

Generation of High-Resolution Orthomosaics from Historical Aerial Photographs Using Structure-from-Motion and Lidar Data

Ji Won Suh and William Ouimet

Abstract

This study presents a method to generate historical orthomosaics using Structure-from-Motion (SfM) photogrammetry, historical aerial photographs, and lidar data, and then analyzes the horizontal accuracy and factors that can affect the quality of historical orthomosaics products made with these approaches. Two sets of historical aerial photographs (1934 and 1951) were analyzed, focused on the town of Woodstock in Connecticut, U.S.A. Ground control points (GCPs) for georeferencing were obtained by overlaying multiple data sets, including lidar elevation data and derivative hillshades, and recent orthomosaics. Root-Mean-Square Error values of check points (CPs) for 1934 and 1951 orthomosaics without extreme outliers are 0.83 m and 1.37 m, respectively. Results indicate that orthomosaics can be used for standard mapping and geographic information systems (GIS) work according to the ASPRS 1990 accuracy standard. In addition, results emphasize that three main factors can affect the horizontal accuracy of orthomosaics: (1) types of CPs, (2) the number of tied photos, and (3) terrain.

Introduction

Land use land cover (LULC) change plays a fundamental role in recording the impact of human activities on earth surface processes and understanding these impacts is one of the grand challenges in environmental science today (National Research Council 2001). Satellite-based data such as Landsat have been widely used to understand LULC in the field of remote sensing (Leh *et al.* 2013; Verbesselt *et al.* 2012; Zhu *et al.* 2016, 2020; Zhu and Woodcock 2014), but are limited in terms of temporal scope (Landsat first launched in 1972) and spatial resolution (with the best available pixel resolution being 15–30 m between the 1970s and early 2000s). To understand historical LULC changes prior to the satellite era and at much higher spatial resolution, historical aerial photography has long been considered an important source of data, adding multiple new time intervals in the study of land use activity in forestry, ecology, urban planning, cultural resources, and geomorphology related studies (Kadmon and Harari-Kremer 1999; Llana *et al.* 2018; Mallinis *et al.* 2011; Nita *et al.* 2018; Rocchini *et al.* 2006; Sevara 2013; Verhoeven *et al.* 2012; Zomeni *et al.* 2008).

Unlike modern remote sensing imagery derived from airplanes, unmanned aerial vehicles (UAV) or satellites, a number of issues need to be taken into account when using historical aerial photographs. These include: (1) low data accessibility, (2) no georeferencing, (3) time-consuming work for mosaicking fragmented aerial photograph campaigns and scaling up the spatial extent, (4) poor information for aligning or calibrating individual aerial photograph such as camera position, flight altitude, yaw, pitch, and roll (Fox and Czipferszky 2008; Frankl *et al.* 2015), (5) no ground control points (GCPs), which are crucial in removing

inherent radial distortion and tilt for georeferencing (Bolstad 1992; Wolf *et al.* 2014), and (6) low image quality due to digital scanning. These problems have generally restricted the application of modern photogrammetric techniques when dealing with historical data (Lingua *et al.* 2009; Turner *et al.* 2012; Weng *et al.* 2013). In addition, they can attenuate errors associated with automated image processing in photogrammetric software such as Agisoft Metashape and Pix4Dmapper when it comes to the shortage of metadata or increasing the number of photos stitched.

Despite these challenges, a number of studies have been conducted to orthorectify historical aerial photographs and reconstruct historical digital elevation models (hDEM) based on photogrammetric techniques such as Structure-from-Motion (SfM). The historical aerial photography considered in these studies ranges from the 1930s (Fox and Czipferszky 2008; Frankl *et al.* 2015; Geyman *et al.* 2022) to the 1990s (Arnaud *et al.* 2015) and most of the research has been concentrated on images from the 1950s onward (Comiti *et al.* 2011; Gennaretti *et al.* 2011; Gomez 2014; Kadmon and Harari-Kremer 1999; Llana *et al.* 2018; Marignani *et al.* 2008; Maurer and Rupper 2015; Nebiker *et al.* 2014; Rocchini *et al.* 2006). To overcome the lack of external information of historical data, creating GCPs from reference data such as DEMs is essential during SfM processing. Depending on the spatial resolution of reference data, the quality of output orthomosaics or DEMs varies. The quality of orthomosaics has been evaluated by root-mean-square error (RMSE) and the RMSE results with coarse reference data (e.g., ~10–40 m) ranges from 5 m to 15 m (Baker *et al.* 1995; Gennaretti *et al.* 2011; Marignani *et al.* 2008; Rocchini *et al.* 2006).

The advent of Light Detection and Ranging (lidar) data and high-resolution DEMs (e.g., 1 m) can lead to increasing the accuracy of historical orthomosaics and hDEMs given that lidar derivatives such as hillshades allow for visualizing of small-scale features and greater potential for choosing GCP with accurate coordinates and elevation values. To date, however, the application of high-resolution topographic data from lidar in SfM photogrammetry processing has been focused on assessing vertical accuracy of hDEM products (Child *et al.* 2021; Nebiker *et al.* 2014) rather than extracting GCPs for the orthomosaic. One of reasons for this is that high-resolution historical imagery needs high-resolution topographic data as a reference to build orthomosaics with less horizontal error. Furthermore, there is a lack of research focused on expanding the spatial extent of orthophotos made from high-resolution historical aerial photos (i.e., less than 1:20 000 scale, less than 1 m pixel resolution) and lidar data. This is because a large number of historical photos need to be aligned and orthomosaicked to cover larger spatial extent, in contrast to the aforementioned previous studies that use less than 40 photos. In other words, the paucity of such studies underlines the need for an improved methodological approach. Lidar data and its derivatives provide an opportunity to fill in the gap between the low resolution of reference topographic data (e.g., 30 m DEMs) and the resolution of input from historical imagery (e.g., 1 m) (Llana *et al.* 2018).

Ji Won Suh and William Ouimet are with the Department of Geography, University of Connecticut, Storrs, CT 06269 (ji.suh@uconn.edu).

William Ouimet is also with the Department of Earth Sciences, University of Connecticut, Storrs, CT 06269.

Contributed by Dorota Iwaszczuk, March 17, 2022 (sent for review May 6, 2022; reviewed by Michael Cramer).

Photogrammetric Engineering & Remote Sensing
Vol. 89, No. 1, January 2023, pp. 37–46.

0099-1112/22/37–46

© 2023 American Society for Photogrammetry
and Remote Sensing
doi: 10.14358/PERS.22-00063R2

The goal of this paper is to present the generation of high-resolution historical orthomosaic over a broad area using SfM photogrammetry combined with GCPs from lidar derivatives. Our objectives are to (1) provide a method to build high-resolution historical orthomosaics by using 1934 and 1951 black and white aerial photographs focused on the town of Woodstock, Connecticut, United States; (2) evaluate the horizontal accuracy of orthomosaics based on RMSE values for overall assessment and residual error to understand spatial distribution of errors, and (3) analyze factors that can affect the quality of these historical orthophotos by comparing orthomosaics from 1934 and 1951.

Data and Methods

Study Area

This study was conducted focusing on the town of Woodstock in north-eastern Connecticut. In order to cover the entirety of Woodstock, our study area is a rectangle and includes adjacent towns described in Figure 1. To be

specific, Figure 1B shows the DEM ranging from 64 m to 335 m and topography in the west part comprised mainly of hilly uplands with mixed coniferous-deciduous forests and agricultural lands. On the other hand, the east part consists of lowland used for agricultural lands as well as mixed forests and muddy brook which flows south into Roseland lake. Distributed over the entire study area are stone walls, stacked around agricultural lands that indicate an anthropogenic legacy of English-style agriculture during the 17th to 20th centuries (Cronon 1983; Thorson 2002) (Figure 1C and 1D). They are easily detected in a hillshade map derived from lidar in open area and deciduous forest (Johnson and Ouimet 2014, 2016).

