quad (7)$$

where f_i is the score for pixel i , d_i is the pixel value in the distance raster, and I_j is the viewing angle (in degrees) of image j .

When the score for each pixel is calculated, the procedure generates a raster in which each pixel belongs to one of the orthoimages in the block. As the master image has the lowest viewing angle it forms the largest part of the true orthoimage. Besides the areas with the highest score function value, it is also used to fill in the patches far from occlusions and the areas where the occlusions are present in all images (patches with no data), if these exist. Slave images usually form a smaller part of the final image. If the master image is close to nadir, their percentage is truly small.

4) *Feathering and Mosaicing*: Although the images used in mosaicing are radiometrically balanced, there are still visible differences in the radiometric values on both sides of the seamlines. These differences are usually small, but can be extensive when the gradient between the values is steeper. The radiometric differences close to the seamlines can be minimized with feathering.

Feathering aims to eliminate the seamlines between adjacent images by applying a weighted average method [43]. A simple and efficient method of feathering (which was used in our procedure) is to apply a low-pass (i.e., smoothing, averaging) filter. Such filters are often used for removing noise from images and are therefore suitable for smoothening values around the seamlines. Even though using a small kernel multiple times is usually faster than applying a large kernel, both methods produce similar results.

The size of the kernel and the number of times it is applied defines the region of the feathering. For example, when using a size 5 low-pass kernel, the feathering will be 4 pixels wide when applied once. When the filtering is performed 5 times, the feathering will become 20 pixels wide. Fig. 9 shows the percentage of each image pixel value that forms the final pixel, when the kernel is applied once or 5 times. It is clearly visible that multiple filtering produces a wider feathering size with smoother blending between pixels. Whenever possible, the use of a larger feathering zone is recommended, as this yields better results. Thus, the final true orthoimage pixel values are a combination

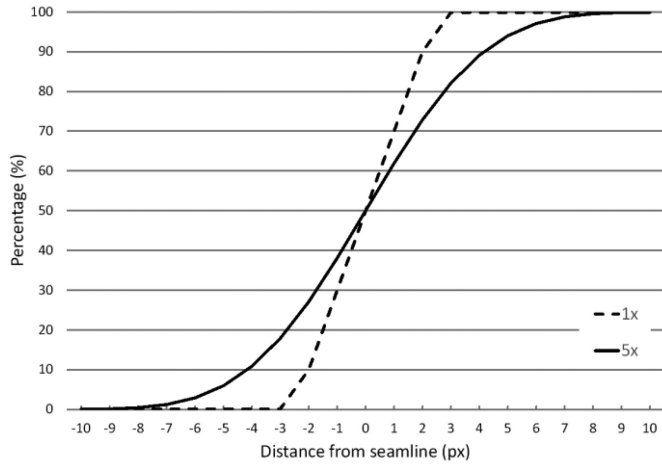


Fig. 9. Curves represent the percentage of an image pixel value in regard to the distance from the seamline. This example presents the situation for one and five kernel applications. If the distance from the seamline is 0, the output pixel value is composed equally from both images. Based on the number of applications, the percentages differ when moving from the seamline. Both images have the same percentages in pixels, but mirrored (left side of the curves contain a higher percentage of the left image and vice versa).

of two images, but only when the pixels are located within the feathering zone.

Feathering can reduce the effect of seamlines and smooth the transition between images, but has problems when images have shadows at different positions or are not properly radiometrically balanced.

In this procedure, feathering is applied to the raster that was determined by the score function. The raster is filtered with the low-pass kernel for each orthoimage. The obtained mask is then used in the mosaicing process which is the final step in the procedure.

The last step uses the available orthoimages and other data obtained in the previous steps to generate the true orthoimage. First, an empty raster, the size of the usable area shared by all orthoimages, is created. Then, the value of each pixel is computed as the sum of weighted values from all images

$$t_{ij} = \sum_{k=1}^n m_{ij}^k \cdot o_{ij}^k \quad (8)$$

where t_{ij} is the value of the pixel at position i, j in the true orthoimage, n is the number of input images, m_{ij}^k is the weight of pixel i, j in mask n , and o_{ij}^k is the value of pixel i, j in the radiometrically balanced orthoimage n . The mask of each orthoimage is the product of the distance raster, the viewing angle and the feathering. Most pixel values will derive from a single image, whereas the pixels close to the seamlines will be composed from several images.

With this step the true orthoimage has been generated and the automatic process writes the product to the hard drive and generates a log file with all the information on the processing and the properties of the product.

Most true orthoimages generated from satellite images still contain occluded pixels with no values (empty pixels). The

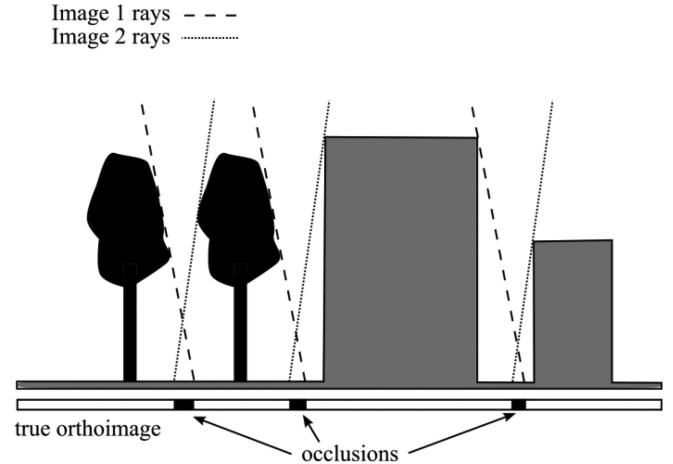


Fig. 10. When satellite images with higher viewing angles are used, the true orthoimages may still contain occlusions even after the mosaicing process has been applied. In this example, three small occlusions remain even after the two images have gone through the mosaicing process.

TABLE I
SELECTED METADATA PARAMETERS FOR THE SIX PLÉIADES
PANCHROMATIC IMAGES

Image	Date	Samples (px)	Lines (px)	Viewing ang. (°)
1	22.7.2013	13,981	16,976	19.5
2	3.8.2013	14,070	15,119	23.6
3	3.8.2013	15,483	17,448	6.8
4	3.8.2013	15,597	17,648	2.2
5	3.8.2013	15,594	17,628	2.7
6	9.9.2014	14,807	16,928	16.2

number of pixels depends on the combination of viewing angles from the orthoimages that compose the true orthoimage and the height and distribution of the objects (e.g., trees, buildings) present in the AOI. Fig. 10 presents the reason for the presence of occluded pixels in the final image. Although images are taken from different angles, the height of the objects that are close to each other prevent the rays from reaching all areas. Only nadir images can cover the entire scene.

