Undistorting the Past: New Techniques for Orthorectification of Archaeological Aerial Frame Imagery

3

Geert Verhoeven, Christopher Sevara, Wilfried Karel, Camillo Ressl, Michael Doneus, and Christian Briese

Contents

3.1 Aerial Archaeological Frame Footage: An Introduction and Overview

3.1.1 One Hundred Years of Status Quo

Since Joseph Nicéphore Nièpce (1765–1833) invented 'drawing with light' in the 1820s, photography can almost celebrate its second centenary. Archaeological aerial photography covers approximately one half of that time

G. Verhoeven (✉)
Department of Archaeology, Ghent University, Ghent, Belgium

VIAS – Vienna Institute for Archaeological Science, University of Vienna, Vienna, Austria

LBI for Archaeological Prospection and Virtual Archaeology, Vienna, Austria
e-mail: geert.verhoeven@ugent.be

C. Sevara
Initiative College for Archaeological Prospection, VIAS – Vienna Institute for Archaeological Science, University of Vienna, Vienna, Austria

W. Karel
VIAS – Vienna Institute for Archaeological Science, University of Vienna, Vienna, Austria

Department of Geodesy and Geoinformation, Vienna University of Technology, Vienna, Austria

C. Ressl
Department of Geodesy and Geoinformation, Vienna University of Technology, Vienna, Austria

M. Doneus
VIAS – Vienna Institute for Archaeological Science, University of Vienna, Vienna, Austria

LBI for Archaeological Prospection and Virtual Archaeology, Vienna, Austria

C. Briese
LBI for Archaeological Prospection and Virtual Archaeology, Vienna, Austria

Department of Geodesy and Geoinformation, Vienna University of Technology, Vienna, Austria

C. Corsi et al. (eds.), *Good Practice in Archaeological Diagnostics*, Natural Science in Archaeology,
DOI 10.1007/978-3-319-01784-6_3, © Springer International Publishing Switzerland 2013

span. The first aerial image was taken in 1858 from a tethered hot-air balloon by Gaspard-Félix Tournachon – also known as Nadar – from the village of Petit Bicêtre (Colwell 1997; Newhall 2006). It was not, however, until June 1899 that the first (European) archaeological photograph, of the forum in Rome, was taken from a balloon by Giacomo Boni (Castrianni 2008). Despite the first flight of a manned, motor-driven machine built by Orville and Wilbur Wright in 1903, archaeologically significant pictures were not captured from an aeroplane until World War I (Barber 2011). In this first phase of archaeological aerial reconnaissance, much credit must be given to O.G.S. Crawford (1886–1957). This Englishman is considered to be the inventor of scientific aerial reconnaissance, and his work in the 1920s and beyond was the basis for the future development of aerial archaeology (e.g. Crawford 1924, 1929, 1933; Crawford and Keiller 1928).

Since Crawford and other pioneers of aerial archaeology such as Antoine Poidebard (1878–1955) and Theodor Wiegand (1864–1936), it has been recognised that archaeological remains can show up on the earth's surface in a number of ways. Aside from standing structures (e.g. bridges, theatres, fortifications) which are directly visible from the ground as well as the air, most archaeological remains are partly eroded or only exist as sub-surface archaeological features, showing up on the surface under certain conditions as *visibility marks*: i.e. indirect indicators of archaeological residues due to the changed properties of the soil matrix or the local topography (Crawford 1924; Scollar et al. 1990; Wilson 2000; Bewley and Rączkowski 2002; Brophy and Cowley 2005; see Chap. 2 in this volume). Apart from the less frequent flood and wind marks, archaeologists generally differentiate between four main types of marks:

- Soil marks – due to varying chemical and physical properties affecting the soil colour on the surface
- Crop/vegetation marks – due to variable growth and vigour of the vegetation
- Shadow marks – when earthworks are thrown into relief by low slanting sunlight
- Snow/frost marks – due to differential snow accumulations and differential melting of snow or frost

To date, the common practice of active archaeological aerial photographic reconnaissance is quite straightforward and seems not to have changed over the past century (Verhoeven 2009a). In general, images are acquired from the cabin of a low-flying aircraft using a small- or medium-format hand-held photographic/still frame camera equipped with a lens that is commonly uncalibrated (Wilson 1975; Crawshaw 1995). Once airborne, the archaeologist flies over targeted areas, trying to detect possible archaeologically induced anomalies in the landscape. Once an archaeological feature is detected, it is orbited and documented from various positions (generally from an oblique point of view) on the digital camera sensor or a specific panchromatic, true colour, monochromatic infrared or (false-) colour infrared film. This type of aerial photographic reconnaissance has been the workhorse of all archaeological remote-sensing techniques since it is one of the most cost-effective methods for site discovery and the non-invasive approach yields easily interpretable imagery with abundant spatial detail (Wilson 2000; Palmer 2005).

However, no matter how efficient this reconnaissance approach can be in certain areas and periods, its main disadvantage is the fact that the whole flying strategy is *observer directed* (Palmer 2005) and generates extremely selective (i.e. biased) data that are totally dependent on an airborne observer recognising archaeological phenomena. Thus sub-surface soil disturbances that are visually imperceptible at the time of flying (e.g. Verhoeven 2009a), or those that are simply overlooked, will not make it into a photograph. To counteract this, several authors have already questioned this strategy of observer-directed survey and pointed out the advantage of a so-called unbiased, vertical approach (Palmer 1996, 2007; Doneus 1997, 2000; Mills 2005; Coleman 2007). Although the observer-directed flying method might yield vertical photographs as well, the vast majority of the photographs will be oblique in nature. This means that the optical axis of the imager intentionally deviates more than 3° from the vertical to the earth's surface (Schneider 1974). Depending on the visibility of the horizon, the image is then further classified as low oblique (i.e. horizon is not included) or high oblique (Harman et al. 1966).

3.1.2 The Vertical Debate

In a strictly vertical sortie, every parameter is set to make sure that all photographs are nadir/vertical images. In effect, this means that photographs will be acquired with expensive, accurately calibrated, built-in (versus handheld), gyro-stabilised and low distortion mapping frame cameras (often referred to as metric or cartographic cameras – Slater et al. 1983). These cameras are solidly housed and operated in bigger and higher-flying aeroplanes. Images are acquired in parallel strips at regular intervals, generally with a large frame overlap: in one flight strip, each photograph has a generally accepted degree of overlap of circa 60 % ± 5 % (figures to 90 % can be found as well, see Schneider 1974) with the following and preceding image (longitudinal overlap). Adjacent strips have on average an overlap of 25–40 % (lateral overlap) (Read and Graham 2002). The camera is pointing directly down to the earth to acquire (near) nadir photographs. Because a perfect vertical is almost never achieved, an image with an angle of less than or equal to 3° is called vertical (Estes et al. 1983).

Archaeological resources often appear on verticals through what has been termed the *serendipity effect* (Brugioni 1989): a circumstance in which photosets yield unanticipated or 'bonus' material which was not the primary objective during original data collection. Unlike oblique aerial photography for archaeological purposes, those vertical surveys are generally executed to acquire basic material for (orthophoto) map generation (Falkner and Morgan 2002). Although this approach generates geographically unbiased photographs of large areas in a very fast manner, its adversaries remark that several issues militate against the effective use of those vertical photographs for archaeological purposes. Of those, the fact that imagery is not captured at the perfect oblique angle to maximise the visibility of archaeological information (Crawshaw 1997) is often seen as the strongest argument to not fly (or even use) verticals. On the other hand, vertical footage offers an advantage in mapping, as the induced geometrical distortions are much less than those embedded in oblique footage (Imhof and Doolittle 1966; see part 2). Since the data are by default captured in stereo pairs, they are also perfectly suited to create analogue or digital 3D stereo models. Additionally, the high spatial resolution and comparatively broad coverage of standard vertical mapping images make them valuable for a holistic view of the landscape as well as for the primary discovery of individual archaeological features.

As a result, many aerial archaeologists have extracted much valuable information from verticals (e.g. Moscatelli 1987; Kennedy 1996; Doneus 1997), and a few studies have proven the undeniable and often complementary value of verticals after a thorough comparison with obliques from the same area (e.g. Zantopp 1995; Doneus 2000; Palmer 2007). In reality, even those archaeologists that favour obliques over a blanket vertical coverage will incorporate verticals into their research, simply because many valuable historic aerial photographs were acquired with a (near-)vertical approach (Stichelbaut et al. 2009; Hanson and Oltean 2013). Since these photograph series are able to illustrate change through time, they provide valuable data regarding landscape change and indirect land use impact on archaeological resources (Cowley and Stichelbaut 2012).

3.1.3 The Rise of the Unmanned Machines

Finally, it needs to be mentioned that both oblique and vertical frame images can also be acquired from low-altitude unmanned platforms. Since the beginning of aerial photography, researchers have used all kinds of devices (from pigeons, kites, poles and balloons to rockets) to take still cameras aloft and remotely gather aerial imagery (see Verhoeven 2009b for an archaeological overview). To date, many of these unmanned devices are still used for what has been referred to as low-altitude aerial photography or LAAP (Schlitz 2004). In addition to these more traditional camera platforms, radio-controlled (multi-) copter platforms have recently added a new aspect to LAAP (Fig. 3.1). The overwhelming amount of

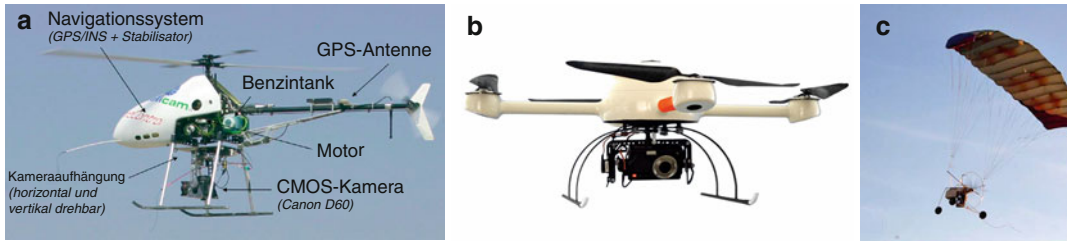

Fig. 3.1 (**a**) Example of a remotely controlled helicopter to acquire digital aerial imagery (Reproduced from Eisenbeiss et al. 2005) (**b**) The Microdrone MD4-200 quadcopter (Microdrones GmbH 2008) (**c**) Remotely controlled paraglider (Krijnen 2008)

brands and types (heli-, dual-, tri-, quad-, hexa- and octocopters), together with the wide variety of navigation options (e.g. altitude and position hold, waypoint flight – Eisenbeiss 2009; Eisenbeiss and Sauerbier 2011) and camera mounts, indicate that these platforms are here to stay for some time. Given the multitude of still camera types and the image quality they are currently capable of, endless combinations of low- and high-cost LAAP solutions are available. In addition, LAAP allows for the exploitation of new imaging techniques, as it is often only a matter of lifting the appropriate device. In this way several archaeological studies have utilised close-range near-infrared photography (e.g. Whittlesey 1973; Aber et al. 2001; Verhoeven et al. 2009a; Wells and Wells 2012) or even less straightforward near-ultraviolet imaging (e.g. Verhoeven 2008a; Verhoeven and Schmitt 2010; Wells and Wells 2012).

3.1.4 The Mapping Paradigm

Despite this large variety of still frame images and means to acquire them (actively or passively), their archaeological information cannot (or will not) be exploited efficiently as long as the images are not thoroughly interpreted (i.e. interpretatively mapped – cf. Doneus et al. 2001) and integrated with other data sources. The lack of this interpretative mapping is often encountered and may have multiple causes. Availability of resources may be one cause, but one of the most important ones is likely the time-consuming (and often difficult) georeferencing process (Verhoeven et al. 2012a). As a result, millions of aerial photographs are just stored in archives, waiting for their archaeological potential to be explored. Obviously, aerial archaeology is in need of fast, straightforward and accurate georeferencing approaches that allow orthophoto production of a wide variety of images: old or new, acquired in a vertical or oblique manner from low or high altitudes.

This chapter elaborates on such an approach and presents a method to automate the important but recurring task of orthophoto generation. The approach proposed here attempts to overcome the conventional georeferencing problems related to archaeological aerial frame images, which means that in this chapter imagery resulting from panoramic and line cameras is not included. To this end, the methodology exploits some of the technological improvements in hardware configurations as well as state-of-the-art algorithms mainly developed in the fields of computer vision and photogrammetry: two disciplines that research the recovery of 3D content from 2D imagery using – to a certain extent – their own specific approaches (Hartley and Mundy 1993). Before outlining the method (Sect. 3.3), the concept of georeferencing and all the sources of geometrical image deformations that have to be taken into account will be outlined in Sect. 3.2. Section 3.4 will illustrate these concepts with several case studies. In addition to this illustrative purpose, these case studies will also provide some more in-depth knowledge about specific aspects of particular aerial image types. A conclusion, presenting some future aims and remarks, will then finalise this chapter.

3.2 Aerial Frames Offer Deformed Views

3.2.1 (Digital) Aerial Images

Aerial imaging is facilitated by the use of an airborne remote-sensing instrument that gathers the earth's spatially, temporally, radiometrically and spectrally varying upwelling electromagnetic radiation and uses this to generate (digital) images (see Schott 2007 for a good treatise of this subject). In past decades, this detection of radiation was usually accomplished by a photographic emulsion sensitised into one or more spectral regions of the visible and near-infrared electromagnetic spectrum. Although geometrical processing of these film frames was performed for decades in an analogue – and later analytical – way, they are normally scanned now to enable a digital processing of the aerial image.

To date, most airborne imaging devices provide digital products directly since the detection is usually accomplished by the conversion of incoming electromagnetic radiation (expressed as *at-sensor radiance*) into an electrical output signal which is subsequently digitised into digital numbers (*DNs*). Most digital image capture systems comprise optical elements such as lenses, mirrors, prisms, gratings and filters that gather the radiation and focus it onto an imaging sensor. This imaging sensor itself consists of several (often millions) of individual optical detectors (also called *photodetectors* – Norton 2010) that can detect the incoming radiation and generate a signal in response to it (Verhoeven 2012a). In this chapter, all imaging sensors are considered to be frame sensors, since they consist of an array of individual photodetectors arranged in a rectangular frame. Moreover, they are assumed to work in the optical radiation spectrum, commonly accepted to reach from the ultraviolet to the infrared (Ohno 2006; Palmer and Grant 2010). Additionally, for the remainder of this chapter, image and photograph are assumed to mean digital image.

Whether they are generated by scanning the analogue film frame or directly produced by the digital imaging sensor, the fundamental building blocks of any digital image are called pixels or pels, coined terms for picture elements (see Billingsley (1965) and Schreiber (1967), respectively, for the first use of these terms). In the case of a digital imaging sensor, each photodetector commonly produces one pixel. An array of pixels is called a digital image, which can be mathematically represented as an $M \times N$ matrix of numbers, M and N indicating the image dimensions in pixels. Pixels are thus determined by a pair of pixel coordinates (r, c indicating row and column) and a certain value or grouping of values that contains information about its measured physical quantity (Smith 1997). Just as a pixel of a common digital colour photograph contains three samples or DNs at the same location to represent the amount of radiation captured in three individual spectral bands, a greyscale image consists of one DN per pixel. Images can thus be represented by O matrices of $M \times N$ elements, in which O equals the amount of spectral bands that are sampled (Bernstein 1983). Every image is also characterised by a certain bit depth, which determines the resolution by which the amplitudes of the continuous analogue radiation signal can be mapped onto a discrete set of digital values. Consider an 8 megapixel digital image, 4,000 pixels in width and 2,000 pixels in height. If the image is an 8-bit greyscale image, every pixel has an integer DN between 0 and 255. 16-bit integer pixels could contain values between 0 and 65 535. Digital images are thus said to be sampled (spatially, spectrally and temporally) and quantised (radiometrically, defined by the number of bits) representations of a scene, defined by a multidimensional matrix of numbers.