Data

Collected data for this study can be divided into two sets in terms of usages: to produce historical orthomosaics and to create GCPs (Table 1). In particular, the first data set includes high-resolution black and white air photos of 1934 and 1951 provided from the Connecticut State Library (1934, 1951). For 1934, 141 air photos (spatial scale 1:12 000) were used for an orthomosaic covering about 264 Km² area. The

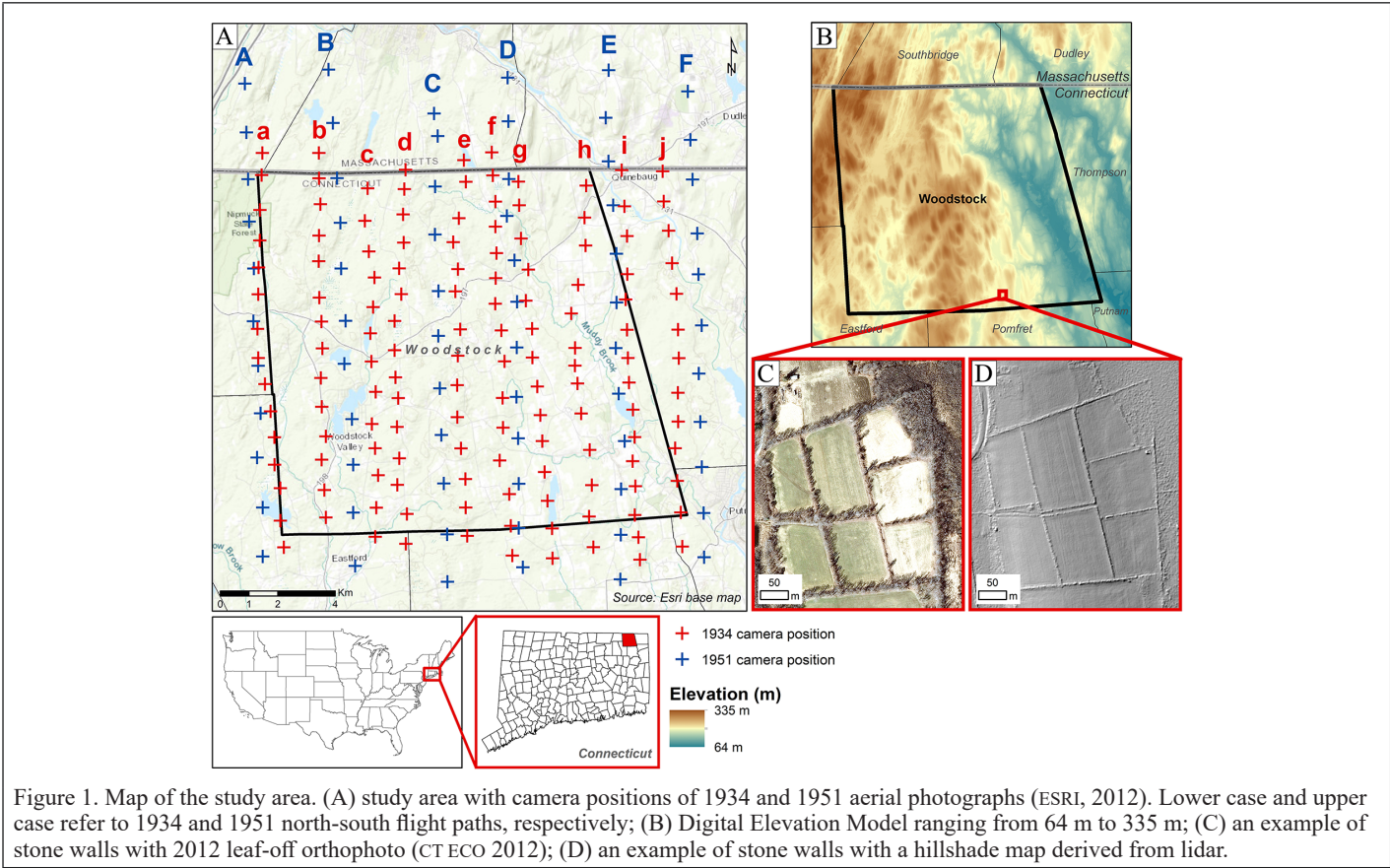


Table 1. List of maps and aerial photographs used for this study.

Data Usage	Data Type	Year	Date	Resolution/Map Scale	# of Images (Covering Area)	Source
Historical orthomosaics	A/BW	1934	April 1934	1:12 000	141 (264 Km²)	Connecticut State Library ¹
			5 September 1951			
	A/BW	1951	13 October 1951	1:20 000	68 (380 Km²)	Connecticut State Library ²
			25 November 1951			
			27 November 1951			
Reference data	A/O/C	2012		0.3 m		CT ECO
	A/O/C	2016		0.07 m		CT ECO
	DEM	2016		1 m		CT ECO
	Hillshade	2016		1 m		2016 DEM

A: aerial photographs; O: ortho-rectified aerial photographs; BW: black and white; C: color. CT ECO = Connecticut Environmental Conditions Online; DEM = digital elevation model.
¹ Connecticut State Library, 1934 Aerial Surveys, Record Group 089:011, Department of Transportation, State Archives.
² Connecticut State Library, 1951 Aerial Surveys, Record Group 089:011b, Department of Transportation, State Archives.

vertical photography was taken in April 1934 (note that specific dates were not available) by Fairchild Aerial Survey Corporation using a K-3 aerial survey camera with a 241.3 mm focal length. The survey flights were flown at an altitude of around 3500 m. The overlap rate between two images were approximately 50%. This historical photography was digital-scanned as a 1270 dots-per-inch in 2006. On the other hand, 68 air photos of 1951 (spatial scale 1:20 000) were used covering about 380 Km². Robinson Aerial Surveys took 1951 imagery on the following dates: 5 September 1951, 13 October 1951, 25 November 1951, and 27 November 1951. Unlike 1934 photography, detailed information about survey and scanning process for 1951 was not available. Regarding the number of photos of 1934 and 1951, each 1934 photo covers small extent compared to 1951 photo so that the more photos of 1934 were collected.

As reference data to create GCPs, the second data set includes recent orthophotos (2012 and 2016), DEM derived from 2016 lidar, and a hillshade map produced from lidar DEM. First, 2012 and 2016 high-resolution orthophotos were streamed from Connecticut Environmental Conditions Online (CT ECO) image services in ArcGIS Desktop 10.5 (CRCoG 2016; CTECO 2012). The spatial resolution of 2012 and 2016 orthoimagery is 0.3 m and 0.07 m, respectively (Table 1). Next, high-resolution DEM from lidar is used to get elevation values of GCPs with fewer errors. Lastly, lidar DEM is also a source to produce a hillshade map that is able to detect surface features such as stone walls (Johnson and Ouimet 2014), which ensures the horizontal position of target locations.

A Workflow for Generating Historical Orthomosaics

In order to build 1934 and 1951 historical orthomosaics and conduct horizontal accuracy assessment, we provided a workflow from steps 1 through 11 by using Agisoft Metashape 1.6 and ArcGIS Desktop 10.5,

described in Figure 2. Agisoft LLC (2019) provides a general workflow to build an orthomosaic or DEMs and a number of studies have followed it (Ajayi *et al.* 2017; Midgley and Tonkin 2017; Nita *et al.* 2018; Riquelme *et al.* 2019). However, it needs to be improved for applications involving historical data in order to overcome a lack of photo information and build high-quality georeferenced orthophotos that can be used for mapping and GIS purposes. A workflow consists of three general stages, (1) preprocessing (steps 1 and 2), (2) georeferencing and orthomosaicking (steps 3 to 10), and (3) horizontal accuracy assessment (step 11). Details on each step will be addressed in the following sub-sections: “Pre-Processing Stage (Steps 1 to 2)”, “Photo-Alignment/Orthomosaicking Stage (Steps 3 to 10)”, and “Horizontal Accuracy Assessment Stage (Step 11)”. The data were processed with Intel Xeon CPU E5-2687W v3 at 3.10 GHz with 10-core, 128 GB RAM, 20 processors, AMD FirePro W7100 graphics card, and operating on Windows 10 64-bit.

Pre-Processing Stage (Steps 1 to 2)

As a preprocessing stage, input historical photos were masked (step 1) to eliminate unnecessary information such as frame and letters and then the image quality of these photos was estimated (step 2) based on the sharpness value of images that Agisoft image quality tool provides (Agisoft LLC 2019). A blurry image can reduce an orthophoto quality at the final step so images below 0.5 sharpness value out of 1.0 were discarded during the orthomosaicking process (step 10 in Figure 2). The quality test result of 1934 air photos ranges from 0.47 to 0.66 and two photos below the 0.5 quality threshold were disabled. Figure 3 represents an image quality comparison between 0.47 (A) and 0.66 (B) snapped in the same spatial extent. On the other hand, all 1951 air photos were used since they meet the requirement by ranging between 0.74 and 0.88.