III. EXPERIMENTAL RESULTS AND PERFORMANCE ASSESSMENT

A. Test Data and Performance Assessment Methodology

The developed processing workflow is designed to work with VHR satellite images representing the same area. In order to test the procedure, we focused on the area covered by the city of Ljubljana. We used a dataset of six panchromatic Pléiades images acquired in 2013 and 2014 (Note: Four images were acquired during the same pass, whereas two images were acquired individually) with different off-nadir viewing angles. The level of processing of the images is Primary with a spatial resolution of 0.5 m and an image area of approximately 68 km².

The AOI mainly consists of urban areas which are ideal for testing the proposed algorithm for generating true orthoimages. The properties of the test images can be seen in Table I. They exhibit an array of off-nadir viewing angles, which range from 2.2° to 23.6° (see Fig. 11). The images were acquired from various

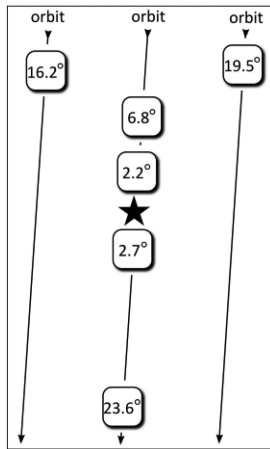


Fig. 11. Test dataset is composed from six Pléiades images that were acquired from three orbits. The figure shows off-nadir viewing angles and the approximate locations of image acquisitions, as well as the AOI (marked with a star).

TABLE II
NUMBER OF USED GCPs AND RMSES AT 15 ICPs FOR
EACH IMAGE (UNIT: PX)

Image	No. of GCPs	RMSE X	RMSE Y	RMSE XY
1	158	0.82	0.64	1.04
2	150	0.79	0.84	1.16
3	173	1.13	0.79	1.38
4	198	0.64	0.47	0.80
5	185	0.93	0.56	1.09
6	124	1.01	0.51	1.13

sides (relative to the AOI). Although only six images are present in the block, the procedure can be tested in various scenarios with two or more images. Since the images were acquired over three distinctive days, the generated true orthoimage is expected to exhibit artifacts as a result of the moving objects and changing light conditions.

The efficiency and behavior of the described workflow for the automatic generation of true orthoimages from satellite data is verified by several tests. First, the adjustment results that determine the geolocation accuracy are analyzed. Then, the occlusion exactness is checked on individual orthoimages. The performance of the procedure was tested with different image combinations, whereas the resulting true orthoimages were analyzed based on their visual accuracy and usefulness in further image processing. Aside from the performance assessment tests we tried to determine the value of the processing parameters (e.g., maximum distance) that are embedded in the system and influence the quality and appearance of the final product.

B. Results

1) *Adjustment Results*: The geometric model results for the six individual images are presented in Table II. The number of GCPs represents the final number of points that were used in the adjustment. They were obtained once the accuracy check and the gross error removal techniques have been performed.

Following the adjustment, the absolute precision of the solution was checked with 15 evenly distributed independent check

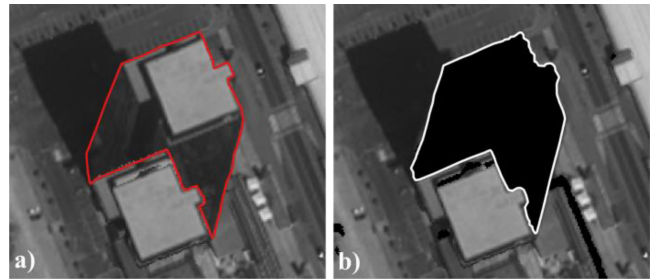


Fig. 12. Example of the completeness evaluation. (a) Orthoimage with a vectorized double mapped area. (b) Orthoimage with a vectorized occlusion. The polygons cover an almost identical area.

TABLE III
OCCLUSION COMPLETENESS VALUES FOR FOUR TEST BUILDINGS IN
TWO PLÉIADES IMAGES

Building	Completeness (%)	
	Image 2	Image 1
1	95.6	93.5
2	94.3	95.0
3	94.3	95.4
4	93.1	95.5
All	94.3	95.0

points (ICPs) that were manually measured on each image. The reference coordinates of the ICPs were manually measured on the national aerial orthophotos with a 0.5 m spatial resolution. The achieved root mean square errors (RMSE), computed from the differences between the measured coordinates and the coordinates obtained with the geometric model, usually reach subpixel values at each axis and around 1 pixel in the XY plane.

2) *Occlusion Completeness Assessment*: Completeness is a quality index that evaluates the share of the reference polygon detected by the occlusion detection method as an occlusion area [39]. Completeness provides us with the reliability of the used method to detect the occluded area. If the occlusions are not precisely identified, the true orthoimage will contain patches in which the objects are wrongly represented.

In our case the occlusion completeness was assessed in four areas. Each area is represented by a tall building in our AOI (see Fig. 12). Two images with the highest off-nadir viewing angle were used in the experiment. The reference polygon was delineated manually from the double-mapped representations that were found in the orthoimages generated with the indirect method. The double-mapped area is easy to extract and should represent the true size of the occluded areas.

The completeness evaluation tried to assess the performance level of the used occlusion detection method. The obtained values (see Table III) reveal that the percentages are all above 93%, which indicates that the used method performed very well and is comparable to other methods [39].

3) *Maximum Distance Parameter*: The following test tried to assess the impact of the maximum distance parameter on the mosaicing process. This parameter that defines the maximum distance (in pixels) to the occlusions is applied to the distance raster. When limiting the distance value, we increase the influence of the viewing angle parameter in the mosaicing process.