However, the analogue real-world signal (in the form of electromagnetic radiation arriving at the imaging sensor) is degraded in various ways. As a result, the final digital image is never a faithful reproduction of the real-world scene. Aside from the spectral and radiometric transformations that occur, the geometric three-dimensional (3D) properties are mapped to a two-dimensional (2D) plane (Fig. 3.2). This mapping result (i.e. the final image) is influenced by a wide variety of factors such as earth curvature, film and paper shrinkage, nonplanar image film plane, atmospheric refraction effects, optical distortions, tilt and relief displacements (Imhof and Doolittle 1966). Not only

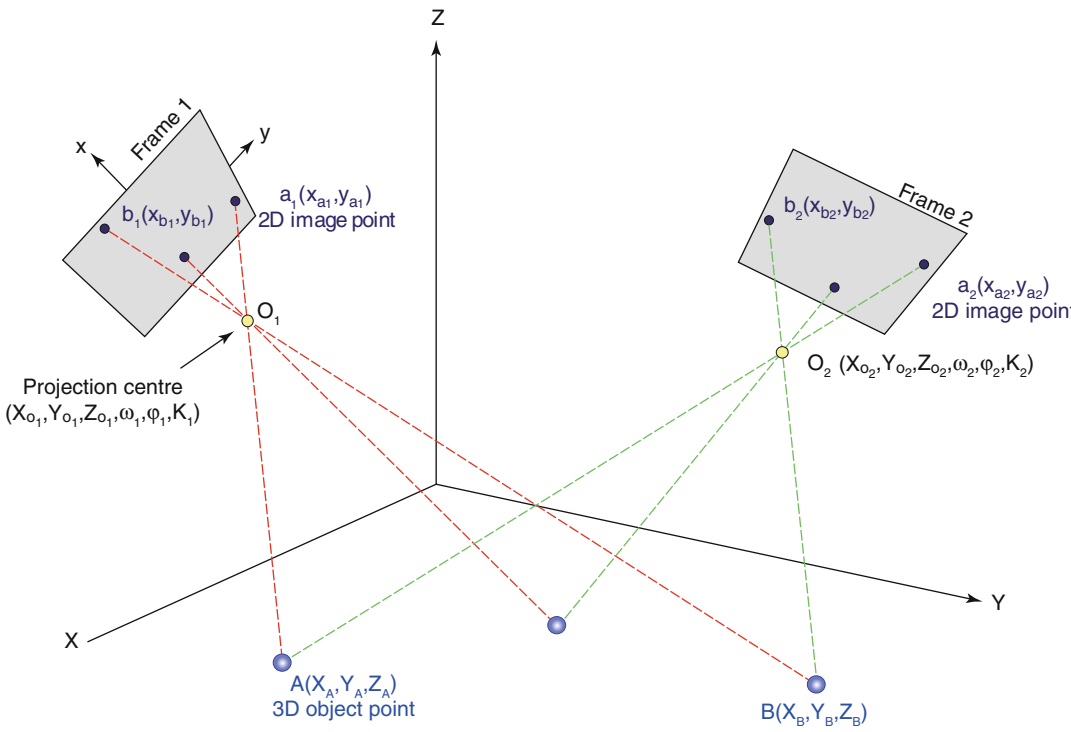

Fig. 3.2 Mapping of 3D object points onto 2D points in two aerial frame images

does every individual aerial image suffer from these geometrical deformations, but they also vary from frame to frame due to variations in the flying height and camera tilt. Compensating for them through some kind of geometric correction is essential for accurate mapping and information extraction. Since the geometric errors induced by the optics, the topographical relief and the tilt of the camera axis contribute most to image deformations; they will be shortly reviewed below.

3.2.2 Optical Distortions

In photogrammetry and computer vision, the geometry of central perspective projection is used to model the formation of an image mathematically (Mundy and Zisserman 1992; Buchanan 1993). In the field of photogrammetry, this is expressed by the *collinearity equation* which states that the object point, the camera's projection centre and the image point are located on a straight line and the image is formed on an exact plane (Fig. 3.2). Lens distortions (radial and decentring), atmospheric effects (mainly refraction) and a nonplanar image sensor are factors which prevent this. Since digital image sensors are by default treated as perfectly planar surfaces (Wolf and Dewitt 2000) and refraction is a very specific topic that is only of major importance when imaging from rather high altitudes and off-nadir angles (Hallert 1960; Gyer 1996), only lens distortions will be considered here.

In the case of an ideal camera, which would be a perfect central projection system in which projection implies a transformation of a higher-dimensional 3D object space into a lower-dimensional 2D image space (Mikhail et al. 2001), the lens imaging system would be geometrically distortionless (Billingsley et al. 1983). The mathematical parameters describing this ideal situation are the principal distance and the principal point (forming the so-called *interior/inner orientation*; see below). However,

since optical distortions are always present in real cameras, the image points are imaged slightly off of the location they should be at according to the central projection. To metrically work with airborne images, every image point must be reconstructed to its location according to this ideal projective camera (Gruner et al. 1966). Therefore, the deviations from the perfect situation are modelled by suitable distortion parameters, which complete the interior orientation. All the parameters of the interior orientation (also called *camera intrinsics*) are determined by a *geometric camera calibration* procedure (Sewell et al. 1966). After this geometric camera calibration, all parameters that allow for the building of a model that can reconstruct all image points at their ideal position are obtained, thereby fulfilling the basic assumption used in the collinearity condition. More specifically, the main elements of interior orientation which camera calibration should determine are the following:

- *Principal distance* (*PD*): the distance measured along the optical axis from the perspective centre of the lens (more exactly the rear nodal point of the optical system) to the image plane (more exactly the principal point of the image) (Mikhail et al. 2001). When the camera is focused at infinity, this value equals the focal length f of the lens (Wolf and Dewitt 2000). For close-range focusing this is no longer the case and the principal distance will increase. This means that any change in focus or zoom produces a new calibration state. In aerial mapping cameras applied for vertical surveys, the calibrated focal length f_c is often given, which equals the principal distance that produces an overall mean distribution of lens distortion (Slater et al. 1983).
- *The location of the principal point* (x_p, y_p): this is the second essential quantity to adequately define the internal camera geometry. It can be defined as the intersection of the optical axis of the lens system with the focal plane (Mikhail et al. 2001). This means that the location of the principal point can change with different zoom settings, but it will always be close to the image centre. In an ideal camera the principal point location would coincide with the origin of the image coordinate system.
- *Radial lens distortion parameters* (k_1, k_2, k_3, k_4): in optics, distortion is a particular lens aberration, but one that does not reduce the resolution of an image (Gruner et al. 1966; Slater et al. 1983). Radial lens distortion is the central symmetrical component of lens distortion and occurs along radial lines from the principal point. Although the amount may be very small in aerial mapping cameras, this type of distortion is unavoidable (Brown 1956). In consumer lenses, radial distortions are usually quite significant. Generally, one to four k parameters are provided to describe this type of distortion. Radial distortion can have both positive (outward, away from the principal point) and negative (inward) values. Negative radial distortion is denoted as pincushion distortion (since an imaged square will appear to have its sides bow inward), while positive distortion is termed barrel distortion (because straight lines bow outward) (Gruner et al. 1966). Either positive or negative radial distortion may change with image height (Fig. 3.3), and its amount is also affected by the magnification at which the lens is used. It can also occur that one lens system suffers from both negative and positive distortion (Kraus 2007). Figure 3.3 depicts a typical distortion curve. On the left, the distortion scale is indicated in micrometres. In the graph, the distortion is plotted as a function of the radial distance r from the principal point.
- *Decentring lens distortion parameters* (p_1, p_2): this distortion can be broken down into asymmetric radial distortion and tangential lens distortion. Both distortions are caused by imperfections in the manufacture and alignment of individual lens elements during the construction of the lens (Brown 1966). Their magnitude is typically much smaller than that of radial lens distortion (Fig. 3.3) and conventionally described by two parameters p_1 and p_2 (Burnside 1985). Although it is generally not significant in aerial mapping lenses, decentring distortion is common in commercial lenses with variable focus or zoom.

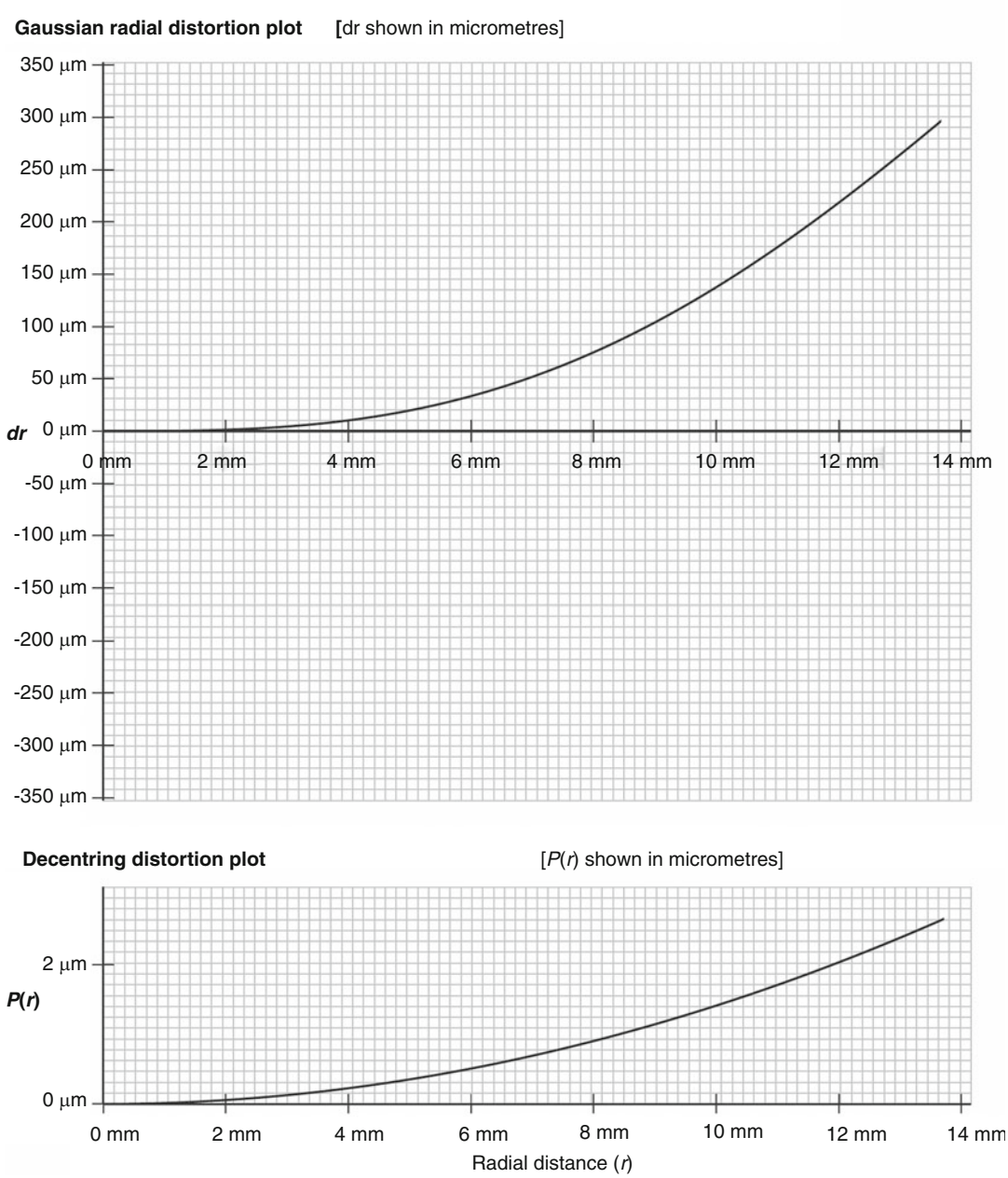

Fig. 3.3 Radial and decentring distortion plots of the AF Nikkor 24 mm f/2.8D (infinity focus). The radial distortion *dr* (expressed in micrometres) and decentring distortion $P(r)$ are given as a function of radial distance r (mm)

In addition to the abovementioned parameters, several other camera characteristics can be calibrated: affinity in the image plane (consisting of aspect ratio (or squeeze) and skew (or shear)), unflatness of the film plane and the coordinates of the fiducial marks. The latter are used in analogue systems and provide a coordinate reference for the principal point and all image points, while also allowing for the correction of film distortion (Kraus 2007). Calibrating a digital frame camera is in many ways more straightforward than calibrating film cameras, since the individual sensor

photodetectors are essentially fixed in position, which practically eliminates film distortion considerations (Wolf and Dewitt 2000). Fiducials are therefore not needed in digital cameras (Graham and Koh 2002). Moreover, zero skew (i.e. perpendicular axis) and a unit aspect ratio (i.e. photodetector width to height equals 1) can be assumed for digital frame cameras as well (Remondino and Fraser 2006; Xu et al. 2000; Szeliski 2011).

From the previous paragraphs, it should now be obvious that the nonmetric cameras conventionally used in archaeological oblique aerial reconnaissance are characterised by an adjustable principal distance, varying principal point and high-distortion lenses, while lacking film flattening and fiducial marks (in the case of analogue devices). Finally, it can be mentioned that there exists a wide variety of digital camera (auto-) calibration methods (see Remondino and Fraser (2006) for an overview). Although exceptions exist, the calibration methods applied in photogrammetry are tailored towards high accuracy and try to recover at least ten interior orientation parameters. Current computer vision methods (see Sect. 3.3) generally use camera models described by only four to five interior orientation parameters.

3.2.3 Tilt Displacement

A camera is placed at a certain location in space (in the air or on the ground) and is pointed in a certain direction. The location defines the projection centre O with three coordinates (X_O, Y_O, Z_O), and the direction is defined by three rotation angles roll, pitch and yaw (ω, φ, κ). Together, these six parameters establish the so-called *exterior/outer orientation* (Fig. 3.2) (Kraus 2007). Other terms for that are *camera extrinsics* or simply *pose*. Together with the interior orientation the position of the image is unequivocally defined. During a vertical photography flight, φ and ω are near to zero. When they equal zero, the result is a perfect nadir/vertical photo that does not need any correction for tilt displacement. The more tilted the photographic axis with respect to the ground surface, the more corrections need to be dialled in (Tewinkel et al. 1966).

These effects may be illustrated most clearly by considering the appearance of a regular grid and a circle on a completely flat terrain in both a vertical and a tilted photograph (Fig. 3.4, lens distortions are excluded for the sake of illustration). A vertical optical axis images the circle as a circle, while the net of squares remains unaltered as well. The same features photographed with a non-zero angle of tilt result in a distorted square net as well as an ellipse-like feature. The difficulty inherent to tilt displacements is the fact that it is often hard to detect while it yields constantly varying scale changes across the image (Dickinson 1969). When dealing with vertical photographs, there is just one nominal scale S that can be calculated by $S = PD/H$ (i.e. the ratio of the principal distance to the flying height H above the terrain) (Tewinkel et al. 1966). In this case, the scale is completely independent of the measurement direction. For tilted images, the scale will vary with direction (Estes et al. 1983). In the background of a tilted photograph, the scale is smaller than the scale in the foreground. The projective transformation of a tilted aerial image to a horizontal plane to remove these tilt displacements (and thus scale differences) is called (*planar*) *rectification* (Spurr 1960; Altenhofen and Hedden 1966; Dickinson 1969).

For convenience, the tilt in Fig. 3.4 is considered to be acting only along the direction of flight (φ). In practice, tilt will act in random directions due to a combination of non-zero φ and ω angles and rectification will be needed to correct for these displacements. That is why rectification is also said to transform an oblique aerial photograph to an equivalent vertical image (Wolf and Dewitt 2000). However, the rectified image will only be completely identical to the vertical image geometry in absence of lens distortions and perfect flatness of the imaged scene, since any terrain undulation will cause so-called relief (or topographic/elevation) displacements and those even affect perfect nadir images.

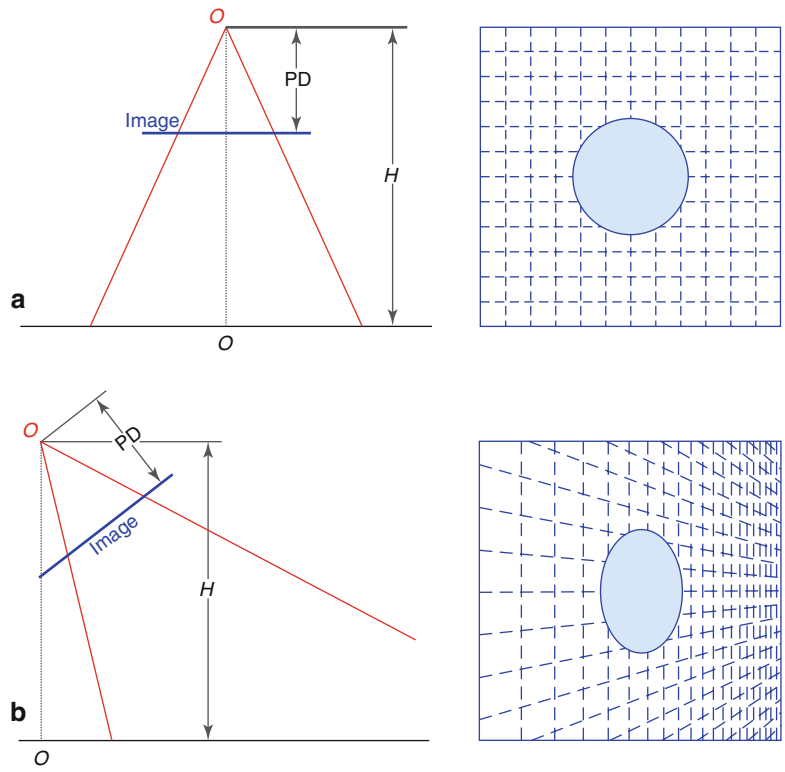

Fig. 3.4 (**a**) Vertical image, (**b**) Oblique image with resulting tilt displacement (*O* denotes the projection centre, *o* the nadir or plumb point, *PD* the principal distance and *H* the flying height above the terrain. The camera's field of view is indicated by the *red lines*.)

3.2.4 Relief Displacement

Image displacements are not only caused by tilt. Any (even tilt-free) aerial photograph will contain displacements due to topographic relief and other height differences (Tewinkel et al. 1966). Thus any feature lying below or above the horizontal reference surface will be misplaced in a planar rectification (Estes et al. 1983) due to the central perspective of the air photo (Hallert 1960).