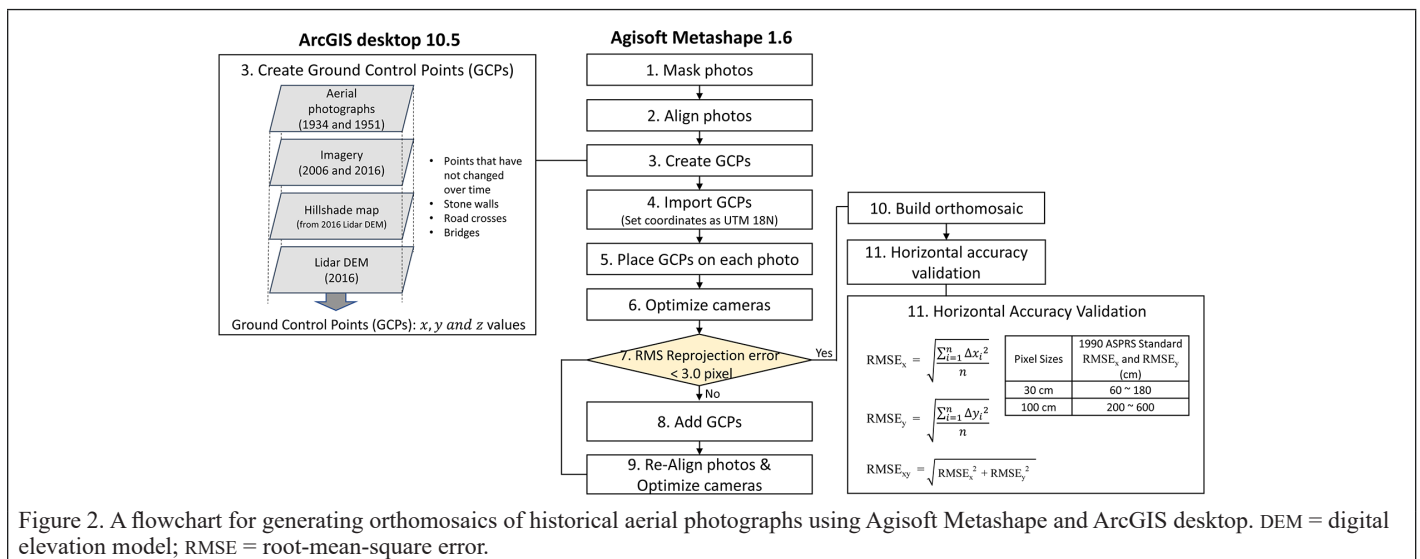


Figure 2. A flowchart for generating orthomosaics of historical aerial photographs using Agisoft Metashape and ArcGIS desktop. DEM = digital elevation model; RMSE = root-mean-square error.



Figure 3. An image quality comparison between 0.50 (A) and 0.61 (B) snapped from 1934 air photos in the same spatial extent. 1934 Aerial Surveys, State Archives Record Group 089:011, State Archives, Connecticut State Library.

Photo-Alignment/Orthomosaicking Stage (Steps 3 to 10)

In this stage, photo-alignment (step 3) was conducted with high accuracy and the generic pair preselection option in Agisoft Metashape. The key point limit and tie point limit was set as 100 000 and 8000, respectively. Given that no information about camera position is used during the photo-alignment process, GCPs with northing, easting, and elevation values should be placed on photos to better alignment as well as georeferencing.

However, there are no GCPs available for historical data, so we defined GCPs (step 4) by selecting points that have not changed over time in a comparison of historical air photos, recent orthophotos, and lidar hillshade maps. Three aspects were taken into account when selecting GCPs. First, the priority of point selection was set considering

frequency and stability of features through time. Therefore, stone walls were the first priority since they existed in the entire study area and were relatively stationary features pre-1934. Then it is followed by road crosses or edges, fixed structures (e.g., bridge, dam, etc.), and the attributes of natural landscape (e.g., creek crosses) (Figure 4). However, the attributes of natural landscape such as creek crosses can be relatively easily changed over time so it was only considered for GCP selection when the rest of potential features (e.g., stone walls, road crosses or edges, fixed structures) were unavailable.

The second aspect was the distribution of GCPs. Given that GCPs with three-dimensional (3D) coordinates strongly control error behavior (Wolf *et al.* 2014), the optimal distribution of GCPs is a point on each corner and additional points uniformly and densely distributed

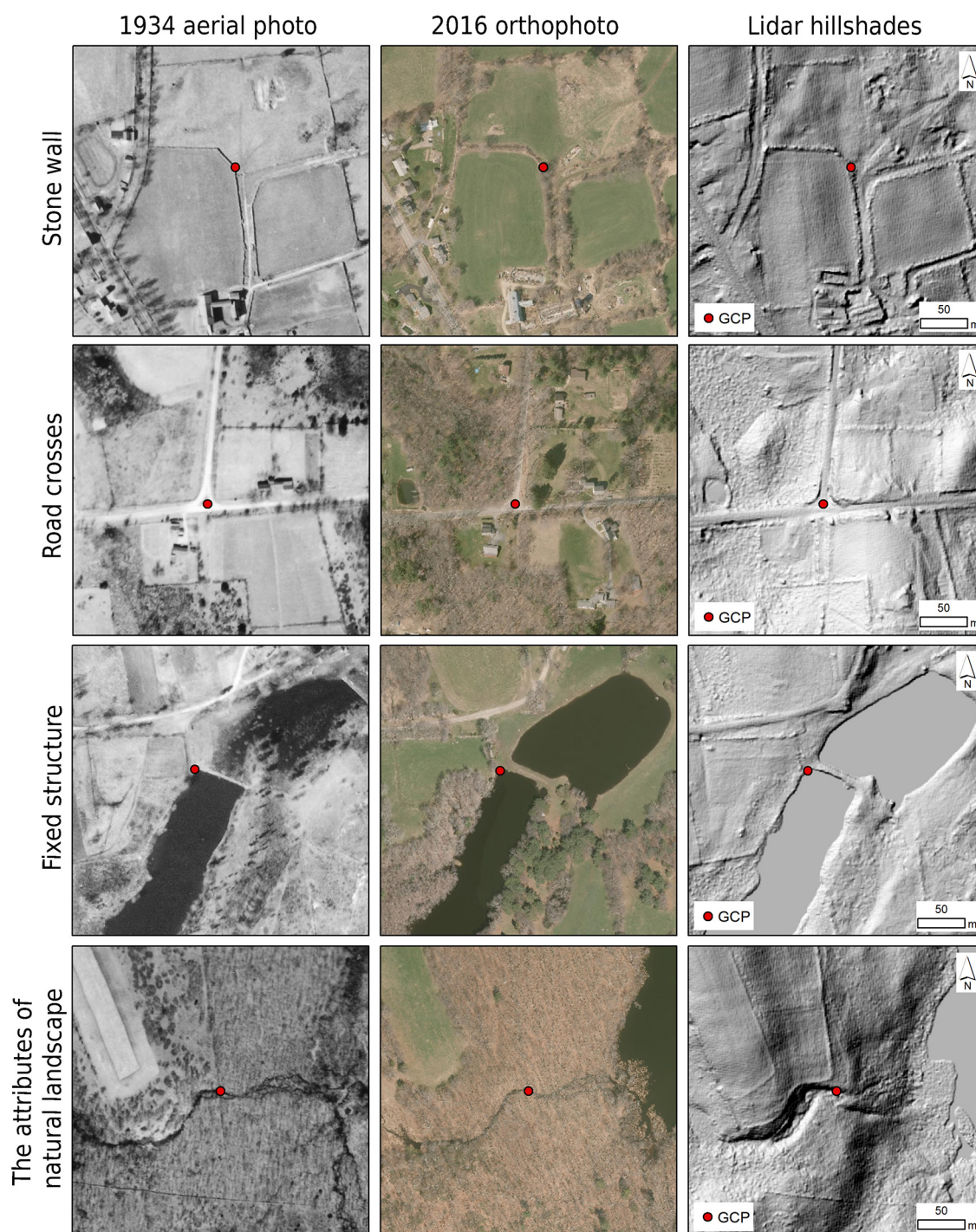


Figure 4. Examples of ground control points (GCPs) on 1934 aerial photo, 2016 orthophoto, and lidar hillshade. 1934 images from 1934 Aerial Surveys, State Archives Record Group 089:011, State Archives, Connecticut State Library.

throughout an image. We tried to place our GCPs as uniformly as possible but were limited by the fact that landscape changes between the images used are often not uniform. When using historical aerial photograph, with 50–90 years between original images and modern lidar and orthophoto data sets, this will likely also be a limitation contributing to error.