TABLE IV
COVERAGE PERCENTAGE OF A SET OF TWO IMAGES FOR DIFFERENT VALUES OF THE MAXIMUM DISTANCE PARAMETER

Maximum dist. (px)	Coverage (%)	
	Image 3	Image 2
20	97.3	2.7
15	97.6	2.4
10	98.0	2.0
5	98.6	1.4

TABLE V
COVERAGE OF IMAGES IN THE TRUE ORTHOIMAGE FOR DIFFERENT IMAGE SETS AND COST FUNCTIONS

No. of images	Cost function	Coverage (%)				Occlud. pixels (%)
		Image 3 (6.8°)	Image 6 (16.2°)	Image 1 (19.5°)	Image 2 (23.6°)	
2	None	95.7	n/a	n/a	4.3	0.32
	Linear	96.9	n/a	n/a	3.1	
	Power	97.6	n/a	n/a	2.4	
3	None	n/a	69.7	16.1	14.2	0.87
	Linear	n/a	72.1	15.6	12.3	
	Power	n/a	73.6	15.0	11.4	
4	None	94.4	1.2	0.5	3.9	0.25
	Linear	96.1	0.9	0.3	2.8	
	Power	96.9	0.7	0.2	2.2	

For example, if the maximum distance is high, the effect of the master image would be much lower. Although the maximum distance should be kept low, it still has to be reasonably distant from the occlusions and large enough to enable feathering.

Table IV shows the effects of the maximum distance parameter when generating a true orthoimage from two images acquired from different angles (6.8° and 23.6°). Four values were tested, ranging between 5 and 20 pixels. The table shows that although we used different values, the changes in the coverage percentages are minimal. The coverage of the master image decreases with higher maximum distance values. The change is rather small and does not have a significant effect on the true orthoimage. Therefore, the value in our implementation was set to 15, mostly because of the feathering.

4) *Different Image Sets and Cost Functions*: This test evaluated the coverage percentage for each individual image that composes the true orthoimage. For this test various numbers of images were used in combination with three different cost functions. Three image sets were prepared, containing two, three, and four images, respectively. Only images with higher viewing angles were taken into account as these are better at illustrating the behavior of the results and since the sets are often composed from such images (mainly for stereoscopic processing). Each test set was processed with three different cost functions: linear, power, and no cost function (value is 0).

Table V presents the results of the test with different image sets and cost functions. The table presents the images in an ascending order as regards the viewing angle. The image with the smallest viewing angle covered the majority of the true orthoimage in all tests. Even when the images had a similar angle (see set with three images) the master image covered at least 70% of the resulting image. When an image in the set had a much lower

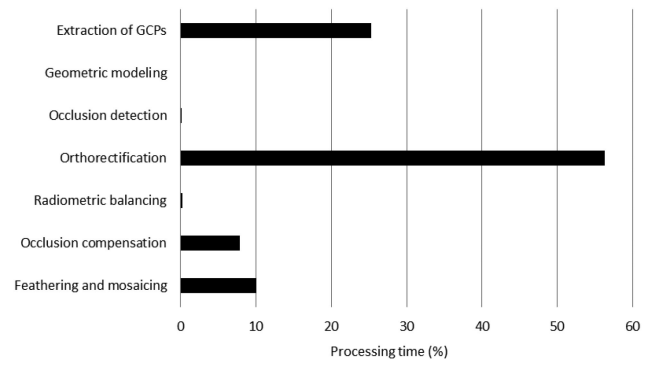


Fig. 13. Processing time for each module.

angle than the others, it covered about 95% of the image (see sets with two and four images).

The results also show the percentage of occluded (empty) pixels still present in the final true orthoimage. When images with angles above 16° were used, the percentage reached almost 1%. In sets in which an image with a low angle was utilized, the percentage was around 0.3%.

The utilized cost function has a minor impact on the image coverage. As expected, if the cost function equals 0, the master image has smaller coverage than when the linear or power functions are used. In most cases the use of cost functions impacts the coverage of images with higher viewing angles.

The processing time for each module depends on the number of images involved in generating the true orthoimage. On a personal workstation (3.2 GHz CPU, 6 cores, 64 GB RAM) the whole procedure took almost 380 min for three images. Most of the time is spent on orthorectification, which takes up 56% of the total processing time (see Fig. 13). The processing time required for the extraction of GCPs depends on the number of GCPs, which can be reduced programmatically. Each additional image extends the duration of the procedure by approximately 100 min, which is essentially the time required for extracting the GCPs and orthorectifying the image. The processing times for occlusion detection are about 13 s per image, which is comparable to the height based ray-tracing method used for similar satellite images [16]. In the referenced study the height based ray-tracing method is faster than the modified Z-buffer method and almost as fast as the sorted DSM method.

It is important to note that only the orthorectification module is optimized for execution on multiple cores. With additional optimization, the entire processing would be faster.

5) *True Orthoimage Quality*: The final quality assessment was achieved by primarily analyzing the orthoimages and true orthoimages, and then comparing the final results to the national reference data.

The orthoimages (see Fig. 14) were checked for the number of occluded pixels that influence the mosaicing process and quality of the results. The share of occluded pixels in selected images is shown in Table VI. It is close to 0% in the image with a 2.2° viewing angle, whereas in the image with a 23.6° angle (which is common in stereoscopic processing) it is over 9%.

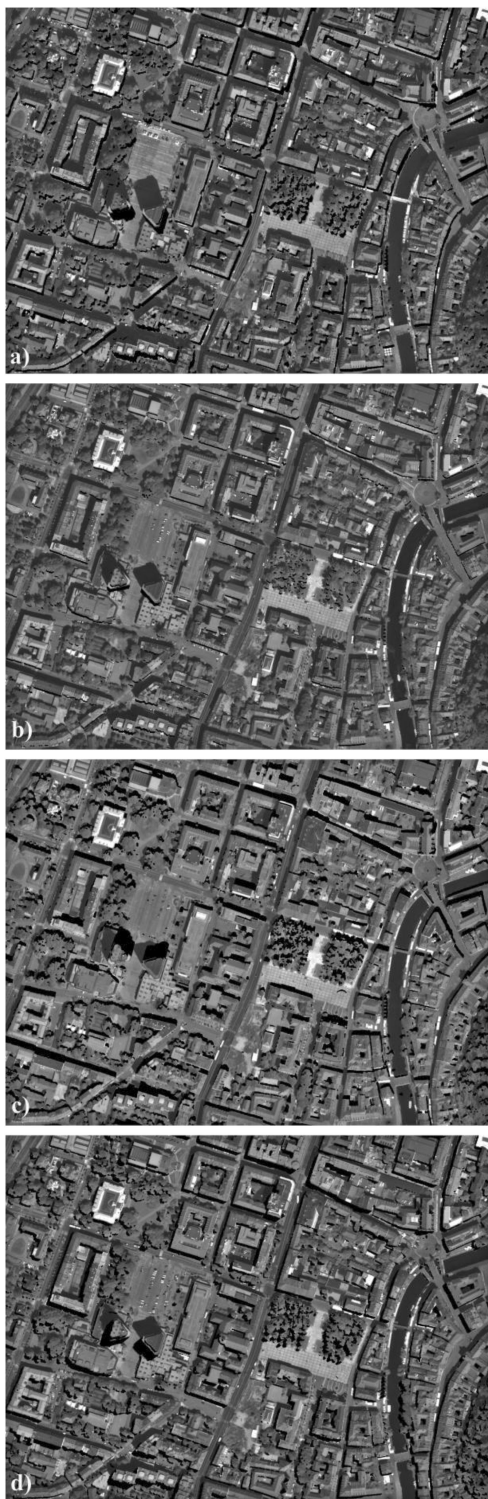


Fig. 14. Example of four orthoimages with different viewing angles. (a) Im. 6 (16.2°). (b) Im. 3 (6.8°). (c) Im. 2 (23.6°). (d) Im. 1 (19.5°).