In Fig. 3.5, the acquisition of a perfect vertical photograph is depicted. KK' indicates the average terrain height but can also be seen as any reference horizontal plane (called a datum surface). On the right, a green tower is shown. If the left top of this tower was to be depicted in a map, the orthogonal projection used to create maps would make it fall in point z, the same point which indicates the foot of the tower. In the aerial image, one also notices point z. Nevertheless, due to the central projection, the top is depicted in z' instead of point z. Consequently, the top of the tower has undergone a displacement of magnitude p'', resulting in a tower whose side is visible in the aerial image.

Although it may not be as visually obvious as in the case of buildings, imaged relief also suffers from this (Falkner and Morgan 2002). Consider the hill in the middle of Fig. 3.5. The top y' should normally be projected in point y, like on a map. However, in this case the projection also causes a displacement p' and instead of being depicted in y, the top is projected onto y' in the image. Following the same principle, the valley on the left also suffers from relief displacement (of magnitude p). In this case, it is not a displacement away from the centre, but towards it. Without regard to direction, this distance of such displacement is called *parallax*. In this respect, parallax gives a numeric value for the relief or topographic displacement.

Although this phenomenon complicates the mapping and interpretation of aerial imagery, it also enables humans to perceive three dimensions and calculate the height of objects from images (Spurr 1960). As the location of the nadir point does not suffer from this displacement

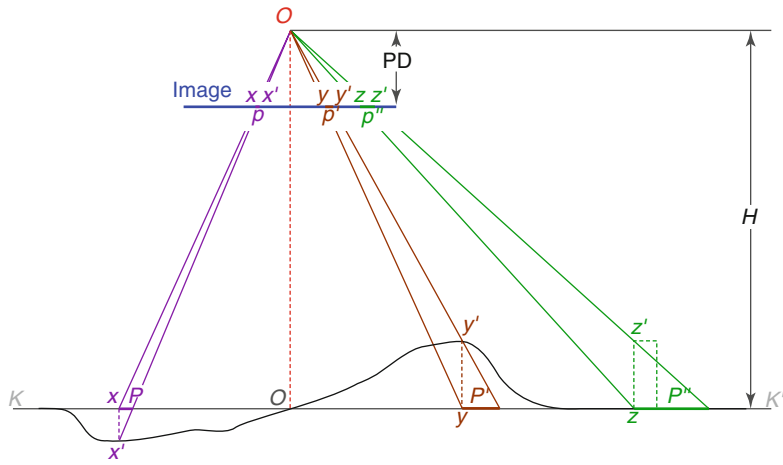

Fig. 3.5 The phenomenon of relief displacement and how it influences the geometry of a vertical image (*symbols* are explained in the main text)

(because its projection is a perfect orthogonal projection), relief displacement is always radial from the nadir or plumb point *o*. This is determined by the intersection of a vertical, constructed from the optical centre *O* towards the ground, and the image plane; this vertical axis is equal to the optical axis of the whole system in the case of a perfect vertical photograph – such as Fig. 3.5 (Tewinkel et al. 1966).

Geometric correction aims to compensate for most of these deformations. The result of such a correction must be an image with a geometric integrity like a map, i.e. an orthogonal projection to the horizontal reference plane. Just as rectification denotes the process of removing tilt from a photograph, relief displacements and other geometrical deformations (such as optical distortions) can be corrected through the process of *orthorectification* or *differential rectification* (Hassett et al. 1966; Turpin et al. 1966; Wolf and Dewitt 2000).

3.2.5 Georeferencing and Geometric Correction

Aerial photography provides a basis for gathering spatial data. Before archaeological information can be extracted from these sources in a way that is useful for mapping and further analysis, the aerial images must be *georeferenced* in an absolute manner. This process, which is also known as *ground registration*, assigns spatial information to any kind of spatial data (raster data such as imagery as well as vector data) to explicitly define their location and rotation in respect to a specific Earth-related coordinate frame.

Often, the geometry of these data is already corrected for any possible deformation. However, the process of georeferencing is often applied to geometrically distorted data as well. Although it is *sensu stricto* not covered by its definition, georeferencing can thus also involve the necessary steps to remove the optical distortions as well as tilt and relief displacements of the aerial image in order to place each image pixel on its true location on the Earth's surface. To do this, a wide variety of approaches and software solutions exist. In many cases, archaeologists fit tilted images to a flat surface by means of a projective transformation, a process introduced in the previous sections and denoted (planar) rectification (Hallert 1960; Altenhofen and Hedden 1966; Wolf and Dewitt 2000). Although these rectified images no longer suffer from tilt displacements, they still contain scale variations and displacements due to topographic relief (hills, buildings etc.). Consequently, projective transformations can only be considered 'archaeologically sufficient' when dealing with completely flat areas. If the aerial view suffers from relief displacements, georeferencing often employs polynomial corrections, spline algorithms or piecewise affine warpings embedded in archaeologically dedicated tools such as AirPhoto SE (Scollar 2002) and AERIAL (Haigh 2005). Although these approaches are very popular and might deliver

fairly good metrical information when the terrain variations are quite moderate, the methods are often suboptimal because they do not (or only partly) eliminate all the image displacements, the distortion of the optics and – to a lesser extent – the atmospheric refraction. Consequently, this image georeferencing is well suited for rather small-scale mapping but inadequate for a detailed multi-temporal and multi-method analysis.

When one needs to mosaic several multi-temporal aerial observations into an extensive overall view of an archaeological region – hence serving as a basic information layer for further prospection and excavation, protection measures and heritage management – the aforementioned issues need to be dealt with. Therefore, planimetrically correct true orthophotographs are of the utmost value. However, these can only be achieved when more advanced ortho-correction approaches embedded in programs such as Leica Photogrammetry Suite or Trimble INPHO photogrammetric system are utilised. Although these more expensive packages offer rigorous orthorectification algorithms to produce superior geometric quality, they are limited by the fact that photogrammetric skills, interior orientation parameters and an accurate, high-resolution digital surface model (DSM) are essential, three conditions that are generally not met in aerial archaeology.

Irrespective of the method applied, the georeferencing of (individual) images is commonly determined with ground control points (GCPs), whose manual measurement and identification is a time-consuming operation that requires experience while being bound to certain prerequisites. As a result of all these issues, many archaeologically valuable aerial images never get properly georeferenced and stay hidden on local hard drives or in image archives.

3.3 A New Workflow

Since a variety of factors contribute to image deformation, imagery needs to be geometrically corrected in order to correspond as closely as possible to a map. At the same time, the workflow should be as straightforward and generally applicable as possible. Currently, cost-effective means are available for orthorectification of a wide variety of (archaeological) aerial frame imagery. These became possible due to the ever increasing technological improvements in computer hardware and the serious advances made the past 15 years in the scientific field of computer vision, which is often defined as the science that develops mathematical techniques to recover information from images. This image data can take many forms, such as multidimensional imagery from medical scanners, stereo photographs, video sequences or views from multiple still cameras. Initially, many computer vision applications were focused on robotic vision and inspection. As a result, the methods were characterised by few constraints and focused on a high degree of automation rather than the accuracy and reliability characteristic of photogrammetry (Remondino et al. 2012). However, the last decade has witnessed a shift of focus to more accurate 3D visualisations and virtual reality, along with many new insights in the geometry of multiple images (see Faugeras et al. 2001 or Hartley and Zisserman 2003 for a good overview).

Using techniques such as triangulation, an image point occurring in at least two views can be reconstructed in 3D (Fig. 3.2). However, this requires the knowledge of the interior and exterior orientations of the images. In computer vision, these orientation parameters are usually combined in the so-called *projection matrices* of the images (Robertson and Cipolla 2009), which can be determined by an approach called *structure from motion* (*SfM*; Ullman 1979). During this approach the relative projection geometry of the images is computed along with a set of 3D points that represent the scene's structure. SfM only requires corresponding image features occurring in a series of overlapping photographs captured by a camera moving around the scene (Fisher et al. 2005; Quan 2010; Szeliski 2011). Sometimes, this approach is also referred to as *structure and motion* (*SaM*), since both the structure of the scene and the motion of the camera (i.e. the different camera positions during image acquisition) are recovered.

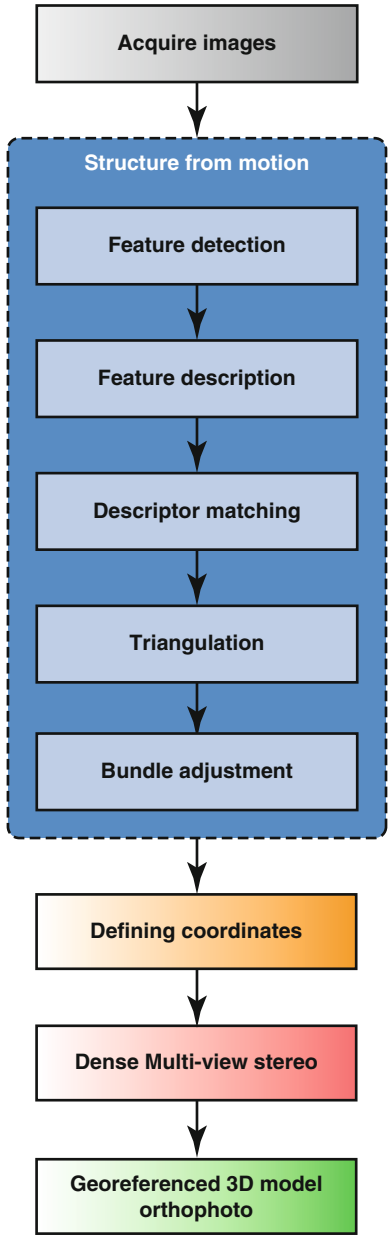

Fig. 3.6 The individual steps of the SfM + MVS processing pipeline (terminology is explained in the text)

In order to achieve this, SfM relies on algorithms that detect and describe local features for each image and then match those 2D points throughout the multiple images. Using this set of correspondences as input, SfM computes the locations of those interest points in a local coordinate frame (also called model space) and produces a sparse 3D point cloud that represents the geometry/structure of the scene. As mentioned previously, the camera pose and internal camera parameters are also retrieved (Hartley and Zisserman 2003; Szeliski 2011). Below, the SfM approach and the individual steps (Fig. 3.6) essential for its execution are outlined in greater detail. Afterward, some details are given about the subsequent process, *multi-view stereo* (*MVS*), as this last stage uses the SfM output to generate a dense 3D model needed for accurate image orthorectification.

3.3.1 SfM + MVS Pipeline

3.3.1.1 Image Acquisition

For an SfM + MVS approach, it does not matter if the images are acquired with a metric camera or not, or whether they are shot in a vertical or oblique pose. Attention should, however, be paid to the angular separation of images in order to ensure that it is not too large. This will maximise the likelihood that a stable image network can be achieved. Although several feature point extraction algorithms (see the next part) with particular strengths and weaknesses have since been developed, Moreels-Perona found out that no detector/descriptor combination performs well with viewpoint changes of more than 25–30° (Moreels and Perona 2007). Therefore, a sufficient image overlap is advised (around 60–80 % for vertical images), and it is preferable for every image to be captured from a unique location. Panning from the same location should thus be avoided (Tingdahl et al. 2012). Moreover, the objects being photographed need to possess sufficient unique texture. In general, all these assumptions can be met in aerial archaeological imaging.

Once the images are acquired, the second stage of the pipeline can be executed. This stage is denoted the SfM algorithm and consists of several individual processing steps (some authors consider only the last two steps in this stage as the SfM algorithm, but Fig. 3.6 groups all these individual computing steps into one SfM stage). For the sake of clarity, all the individual steps will be defined below.

3.3.1.2 Feature Detection

Feature detection is the first step of many computer vision and photogrammetry-related applications, such as panorama stitching, object recognition, camera calibration, robot localisation and SfM. In past decades, a wide variety of feature detectors have been developed. Aside from their effectiveness, they vary widely in computational complexity and the type of features they detect. Although approaches exist that detect edges, ridges and regions of interest (e.g. Kadir and Brady 2001; Jurie and Schmid 2004; Matas et al. 2004; Deng et al. 2007), the image features used in most SfM approaches comprise interest points (IPs).

IPs represent image locations that are in a certain way exceptional and are locally surrounded by distinctive texture. Additionally, they should be stably defined in the image and scale spaces and reproducible under different imaging conditions. In technical jargon, it is said that IPs should have a high *repeatability*, which means that they should be invariant to any change in illumination, image noise and basic geometric transformations such as scaling, translation, shearing and rotation. In the last 10 years, several new algorithms have been proposed to compute such IPs (e.g. Features from Accelerated Segment Test or FAST (Rosten and Drummond 2005)). However, most detector techniques are based on:

- Hessian-based detectors (Lindeberg 1998)
- Harris-based detectors (Harris and Stephens 1988)

This means that frequently mentioned algorithms such as SIFT (Scale Invariant Feature Transform (Lowe 2004)), SURF (Speeded-Up Robust Features (Bay et al. 2006, 2008)) and ASIFT (Affine-SIFT (Morel and Yu 2009; Yu and Morel 2011)) use variants of the abovementioned detectors (the popular SIFT and SURF detectors both rely on Hessian-based detectors). Figure 3.7a shows IPs computed with SURF. The airborne image in the figure was acquired on the 4th of September 2012 at around 11.00 h using an Olympus PEN E-P2 (a 12.3 megapixel mirrorless Micro Four Thirds camera) equipped with an Olympus M. Zuiko Digital 17 mm f/2.8 lens, mounted on a radio-controlled Microdrone MD4-1000 quadcopter. The aerial frame depicts a part of the excavated Roman city wall of *Carnuntum* (Austria).

3.3.1.3 Feature Description

Since the aim is to find correspondences between these IPs – which means that an algorithm has to find out which IPs are a 2D representation of the same physical 3D point – the IPs have to be described. This task is fulfilled by so-called *feature descriptors* or *feature vectors*. Such a descriptor computes a feature vector with local characteristics to describe a local patch (whose size can vary – Fig. 3.7b) of pixels around each IP (Schmid and Mohr 1996). Just as the IP, this vector should be invariant (i.e. robust to detection displacements, image noise and photometric plus geometric deformations). Various methods also exist to describe the patch around each IP:

- Gradient Location and Orientation Histogram (GLOH) (Mikolajczyk and Schmid 2005)
- Speeded-Up Robust Features (SURF) (Bay et al. 2006, 2008)
- Scale Invariant Feature Transform (SIFT) (Lowe 2004)
- Local Energy based Shape Histogram (LESH) (Sarfraz and Hellwich 2008)
- ASIFT (Affine-SIFT) (Morel and Yu 2009; Yu and Morel 2011)
- Histogram of Oriented Gradients (HOG) (Dalal and Triggs 2005)

In the end, an image feature can be defined as an IP and its descriptor. Note that several IP detectors also define their descriptor (e.g. SIFT, SURF, ASIFT). As can be expected, several authors have tried to compare the performance of various detector and descriptor combinations (e.g. Mikolajczyk and Schmid 2003, 2005; Mikolajczyk et al. 2005; Moreels and Perona 2007; Tuytelaars and Mikolajczyk 2007; Juan and Gwon 2009).