The last aspect was the number of photos that each GCP stitched. Each GCP should be located in the place where at least two photos were overlapped. Once GCPs are created at least 1:500 scale level through ArcGIS desktop, northing and easting values were extracted from created GCPs, and elevation values were extracted from lidar DEM (Llena *et al.* 2018) and exported as CSV format, herein we used UTM 18N projection (EPSG: 6347) since measuring horizontal distances was required in accuracy assessment (step 11). After creating GCPs, these GCPs were imported in Agisoft and placed on photos (step 5).

Steps 6 to 9 were repeated until RMS reprojection error was less than 3.0 pixels (see step 7). Camera optimization (step 6) was undertaken using Brown's distortion model (Agisoft LLC 2019) to adjust photo-alignment considering lens distortion. In particular, the following 11 variables were used; focal length (f), principal point offset (c_x and c_y), radial distortion coefficients (K_1 , K_2 , K_3 , and K_4), affinity and skew transformation coefficients (B_1 and B_2), and tangential distortion coefficients (P_1 and P_2). This was an alternative way to overcome a lack of camera calibration information for historical imagery. The quality of photo alignment step is evaluated by RMS reprojection error (in pixel) that is a geometric error associated to the distance between a reconstructed 3D point and an original 3D point detected on the photo. If RMS reprojection error was greater than 3.0 pixel, more GCPs were created and placed on photo. If RMS reprojection error was less than 3.0 pixel, building an orthomosaic (step 10) was conducted. As stated in the description of step 2, images above 0.5 quality threshold were used during an orthomosaicking step and surface parameter was set as 1 m resolution lidar DEM instead of hDEM derived in Metashape due to low resolution of hDEM (e.g., 40 m). Then the orthophoto product was exported and assessed for horizontal accuracy (step 11).

Horizontal Accuracy Assessment Stage (Step 11)

To assess horizontal accuracy of the 1934 and 1951 orthophotos, check points (CPs) were placed on the photo to calculate the residual error of each point and the standard deviation of all residual errors (RMSE). Like GCPs, CPs were typically one of four types: (1) stone walls (SW), (2) road crosses and edges (Rd), (3) natural landscape feature (NL) (e.g., creek crosses), and (4) fixed structures (FS) (e.g., bridge, dam, etc.). In general, RMSE values (i.e., $RMSE_x$, $RMSE_y$, and $RMSE_{xy}$) of CPs are widely used to evaluate the quality of an orthophoto product (American Society for Photogrammetry and Remote Sensing 1990, 2014; Congalton and Green 2009; Tomaščík *et al.* 2019) since they are not used during photo-alignment process. In this study, they are calculated as:

$$RMSE_x(m) = \sqrt{\frac{\sum_{i=1}^n (x_i - \hat{x}_i)^2}{n}} \quad (1)$$

$$RMSE_y(m) = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (2)$$

$$RMSE_{xy}(m) = \sqrt{RMSE_x^2 + RMSE_y^2} \quad (3)$$

where:

x_i is the easting value of CPs from a lidar hillshade map; $\hat{x}_i$ is the estimated easting value of CPs in an orthomosaic product; y_i is the northing value of CPs from a lidar hillshade map; and $\hat{y}_i$ is the estimated northing value of CPs in an orthomosaic product.

The results of RMSE values were compared to American Society for Photogrammetry and Remote Sensing (ASPRS) 1990 standard shown in Table 2. Even though the 1990 standard is regarded as a legacy, we used this standard instead of the recent RMSE standard (e.g., less than 1.4 cm * resolution of geospatial data) given that our inputs were 1934 and 1951 historical data.

Table 2. ASPRS 1990 horizontal accuracy standard information including classes, RMSE values, and recommended uses for each class.

Class	1990 Standard RMSE _x and RMSE _y (m) based on pixel size		Recommended Use
	1934	1951	
1	0.6 m	2.0 m	Highest accuracy work
2	1.2 m	4.0 m	Standard mapping and GIS work
3	1.8 m	6.0 m	Visualization

ASPRS = American Society for Photogrammetry and Remote Sensing; RMSE = root-mean-square error; GIS = geographic information systems.

Results

By following all steps provided in the orthomosaicking workflow (Figure 2), two sets of historical orthomosaics were produced. One is 1934 orthomosaic (ground resolution: 0.3 m/pixel) mosaicked with 141 photos covering 264 km² and the other is 1951 orthomosaic (ground resolution: 0.9 m/pixel) stitched with 68 photos covering 380 km². To meet RMS reprojection error condition (<3.0 pixel), 237 of GCPs was used for 1934 and 234 of GCPs was used for 1951. During the photo-alignment step, the total number of valid tie points for 1934 was 185 293 out of 474 139 and the RMS reprojection error was 2.93 pixel. After bundle adjustment, the mean residual errors of X, Y, Z, and total three coordinates for 1934 were 1.15 m, 1.26 m, 5.56 m, and 5.82 m, respectively. For 1951, the total number of valid tie points was 221 475 out of 464 127 and the RMS reprojection error of 1951 point cloud was 1.1 pixel. After bundle adjustment, the mean residual errors of X, Y, Z, and total three coordinates were 1.59 m, 1.54 m, 9.2 m, and 9.46 m. The bundle adjustment result shows that a large residual error in vertical value (Z) occurred compared to horizontal value (X and Y). This result supports the use of high-resolution lidar DEM as a better resource for the orthomosaicking process instead of the hDEM reconstructed from the historical aerial photographs.

Table 3 and Table 4 shows estimated distance and overlap information for the 1934 and 1951 results based on the estimated camera position by Metashape. The range of side overlap slightly changes depending on the north-south flight path; overall, the survey was flown as regular.

Table 3. Distance and overlap information between camera flight paths of 1934 aerial photographs. The lower case refers to the north-south flight path IDs shown in Figure 1.

1934 Camera Paths	a-b	b-c	c-d	d-e	e-f	f-g	g-h	h-i	i-j
Avg. overlap (km)	1.8	1.7	1.1	2.1	1.4	1.1	1.8	1.7	1.6
Max. overlap (km)	2	1.8	1.3	2.2	1.7	1.4	2.3	1.9	1.9
Min. overlap (km)	1.6	1.6	0.9	2	1	0.8	1.4	1.5	1.2
Side overlap (%)	35-47	35-40	54-67	27-35	41-62	51-73	24-55	45-37	41-60
Forward overlap (%)	45-60								

Table 4. Distance and overlap information between camera flight paths of 1951 aerial photographs. The upper case refers to the north-south flight path IDs shown in Figure 1.

1951 Camera Paths	A-B	B-C	C-D	D-E	E-F
Avg. overlap (km)	3.35	3.66	2.98	3.71	3.98
Min. overlap (km)	2.9	3.15	2.49	3.38	2.83
Max. overlap (km)	3.13	3.37	2.67	3.53	3.05
Side overlap (%)	28-33	21-25	39-40	19-25	29-36
Forward overlap (%)	60-80				

Horizontal accuracy assessment was then conducted to validate the quality of final orthophoto products by using RMSE values. To do this, a total of 287 CPs for the 1934 orthomosaic and 182 CPs for the 1951 orthomosaic were established for the accuracy evaluation. More CPs for the 1934 orthomosaic were required compared to those for the 1951 orthomosaic due to the number of CPs per image. In addition, the 1951 aerial photographs were taken during leaf-on conditions, which led to additional challenges associated with identifying reference objects that have not changed over and were not forested.

Table 5. RMSE results of 1934 and 1951 orthomosaics. *n* is the number of CPs.

	1934	1951
<i>n</i>	287	182
RMSE _x (m)	0.94	1.28
RMSE _y (m)	1.08	1.28
RMSE _{xy} (m)	1.43	1.82
RMSE _{xy} (m) without extreme outliers (>3 m)	0.83 (<i>n</i> = 277)	1.37 (<i>n</i> = 173)

RMSE = root-mean-square error; CPs = check points.

RMSE Results of 1934 and 1951 Orthomosaics

Table 5 shows the RMSE results of the 1934 and 1951 orthomosaics. According to the 1990 ASPRS horizontal accuracy standard (Table 2), the 1934 orthomosaic can be used for standard mapping and GIS work in that the RMSE result without outliers is less than 1.2 m and 1951 orthophoto products can be used for highest accuracy work considering the RMSE without outliers is less than 2.0 m (Table 5). It was found that high-resolution historical orthomosaics can be constructed to town-scale with high accuracy and the orthomosaic procedure can be further applied to expand the spatial scale of study area (i.e., state-scale).