Fig. 15 presents the results in Table VI graphically. It can be noted that the percentage of occluded pixels grows almost linearly with higher viewing angles.

An example of a section of a true orthoimage generated from three orthoimages is shown in Fig. 16. The quality of the mosaicing process was first assessed in terms of the radiometric

TABLE VI
NUMBER AND SHARE OF OCCLUDED PIXELS FOR FOUR ORTHOIMAGES WITH DIFFERENT VIEWING ANGLES (ASCENDING ORDER)

Image (view. angle)	No. occlud. pixels	Occlud. pixels (%)
4 (2.2°)	20872	0.01
3 (6.8°)	1540302	0.75
6 (16.2°)	10146475	4.95
2 (23.6°)	18721209	9.13

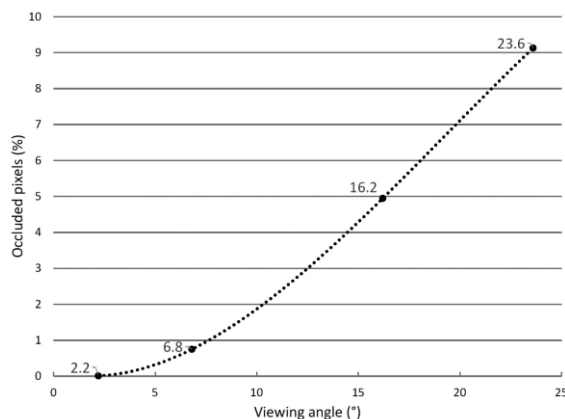


Fig. 15. Percentage of occluded pixels in respect to the viewing angle.

fidelity of the true orthoimage and its suitability to be used as input data for various geospatial analysis. For this test we changed the roles of the images in the mosaicing process in order to emphasize the correctness of the processing chain results: the one with the highest viewing angle was set as the master image. The true orthoimage produced in this forced setting contained larger patches of the three images that were used in the process. The radiometric balancing was tested by segmenting the final true orthoimage into areas with homogeneous spectral values (a procedure usually involved in image classification) which were then compared to the seamlines. Fig. 17 shows the results of the test. A subset of the true orthoimage containing a high building is presented in Fig. 17(a). Fig. 17(b) shows the share of the master image and both slave images in the final true orthoimage. The vectorized Fig. 17(b) is presented in Fig. 17(d), which shows the seamlines. Fig. 17(c) presents the vector of the homogeneous segments obtained from the true orthoimage. When the segmentation and the seamline vectors were compared it was clearly visible that there was no correlation between them. This led to the conclusion that the radiometric balancing and mosaicing process were successful and that the resulting true orthoimage can be used in the extraction of spatial information through classification.

The geometric quality of the true orthoimages was assessed by comparing them with external reference data. For this purpose, we decided to use data from the national building cadaster. In order to visually assess the geometric accuracy, the true orthoimages were overlaid with cadastral data. A few examples of the evaluation are presented in Fig. 18. Examples (a) to (f) show different buildings in the city of Ljubljana—from residential blocks and business buildings to old town apartment buildings



Fig. 16. Section of a true orthoimage.

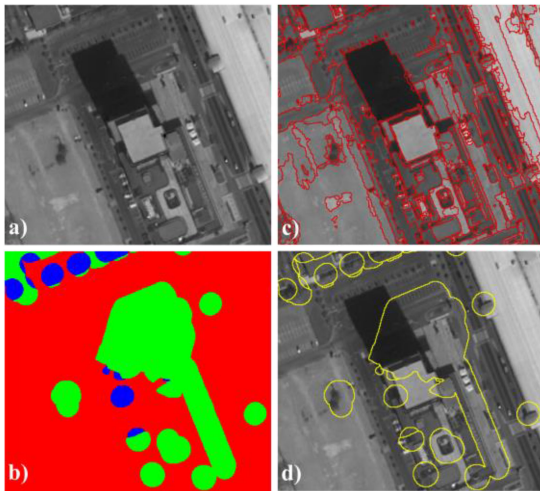


Fig. 17. Radiometric balancing and mosaicing assessment of a true orthoimage with changed image roles setting; a subset is shown. (a) True orthoimage created from three overlapping images. (b) Color coded share of images in the true orthoimage (red: master image; green: slave image 1; blue: slave image 2). (c) Vector of the homogeneous segments. (d) Vectorized color coded image.

and industrial towers. The results show a good match between the datasets. The minor discrepancies that were discovered are

TABLE VII
COMPLETENESS VALUES FOR SIX TEST BUILDINGS OF A TRUE ORTHOIMAGE
GENERATED FROM THREE INPUT IMAGES

Building	Area (m ²)	Accuracy (%)
a)	839.9	92.9
b)	3104.8	93.4
c)	8702.1	98.0
d)	1068.6	96.0
e)	1100.2	95.9
f)	427.8	91.3
All	15243.4	96.2

mainly a result of the nonupdated reference data, adjustment accuracy (see Table II) and DSM artifacts.

The accuracy of the buildings in Fig. 18 was also assessed quantitatively with the testing of their completeness. For each building, a reference polygon was delineated manually from the true orthoimage. The polygons were then compared to the reference polygons from the building cadaster. The completeness evaluation assessed the performance level of the occlusion compensation and mosaicing method.

The accuracy of the results can be seen in Table VII. The results which show an average percentage for all buildings above 96% indicate that the used methods performed well. Overall, the quality of the generated true orthoimages appears suitable for various geospatial analysis.