3.3.1.4 Descriptor Matching and Pairwise Image Orientation (Fundamental Matrices)

Finally, all descriptor vectors are matched between different images by associating each IP from one image to the other IPs of the remaining

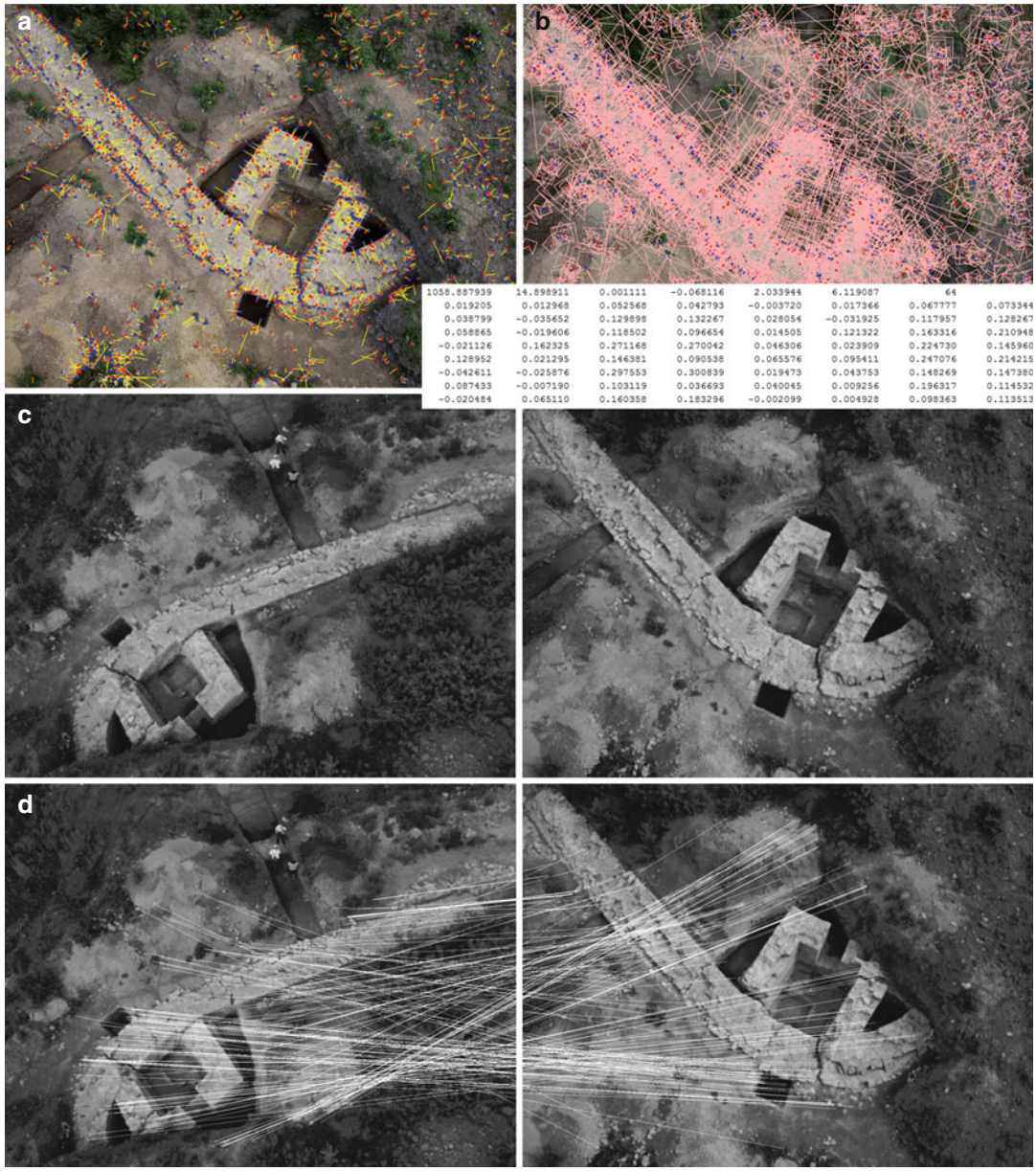

Fig. 3.7 (**a**) SURF IPs computed from an airborne image using ImageJ SURF (Labun 2009). The 1,852 IPs are accompanied by their orientation vectors whose lengths indicate the strength of the computed IPs. (b) The 1,852 SURF IPs with their descriptor windows. The inset shows the vector describing one of these IPs. Computations were performed with ImageJ SURF (Labun 2009). (**c, d**) The difference between two image matching routines. While (**c**) used the SIFT detector and was unable to find any matching points, ASIFT was applied for (**d**). This test was performed using the ASIFT online demo application at http://demo.ipol.im/demo/my_affine_sift/ (Yu and Morel 2011)

images. To compute a match, a distance between the descriptors is generally used (e.g. the Euclidean distance). The dimension of the descriptor has a direct impact on the time this takes, and fewer dimensions are desirable for fast IP matching. However, lower-dimensional descriptor vectors are generally less distinctive than their high-dimensional counterparts. Besides,

approximate but fast methods exist (e.g. approximate nearest neighbour searches in kd-trees), while slow but rigorous matching procedures such as quadratic matching can also be applied.

A robust outlier detection algorithm such as RANSAC (RANdom SAmple Consensus (Fischler and Bolles 1981)), ORSA (Optimized Random Sampling Algorithm (Moisan and Stival 2004)), LMedS (Least Median of Squares (Rousseeuw 1984)) or MAPSAC (Maximum A Posteriori SAmple Consensus (Torr 2002)) will ensure the rejection of probable false matches by testing them for consistency. This is done for all possible image pairs by checking if their putative matches fulfil the so-called *epipolar geometry constraint*: i.e. that the displacements of IPs are a possible result solely of the motion of the camera between both images. At the end of this process, the *fundamental matrices F* of the image pairs are obtained: each of them is a 3×3 matrix depending on seven parameters that describes the motion (i.e. *relative orientation*) from the first to the second image. When dealing with calibrated cameras or pinhole camera models, the *essential matrix E* is used; in case of an image triplet, the trifocal tensor *T* can be applied (Robertson and Cipolla 2009). Because the fundamental matrix describes the correspondences in more general terms, it is used with uncalibrated cameras. This has very important implications as the matching can be performed without initially calibrating the cameras. Finally, the complete set of image correspondences (called tie points in photogrammetry) for the whole image sequence is obtained after considering all meaningful image pairs. The set of corresponding IPs thus obtained functions, together with the fundamental matrices, as input for the last steps of the SfM computation.

Figure 3.7c–d shows, however, that this input varies widely according to the algorithms applied to obtain this set of image correspondences. The differences between two image matching routines are illustrated, both of them trying to reliably identify and match two aerial images. In Fig. 3.7c, the SIFT detector is used while Fig. 3.7b uses the ASIFT approach. All IPs are then coded with the SIFT descriptor. The matching process first computes the Euclidean distance between an IP descriptor in the first image with all the descriptors found in the second image and uses its values to define whether IPs are considered as matched. Afterward, the ORSA algorithm is applied to filter out the false matches. The example shows that ASIFT retrieves the matches – indicated by the white lines – even under large changes of viewing angle, while there is a total failure in finding image correspondences using SIFT IPs. This is due to the nature of the algorithms used. While SIFT can only deal with a similarity invariance (i.e. invariant to four parameters describing translation, rotation and zoom) and less viewpoint change from one image to another, ASIFT is fully affine invariant. This means that ASIFT possesses invariance for the four similarity degrees of freedom as well as for the two angles defining the camera axis orientation. To achieve this, it simulates rotation and tilt on the images and can therefore deal with frames whose viewing angle is very different (Morel and Yu 2009; Yu and Morel 2011).

3.3.1.5 Triangulation

Relying on the algorithms that detect, describe and match local feature points throughout the multiple images, SfM computes the locations of those feature points in a local coordinate frame, creating a sparse 3D point cloud that represents the geometry/structure of the scene. This determination of a point's 3D position when observed from two or more cameras (Fig. 3.2) is called *image triangulation* (Szeliski 2011). However, image triangulation requires the knowledge of the images' interior and exterior orientation. These are obtained after combining all the relative orientations of the image pairs in form of their fundamental matrices.

SfM can accomplish this as it is based on the *projective reconstruction theorem*, which states: given a set of point correspondences in two views defined by the fundamental matrix, the 3D scene geometry and images' projection matrices (which comprise all the orientation parameters) may be reconstructed from these correspondences alone, and any two such reconstructions from these correspondences are projectively equivalent (Hartley 1994; Szeliski 2011). However, rather than a

projective reconstruction, a metric reconstruction is wanted: i.e. one in which orthogonal planes are at right angles, parallel lines stay parallel and the reconstructed 3D model is a scaled version of reality. This can be accomplished by running a simultaneous *self-calibration/auto-calibration* to define the camera's interior orientation. The latter is stored for each image in the *intrinsic parameter matrix K* (Hartley and Zisserman 2003; Moons et al. 2008).

3.3.1.6 Bundle Adjustment

Up to now all images were dealt with in pairs, for each of which a fundamental matrix was computed (in a linear way by minimising a physically non-meaningful quantity – the so-called *algebraic error*). Afterwards the oriented image pairs were combined to form the complete block of images and to yield the structure of the scene. The results obtained this way are, however, sub-optimal because not all overlapping images are used at the same time and the discrepancies in the structure (caused by small errors during the feature measurement phase) are not optimally minimised. To overcome these problems, the final stage of most SfM algorithms is bundle adjustment. Bundle adjustment iteratively optimises the 3D structure and the projection matrices of all images simultaneously by performing a robust non-linear minimisation of the actual measurement errors, also known as *re-projection errors* (Triggs et al. 2000). The technique was developed half a century ago in the field of photogrammetry but is now also largely applied in the computer vision community. The term bundle adjustment comes from the fact that the bundles of rays connecting camera/projection centres to 3D scene points are adjusted to minimise the sum of squared differences between the observed and re-projected image points (Szeliski 2011).

This means that an SfM approach can recover the scene structure and camera projection matrices from image correspondences alone without prior knowledge about camera poses or interior orientation (Hartley and Zisserman 2003; Szeliski 2011). There is thus no real need to use calibrated cameras and optics during the image acquisition stage (Quan 2010), which makes the procedure very flexible and well suited for almost any kind of imagery, particularly for completely unordered photo collections such as those that can often be found in aerial archives. It needs to be noted, though, that it is still more accurate to recover the significant interior orientation parameters in a separate calibration routine using a dedicated image network geometry (Remondino and Fraser 2006).

3.3.1.7 Defining a Coordinate Reference System

It is essential to understand that the SfM output is characterised by a scale ambiguity. This means that if the entire scene is scaled by some factor and the distance between the camera positions is simultaneously scaled by the same scale factor, the projections of the scene points in the image will remain exactly the same. The reconstructed 3D scene obtained after a standard SfM approach is thus expressed in a local coordinate framework and equivalent to the real-world scene up to a global scaling, rotation and translation. These parameters can only be recovered via the use of additional data, which in turn define a coordinate reference system (CRS). According to Barazzetti et al. (2011), this can be achieved in two ways:

- Import at least three spatially well-distributed GCPs with known altitude values and transform the complete model into an absolute CRS with a Helmert similarity transformation. Although more GCPs are advisable, three is the minimum since seven parameters (three translations, one scale and three rotations) must be determined for this spatial transformation. Since this operation is performed after the SfM computation and does not introduce any external constraint, it will not improve the initially obtained SfM result.
- Import highly accurate camera positions or a minimum of three GCPs and use them as constraints in the bundle adjustment. This rigorous approach is a better solution as it can correct for errors such as drift in the recovered camera and point locations (Snavely et al. 2006), avoids instability of the bundle solution (Remondino et al. 2012) while the SfM output is directly georeferenced (Verhoeven et al. 2012a).

3.3.1.8 Dense Multi-view Stereo (MVS)

At this stage a georeferenced sparse 3D reconstruction of the scene is available. 'Sparse' because it is only based on the reconstructed set of IPs. However, with the now known orientation of the images, it becomes possible to create a texture-mapped dense 3D model and compute orthophotographs. The essential step in this process is the computation of this denser 3D model. Alternatively, one could interpolate the sparse set of 3D points, but this would yield a far from optimal result. Therefore, it is better to run a multi-view stereo (MVS) algorithm to compute a dense estimate of the surface geometry of the observed scene. Because these solutions operate on pixel values instead of on feature points (Scharstein and Szeliski 2002; Seitz et al. 2006), this additional step enables the generation of detailed 3D meshed models (or dense point clouds) from the initially calculated sparse point clouds, hence reproducing fine details present in the scene.

Just as in all previous stages, MVS comes in many variants and a comparison of several approaches can be found in Seitz et al. (2006). However, since the publication of this paper by Seitz and his colleagues, many new algorithms have been developed. Although elaborating on them is outside the scope of this text, it might be worthwhile to notice that the most common algorithms can be divided into region growing patch-based approaches (e.g. Lhuillier and Quan 2005; Habbecke and Kobbelt 2006; Furukawa and Ponce 2010) and depth-map fusion pipelines (e.g. Mellor et al. 1996; Pollefeys et al. 2004; Goesele et al. 2006; Strecha et al. 2006; Bradley et al. 2008; Hirschmüller 2008). Obviously, each of those has its own specific pros and cons, generally striking a balance between accuracy and consistency (region growing approaches) versus a fast and elegant pipeline (depth-map fusion).

3.3.1.9 Georeferenced 3D Model and Orthophoto

The final georeferenced dense 3D model generated from these aerial images can be considered a DSM: a numerical representation of the topography and all its imposed structures such as trees and houses. As is known from conventional orthorectification (Manzer 1996), such a dense DSM is elementary when one wants to generate a so-called true orthophoto in which all objects with a certain height (such as houses, towers and trees) are also accurately positioned (Kraus 2002; Braun 2003). When combined with the previously calculated camera poses and interior orientation parameters, this dense DSM thus enables the generation of true orthophotos. Because the whole process takes most relevant geometrical degradations into account, the orthographic image is perfectly suited for archaeological purposes. For visualisation purposes, one could also export a textured 3D mesh which could be created by a texture mapping using a particular selection of the initial images.

3.3.2 Tools

3.3.2.1 Software

In recent years, SfM has received a great deal of attention due to Bundler (Snavely 2010) and Microsoft's Photosynth (Microsoft Corporation 2010): two SfM implementations that are freely available on the Web. To date, several SfM-based packages can be applied to obtain a (semi-) automated processing pipeline for image-based 3D visualisation. Often, these packages are complemented by an MVS approach (see Table 3.1). An overview of the accuracies that can be obtained in automated image orientation and camera calibration parameters with some of these packages is detailed in Remondino et al. (2012).

3.3.2.2 Hardware

Besides novel algorithms, the routine outlined above exploits some of the technological improvements in hardware configurations. Obviously, high-quality reconstructions with large image files are very resource intensive. All processing should therefore be undertaken on a multicore computer (or computing grid) with a 64-bit operating system and a large amount of RAM. Additionally, the graphics processing unit (GPU) can be considered one of the crucial hardware elements, as a high-performance GPU can greatly shorten processing times. Many

Table 3.1 Some commercial and freely available SfM and MVS packages

Company	Software	Free	SfM	MVS	Web	Orthophoto
Agisoft LLC	PhotoScan standard		X	X		
Agisoft LLC	PhotoScan professional		X	X		X
Matis laboratory (I.G.N.)	Apero	X	X			
Matis laboratory (I.G.N.)	MicMac	X		X		X
University of Washington and Microsoft Corporation	Bundler	X	X			
Microsoft Corporation	PhotoSynth	X	X		X	
University of Washington	VisualSFM	X	X	X		
AutoDesk	123D Catch	X	X	X	X	
KU Leuven	Arc3D	X	X	X	X	
Eos Systems Inc.	PhotoModeler Scanner		X	X		X
University of Illinois and University of Washington	PMVS2	X		X		
3Dflow SRL	3DF Samantha	X	X			
Henri Astre and Microsoft Corporation	PhotoSynth Toolkit	X	X	X		
CTU Prague	CMPMVS	X	X	X	X	X
Acute3D	Smart3DCapture		X	X		

SfM+MVS applications support the OpenCL (Open Computing Language) programming platform and can therefore access the GPU for executing very intensive computing during specific steps in the pipeline, although the steps that can be accelerated depend on the software. Still, better and more optimised algorithms are needed before time-efficient processing of large image sets on standard computers can take place (Verhoeven et al. 2012a).

3.4 Case Studies

SfM-based applications started to find their way into archaeological research about 10–15 years ago (e.g. Pollefeys et al. 1998, 2000, 2001, 2003, 2004; Pollefeys and van Gool 2002; El-Hakim et al. 2003). During the decade that followed, the SfM concept and dense matching techniques made great improvements and became capable of orienting very large datasets and delivering satisfactorily accurate dense 3D models (Barazzetti et al. 2011). Nowadays, an SfM and MVS pipeline can almost be considered a standard tool in many aspects of archaeological research (e.g. Ludvigsen et al. 2006; Lerma et al. 2011; Appetecchia et al. 2012; Bezzi 2012; Forte et al. 2012; Kersten and Lindstaedt 2012; Lo Brutto and Meli 2012; Opitz and Nowlin 2012). Although most of these studies use terrestrial images, there are some papers in which archaeological aerial frame images have also been used (e.g. Doneus et al. 2011; Verhoeven 2011, 2012c; Lo Brutto et al. 2012; Reinhard 2012; Remondino et al. 2011; Scollar and Giradeau-Montaut 2012; Verhoeven et al. 2012a).

The three case studies described below show the potential of this combined SfM+MVS method using diverse imagery (oblique and vertical, old and new, acquired in the visible and near-infrared spectral domain from manned and unmanned platforms) covering a variety of topographic settings. As these image sets predate the development of SfM-based approaches, they provide a perfect opportunity to evaluate the applicability of the method to older datasets. The case studies are presented in a common format: first, a short introduction to the site and the acquisition of the photographs are presented; secondly, the building of the orthophoto and possible drawbacks are addressed; and thirdly, each case study will also highlight some very specific advantages of this approach.

All 3D models and orthophotographs were computed using PhotoScan Professional edition (v. 0.8.1, build 877 and later) from Agisoft LLC. The choice for this software was based on its features, cost and completeness: it is currently the only commercial, frequently updated package

that combines both SfM and MVS algorithms while additionally offering tools for generating orthophotographs, texture mapping and post-processing 3D models (Agisoft LLC 2012). Concerning the MVS stage, PhotoScan uses a pairwise binocular stereo approach to compute a depth estimate for almost every image pixel of each view. Afterward, several dense 3D reconstruction methods are provided, each differing in the way these individual depth maps are merged.

3.4.1 *Trea* (Italy)

Generally, the advised strategy when using PhotoScan is to solve the complex SfM math of as large as possible a set of images, without having to rely on virtual memory. Later, one can 'disable photos' and perform the subsequent dense reconstruction in parts (Verhoeven 2011). Although this approach is meant to tackle limited hardware resources, it opens up a completely new application field for aerial archaeologists. To illustrate this, a time series covering 6 years of aerial research on the Roman town of *Trea* (central Adriatic Italy, 43°19′ 06″ N, 13° 17′′ 31″ E – WGS84) will be used.