Spatial Pattern of Residual Errors

Figure 5 shows the spatial distribution of residual errors of CPs regarding their location: inside and margin area. Figure 6 shows the residual error boxplots of each CPs for inside area and margin area in both 1934 and 1951 orthophoto results. The results demonstrate that (1) the mean residual errors of 1951 orthophoto are larger than those of 1934 orthoimage and this is caused by spatial resolution differences between two periods, (2) the residual errors from the margin area are larger than those from inside area, and (3) extreme outliers tend to be observed in edge part of the margin area. These results of (2) and (3) supported an edge effect reported in previous studies (Hung *et al.* 2019; Khan and Miklavcic 2019; Nogueira and Roberto 2017). It is partly due to the

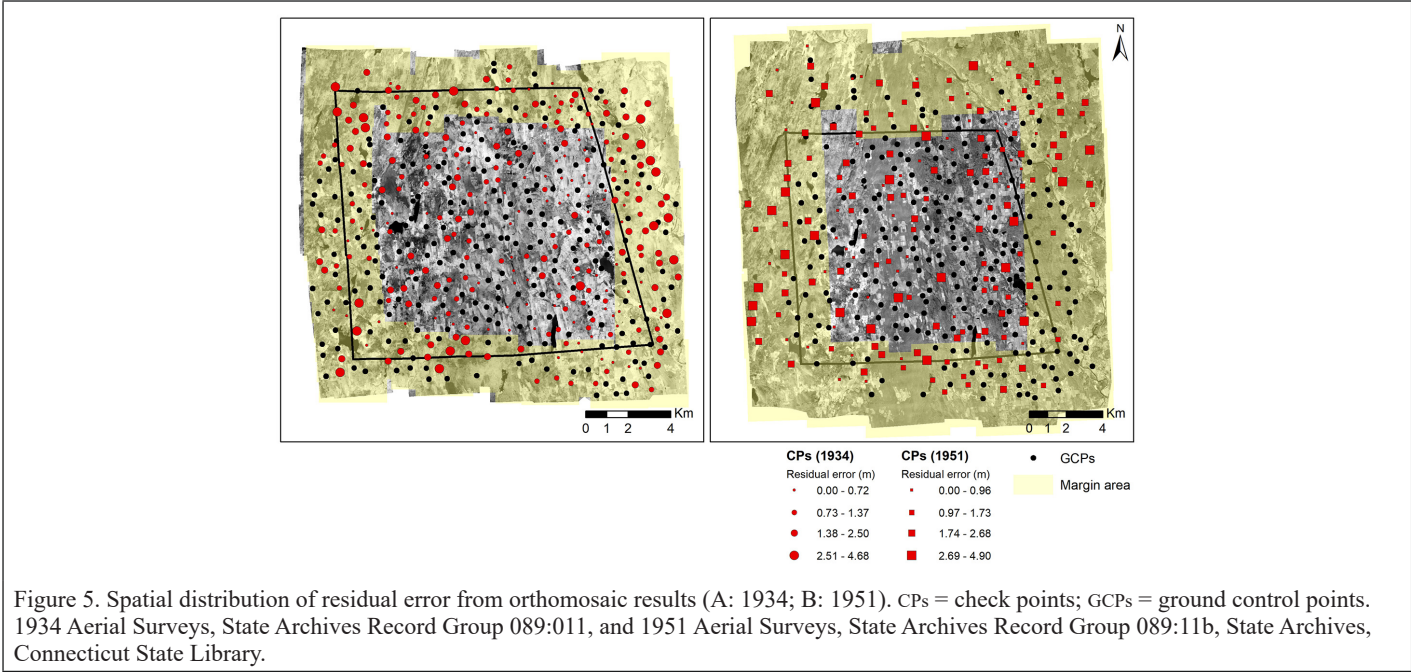


Figure 5. Spatial distribution of residual error from orthomosaic results (A: 1934; B: 1951). CPs = check points; GCPs = ground control points. 1934 Aerial Surveys, State Archives Record Group 089:011, and 1951 Aerial Surveys, State Archives Record Group 089:11b, State Archives, Connecticut State Library.

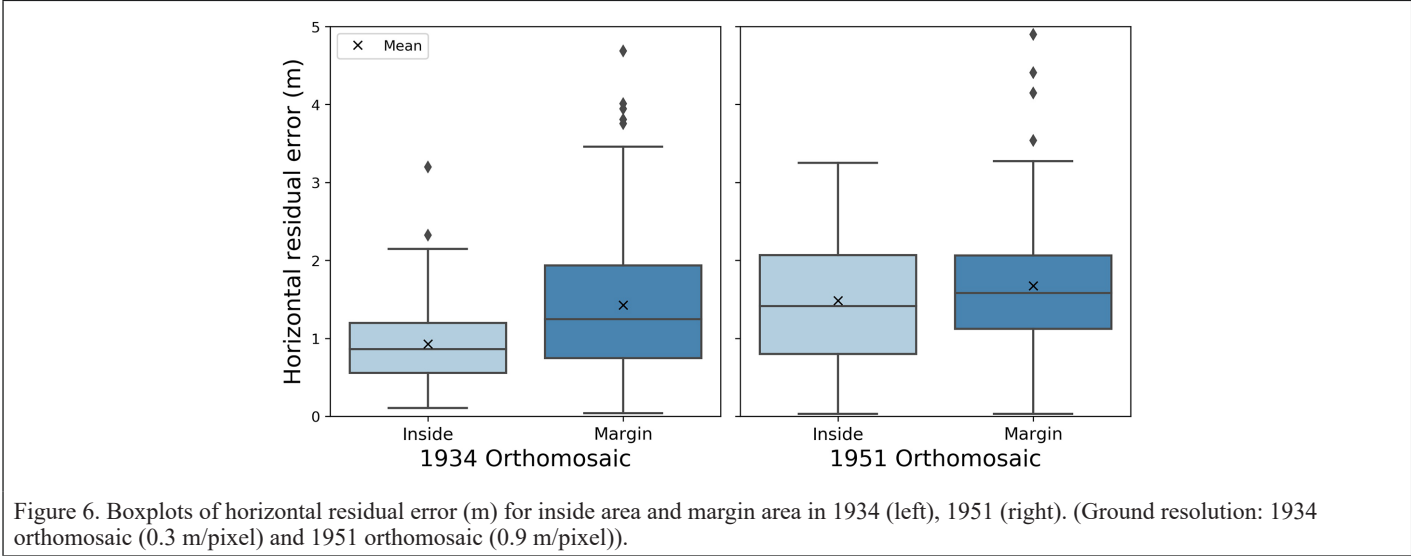


Figure 6. Boxplots of horizontal residual error (m) for inside area and margin area in 1934 (left), 1951 (right). (Ground resolution: 1934 orthomosaic (0.3 m/pixel) and 1951 orthomosaic (0.9 m/pixel)).

fact that the smaller number of key points were extracted and matched from limited images (e.g., two photos) in the margin area.

Discussion

As stated earlier, it was found that the accuracy of a historical orthomosaic can be affected by the spatial resolution of input photos and reference data, and their location (i.e., inside/margin). In other words, the conditions for obtaining high quality of orthomosaics are (1) to use high-resolution input photos (American Society for Photogrammetry and Remote Sensing 2014), (2) to extract reference points from high-resolution reference data such as lidar DEMs and hillshade, and (3) to use additional input photos by buffering a target area at least one flight strip surrounding it to minimize edge effect. However, the results (Figure 5) also demonstrate that a couple of major residual errors occur locally unlike an orthomosaics derived from modern UAV such as a drone. Therefore, we herein discussed additional factors influencing the quality of an orthomosaic from historical aerial photos over a broad area based on a spatial join between CPs' residual errors and map of each factor.

Types of CPs

As aforementioned in section "Horizontal Accuracy Assessment Stage (Step 11)", CPs fall into four types: SW, Rd, NL, and FS. The number of CPs categorized as SW or Rd accounts for ~80% since their locations are relatively easy to be identified between orthomosaic and lidar hillshade and are common landscape features that are static over time.

The occurrence of each CP was the close to the same for each time frame: for the 1934 orthomosaic, the number of CPs categorized as SW accounts for 56%, Rd is 22%, NL is 17%, and FS is 5%; for the 1951 orthomosaic, SW accounts for 53%, Rd is 24%, NL is 19% and FS is 4%. Figure 7 is a boxplot representing the distribution of residual error for each type of CP. CPs of SW and Rd are widely distributed throughout the orthomosaics and tend to show lower residual error compared to NL and FS—although it is difficult to generalize this result because the number of points for each type varies and the accuracy may depend on the location of points (e.g., inside/margin area). In addition, SW and Rd types were easily detectable and more easily delineated on the hillshade maps derived from high-resolution lidar DEM.