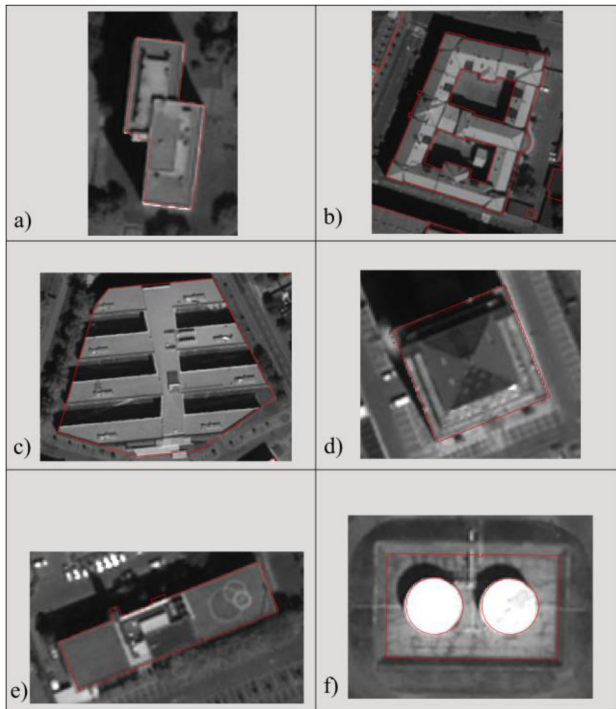


Fig. 18. Examples of geometric assessment. (a) Apartment block. (b) Old town apartment building. (c) Business building. (d) High business building. (e) Elongated business building. (f) Industrial towers.

The accuracy of occlusion detection showed promising results compared to other methods. It achieves better results than methods used with comparable satellite imagery, where the values ranged from 0% for differential rectification to 73.7% and 86.7% for the sorted DSM method and the modified Z-buffer method, respectively, and finally around 94% for the height based ray-tracing method [16]. The completeness is slightly higher even compared to methods used with UAV images, where the best performing surface-gradient-based method achieved an accuracy of about 94.2% [39].

Most of the artifacts that can be found in true orthoimages are caused by the imperfections of the lidar DSM used for their generation. Fig. 19 shows a typical example of the encountered artifacts. The overlaid building cadastral data shows the imperfections of the DSM that are in most cases present at the edges of the building see Fig. 19(a). These artifacts are reflected in the true orthoimages as shown in Fig. 19(b) and (c). Subsets are taken from different true orthoimages that were generated from different sets of three images. Although different images were used for the true orthoimage generation, the artifacts are located on the same spots and reflect the same DSM imperfections.

C. Discussion

The automatic generation of true orthoimages from satellite data encompasses various steps and the quality of the final result often depends on the used ancillary data. Geometric processing that uses vector roads as the main reference data is an important

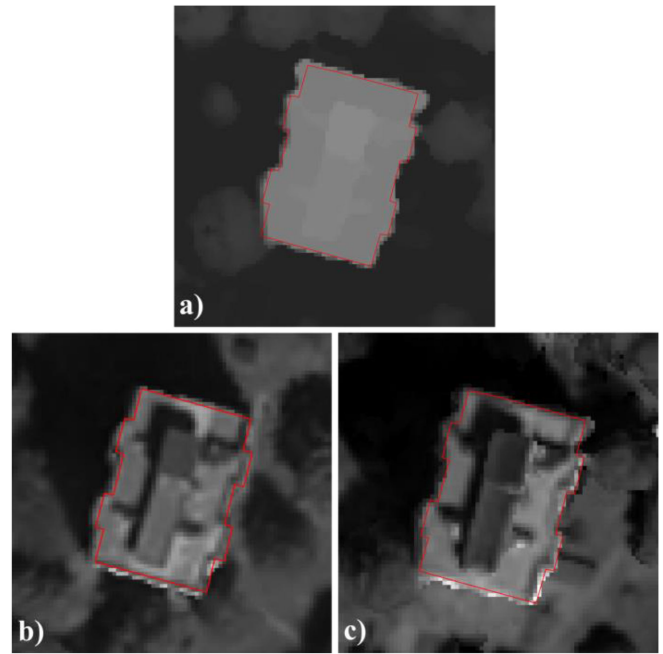


Fig. 19. Artifacts in true orthoimages caused by the DSM. Two different true orthoimages—(b) and (c)—generated with the same DSM—(a)—with overlaid cadastral data. Most imperfections that appear at the edge of the apartment block are a result of the imprecise DSM.

step that relies heavily on ancillary data but was not thoroughly described in this article. The accuracy of the geometric results has a high impact on the mosaicing process—poor geometric correction results can lead to a misalignment of objects (e.g., roads, buildings) or wrong shapes of objects in the final true orthoimage. The results from our experiments show that a planar accuracy around one pixel can be achieved for images with a 0.5 m resolution. A visual inspection confirms that this accuracy is appropriate for the generation of true orthoimages. It needs to be pointed out that the geometric model uses block adjustment with GCPs that are shared between multiple images. Such GCPs can be automatically extracted because the algorithm searches for matches only on predefined crossroads with known spatial coordinates.

The results of the geometric model are further used in the orthorectification with the indirect model. For this procedure, the occlusion mask has to be generated beforehand, and this is based on the satellite incidence angle. The mask represents the occluded areas that appear in the image as nodata pixels. The assessment of the occlusion completeness gave excellent results for both tested images (both having high off-nadir angles). The high accuracy proves the fidelity of the incidence angles reported in the satellite image metadata file, which left us merely with the task of choosing the method for generating the mask. The used ray-tracing method is relatively slow, however, its upside lies in its high precision. The final results need some postprocessing, especially when the used DSM has many artifacts (which are frequent when the DSM is generated from lidar data or through image stereo matching).

The true orthoimage generation procedure involves the choosing of the final values for several processing parameters. The maximum distance parameter was tested to evaluate its impact on the mosaicing process. In the test only lower values were tested as a high value would excessively minimize the effect of the viewing angle parameter during mosaicing. Selecting a small maximum distance retains a larger portion of the master image, which usually contains smaller errors (caused by the DSM, GCPs, etc.) than the other images in the block.

Unlike aerial images, satellite images have a steady viewing angle with negligible variation throughout the image. Thus, one does not need to use the distance to the nadir raster in the computation. Although such a raster is not feasible, we still needed to take the viewing angle into account. This was introduced in the cost function that adds an additional weight to the images during the mosaicing process. Images with larger weights cover larger parts of the mosaic. Even though the tested cost functions did not have a great influence on the share of the image within the mosaic, it is important when large gaps are present in the angle values of the images in the block.