In January 2000, Ghent University initiated the Potenza Valley Survey (PVS) project in the central Adriatic Region of Marche. This interdisciplinary geoarchaeological project has mainly been aimed at reconstructing the changing physical and human landscape along the Potenza River, one of Marche's major rivers. Aerial archaeological reconnaissance was identified from the start as one of the main survey techniques to be used due to its cost-effectiveness (Vermeulen 2002, 2004). Along the Potenza Valley lies the former Roman town of *Trea*, located on a hill surrounded by the heavily undulating landscape of the middle Potenza valley. The scene can thus be considered quite complex and the relief displacement in the aerial images very substantial. Although there have been a series of investigations into the character and extent of this city, almost nothing was known about its general layout and organisation before the systematic aerial campaigns of the PVS (see Moscatelli 1985). The survey results now allow for a near complete mapping of the main urban structures of this abandoned Roman city, such as the town defences, the internal street network and the main public and private buildings.

From the 208 images initially selected, 203 were aligned correctly in PhotoScan (Fig. 3.8a). This number is extremely high given the circumstances: a wide variety of cameras and lenses were used during the reconnaissance flights; the land cover varied from bare soil to crops in various phenological states; 39 images only recorded the radiance in the near-infrared (NIR) spectral band (see Verhoeven 2008b, 2012b; Verhoeven et al. 2009b for details on this). Unquestionably, this alignment result was facilitated by the fact that all images still had information about the focal length embedded in the Exif (exchangeable image file format) metadata tags, so that these values could be used to initialise the SfM step. To execute the dense reconstruction stage, a subset of 143 suitable images was used as input. The selection criteria for this were largely based on image scale, scene coverage and sharpness. This does not render the remaining images unusable, however. Once an accurate 3D model of the terrain is generated (Fig. 3.8b), every image or combination of images in the project can be transformed into an orthophoto through the use of the DSM for correction.

This way, it is possible to use only the NIR images (Fig. 3.8c-3) or those that best illustrate the crop marks (Fig. 3.8c-1) or soil mark state (Fig. 3.8c-2) or to generate a bespoke coverage. Not only does this approach speed up the processing of individual images (or related photo sets) considerably, but the final interpretation is more trustworthy as well: due to the heavy undulating nature of the terrain and the very steep slopes bordering the central plateau, most GIS packages and tools specifically developed for archaeological research (such as AERIAL or AirPhoto SE) will typically fail to accurately georeference these images. Although this might not seem to be a big issue when dealing with vague soil marks, the nature of the crop marks (faint and small) as well as the type of site (a complex Roman town with different phases) makes the accurate mapping of

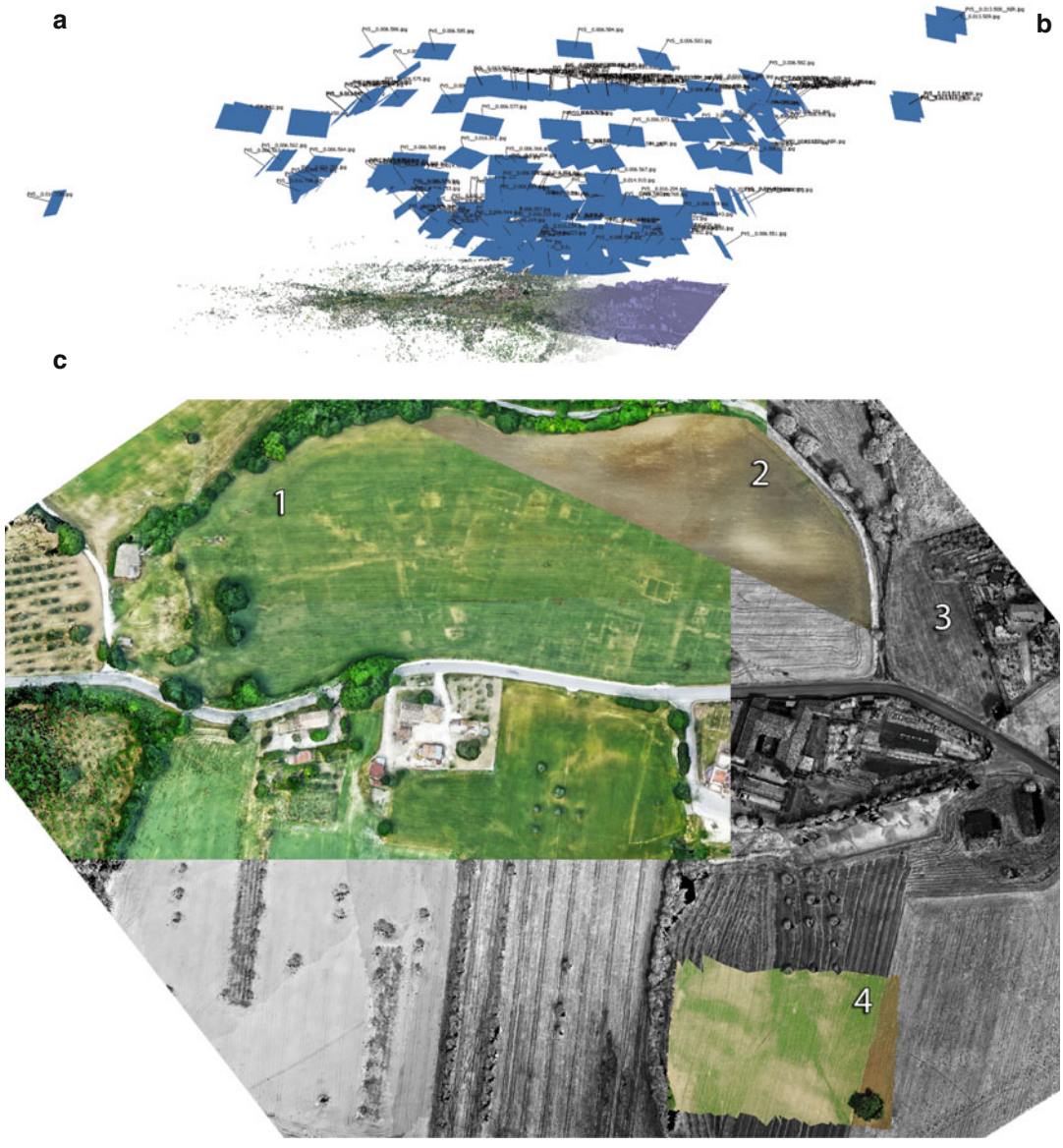

Fig. 3.8 (**a**) The relative position of all 203 camera stations. (**b**) The extracted DSM of *Trea*. (**c**) The integration of several orthophotos, showing crop marks (*1*) and soil marks (*2*) in the visible domain, the NIR terrain reflectance (*3*) and the orthorectification of an image (*4*) without any useable GCP

the features of the utmost importance for comparison of aerial footage from different years or to interpret the data with respect to a geophysical survey (for this case study, the georeferencing delivered a planar RMSE of 6.2 cm and an RMSE of 4.6 cm for the altitude component). Additionally, the whole process of orthophoto production is straightforward, fast and can deal with a variety of frame imaging sensors from which no calibration parameters need to be supplied. Moreover, as Fig. 3.8c-4 indicates, even individual images without any GCP can be transformed into orthophotos. The combination of these advantages largely overcomes the current drawbacks that archaeologists encounter in most (ortho)rectification approaches, certainly when dealing with larger areas (features of a palaeolandscape, extensive sites) or terrain undulation.

However, it should be noted that such an integrated approach only works when no major scene changes have taken place during the years of image acquisition. In the case study of *Trea*, the biggest surface difference was related to the phenological state of the vegetation: sometimes the fields were just harvested, while at other times the camera recorded the full canopy. Although it did not hamper the SfM stage, the DSM will obviously be influenced by this. Therefore, one can best use a set of images displaying the most common surface condition, after which a numerical form of the latter can be used to compute the orthophotos of more or less all images. This approach was used in this case study and did not result in archaeologically relevant positional differences of the computed orthophotos. In case the difference between different topographical conditions is too big, a multitude of DSMs should be computed to cover all possible surface states. In the worst case scenario, the landscape can have changed so drastically over time that image alignment will fail.

3.4.2 Kreuttal Region (Austria)

The acquisition of oblique aerial photographs is well suited for a computer vision approach. However, very ordered collections of vertical imagery can also be successfully processed into true orthophotos. Their high longitudinal and lateral overlap makes them very useful for 3D data extraction via photogrammetric means, but this also translates to high usability, automation and accuracy in an SfM-driven environment. This is not limited strictly to modern air photos, but can be used on high-quality historical air photo datasets as well. Furthermore, due to the high overlap of imagery, SfM-based data processing methodologies are able to extend the usability of these types of datasets into the 3D realm, allowing for the creation of not only 2D orthomosaics but 3D historical digital elevation models (hDEMs). Therefore, historic land use and land change can be evaluated from a topographic perspective, bringing a new dimension to archaeological landscape analysis (cf. Pérez Álvarez et al. 2013).

Of the many archives of vertical historical aerial images that exist, perhaps some of the most well known are The Aerial Reconnaissance Archives (TARA) and the National Archives and Records Administration (NARA) holdings. Located in Edinburg and Washington D.C., respectively, the total number of photographs in these archives is ca. 21 million (Cowley and Stichelbaut 2012; Cowley et al. 2013) dating from as early as 1918. Numerous national and regional archives also exist, of which a number are further detailed in Wilson (2000), Cowley et al. (2010) and Hanson and Oltean (2013). While the condition of materials in these archives can be highly variable, they are nevertheless vast and largely unique sources of information, and lack of proper camera and lens data for many of the photos contained therein is not necessarily an obstacle to successful reconstruction with SfM-based approaches.

The case study presented here examines the use of historical vertical datasets in the Kreuttal region of Lower Austria (48° 26′ 40″ N, 16° 27′ 01″ E – WGS84). Situated roughly 25 km north of Vienna, the Kreuttal contains traces of past land use from the Neolithic to the Modern Historic eras. Archaeological sites in this topographically varied region manifest themselves on aerial photographs in the form of vegetation marks, soil marks and shadow marks, with a

number of upstanding and particularly well-preserved hill forts from the Bronze and Iron Age visible in the forest during off-leaf seasons. Two vertical datasets, acquired in March of 1945 and 2010, have been chosen from among the large archive of air photographs of the region to showcase the uses and issues involved in the processing of historic vertical datasets with SfM applications.

Sortie 15SG-1374, acquired on 23 March 1945, consists of 20 images acquired as part of an allied sortie over Lower Austria at the end of World War II. Images were acquired stripwise, west–east then east–west, at a scale of ca. 1:10 500 (Fig. 3.9a). Acquired from TARA through a local Austrian partner, the images came with no other camera or mission information. All images were 1,200 spi (samples per inch) scans of prints, many of which contain significant localised error due to warping and other degradation as a result of age and possibly improper storage before being acquired by TARA (Fig. 3.9b). Images were not

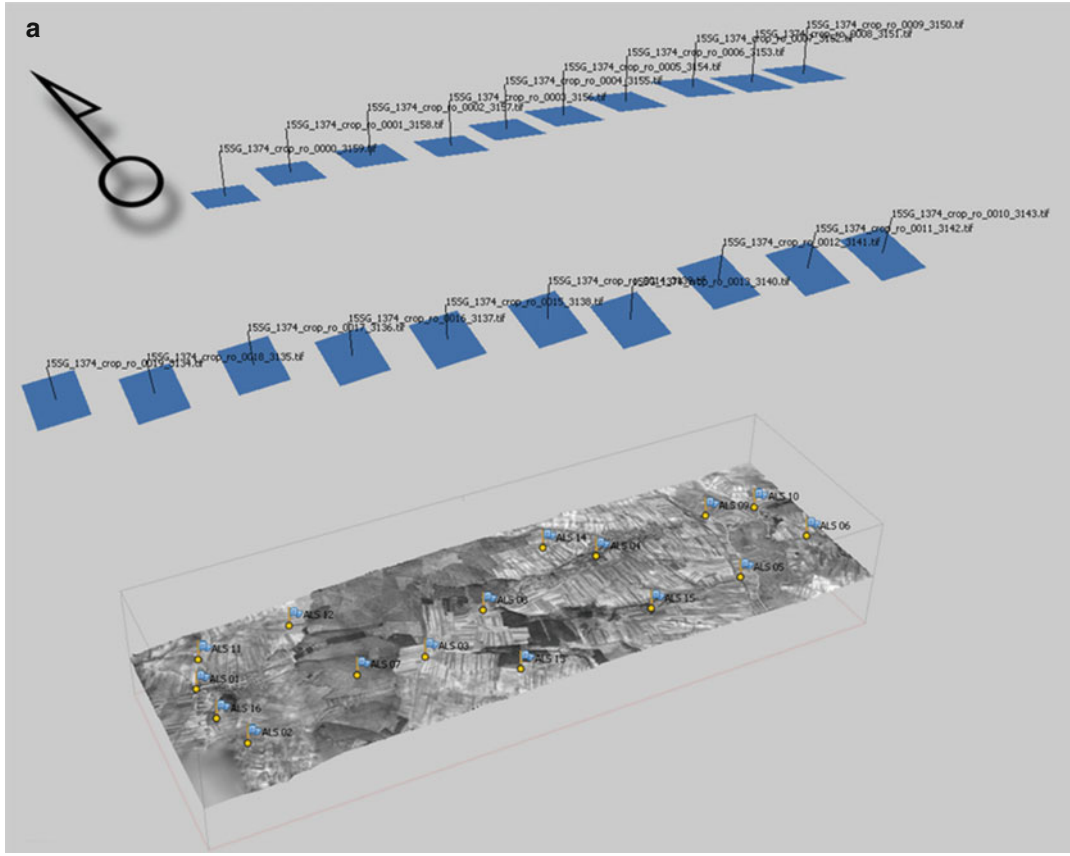

Fig. 3.9 (**a**) Reconstruction of flight path for sortie 15SG-1374. (**b**) Sample image from sortie 15SG-1374. (**c**) Reconstruction of flight path for flight 02100301. (**d**) Sample image from flight 02100301

Fig. 3.9 (continued)

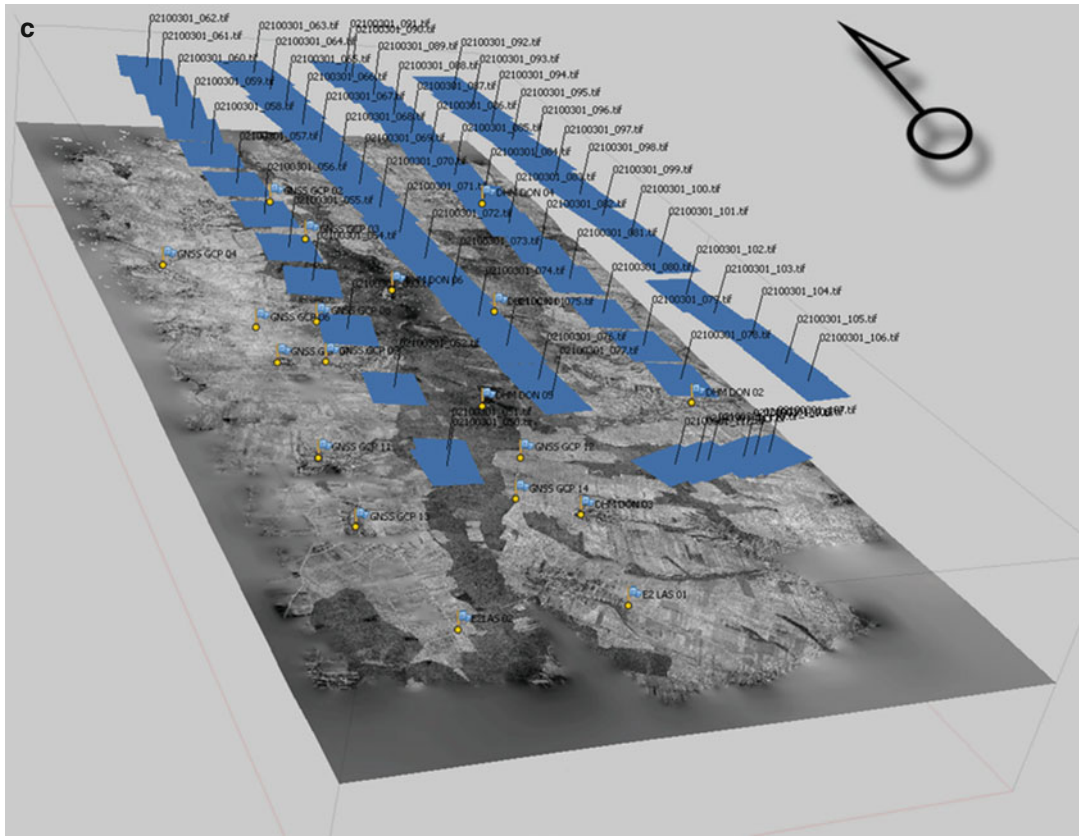

Fig. 3.9 (continued)