The Number of Tied Photos

The number of tied photos was considered as a second factor affecting the quality of orthomosaic because, theoretically, the more photos tied to specific reference points, the higher the accuracy of projected point placement (Agisoft LLC 2019). Figure 8 shows boxplots of CPs' residual errors according to the number of tied photos in 1934 and 1951 orthomosaics. Our results do not completely support the theoretical assumption. In the 1934 orthomosaic, mean residual error tends to decrease as the number of tied photos goes from 1 to 5, but in the 1951 orthomosaic, mean residual error does not decrease as the number of tied photos increases. This is likely due to the fact that it is difficult to place reference points in the exact same location in all photos because some photos have poor sharpness or bad conditions.

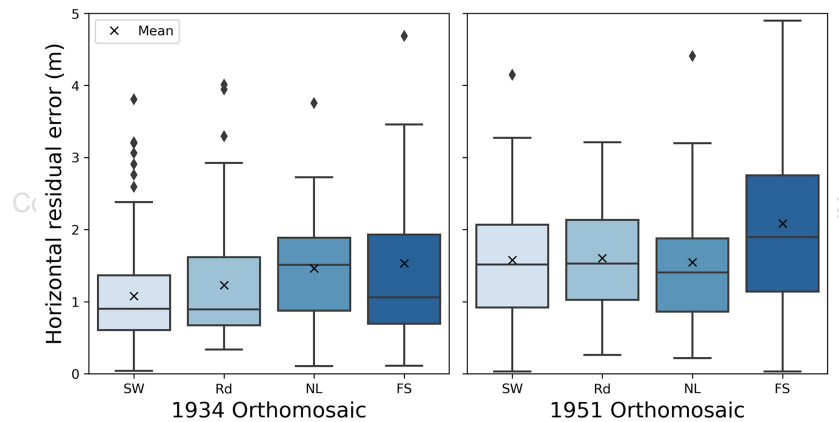


Figure 7. Residual error boxplots of four types of CPs in 1934 (left), 1951 (right). (SW = stone walls, Rd = road crosses or edges, NL = natural landscape features such as creek intersections, FS = fixed structures such as dams or bridges). (Ground resolution: 1934 orthomosaic (0.3 m/pixel) and 1951 orthomosaic (0.9 m/pixel).)

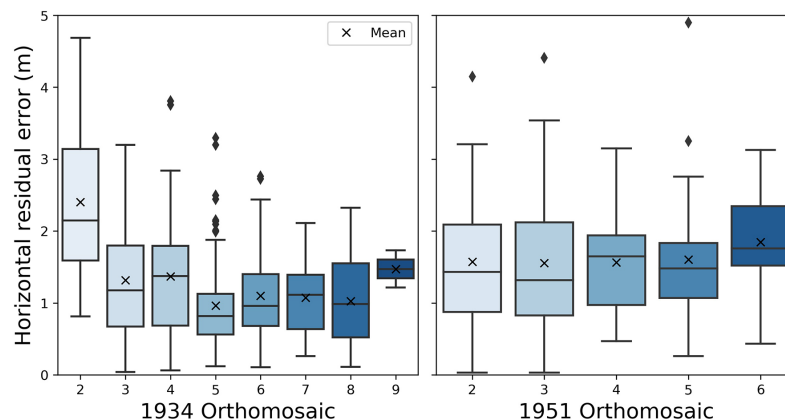


Figure 8. Residual error boxplots as a function of the number of tied photos used in 1934 (left) and 1951 (right). (Ground resolution: 1934 orthomosaic (0.3 m/pixel) and 1951 orthomosaic (0.9 m/pixel).)

Terrain

As the last factor that affects the accuracy of an orthomosaic result, three terrain elements (elevation, relief, and slope) were taken into account (Figure 9). Relief indicates the topographic range in elevation that exists over a specified window; we used a window size of 100 m. Elevation in the study area ranges from 64 m to 335 m, relief ranges from 0 to 71 m, and slope ranges between 0 to 40 degrees. From the perspective of spatial variation, the northwest portion of the study area shows the highest elevation, relief, and slope values. To figure out the relationship between terrain elements and positional accuracy of orthomosaic, each terrain element was reclassified into three classes, which are low (1), medium (2), and high (3) based on natural break and residual error boxplots of terrain conditions were created for both the 1934 and 1951 orthomosaic (Figure 10).

Mean positional residual error of historical orthomosaics increases slightly in areas characterized by higher elevations and higher slope values (Figure 10). One factor contributing to this trend is that fewer GCPs were used in areas with higher elevations and slope values

because these areas tend to have high forest cover in our study area, which poses a challenge to identifying GCPs in lidar hillshades and historical aerial photographs. Therefore, it could be the coupling of terrain characteristics with vegetation differences, rather than strictly terrain, that prevents accurate matching of key points in the aerial photos and leads to larger residual errors.

Conclusion

This study presents a methodological procedure for generating high-resolution historical orthomosaic over a broad area using SfM software Agisoft Metashape and ArcGIS desktop, which allows for the incorporation of lidar data. Among 11 steps in the procedure, creating GCPs with lidar data is the most important step in order to align photos and build historical orthomosaics with high horizontal accuracy. We produced two town-scale high-resolution historical orthomosaics from different timeframes (1934 and 1951) that vary in terms of spatial resolution—1934 (0.3 m/pixel) and 1951 (0.9 m/pixel). In terms of

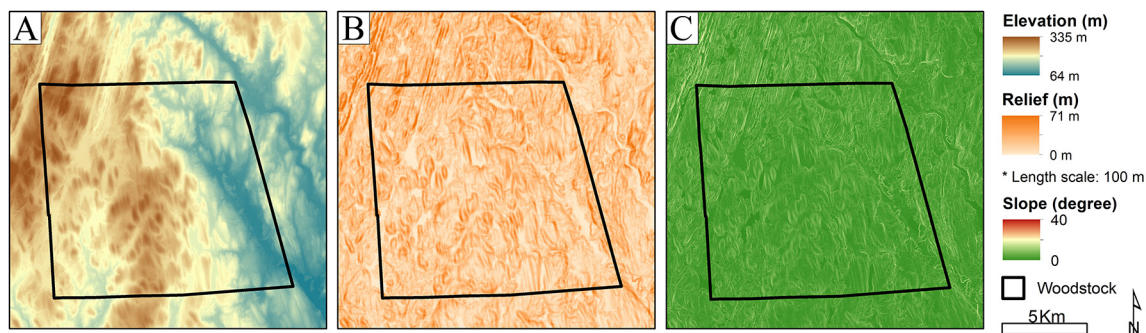


Figure 9. Map of three terrain elements (elevation, relief, and slope) in the study area.(A) elevation map; (B) relief map; (C) slope map.

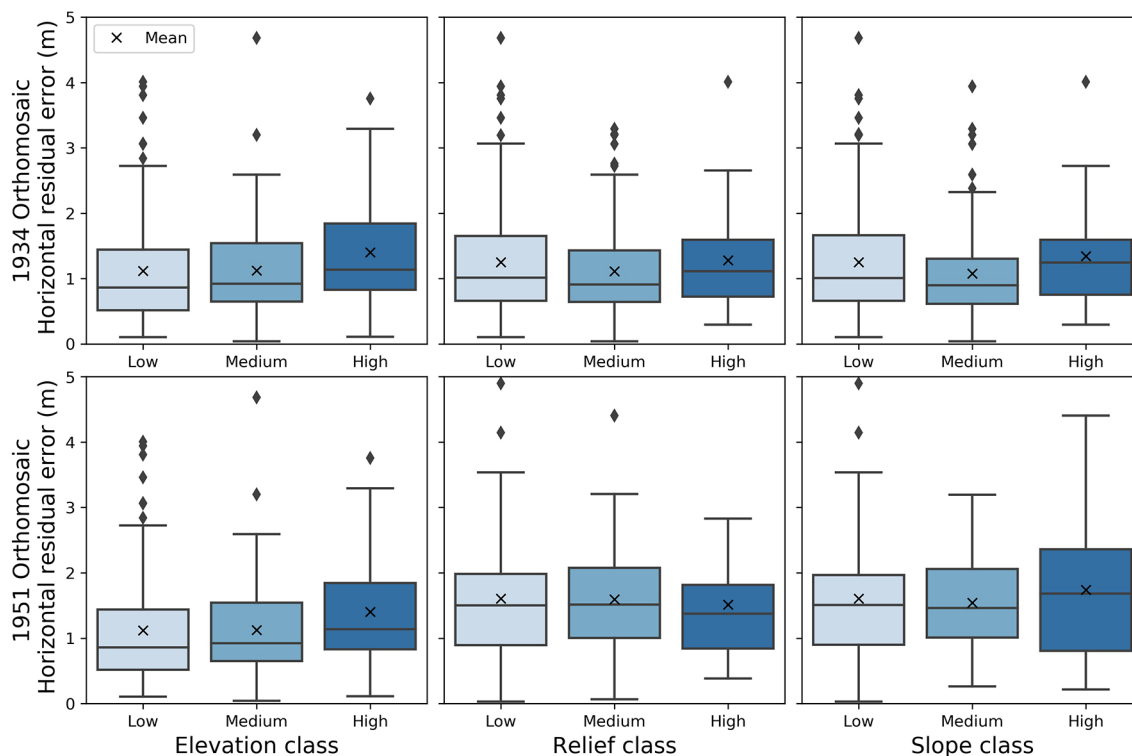


Figure 10. Residual error boxplots of three terrain elements (elevation, relief, and slope) reclassified into low, medium, and high classes based on a natural break. The top row shows the residual error results for the 1934 orthomosaic; the bottom row indicates that of the 1951 orthomosaic.