The tests on different sets of images showed that the master image prevails and always covers a major part of the true orthoimage. In sets in which the master image was acquired with a very small viewing angle (compared to the other images in the set), the master image took over most of the mosaic. A nadir image would probably be sufficient to generate the true orthoimage without occluded pixels, at least in areas without extremely high objects.

Even when the block contains images relatively close to nadir, we can still expect that the true orthoimage will contain occluded pixels. Their numbers depend on the distribution of the images in respect to the AOI and the height and distribution of objects within the image. The remaining occlusions are usually located close to or between high objects. If a stereopair is used to generate a true orthoimage of a city it can be expected that over 1% of all pixels will remain occluded after mosaicing. If we want to generate mosaics without occlusions, we need to use additional images from different orbits. In this case the time gap between the acquisitions can cause inconsistencies within the mosaic.

As demonstrated in the tests, the produced true orthoimages can be used for extracting spatial information. The results were tested with a segmentation procedure that was not biased by the seamlines and clearly delineated the objects in the image. The national building cadaster data confirmed these findings. Due to the time difference between the images and the reference data it was impossible to perform extensive testing.

An important aspect not highlighted in this article are the atmospheric conditions that could influence the radiometric balancing between satellite images. Images acquired under similar conditions (as was the case in our example) can be balanced by histogram matching, whereas appropriate atmospheric correction is advisable when the images are severely contaminated by aerosols, clouds, and cloud shadows [44]. In this case, one should apply a procedure that retrieves surface reflectance before mosaicing.

Another issue that was not properly addressed in the research were object (building) shadows. This shadows can hide features within them and affect the appearance of the orthoimages. Within our study we developed an automatic procedure for detecting shadows. It is similar to the occlusion detection algorithm, but instead of using the satellite incidence angle it uses the sun angle at the time of acquisition. This procedure enabled us to detect the shadows precisely, however, inconsistent results were obtained during the radiometric balancing of the shadowed areas. The experiments were focused on adjusting the histogram of the shadowed areas to match the areas with no shadows. As the shadows usually contained little information, the balancing results failed to significantly improve the usefulness of the orthoimages. Thus, the shadow correction method was not implemented in the true orthoimage generation procedure.

A detailed inspection of the true orthoimages revealed minor errors that can be attributed to various factors. One of them lies in the different sources of ancillary data that is used in the process. The data come from various databases and has different accuracies. GCPs depend mainly on vector roads. On the other hand, the occlusion masks and true orthoimages depend on the DSM that was generated from lidar data. Combining this data can lead to small errors that are visible on the edges of the buildings after orthorectification. Although the data provenience is different, some orthoimages show no artifacts at all.

The geometric model is another important factor for the results. As a result of the block adjustment, the alignment between the images is appropriate and without notable artifacts. However, the absolute accuracy of the automatic process may exceed 1 pixel which can cause a small misalignment with the DSM.

Due to the changes that appear at different time frames, inconsistencies appear when generating a true orthoimage from images acquired on different dates. The most obvious are the moving objects (e.g., cars), but large changes can also occur between images acquired several days apart. These mosaics can exhibit discrepancies especially in object shadows or newly adapted or constructed buildings.

Probably the most important factor that causes artifacts in the mosaic is the DSM. Alongside the aforementioned alignment of data, the DSM should perfectly resemble the shape of the objects if it is to produce a flawless mosaic. Such DSM is hard to generate if we do not have extremely good data. Usually only precise digital building models (DBMs) or lidar DSMs generated from very dense point clouds can satisfy all requirements. In our case the DSM was created with relatively sparse lidar data which led to imperfections in the shapes of the objects. The DSM was visualized with different methods [45], [46] which identified its irregularities and made it possible for us to find the possible solutions. Although the model was generated with an advanced algorithm, additional manual editing would still be needed at least for the model update.

The use of a DSM generated from stereo satellite imagery can represent a viable option to manually updating the DSM generated with lidar data. However, such DSMs often suffer from mismatches, missing values, or blunders, resulting in coarse building shape representation. During the last few years, some

studies tried to correct the shortcomings of the DSM generated from stereopairs and remarkable improvements have been made [47]. We can expect that in the near future advanced methods (e.g., machine learning) will be able to provide precise DSMs directly from imagery that will match the precise lidar DSMs in terms of accuracy and reliability.

IV. CONCLUSION

In this study, we proposed a completely automatic procedure for the generation of true orthoimages from optical very high-resolution satellite images. As the workflow utilizes an RFM, we introduced a special procedure that varies from the workflows that employ aerial or UAV images. In order to make the procedure as simple and efficient as possible it was divided into consecutive steps. In the first step, individual images were geometrically corrected and orthorectified with the use of an occlusion mask that was generated with the ray-tracing method. The slave orthoimages were then radiometrically balanced so that they matched the master image that was selected on the basis of the viewing angle. The mosaicing of the images was performed on the basis of the viewing angle and the distance to the occlusions.

This original workflow is capable of generating true orthoimages with an automatic procedure that uses a very effective method of occlusion detection. Therefore, the automatic workflow using VHR satellite images as input data and precise occlusion detection are the main contributions in this article. Other contributions are the robustness and accuracy of the workflow compared to state-of-the-art methods and the fast processing times for occlusion detection.

The performance of the procedure was tested with six Pléiades images. The results indicate the capability of this fully automatic process to compensate for occlusions and to successfully generate accurate true orthoimages from two or more satellite images. The limiting factors of the workflow are usually connected to inappropriate ancillary data (e.g., DSM) that can cause that various artifacts appear, usually on the edges of high objects. If merely images with high viewing angles are used, the generated true orthoimage can still exhibit occluded areas, most of which occur between adjacent high objects.

In order to streamline the process, the workflow during the occlusion compensation utilizes a single cost function based on the viewing angle. There are many other possibilities of selecting the best orthoimage to fill the chosen pixel in the final true orthoimage. Larger weights could be placed on images that were acquired closer to the acquisition date (ideally on the same day) of the master image or on images that were acquired in similar light conditions.

In order to achieve more harmonized mosaics, a change in the selection of the master image could be introduced. Instead of selecting the image with the lowest viewing angle from all of the images, the image with the lowest angle amongst the images taken from the same orbit (if such images exist) could be selected.

There are several sections in which the developed workflow could be upgraded. However, the current procedure is robust and

reliable and can serve as the backbone for further development. Additionally, the presented procedure is fully functional and, as has been proven by the tests, it can already generate true orthoimages that can be used for the extraction of various spatial information.

ACKNOWLEDGMENT

The author would like to thank Ž. Kokalj for the lidar DSM and P. Pehani for his help and support in reviewing the article. The constructive comments and suggestions from anonymous reviewers have considerably improved this article.