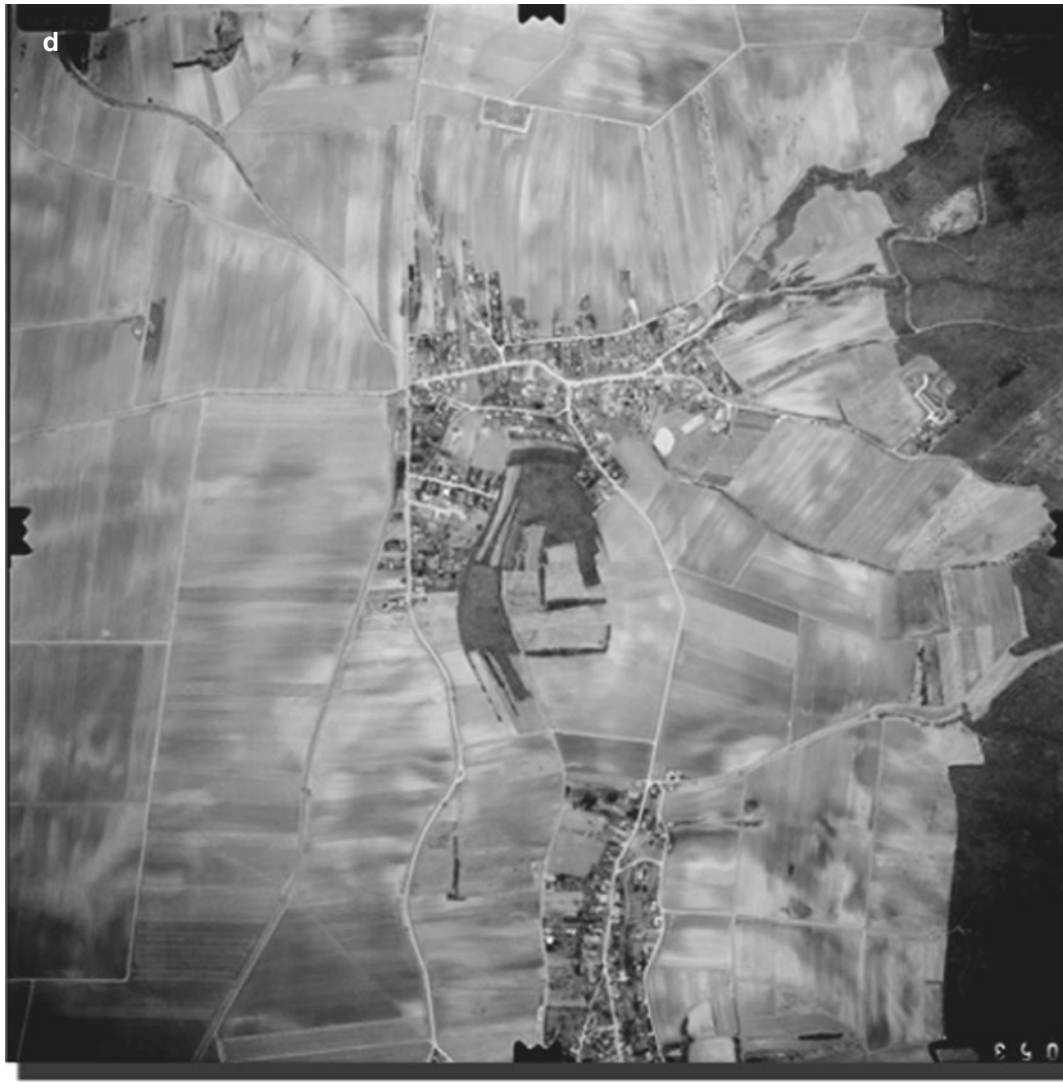

Fig. 3.9 (continued)

uniformly sharp, and many areas, including borders and fiducial marks, had to be masked so as not to interfere with reconstruction. Furthermore, as the images are scans of 'predigital' photographs, they contain no Exif data or calibration data which the software could use in the SfM phase.

Despite all this, PhotoScan was able to align and match all 20 images as delivered by the archive. However, there were significant issues with camera pose estimation. This was due to the fact that, as a by-product of the scanning process, all images had different pixel dimensions. This issue was resolved by loading all of them into a photo editor, aligning them via their fiducial marks and cropping them to identical dimensions. Once this was completed, camera pose estimation improved significantly. GCPs were then placed in order to georeference the dataset while further refining camera calibration and pose by treating the GCPs as constraints in a subsequent bundle adjustment. This presented its own obstacles as landscape change was significant enough over the intervening 58 years as to make it extremely difficult to locate unchanged reference points. Through extensive comparison with other datasets a number of GCPs were eventually identified, with a 50 cm spatial resolution DSM generated from airborne laser scanning (ALS) data used to acquire GCP coordinates.

After masking, GCP placement and several bundle adjustments, the final model was able to achieve a total distributed georeferencing error of 3.02 m, the majority of that being in the RMSE(Z) (*X* error 0.562 m, *Y* error 0.854 m, *Z* error 2.842 m). This was largely due to the degraded quality of the prints causing excessive localised distortion in the 3D reconstruction. In this instance, 2D orthomosaics proved the most useful output as the 3D hDEM was extremely noisy and still contained significant local error. This could be corrected by further post-processing methods to reduce noise and correct for residual local distortion (Sevara 2013).

Flight 02100301 was acquired on the 1st of March 2010 by the Austrian Military at the request of the Aerial Archive at the University of Vienna (Doneus et al. 2001). This flight consisted of 63 images and was flown stripwise north–south to south–north at a scale of 1:10,000 (Fig. 3.9c). Unlike sortie 15SG-1374, all camera parameters for this flight are known and interior orientation data were readily available. Images were scanned from negatives using a Vexcel UltraScan 5000 photogrammetric scanner (Doneus et al. 2007) at a resolution of 5,080 spi. As a result, the images from flight 02100301 are of a significantly higher quality than those of sortie 15SG-1374 (Fig. 3.9d). Images still needed to be masked and the same issues were still present with regard to lack of Exif data as with 15SG-1374. However, since all camera parameters were known, these could be entered manually into PhotoScan.

With all of these factors significantly improving alignment and pose estimation, initial results were already far more accurate. Due to the high quality of the scan process, all images were the same dimensions, obviating the need to manually crop them. GCP placement was also significantly easier, due to the recent nature of the dataset. GCPs were acquired from the same DSM as for sortie 15SG-1374. Once GCPs were placed and the model was cleaned and optimised by an additional bundle adjustment, re-projection error dropped to below 1 pixel. The total distributed error for this dataset was 0.89 m utilising 17 of the 19 GCPs, the error being more evenly distributed this time (*X*, 0.49 m; *Y*, 0.59 m; *Z*, 0.44 m).

In this instance, both 2D and 3D products generated from flight 02100301 were of extremely high quality. The 2D orthomosaic corresponded in horizontal quality to that of orthomosaics generated in Leica Photogrammetry Suite (LPS) using the same dataset, with significant improvement over the LPS dataset in heavily wooded and variegated terrain due to the high accuracy of the hDEM used for orthorectification. The hDEM provided a correspondence of <50 cm when analysed against independently collected ground control using a Leica GPS 500 RTK receiver. Furthermore, accurate 3D data could also be acquired for upstanding prehistoric earthworks in the area.

As can be seen from this case study, SfM-based approaches to orthomosaic generation and

terrain reconstruction also work with historic datasets in a way that far exceeds the original intended use of the data. However, results can be highly variable and depend heavily on both the quality and quantity of original photographs, much as the other case studies in this section illustrate. Further information regarding this case study can be found in Sevara (2013).

3.4.3 Pitaranha (Portugal-Spain)

Ancient quarry sites are a good example of the multifaceted nature of certain archaeological sites. The often complex morphological and topographical characteristics of quarry landscapes, as well as the severe modification of the terrain configuration by both intensive quarrying and the intricate logistical extraction infrastructure complicate their survey. Since an accurate digital representation of the topographical surface is elementary to the spatial analysis of quarry sites and the availability of an orthophoto map a necessary prerequisite for fast and effective site navigation, the acquisition of such information is a crucial component of efficient quarry research. To this end, a cost-effective technique was developed to map the Roman quarry of Pitaranha, located on the present-day border between Portugal and Spain, some 200 m northeast of the village of Pitaranha (Alentejo, Portugal; 39º 22′ 13″ N, 07º 18′ 49″ W – WGS84). Historically, the quarry mainly provisioned the nearby Roman town of *Ammaia* (Vermeulen and Taelman 2010). Several periods of intensive building in the Roman town suggest large-scale quarrying at Pitaranha during the first centuries AD (Taelman et al. 2009). A thorough mapping of the site was deemed necessary in order to fully comprehend the particular mechanisms of the quarry.

After establishing a dense network of well-distributed GCPs (Fig. 3.10a), an unmanned low-altitude Helikite-based aerial system was used (Fig. 3.10b) to acquire aerial still imagery (detailed information on the development and construction of the Helikite platform can be found in Verhoeven et al. 2009a). For this case study, the Helikite platform was equipped with a 10 megapixel Nikon D80 reflex camera fitted with a Nikkor 20 mm f/3.5 AI-S. Although this lens suffers from quite some optical distortions, its resolving power – certainly in the centre of the image – is great, while it also offers a large angular field of view (61° by 43°) and is very light (235 g).

As a result of unstable wind conditions (i.e. thermal airstreams alternated with windless areas) and strong electromagnetic interference during camera and platform control, an unstructured collection of about 1,400 digital photographs was necessary to cover almost the entire quarry site. The scales of these images varied enormously, while the camera orientations – and to a certain extent the flight path – were almost random and certainly not as structured as initially intended. Since the ground-sampling distance (GSD) varied between approximately 3 and 8 cm, this variation was expected to be challenging because high-resolution detail would be attenuated with low-resolution geometries extracted from the images taken at high altitudes. Obviously, all these factors are normally not encountered in the highly structured datasets acquired by conventional aerial survey, such as those of the previous example.

In a first step, the complete image dataset was reduced to a more manageable photo collection of 377 sharp and well-exposed images. Altering the parameters resulted in different SfM solutions of which only the most accurate one was retained for subsequent MVS processing. After the calculation of a detailed continuous 3D surface, the final orthophotograph (Fig. 3.10c) was computed and its positional accuracy determined. To incorporate all possible uncertainties in the computed dataset (including those introduced by the control coordinates), the 95 % confidence interval was calculated and expressed according to the NSSDA standard (Federal Geographic Data Committee – Subcommittee for Base Cartographic Data 1998). In the end, the horizontal accuracy turned out to be 13.7 cm, while the overall absolute vertical accuracy value was 31 cm. Given that the source material consisted of an extremely unordered image collection of vertical, low and high oblique aerial photographs

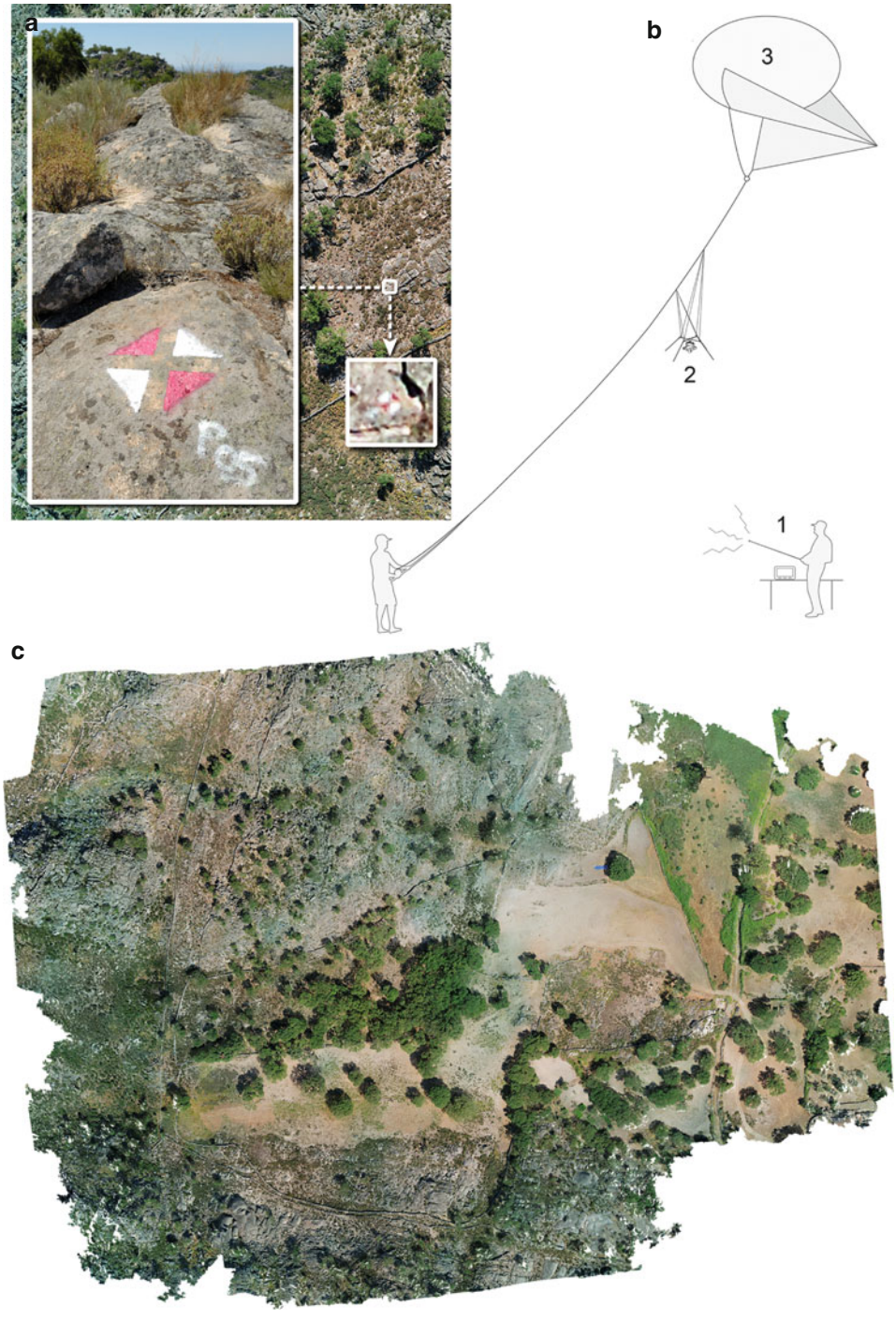

Fig. 3.10 (**a**) One of the oblique aerial images taken with the Helikite platform. The insets show one of the applied ground targets and how it is rendered in the final aerial photograph (**b**). A schematic overview of the Helikite aerial photography system, consisting of a Helikite (*1*), a digital still camera (*2*) and a camera operator with live video (*3*). (**c**) The final orthophotograph of the quarry

that are characterised by a GSD of minimum 8 cm, all acquired with non-metrical lens that suffered from a good deal of distortion, the reported positional accuracy of these datasets is considered very good (planimetric) to good (altimetric) and certainly better than initially expected. Moreover, at the moment of orthophoto production, the version of PhotoScan used did not allow to run a bundle adjustment which included the GCPs. As a result, the GCPs could not be applied to further optimise the SfM output but only to transform the complete model into an absolute CRS with a Helmert similarity transformation. Following the accuracy guidelines of the American Society for Photogrammetry and Remote Sensing (ASPRS), the RMSE values mean that the orthophoto can be used at a class 1 hard copy scale of 1:200 and contour lines with 50 cm intervals can be derived from the DSM (American Society for Photogrammetry and Remote Sensing 1990). More details on the rigorous assessment of the positional accuracy of this orthophoto and DSM can be found in Verhoeven et al. (2012b).

Conclusion

Straightforward orthophoto production is very important in the discipline of aerial archaeology. In this article, computer vision algorithms (structure from motion and multi-view stereo) complemented by proven photogrammetric principles (such as bundle adjustment) were exploited to present an integrated, cost-effective, semi-automated orthophoto production of archaeological aerial (uncalibrated) frame images. This approach is straightforward and requires no assumptions with regard to the camera projection matrix, extensive photogrammetric and computer vision knowledge of the user or the topography of the scenes. Moreover, simplicity is combined with geometrical quality due to the fact that the inner camera calibration parameters are automatically computed and a dense DSM is extracted and applied in a final phase to generate true orthophotos. As a result, this method largely accounts for most relevant kinds of geometrical degradations and is capable of generating 3D models and orthophotos that are perfectly suited for archaeological purposes. Further, only minimal technical knowledge and user interaction are required. Finally, this approach can also work in the total absence of any information about the instrument the imagery was acquired with, although it is still advised to have at least information on the focal length of the imaging system applied. The extra investments needed for software and computing hardware are recovered easily when taking the time and cost savings of map production into account.

This option of fast and accurate orthophoto production is very welcome for aerial archaeologists, given their current approaches which are not tailored to deal either with individual aerial frame images lacking sufficient ground control or with large amounts of photographs from different cameras shot in different seasons. This newly available method offers the enormous advantage that, besides a handful of GCPs, there are only standard photographic recording prerequisites. One simply needs to make sure that enough overlapping and sharp aerial images are acquired. Even though this might involve flying one or more orbits of the scene of interest (for the oblique approach) or vertical strips with up to 80 % overlap, this method will afterwards prove itself in terms of orthophoto quality and – in most occasions – processing speed, certainly when a larger area must be mapped or uneven terrain is involved. Furthermore, the case studies have shown that a large variety of old and new images can be processed into orthophotos whose accuracy is sufficient for large-scale archaeological photo mapping, as well as being visually appealing.