horizontal accuracy assessment, RMSE values without extreme outliers for both orthomosaics demonstrate they are highly accurate and can be used for standard mapping and GIS work according to 1990 ASPRS horizontal accuracy standard. In addition, the spatial distribution of CP residual errors indicates that an edge effect should be taken into account and enough photos should be included to cover at least one more edge layer larger than study area.

Moreover, we highlight that there are three factors influencing the quality of historical orthomosaics. First, it is important to use CPs not only extracted from high-resolution reference data such as lidar but also selected in terms of being stationary features through time and high frequency across the study area (e.g., stone walls and road crossings). Second, the number of tied photos can increase horizontal accuracy, but more tied photos do not always lead to higher accuracy because external causes such as image quality and stains on images prevent reference points being placed on the exact same locations. Lastly, it was found that the complexity of terrain can also affect the accuracy of orthomosaics.

Limitations to this study include the following: (1) the factors influencing the quality of historical orthomosaics were examined by a spatial join between the map of each factor and a map of CPs' residual errors such that it is not possible to examine the relationships amongst individual factors; (2) there may have been errors associated with orthomosaicking or accuracy assessments results due to digitization and imprecise locations of reference points on both photos and the map; and (3) it can be challenging to apply this method to completely forested areas where, in general, reference points are harder to come by.

Overall, despite these limitations, this contribution provides an important methodological procedure for making high-resolution, historical orthomosaics and for suggesting factors to be considered when applying this method. The procedure presented here can be applied to any study areas where historical aerial photographs and lidar data are available. Future application of this methodology would be to extend beyond the town scale considered here (e.g., 50–200 km²) all the way to state scale or larger (10 000–100 000 km²), such as the state of Connecticut or the northeastern US in general. In addition to this, the outputs of our procedure can be used in various studies based on time-series analysis since the early to mid-20th century when historical aerial photographs become widely available, like detecting environmental disturbances, land use changes, and anthropogenic impacts. These products can also be used as input data for state-of-the-art deep learning algorithms to do image classification to reconstruct and analyze historical LULC.

Acknowledgments

This work was funded by a National Science Foundation grant BCS-1654462 to W. Ouimet. The authors wish to thank Katharine Johnson and Zhe Zhu for comments on an early draft of this contribution, and an anonymous reviewer for their helpful comments and suggestions.

References

- Agisoft LLC. 2019. *Agisoft Metashape User Manual*. <https://www.agisoft.com/pdf/metashape-pro_1_5_en.pdf>.
- Ajayi, O. G., A. A. Salubi, A. F. Angbas and M. G. Odigure. 2017. Generation of accurate digital elevation models from UAV acquired low percentage overlapping images. *International Journal of Remote Sensing* 38(8–10):3113–3134.
- American Society for Photogrammetry and Remote Sensing. 1990. ASPRS accuracy standards for large-scale maps. *Photogrammetric Engineering & Remote Sensing* 56(7):1068–1070.
- American Society for Photogrammetry and Remote Sensing. 2014. ASPRS positional accuracy standards for digital geospatial data, edition 1, version 1.0. *Photogrammetric Engineering & Remote Sensing* 81(3):1–26.
- Arnaud, F., H. Piégay, L. Schmitt, A. J. Rollet, V. Ferrier and D. Béal. 2015. Historical geomorphic analysis (1932–2011) of a by-passed river reach in process-based restoration perspectives: the Old Rhine downstream of the Kembs Diversion Dam (France, Germany). *Geomorphology* 236:163–177.
- Baker, W. L., J. J. Honaker and P. J. Weisberg. 1995. Using aerial photography and GIS to map the forest-tundra ecotone in Rocky Mountain National Park, Colorado, for global change research. *Photogrammetric Engineering and Remote Sensing* 61(3):313–320.
- Bolstad, P. V. 1992. Geometric errors in natural resource GIS data: Tilt and terrain effects in aerial photographs. *Forest Science* 38(2):367–380.
- Capitol Region Council of Governments (CRCoG). 2016. 2016 Aerial Imagery. <<http://cteco.uconn.edu/data/flight2016/index.htm>> Accessed on 29 July 2021.
- Child, S. F., L. A. Stearns, L. Girod and H. H. Brecher. 2021. Structure-from-motion photogrammetry of antarctic historical aerial photographs in conjunction with ground control derived from satellite data. *Remote Sensing* 13(1):1–25.
- Comiti, F., M. Da Canal, N. Surian, L. Mao, L. Picco and M. A. Lenzi. 2011. Channel adjustments and vegetation cover dynamics in a large gravel bed river over the last 200 years. *Geomorphology* 125(1):147–159.
- Congalton, R. G. and K. Green. 2009. The photogrammetric record. In *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, 2d ed. Boca Raton, Fla.: CRC Press.
- Connecticut State Library. 1934. *1934 Aerial Surveys*. Record Group 089:011. College Park, Md.: Department of Transportation, State Archives.
- Connecticut State Library. 1951. *1951 Aerial Surveys*. Record Group 089:011b. College Park, Md.: Department of Transportation, State Archives.
- Cronon, W. 1983. *Changes in the Land: Indians, Colonists, and the Ecology of New England*. New York: Hill and Wang.
- CT Environmental Conditions Online (CT ECO). 2012. 2012 Aerial Imagery. <http://cteco.uconn.edu/help/info_orthos2012.htm> Accessed on 29 July 2021.
- ESRI. 2012. “World Topographic Map”. 1:100 000. <<http://www.arcgis.com/home/item.html?id=30e5fe3149c34df1ba922e6f5bbf808f>> Accessed on 29 July 2021.
- Fox, A. J. and A. Cziferszky. 2008. Unlocking the time capsule of historic aerial photography to measure changes in antarctic peninsula glaciers. *Photogrammetric Record* 23(121):51–68.
- Frankl, A., V. Seghers, C. Stal, P. De Maeyer, G. Petrie and J. Nyssen. 2015. Using image-based modelling (SfM–MVS) to produce a 1935 orthomosaic of the Ethiopian highlands. *International Journal of Digital Earth* 8(5):421–430.
- Gennaretti, F., M. N. Ripa, F. Gobattoni, L. Boccia and R. Pelorosso. 2011. A methodology proposal for land cover change analysis using historical aerial photos. *Journal of Geography and Regional Planning* 4(9):542–556.
- Geyman, E. C., W. van Pelt, A. C. Maloof, H. F. Aas and J. Kohler. 2022. Historical glacier change on Svalbard predicts doubling of mass loss by 2100. *Nature* 601.
- Gomez, C. 2014. Digital photogrammetry and GIS-based analysis of the biogeomorphological evolution of Sakurajima Volcano, diachronic analysis from 1947 to 2006. *Journal of Volcanology and Geothermal Research* 280:1–13.
- Hung, I.-K., D. Unger, D. Kulhavy and Y. Zhang. 2019. Positional precision analysis of orthomosaics derived from drone captured aerial imagery. *Drones* 3(2):46.
- Johnson, K. M. and W. B. Ouimet. 2014. Rediscovering the lost archaeological landscape of southern New England using airborne light detection and ranging (LiDAR). *Journal of Archaeological Science* 43(1):9–20.
- Johnson, K. M. and W. B. Ouimet. 2016. Physical properties and spatial controls of stone walls in the northeastern USA: Implications for anthropocene studies of 17th to early 20th century agriculture. *Anthropocene* 15:22–36.
- Kadmon, R. and R. Harari-Kremer. 1999. Studying long-term vegetation dynamics using digital processing of historical aerial photographs. *Remote Sensing of Environment* 68:164–176.
- Khan, Z. and S. J. Miklavcic. 2019. An automatic field plot extraction method from aerial orthomosaic images. *Frontiers in Plant Science* 10(May):1–13.
- Leh, M., S. Bajwa and I. Chaubey. 2013. Impact of land use change on erosion risk: An integrated remote sensing, geographic information system and modeling methodology. *Land Degradation and Development* 24(2011):409–421.
- Lingua, A., D. Marenchino and F. Nex. 2009. Performance analysis of the SIFT operator for automatic feature extraction and matching in photogrammetric applications. *Sensors* 9(5):3745–3766.