REFERENCES

- [1] G. Zhou, W. Chen, J. A. Kelmelis, and D. Zhang, "A comprehensive study on urban true orthorectification," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 9, pp. 2138–2147, Sep. 2005.
- [2] M. K. Fras, "Vpliv kakovosti vhodnih podatkov na kakovost ortofota : Influence of input data quality on the quality of orthophoto," *Geodetski Vestnik*, vol. 48, no. 2, pp. 167–178, Jan. 2004.
- [3] F. Amhar and R. Ecker, "An integrated solution for the problems of 3D man-made objects in digital orthophotos," *Int. Arch. Photogrammetry. Remote Sens.*, vol. 31, pp. 84–89, 1996.
- [4] F. Amhar, J. Jansa, and C. Ries, "The generation of true orthophotos using a 3D building model in conjunction with a conventional DTM," *Int. Arch. Photogrammetry Remote Sens.*, vol. 32, pp. 16–22, 1998.
- [5] J.-Y. Rau, N.-Y. Chen, and L.-C. Chen, "True orthophoto generation of built-up areas using multi-view images," *Photogrammetry Eng. Remote Sens.*, vol. 68, no. 6, pp. 581–588, Jun. 2002.
- [6] A. Habib, E.-M. Kim, and C.-J. Kim, "New methodologies for true orthophoto generation," *Photogrammetry Eng. Remote Sens.*, vol. 73, no. 1, pp. 25–36, Jan. 2007.
- [7] C. Zhong, H. Li, and X. Huang, "A fast and effective approach to generate true orthophoto in built-up area," *Sens. Rev.*, vol. 31, no. 4, pp. 341–348, Sep. 2011.
- [8] L. Barazzetti, M. Brovelli, and M. Scaioni, "Generation of true-orthophoto with lidar high resolution digital surface models," *Photogrammetry J. Finl.*, vol. 21, no. 1, pp. 26–36, Jan. 2008.
- [9] H. C. de Oliveira and M. Galo, "Occlusion detection by height gradient for true orthophoto generation, using LiDAR data," in *Proc. Int. Arch. Photogrammetry, Remote Sens. Spatial Inf. Sci.*, May 2013, doi: [10.5194/isprsarchives-XL-1-W1-275-2013](https://doi.org/10.5194/isprsarchives-XL-1-W1-275-2013).
- [10] H. Gharibi and A. Habib, "True orthophoto generation from aerial frame images and LiDAR data: An update," *Remote Sens.*, vol. 10, no. 4, Apr. 2018, Art. no. 581.
- [11] L. Barazzetti, R. Brumana, D. Oreni, M. Previtali, and F. Roncoroni, "True-orthophoto generation from UAV images: Implementation of a combined photogrammetric and computer vision approach," in *Proc. ISPRS Ann. Photogrammetry, Remote Sens. Spatial Inf. Sci.*, May 2014, vol. II-5, pp. 57–63, doi: [10.5194/isprsannals-II-5-57-2014](https://doi.org/10.5194/isprsannals-II-5-57-2014).
- [12] Y. Hu, D. Stanley, and Y. Xin, "True ortho generation of urban area using high resolution aerial photos," in *Proc. ISPRS Ann. Photogrammetry, Remote Sens. Spatial Inf. Sci.*, Jul. 2016, vol. III-4, pp. 3–10, doi: [10.5194/isprs-annals-III-4-3-2016](https://doi.org/10.5194/isprs-annals-III-4-3-2016).
- [13] Y. Liu, X. Zheng, G. Ai, Y. Zhang, and Y. Zuo, "Generating a high-precision true digital orthophoto map based on UAV images," *ISPRS Int. J. Geo-Inf.*, vol. 7, no. 9, Sep. 2018, Art. no. 333.
- [14] L.-C. Chen, T.-A. Teo, J.-Y. Wen, and J.-Y. Rau, "Occlusion-compensated true orthorectification for high-resolution satellite images," *Photogrammetric Rec.*, vol. 22, no. 117, pp. 39–52, 2007.
- [15] K. I. Bang, A. Habib, S. Shin, and K. Kim, "Comparative analysis of alternative methodologies for true Ortho-photo generation from high resolution satellite imagery," in *Proc. ASPRS Annu. Conf.*, May 2007.
- [16] K. I. Bang and C. Kim, "A new true Ortho-photo generation algorithm for high resolution satellite imagery," *Korean J. Remote Sens.*, vol. 26, no. 3, pp. 347–359, Jun. 2010.
- [17] Y. Zhang, L. Yu, M. Sun, and X. Zhu, "A mixed radiometric normalization method for mosaicking of high-resolution satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 2972–2984, May 2017.
- [18] K. Liu, T. Ke, P. Tao, J. He, K. Xi, and K. Yang, "Robust radiometric normalization of multitemporal satellite images via block adjustment