Of course, it is not all roses. First of all, it was indicated that the processing is very computer resource intensive, while the method is not applicable for the individual image. At least two – but preferably more images – are needed for accurate DSM computation. In addition, erroneous alignment of the imagery can occur when dealing with very large photo collections, images that suffer from excessive noise or blur, highly oblique photographs or

photographs that have a very dissimilar appearance (e.g. due to major underexposure or changing topographic terrain parameters). Additionally, several authors have already noted that the accuracy of the final products and the recovered camera parameters is often less than results yielded by the expensive and rigorous photogrammetric approaches (Remondino et al. 2012). However, differences are often small, while the approach presented here is superior in versatility and flexibility. The latter point cannot be overestimated, as many archived images do not fulfil the constraints (e.g. camera parameters) that are essential for accurate and straightforward georeferencing using any of the more standard georeferencing approaches by non-photogrammetrists. Currently, the biggest disadvantage of most available SfM-based software packages is the lack of computed metrics and tools in order to inspect the image orientation and matching reliability and accuracy.

Finally, the approach presented here is currently semi-automatic and automation only makes sense when it seriously reduces or completely eliminates steps in a process. In the case of archaeological orthophoto generation, these are the recurring steps of visualising and selection of the images, selecting the essential geodata (GCPs) and setting all the parameters for the subsequent execution of the algorithms. Since this is currently considered to be the bottleneck in large-scale archaeological projects with thousands of images, a project which aims at the creation of completely automatic solutions for orthophoto generation (including the GCP selection) of archaeological aerial photographs was initiated in 2012 (funded by the Austrian Science fund, P 24116-N23). This would offer possibilities for the consistent creation and updating of archaeologically relevant cartographic data in our rapidly changing landscapes.

Acknowledgements This article has been written within the framework of the Austrian Science Fund (FWF): P 24116-N23. The case study from the Potenza Valley Survey project was made possible thanks to support from Belgian Science Policy (Interuniversity Attraction Poles, project P6/22). The Ludwig Boltzmann Institute for Archaeological Prospection and Virtual Archaeology (archpro.lbg.ac.at) is based on an international cooperation of the Ludwig Boltzmann Gesellschaft (A), the University of Vienna (A), the Vienna University of Technology (A), the Austrian Central Institute for Meteorology and Geodynamic (A), the office of the Provincial Government of Lower Austria (A), Airborne Technologies GmbH (A), RGZM (Roman-Germanic Central Museum) Mainz (D), RAÄ (Swedish National Heritage Board) (S), IBM VISTA (University of Birmingham) (GB) and NIKU (Norwegian Institute for Cultural Heritage Research) (N).

References

Aber JS, Aber SW, Leffler B (2001) Challenge of infrared kite aerial photography. Trans Kansas Acad Sci 104:18–27. doi:10.1660/0022-8443(2001)104[0018:COIKAP]2.0.CO;2

Agisoft LLC (2012) Agisoft PhotoScan user manual. Professional edition, version 0.9.0. http://downloads.agisoft.ru/pdf/photoscan-pro_0_9_0_en.pdf. Accessed 13 Feb 2013

Altenhofen RE, Hedden RT (1966) Transformation and rectification. In: Thompson MM, Eller RC, Radlinski WA, Speert JL (eds) Manual of photogrammetry, vol II, 3rd edn. American Society of Photogrammetry, Falls Church, pp 803–849

Álvarez P, Antonio J, Herrera VM, Martínez del Pozo JÁ, de Tena MT (2013) Multi-temporal archaeological analyses of alluvial landscapes using the photogrammetric restitution of historical flights: a case study of Medellin (Badajoz, Spain). J Archaeol Sci 40:349–364. doi:10.1016/j.jas.2012.08.025

American Society for Photogrammetry and Remote Sensing, Specifications and Standards Committee (1990) ASPRS accuracy standards for large-scale maps. Photogramm Eng Remote Sens 56:1068–1070

Appetecchia A, Brandt O, Menander H, Thorén H (2012) New methods for documentation and analysis in building archaeology: prestudy. a project funded by the Swedish National Heritage Board, R & D funds, Lund. http://www.arkeologiuv.se/cms/showdocument/documents/extern_webbplats/arkeologiuv/publikationer_uv/rapporter/uv_rapport/2012/uvr2012_001.pdf. Accessed 4 Feb 2013

Barazzetti L, Remondino F, Scaioni M (2011) Automated and accurate orientation of complex image sequences. In: 3D-ARCH 2011: 3D virtual reconstruction and visualization of complex architectures, Proceedings of the 4th ISPRS international workshop, Trento, Italy, 2–4 Mar 2011. ISPRS

Barber M (2011) A history of aerial photography and archaeology. Mata Hari's glass eye and other stories. English Heritage, Swindon

Bay H, Tuytelaars T, Gool L (2006) SURF: speeded up robust features. In: Aleš L, Horst B, Axel P (eds) Computer vision, 9th European conference on computer vision (ECCV 2006, Graz, Austria, May 7–13, 2006), Proceedings, part I, vol 3951, Lecture notes in computer science. Springer, Berlin, pp 404–417

Bay H, Ess A, Tuytelaars T, van Gool L (2008) SURF: speeded up robust features. Comput Vis Image Underst 110:346–359

Bernstein R (1983) Image geometry and rectification. In: Colwell RN, Simonett DS, Ulaby FT (eds) Manual of remote sensing, vol. 1: Theory, instruments and techniques, 2nd edn. American Society of Photogrammetry, Falls Church, pp 873–922

Bewley R, Rączkowski W (eds) (2002) Aerial archaeology. Developing future practice, vol 337, NATO science series I: life and behavioural sciences. IOS Press, Amsterdam

Bezzi L (2012) 3D documentation of small archaeological finds. http://arc-team-open-research.blogspot.com.br/2012/08/3d-documentation-of-small.html. Accessed 11 October 2012

Billingsley FC (1965) Digital video processing at JPL. In: Electronic Imaging Techniques I, vol 15. SPIE, Bellingham

Billingsley FC, Anuta PE, Carr JL, McGillem CD, Smith DM, Strand TC (1983) Data processing and reprocessing. In: Colwell RN, Simonett DS, Ulaby FT (eds) Manual of remote sensing, vol. 1: Theory, instruments and techniques, 2nd edn. American Society of Photogrammetry, Falls Church, pp 719–792

Bradley D, Boubekeur T, Heidrich W (2008) Accurate multi-view reconstruction using robust binocular stereo and surface meshing. In: CVPR 2008. IEEE conference on computer vision and pattern recognition, 23–28 June 2008. IEEE, Anchorage, pp 1–8. doi:10.1109/CVPR.2008.4587792

Braun J (2003) Aspects on true-orthophoto production. In: Fritsch D (ed) Photogrammetric week '03. Wichmann Verlag, Heidelberg, pp 205–214

Brophy K, Cowley D (eds) (2005) From the air. Understanding aerial archaeology. Tempus, Stroud

Brown DC (1966) Decentering distortion of lenses: the prism effect encountered in metric cameras can be overcome through analytic calibration. Photogramm Eng Remote Sens 32:444–462

Brown DC (1956) The simultaneous determination of the orientation and lens distortion of a photogrammetric camera. Air Force Missile Test Center Technical Report 56–20. Florida

Brugioni DA (1989) The serendipity effect of aerial reconnaissance. Interdiscip Sci Rev 14:16–28. doi:10.1179/030801889789798357

Buchanan T (1993) Photogrammetry and projective geometry: an historical survey. In: Integrating photogrammetric techniques with scene analysis and machine vision, Orlando, FL, USA, 11 Apr 1993. SPIE, Bellingham, pp 82–91. doi:10.1117/12.155817

Burnside CD (1985) Mapping from aerial photographs, 2nd edn. Collins, London

Castrianni L (2008) Giacomo Boni: a pioneer of the archaeological aerial photography. In: Remote sensing for archaeology and cultural heritage management: proceedings of the 1st international EARSeL workshop, CNR, Rome, Arracne, Rome, September 30–October 4, 2008, pp 55–58

Coleman S (2007) Taking advantage: vertical aerial photographs commissioned for local authorities. In: Mills J, Palmer R (eds) Populating clay landscapes. Tempus, Stroud, pp 28–33

Colwell RN (1997) History and place of photographic interpretation. In: Philipson WR (ed) Manual of photographic interpretation, 2nd edn. American Society of Photogrammetry and Remote Sensing, Bethesda, pp 3–47

Cowley DC, Stichelbaut BB (2012) Historic aerial photographic archives for European archaeology. Eur J Archaeol 15:217–236. doi:10.1179/146195711 2Y.0000000010

Cowley D, Standring RA, Abicht MJ (eds) (2010) Landscapes through the lens. Aerial photographs and historic environment, vol 2, Occasional publication of the Aerial Archaeology Research Group. Oxbow Books, Oxford/Oakville

Cowley DC, Ferguson LM, Allan W (2013) The aerial reconnaissance archives: a global aerial photographic collection. In: Hanson WS, Oltean IA (eds) Archaeology from historical aerial and satellite archives. Springer, New York, pp 13–30

Crawford OGS (1924) Air survey and archaeology, vol 7, Ordnance survey professional papers, New series. Ordnance Survey, Southampton

Crawford OGS (1929) Air photographs of the Middle East: a paper read at the evening meeting of the Society on 18 March 1929. Geogr J 73:497–509

Crawford OGS (1933) Some recent air discoveries. Antiquity 7:290–296

Crawford OGS, Keiller A (1928) Wessex from the air. Oxford University Press, Oxford

Crawshaw A (1995) Oblique aerial photography: aircraft, cameras and films. In: Kunow J (ed) Luftbildarchäologie in Ost- und Mitteleuropa/Aerial archaeoloy in Eastern and Central Europe: internationales symposium, Kleinmachnow, Land Brandenburg, 26–30, September 1994, vol 3, Forschungen zur Archäologie im Land Brandenburg. Verlag Brandenburgisches Landesmuseum für Ur- und Frühgeschichte, Potsdam, pp 67–76

Crawshaw A (1997) Letter. AARGnews 14:59

Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: Proceedings of the IEEE Computer Society conference on computer vision and pattern recognition, San Diego, CA, USA, 20–25 June 2005. IEEE Computer Society, Los Alamitos, pp 886–893. doi:10.1109/CVPR.2005.177

Deng H, Wei Zhang, Mortensen E, Dietterich T, Shapiro L (2007) Principal curvature-based region detector for object recognition. In: Proceedings of the 2007 IEEE conference on computer vision and pattern recognition CVPR '07, Minneapolis, MN, USA, 18–23 June. IEEE, Piscataway, pp 1–8. doi:10.1109/CVPR.2007.382972

Dickinson GC (1969) Maps and air photographs. Edward Arnold, London

Doneus M (1997) On the archaeological use of vertical photographs. AARGnews 15:23–27

Doneus M (2000) Vertical and oblique photographs. AARGnews 20:33–39

Doneus M, Eder-Hinterleitner A, Neubauer W (2001) Archaeological prospection in Austria. In: Archaeological prospection: fourth international conference on archaeological prospection, Vienna, 19–23 Sept 2001. Austrian Academy of Sciences, Vienna, pp 11–33

Doneus M, Briese C, Fera M, Fornwagner U, Griebl M, Janner M, Zingerle M-C (2007) Documentation and analysis of archaeological sites using aerial reconnaissance and airborne laser scanning. In: Anticipating the future of the cultural past: proceedings of the XXI international CIPA symposium, Athens, Greece, 1–6 Oct 2007, The ISPRS international archives of the photogrammetry, remote sensing and spatial information sciences. CIPA, Athens, vol XXXVI-5/C53, pp 275–280. ISSN 1682–1750

Doneus M, Verhoeven G, Fera M, Briese C, Kucera M, Neubauer W (2011) From deposit to point cloud: a study of low-cost computer vision approaches for the straightforward documentation of archaeological excavations. In: Geoinformatics 6, XXIIIrd international CIPA Symposium, pp 81–88

Eisenbeiss H (2009) UAV photogrammetry. PhD thesis, ETH Zürich, Zürich. http://e-collection.library.ethz.ch/eserv/eth:498/eth-498-02.pdf#search=%22%28author:henri%20eisenbeiss%29%22. Accessed 11 Feb 2013

Eisenbeiss H, Sauerbier M (2011) Investigation of UAV systems and flight modes for photogrammetric applications. Photogramm Rec 26:400–421. doi:10.1111/j.1477-9730.2011.00657.x

Eisenbeiss H, Sauerbier M, Zhang L, Grün A (2005) Mit dem Modellhelikopter über Pinchango Alto. Geomat Schweiz 9:510–515

El-Hakim, SF, Beraldin J-A, Picard M (2003) Effective 3D modeling of heritage sites. In: Proceedings of the 4th international conference 3-D digital imaging and modeling, Banff, Canada, 6–10 October. IEEE Computer Society Press, Los Alamitos, pp 302–309

Estes JE, Hajic EJ, Tinney LR, Carver LG, Cosentino MJ, Mertz FC, Pazner MI, Ritter LR, Sailer CT, Stow DA, Streich TA, Woodcock CE (1983) Fundamentals of image analysis: analysis of visible and thermal infrared data. In: Colwell RN, Simonett DS, Ulaby FT (eds) Manual of remote sensing, vol. 1: Theory, instruments and techniques, 2nd edn. American Society of Photogrammetry, Falls Church, pp 987–1124

Falkner E, Morgan D (2002) Aerial mapping. Methods and applications, 2nd edn, Mapping sciences series. Lewis, Boca Raton

Faugeras O, Luong Q-T, Papadopoulo T (2001) The geometry of multiple images. The laws that govern the formation of multiple images of a scene and some of their applications. MIT Press, Cambridge

Federal Geographic Data Committee – Subcommittee for Base Cartographic Data (1998) Geospatial positioning accuracy standards. Part 3: National Standard for Spatial Data Accuracy (FGDC-STD-007.3-1998). Federal Geographic Data Committee, Reston

Fischler MA, Bolles RC (1981) Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun ACM 24:381–395. doi:10.1145/358669.358692

Fisher RB, Dawson-Howe K, Fitzgibbon A, Robertson C, Trucco E (2005) Dictionary of computer vision and image processing. Wiley, Chichester

Forte M, Dell'unto N, Issavi J, Onsurez L, Lercari N (2012) 3D archaeology at Çatalhöyük. Int J Herit Digit Era 1:352–378. doi:10.1260/2047-4970.1.3.351

Furukawa Y, Ponce J (2010) Accurate, dense, and robust multiview stereopsis. IEEE Trans Pattern Anal Mach Intell 32:1362–1376. doi:10.1109/TPAMI.2009.161

Goesele M, Curless B, Seitz SM (2006) Multi-view stereo revisited. In: Proceedings of the 2006 IEEE Computer Society conference on computer vision and pattern recognition CVPR'06. IEEE Computer Society Press, Los Alamitos, 17–22 June 2006, vol. 2, pp 2402–2409. doi:10.1109/CVPR.2006.199.