- Llena, M., M. Cavalli, D. Vericat and S. Crema. 2018. Assessing landscape changes associated to anthropic disturbances by means of the application of structure from motion photogrammetry using historical aerial imagery. *Rendiconti Online Società Geologica Italiana* 46(May):74–81.
- Mallinis, G., D. Emmanouloudis, V. Giannakopoulos, F. Maris and N. Koutsias. 2011. Mapping and interpreting historical land cover/land use changes in a natura 2000 site using Earth observational data: The case of Nestos Delta, Greece. *Applied Geography* 31(1):312–320.
- Marignani, M., D. Rocchini, D. Torri, A. Chiarucci and S. Maccherini. 2008. Planning restoration in a cultural landscape in Italy using an object-based approach and historical analysis. *Landscape and Urban Planning* 84(1):28–37.
- Maurer, J. and S. Rupper. 2015. Tapping into the hexagon spy imagery database: A new automated pipeline for geomorphic change detection. *ISPRS Journal of Photogrammetry and Remote Sensing* 108:113–127.
- Midgley, N. G. and T. N. Tonkin. 2017. Reconstruction of former glacier surface topography from archive oblique aerial images. *Geomorphology* 282:18–26.
- National Research Council. 2001. Grand challenges in environmental sciences. In *Grand Challenges in Environmental Sciences*. Washington, DC: The National Academies Press.
- Nebiker, S., N. Lack and M. Deuber. 2014. Building change detection from historical aerial photographs using dense image matching and object-based image analysis. *Remote Sensing* 6(9):8310–8336.
- Nita, M. D., C. Munteanu, G. Gutman, I. V. Abrudan and V. C. Radeloff. 2018. Widespread forest cutting in the aftermath of World War II captured by broad-scale historical corona spy satellite photography. *Remote Sensing of Environment* 204:322–332.
- Nogueira, F. C. and L. Roberto. 2017. Accuracy analysis of orthomosaic and DSM produced from sensor aboard UAV. In *XVIII Simpósio Brasileiro de Sensoriamento Remoto—SBSR*, 5515–5522.
- Riquelme, A., M. Del Soldato, R. Tomás, M. Cano, L. Jordá Bordehore and S. Moretti. 2019. Digital landform reconstruction using old and recent open access digital aerial photos. *Geomorphology* 329:206–223.
- Rocchini, D., G.L.W. Perry, M. Salerno, S. Maccherini and A. Chiarucci. 2006. Landscape change and the dynamics of open formations in a natural reserve. *Landscape and Urban Planning* 77(1–2):167–77.
- Sevara, C. 2013. Top secret topographies: Recovering two and three-dimensional archaeological information from historic reconnaissance datasets using image-based modelling techniques. *International Journal of Heritage in the Digital Era* 2(3):395–418.
- Thorson, R. M. 2002. *Stone by Stone : The Magnificent History in New England's Stone Walls*. New York, N.Y. : Walker & Company.
- Tomaščík, J., M. Mokroš, P. Surový, A. Grznárová and J. Merganič. 2019. UAV RTK/PPK method—An optimal solution for mapping inaccessible forested areas? *Remote Sensing* 11(6):721.
- Turner, D., A. Lucieer and C. Watson. 2012. An automated technique for generating georectified mosaics from ultra-high resolution unmanned aerial vehicle (UAV) imagery, based on structure from motion (SfM) point clouds. *Remote Sensing* 4(5):1392–1410.
- Verbesselt, J., A. Zeileis and M. Herold. 2012. Near real-time disturbance detection using satellite image time series. *Remote Sensing of Environment* 123:98–108.
- Verhoeven, G., M. Doneus, C. Briese and F. Vermeulen. 2012. Mapping by matching: A computer vision-based approach to fast and accurate georeferencing of archaeological aerial photographs. *Journal of Archaeological Science* 39(7):2060–2070.
- Weng, J., T. S. Huang and N. Ahuja. 2013. *Motion and Structure from Image Sequences*. Berlin/Heidelberg, Germany: Springer Science & Business Media.
- Wolf, P. R., B. A. Dewitt and B. E. Wilkinson. 2014. *Elements of Photogrammetry with Applications in GIS*, 4th ed. Columbus, Ohio: McGraw-Hill Education.
- Zhu, Z. and C. E. Woodcock. 2014. Continuous change detection and classification of land cover using all available landsat data. *Remote Sensing of Environment* 144:152–171.
- Zhu, Z., Y. Fu, C. E. Woodcock, P. Olofsson, J. E. Vogelmann, C. Holden, M. Wang, S. Dai and Y. Yu. 2016. Including land cover change in analysis of greenness trends using all available Landsat 5, 7, and 8 images: A case study from Guangzhou, China (2000–2014). *Remote Sensing of Environment* 185:243–257.
- Zhu, Z., J. Zhang, Z. Yang, A. H. Aljaddani, W. B. Cohen, S. Qiu and C. Zhou. 2020. Continuous monitoring of land disturbance based on Landsat time series. *Remote Sensing of Environment* 238:111116.
- Zomeni, M., J. Tzanopoulos and J. D. Pantis. 2008. Historical analysis of landscape change using remote sensing techniques: An explanatory tool for agricultural transformation in Greek rural areas. *Landscape and Urban Planning* 86(1):38–46.

IN-PRESS ARTICLES

- Jordan Bowman, Lexie Yang, Orrin Thomas, Jerry Kirk, Andrew Duncan, David Hughes, and Shannon Meade. UAS Edge Computing of Energy Infrastructure Damage Assessment.
- Firat Erdem, Nuri Erkin Ocer, Dilek Kucuk Matci, Gordana Kaplan, and Ugur Avdan. Apricot Tree Detection from UAV-Images Using Mask R-CNN and U-Net.
- Batuhan Sariturk, Damla Kumbasar, and Dursun Zafer Seker. Comparative Analysis of Different CNN Models for Building Segmentation from Satellite and UAV Images.
- Rabi N. Sahoo, Shalini Gakhar, R.G. Rejith, Rajeev Ranjan, Mahesh C. Meena, Abir Dey, Joydeep Mukherjee, Rajkumar Dhakar, Sunny Arya, Anchal Daas, Subhash Babu, Pravin K. Upadhyay, Kapila Sekhawati, Sudhir Kumar, Mahesh Kumar, Viswanathan Chinnusamy, and Manoj Khanna. Unmanned Aerial Vehicle (UAV)–Based Imaging Spectroscopy for Predicting Wheat Leaf Nitrogen.
- Yunus Kaya, Halil İbrahim Şenol, Abdurrahman Yasin Yiğit, and Murat Yakar. Car Detection from Very High-Resolution UAV Images Using Deep Learning Algorithms.
- Xin Jia, Qing Zhu, Xuming Ge, Ruifeng Ma, Daiwei Zhang, and Tao Liu. Robust Guardrail Instantiation and Trajectory Optimization of Complex Highways Based on Mobile Laser Scanning Point Clouds.
- Shuanggen Jin, Ayman M. Elameen, Daniel Olago. Identification of drought events in major basins of Africa based on weighted water storage deficit index from GRACE measurements.
- Xin Jia, Qing Zhu, Xuming Ge, Ruifeng Ma, Daiwei Zhang, and Tao Liu. Robust Guardrail Instantiation and Trajectory Optimization of Complex Highways Based on Mobile Laser Scanning Point Clouds.
- Zhikang Lin, Wei Liu, Yulong Wang, Yan Xu, Chaoyang Niu. Change Detection in SAR Images through Clustering Fusion Algorithm and Deep Neural Networks.
- Dan Li, Hanjie Wu, Yujian Wang, Xiaojun Li, Fanqiang Kong, and Qiang Wang. Lightweight Parallel Octave Convolutional Neural Network for Hyperspectral Image Classification.
- Tolga Bakirman, Bahadır Kulavuz, and Bulent Bayram. Use of Artificial Intelligence Toward Climate-Neutral Cultural Heritage.