- without master images,” *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 13, pp. 6029–6043, 2020, doi: [10.1109/JSTARS.2020.3028062](https://doi.org/10.1109/JSTARS.2020.3028062).
- [19] X. Li *et al.*, “Generating high-quality and high-resolution seamless satellite imagery for large-scale urban regions,” *Remote Sens.*, vol. 12, no. 1, Jan. 2020, Art. no. 81.
- [20] D. Ibañez, R. Fernandez-Beltran, J. M. Sotoca, R. A. Mollineda, J. Moreno, and F. Pla, “Multitemporal mosaicing for Sentinel-3/FLEX derived level-2 product composites,” *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 13, pp. 5439–5454, 2020, doi: [10.1109/JSTARS.2020.3023593](https://doi.org/10.1109/JSTARS.2020.3023593).
- [21] X. Li, R. Feng, X. Guan, H. Shen, and L. Zhang, “Remote sensing image mosaicking: Achievements and challenges,” *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 4, pp. 8–22, Dec. 2019.
- [22] M. Gianinetto and M. Scaioni, “Automated geometric correction of high-resolution pushbroom satellite data,” *Photogrammetry Eng. Remote Sens.*, vol. 74, no. 1, pp. 107–116, 2008.
- [23] S. Leprince, S. Barbot, F. Ayoub, and J. P. Avouac, “Automatic and precise orthorectification, coregistration, and subpixel correlation of satellite images, application to ground deformation measurements,” *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 6, pp. 1529–1558, Jun. 2007.
- [24] C.-C. Liu, and P.-L. Chen, “Automatic extraction of ground control regions and orthorectification of remote sensing imagery,” *Opt. Exp.*, vol. 17, no. 10, pp. 7970–7984, May 2009.
- [25] J. Oh, and C. Lee, “Automated bias-compensation of rational polynomial coefficients of high resolution satellite imagery based on topographic maps,” *ISPRS J. Photogrammetry Remote Sens.*, vol. 100, pp. 14–22, Feb. 2015.
- [26] H. Kartal, U. Alganci, and E. Sertel, “Automated orthorectification of VHR satellite images by SIFT-Based RPC refinement,” *ISPRS Int. J. Geo-Inf.*, vol. 7, no. 6, Jun. 2018, Art. no. 229.
- [27] A. Marsetič, K. Oštir, and M. K. Fras, “Automatic orthorectification of high-resolution optical satellite images using vector roads,” *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 11, pp. 6035–6047, Nov. 2015.
- [28] P. Pehani, K. Čotar, A. Marsetič, J. Zaletelj, and K. Oštir, “Automatic geometric processing for very high resolution optical satellite data based on vector roads and orthophotos,” *Remote Sens.*, vol. 8, no. 4, Apr. 2016, Art. no. 343.
- [29] A. Marsetič and P. Pehani, “Towards sub-pixel automatic geometric corrections of very-high resolution panchromatic satellite data of urban areas,” *Remote Sens.*, vol. 11, no. 9, Jan. 2019 Art. no. 1097.
- [30] W. Schickler and A. Thorpe, “Operational procedure for automatic true orthophoto generation,” in *Proc. Int. Arch. Photogrammetry Remote Sens.*, 1998, vol. 32, pp. 527–532.
- [31] A. Biasion, S. Dequal, and A. Lingua, “A new procedure for the automatic production of true orthophotos,” in *Proc. Int. Arch. Photogrammetry Remote Sens. Spatial Inf. Sci.*, 2004, vol. 35, pp. 538–543.
- [32] A. Gunay, H. Arefi, and M. Hahn, “Semi-automatic true orthophoto production by using LIDAR data,” in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2007, pp. 2873–2876.
- [33] E. Keinan and Y. Doytsher, “Automatic methods toward generating digital true orthophoto by using dense LiDAR data,” in *Proc. Annu. Conf. ASPRS*, Apr. 2008.
- [34] F. Deng, J. Kang, P. Li, and F. Wan, “Automatic true orthophoto generation based on three-dimensional building model using multiview urban aerial images,” *J. Appl. Remote Sens.*, vol. 9, no. 1, pp. 1–15, Mar. 2015.
- [35] C. S. Fraser and H. B. Hanley, “Bias-compensated RPCs for sensor orientation of high-resolution satellite imagery,” *Photogrammetry Eng. Remote Sens.*, vol. 71, no. 8, pp. 909–915, Aug. 2005.
- [36] M. O’Neill and I. Dowman, “The generation of epipolar synthetic stereo mates for SPOT using a DEM,” in *Proc. 16th ISPRS Congr. Commission*, Jul. 1988, vol. 27, pp. 587–598.
- [37] A. Khosravipour, A. K. Skidmore, and M. Isenburg, “Generating spike-free digital surface models using LiDAR raw point clouds: A new approach for forestry applications,” *Int. J. Appl. Earth Obs. Geoinf.*, vol. 52, pp. 104–114, Oct. 2016.
- [38] T. Kim, D. Shin, and Y.-R. Lee, “Development of a robust algorithm for transformation of a 3D object point onto a 2D image point for linear pushbroom imagery,” *Photogrammetry Eng. Remote Sens.*, vol. 67, no. 4, pp. 449–452, Apr. 2001.
- [39] H. C. de Oliveira, A. P. D. Poz, M. Galo, and A. F. Habib, “Surface gradient approach for occlusion detection based on triangulated irregular network for true orthophoto generation,” *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 11, no. 2, pp. 443–457, Feb. 2018.
- [40] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 4 ed. New York, NY, USA: Pearson, 2017.
- [41] M. O. Nielsen, “True orthophoto generation,” M.S. thesis, Technical Univ. Denmark, Lyngby, Denmark, 2004.
- [42] C. R. Maurer, R. Qi, and V. Raghavan, “A linear time algorithm for computing exact euclidean distance transforms of binary images in arbitrary dimensions,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 2, pp. 265–270, Feb. 2003.
- [43] P. J. Burt and E. H. Adelson, “A multiresolution spline with application to image mosaics,” *ACM Trans. Graph.*, vol. 2, no. 4, pp. 217–236, 1983.
- [44] S. Liang, H. Fang, and M. Chen, “Atmospheric correction of landsat ETM+ land surface imagery. I. Methods,” *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 11, pp. 2490–2498, Nov. 2001.
- [45] Ž. Kokalj and M. Somrak, “Why not a single image? Combining visualizations to facilitate fieldwork and on-screen mapping,” *Remote Sens.*, vol. 11, no. 7, Jan. 2019, Art. no. 747.
- [46] T. Verbovšek, T. Popit, and Ž. Kokalj, “VAT method for visualization of mass movement features: An alternative to hillshaded DEM,” *Remote Sens.*, vol. 11, no. 24, Jan. 2019, Art. no. 2946.
- [47] K. Bittner, P. D’Angelo, M. Körner, and P. Reinartz, “DSM-to-LoD2: Spaceborne stereo digital surface model refinement,” *Remote Sens.*, vol. 10, no. 12, Dec. 2018, Art. no. 1926.



Aleš Marsetič received the B.S. degree in geodetic engineering and the Ph.D. degree in geodetic sciences from the University of Ljubljana, Ljubljana, Slovenia, in 2005 and 2015, respectively.

Since 2006 he has been a Research Fellow with the Research Centre of the Slovenian Academy of Sciences and Arts, Ljubljana. Between 2010 and 2014, he was with the Remote Sensing Department, Centre of Excellence for Space Sciences and Technologies – Space-SI. He has authored one book and more than 40 scientific papers. His main research interests include

photogrammetric applications in satellite remote sensing. Besides photogrammetry (orthorectification, DSM generation), his work also encompasses optical satellite imagery registration and processing, spatial analysis, and lidar. He has vast knowledge in automating satellite data preprocessing, designing small satellites, and developing their ground segment data processing algorithms.

Dr. Marsetič was a recipient of the 3rd Mission Idea Contest for Micro and Nano-Satellites Utilization First Place Award in 2014, and the Research Centre of the Slovenian Academy of Sciences, and Arts Silver Award in 2016.