Graham R, Koh A (2002) Digital aerial survey. Theory and practice. CRC Press/Whittles Publishing, Boca Raton

Gruner H, Pestrecov K, Norton CL, Tayman WP, Washer FE (1966) Elements of photogrammetric optics. In: Thompson MM, Eller RC, Radlinski WA, Speert JL (eds) Manual of photogrammetry, vol I, 3rd edn. American Society of Photogrammetry, Falls Church, pp 67–132

Gyer MS (1996) Methods for computing photogrammetric refraction corrections for vertical and oblique photographs. Photogramm Eng Remote Sens 62:301–310

Habbecke M, Kobbelt L (2006) Iterative multi-view plane fitting. In: Kobbelt L, Kuhlen T, Aach T, Westerman R (eds) Proceedings of the 11th international fall workshop vision, modeling, and visualization 2006, Aachen, Germany, 22–24 Nov 2006. Akademische Verlagsgesellschaft Aka GmbH, Berlin, pp 73–80

Hallert B (1960) Photogrammetry. Basic principles and general survey, McGraw-Hill civil engineering series. McGraw-Hill, New York

Hanson WS, Oltean IA (eds) (2013) Archaeology from historical aerial and satellite archives. Springer, New York

Harman WE Jr, Miller RH, Sidney Park W, Webb JP (1966) Aerial photography. In: Thompson MM, Eller RC, Radlinski WA, Speert JL (eds) Manual of photogrammetry, vol I, 3rd edn. American Society of Photogrammetry, Falls Church, pp 195–242

Harris C, Stephens M (1988) A combined corner and edge detector. In: Proceedings of the fourth Alvey Vision conference AVC88, University of Sheffield Printing Office; Sheffield, 31 August–2 September 1988. BMVA, pp 147–151

Hartley RI (1994) Projective reconstruction and invariants from multiple images. IEEE Trans Pattern Anal Mach Intell 16:1036–1041. doi:10.1109/34.329005

Hartley RI, Mundy JL (1993) Relationship between photogrammetry and computer vision. In: SPIE (ed) Integrating

photogrammetric techniques with scene analysis and machine vision, 11 Apr 1993, Orlando, FL, USA. SPIE, Bellingham, pp 92–105. doi:10.1117/12.155818

Hartley R, Zisserman A (2003) Multiple view geometry in computer vision, 2nd edn. Cambridge University Press, Cambridge

Hassett TJ, Mullen RR, Pilonero JT, Pugh HV, Freeman J, Speert JL (1966) Aerial mosaics and photomaps. In: Thompson MM, Eller RC, Radlinski WA, Speert JL (eds) Manual of photogrammetry, vol II, 3rd edn. American Society of Photogrammetry, Falls Church

Hirschmüller H (2008) Stereo processing by semiglobal matching and mutual information. IEEE Trans Pattern Anal Mach Intell 30:328–341. doi:10.1109/TPAMI.2007.1166

Imhof RK, Doolittle RC (1966) Mapping from oblique photographs. In: Thompson MM, Eller RC, Radlinski WA, Speert JL (eds) Manual of photogrammetry, vol II, 3rd edn. American Society of Photogrammetry, Falls Church, pp 875–917

Juan L, Gwon O (2009) A comparison of SIFT, PCA-SIFT and SURF. Int J Image Process 3:143–152

Jurie F, Schmid C (2004) Scale-invariant shape features for recognition of object categories. In: Proceedings of the 2004 IEEE Computer Society conference on computer vision and pattern recognition, CVPR 2004. IEEE Computer Society Press, Los Alamitos, 27 June–2 July, vol. 2, pp 90–96. doi:10.1109/CVPR.2004.1315149

Kadir T, Brady M (2001) Saliency, scale and image description. Int J Comput Vis 45:83–105. doi:10.1023/A:1012460413855

Kennedy D (1996) Aerial archaeology in the Middle East. AARGnews 12:11–15

Kersten TP, Lindstaedt M (2012) Potential of automatic 3D object reconstruction from multiple images for applications in architecture, cultural heritage and archaeology. Int J Herit Digit Era 1:399–420. doi:10.1260/2047-4970.1.3.399

Kraus K (2002) Zur Orthophoto-Terminologie. Photogramm Fernerkund Geoinf 6:451–452

Kraus K (2007) Photogrammetry. Geometry from images and laser scans, 2nd edn. Walter de Gruyter, Berlin-New York

Krijnen F (2008) A fresh look at aerial photography. http://www.aircatcher.com

Labun E (2009) ImageJ SURF. http://labun.com/imagej-surf/ImageJ_SURF_2009-12-01_08.19.jar

Lerma JL, Navarro S, Cabrelles M, Seguí AE, Haddad N, Akasheh T (2011) Integration of laser scanning and imagery for photorealistic 3D architectural documentation. In: Wang C-C (ed) Laser scanning, theory and applications. InTech, Shanghai, pp 413–430

Lhuillier M, Quan L (2005) A quasi-dense approach to surface reconstruction from uncalibrated images. IEEE Trans Pattern Anal Mach Intell 27:418–433. doi:10.1109/TPAMI.2005.44

Lindeberg T (1998) Feature detection with automatic scale selection. Int J Comput Vis 30:79–116. doi:10.1023/A:1008045108935

Lo Brutto M, Meli P (2012) Computer vision tools for 3D modelling in archaeology. Int J Herit Digit Era 1:1–6. doi:10.1260/2047-4970.1.0.1

Lo Brutto M, Borruso A, D'Argenio A (2012) UAV systems for photogrammetric data acquisition of archaeological sites. Int J Herit Digit Era 1:7–14. doi:10.1260/2047-4970.1.0.7

Lowe DG (2004) Distinctive image features from scale-invariant keypoints. Int J Comput Vis 60:91–110. doi:10.1023/B:VISI.0000029664.99615.94

Ludvigsen M, Eustice R, Singh H (2006) Photogrammetric models for marine archaeology. In: Proceedings of the IEEE/MTS OCEANS'06 conference and exhibition, Boston, MA, 18–21 Sept 2006. IEEE, Piscataway, pp 1–6. doi:10.1109/OCEANS.2006.306915

Manzer G (1996) Avoiding digital orthophoto problems. In: Greve C (ed) Digital photogrammetry: an addendum to the manual of photogrammetry. American Society of Photogrammetry and Remote Sensing, Falls Church, pp 158–162

Matas J, Chum O, Urban M, Pajdla T (2004) Robust wide-baseline stereo from maximally stable extremal regions. Image Vis Comput 22:761–767. doi:10.1016/j.imavis.2004.02.006

Mellor JP, Teller S, Lozano-Pérez T (1996) Dense depth maps from epipolar images, vol 1953, AI Lab technical memo. Massachusetts Institute of Technology/Artificial Intelligence Laboratory, Cambridge

Microdrones GmbH (2008) Key Information for md4-1000 http://www.microdrones.com/products/md4-1000/md4-1000-key-information.php. Accessed 21 April 2008

Microsoft Corporation (2010) Photosynth. Microsoft Corporation, Redmond, http://photosynth.net/

Mikhail EM, Bethel JS, Chris McGlone J (2001) Introduction to modern photogrammetry. Wiley, New York

Mikolajczyk K, Schmid C (2003) A performance evaluation of local descriptors. In: Proceedings of the 2003 IEEE Computer Society conference on computer vision and pattern recognition, CVPR 2003, Madison, WI, USA, 16–22 June 2003, vol. 2. IEEE Computer Society, Los Alamitos, pp 257–263. doi:10.1109/CVPR.2003.1211478

Mikolajczyk K, Schmid C (2005) A performance evaluation of local descriptors. IEEE Trans Pattern Anal Mach Intell 27:1615–1630. doi:10.1109/TPAMI.2005.188

Mikolajczyk K, Tuytelaars T, Schmid C, Zisserman A, Matas J, Schaffalitzky F, Kadir T, van Gool L (2005) A comparison of affine region detectors. Int J Comput Vis 65:43–72

Mills J (2005) Bias and the world of the vertical aerial photograph. In: Brophy K, Cowley D (eds) From the air: understanding aerial archaeology. Tempus, Stroud, pp 117–126

Moisan L, Stival B (2004) A probabilistic criterion to detect rigid point matches between two images and estimate the fundamental matrix. Int J Comput Vis 57:201–218. doi:10.1023/B:VISI.0000013094.38752.54

Moons T, van Gool L, Vergauwen M (2008) 3D Reconstruction from multiple images, part 1: Principles. Found Trends Comput Graph Vis 4:287–404. doi:10.1561/0600000007

Moreels P, Perona P (2007) Evaluation of features detectors and descriptors based on 3D objects. Int J Comput Vis 73:263–284. doi:10.1007/s11263-006-9967-1

Morel J-M, Guoshen Yu (2009) ASIFT: a new framework for fully affine invariant image comparison. SIAM J Imaging Sci 2:438–469. doi:10.1137/080732730

Moscatelli U (1985) Municipi romani della V regio Augustea: problemi storici ed urbanistici del Piceno centro-settentrionale (III – I sec. a.C.). PICUS Studi e ricerche sulle Marche nell'antichità 5:51–97

Moscatelli U (1987) Materiali per la topografia storica di *Potentia*. In: Paci G (ed) Miscellanea di studi marchigiani in onore di Febo Allevi. Facoltà di Lettere e Filosofia/Università di Macerata, Agugliano, pp 429–438

Mundy JL, Zisserman A (1992) Appendix – projective geometry for machine vision. In: Mundy JL, Zisserman A (eds) Geometric invariance in computer vision. MIT Press, Cambridge, pp 463–534

Newhall B (2006) The history of photography. From 1839 to the present, 5th edn. Museum of Modern Art, New York/Boston

Norton PR (2010) Photodetectors. In: Bass M, DeCusatis CM, Enoch JM, Lakshminarayanan V, Li G, MacDonald CA, Mahajan VN, van Stryland EW (eds) Handbook of optics, vol. II. Design, fabrication, and testing; sources and detectors; radiometry and photometry, 3rd edn. McGraw-Hill, New York, pp 24.3–24.102

Ohno Y (2006) Basic concepts in photometry, radiometry and colorimetry. In: Dakin JP, Brown RGW (eds) Handbook of optoelectronics. Taylor & Francis, Boca Raton, pp 287–305

Opitz R, Nowlin J (2012) Photogrammetric modeling + GIS: better methods for working with mesh data. ArcUser Spring:46–49

Palmer R (1996) Editorial. AARGnews 13:3

Palmer R (2005) If they used their own photographs they would not take them like that. In: Brophy K, Cowley D (eds) From the air: understanding aerial archaeology. Tempus, Stroud, pp 94–116

Palmer R (2007) Seventy-five years v. Ninety minutes: implications of the 1996 Bedfordshire vertical aerial survey on our perceptions of clayland archaeology. In: Mills J, Palmer R (eds) Populating clay landscapes. Tempus, Stroud, pp 88–103

Palmer JM, Grant BG (2010) The art of radiometry. SPIE, Bellingham

Pollefeys M, van Gool L, Vergauwen M, Cornelis K, Verbiest F, Tops J (2001) Image-based 3D acquisition of archaeological heritage and applications. In: Proceedings of the 2001 conference on virtual reality, archaeology, and cultural heritage, Glyfada, Greece, 28–30 Nov 2001. Association for Computing Machinery, New York, pp 255–262

Pollefeys M, van Gool L (2002) Visual modelling: from images to images. J Vis Comput Animat 13:199–209. doi:10.1002/vis.289

Pollefeys M, Koch R, Vergauwen M, van Gool L (1998) Virtualizing archaeological sites. In Proceedings of the 4th international conference on virtual systems and multimedia, VSMM 98, Gifu, Japan, 18–20 Nov 1998. IOS Press, Amsterdam

Pollefeys M, Koch R, Vergauwen M, van Gool L (2000) Automated reconstruction of 3D scenes from sequences of images. ISPRS J Photogramm Remote Sens 55: 251–267. doi:10.1016/S0924-2716(00)00023-X

Pollefeys M, van Gool L, Vergauwen M, Cornelis K, Verbiest F, Tops J (2003) 3D recording for archaeological fieldwork. IEEE Comput Graph Appl 23: 20–27. doi:10.1109/MCG.2003.1198259

Pollefeys M, van Gool L, Vergauwen M, Verbiest F, Cornelis K, Tops J, Koch R (2004) Visual modeling with a hand-held camera. Int J Comput Vis 59:207–232. doi:10.1023/B:VISI.0000025798.50602.3a

Quan L (2010) Image-based modeling. Springer, New York

Read RE, Graham R (2002) Manual of aerial survey. Primary data acquisition. CRC Press/Whittles Publishing, Boca Raton

Reinhard J (2012) Things on strings and complex computer algorithms: kite aerial photography and structure from motion photogrammetry at the Tulul adh-Dhahab, Jordan. AARGnews 45:37–41

Remondino F, Fraser C (2006) Digital camera calibration methods: considerations and comparisons. In ISPRS Commission V symposium 'image engineering and vision metrology', 25–27 Sept 2006. International Society for Photogrammetry and Remote Sensing, Dresden, pp 266–272

Remondino F, Barazzetti L, Nex F, Scaioni M, Sarazzi D (2011) UAV photogrammetry for mapping and 3d modelling: current status and future perspectives. In: Proceedings of the international conference on unmanned aerial vehicle in geomatics UAV-g, Zurich, Switzerland, 14–16 Sept 2011, vol 38(1/C22). International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Zürich

Remondino F, Del Pizzo S, Kersten TP, Troisi S (2012) Low-cost and open-source solutions for automated image orientation: a critical overview. In: Progress in cultural heritage preservation. In: Proceedings of the 4th international conference Euromed 2012, Lemessos, Cyprus. October 29–November 3, 2012. Springer, Berlin/Heidelberg, pp 40–54

Robertson DP, Cipolla R (2009) Structure from motion. In: Varga M (ed) Practical image processing and computer vision. Wiley, New York

Rosten E, Drummond T (2005) Fusing points and lines for high performance tracking. In: Proceedings of the tenth IEEE international conference on computer vision ICCV'05. IEEE Computer Society Press, Los Alamitos, 17–21 Oct 2005, vol 2, pp 1508–1515. doi:10.1109/ICCV.2005.104

Rousseeuw PJ (1984) Least median of squares regression. J Am Stat Assoc 79:871–880. doi:10.2307/2288718

Sarfraz MS, Hellwich O (2008) Head pose estimation in face recognition across pose scenarios. In: Proceedings of the third international conference on computer vision theory and applications VISAPP 2008, Funchal,

Portugal, 22–25 Jan 2008, vol 1. INSTICC, Setúbal, pp 235–242
Scharstein D, Szeliski R (2002) A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. Int J Comput Vis 47:7–42
Schlitz M (2004) A review of low-level aerial archaeology and its application in Australia. Aust Archaeol 59:51–58
Schmid C, Mohr R (1996) Combining grey value invariants with local constraints for object recognition. In: Proceedings of the 1996 IEEE Computer Society conference on computer vision and pattern recognition CVPR '96, San Francisco, California, 18 June–20 June 1996. IEEE Computer Society Press, Los Alamitos, pp 872–877. doi:10.1109/CVPR.1996.517174
Schneider S (1974) Luftbild und Luftbildinterpretation, vol 11, Lehrbuch der allgemeinen Geographie. Walter de Gruyter, Berlin/New York
Schott JR (2007) Remote sensing. The image chain approach, 2nd edn. Oxford University Press, New York
Schreiber WF (1967) Picture coding. Proc IEEE 55:320–330. doi:10.1109/PROC.1967.5488
Scollar I, Giradeau-Montaut D (2012) Georeferenced orthophotos and DTMs from multiple oblique images. AARGnews 44:12–17
Scollar I, Tabbagh A, Hesse A, Herzog I (1990) Archaeological prospecting and remote sensing, vol 2, Topics in remote sensing. Cambridge University Press, Cambridge
Seitz SM, Curless B, Diebel J, Scharstein D, Szeliski R (2006) A comparison and evaluation of multi-view stereo reconstruction algorithms. In: 2006 IEEE Computer Society conference on computer vision and pattern recognition CVPR'06, vol. 1. IEEE, Washington, DC, pp 519–528
Sevara C (2013) Top Secret Topographies: Examining the potential for recovering two and three-dimensional archaeological information from historic reconnaissance datasets using image-based modelling techniques. Inl J of Heritage in the Digital Era 2:3
Sewell ED, Livingston RG, Quick JR, Norton CL, Case JB, Sanders RG, Goldhammer JS, Aschenbrenner B (1966) Aerial cameras. In: Thompson MM, Eller RC, Radlinski WA, Speert JL (eds) Manual of photogrammetry, vol I, 3rd edn. American Society of Photogrammetry, Falls Church, pp 133–194
Slater PN, Doyle FJ, Fritz NL, Welch R (1983) Photographic systems for remote sensing. In: Colwell RN, Simonett DS, Ulaby FT (eds) Manual of remote sensing, vol. 1: Theory, instruments and techniques, 2nd edn. American Society of Photogrammetry, Falls Church, pp 231–291
Smith SW (1997) The scientist and engineer's guide to digital signal processing, 1st edn. California Technical Publishing, San Diego
Snavely N (2010) Bundler: structure from motion for unordered image collections. Software
Snavely N, Seitz SM, Szeliski R (2006) Photo tourism: exploring photo collections in 3D. ACM Trans Graph 25:835–846
Spurr SH (1960) Photogrammetry and photo-interpretation. With a section on applications to forestry, 2nd edn. The Ronald Press Company, New York
Stichelbaut B, Bourgeois J, Saunders D, Chielens P (eds) (2009) Images of conflict. Military aerial photography and archaeology. Cambridge Scholars Publishing, Newcastle upon Tyne
Strecha C, Fransens R, van Gool L (2006) Combined depth and outlier estimation in multi-view stereo. In: Proceedings of the 2006 IEEE Computer Society conference on computer vision and pattern recognition, CVPR'06. IEEE Computer Society Press, Los Alamitos, 17–22 June 2006, vol. 2. pp 2394–2401. doi:10.1109/CVPR.2006.78
Szeliski R (2011) Computer vision. Algorithms and applications, Texts in computer science. Springer, New York
Taelman D, Deprez S, Vermeulen F, De Dapper M (2009) Granite and rock crystal quarrying in the *Civitas Ammaiensis* (north-eastern Alentejo, Portugal): a geo-archaeological case study. BABesch – Bulletin Antieke Beschaving 84:171–182
Tewinkel GC, Schmid HH, Hallert B, Rosenfield GH (1966) Basic mathematics of photogrammetry. In: Thompson MM, Eller RC, Radlinski WA, Speert JL (eds) Manual of photogrammetry, vol I, 3rd edn. American Society of Photogrammetry, Falls Church, pp 17–65
Tingdahl D, Maarten V, van Gool L (2012) ARC3D: a public web service that turns photos into 3D models. In: Stanco F, Battiato S, Gallo G (eds) Digital imaging for cultural heritage preservation: analysis, restoration, and reconstruction of ancient artworks, Digital imaging and computer vision series. CRC Press, Boca Raton, pp 101–125
Torr PHS (2002) Bayesian model estimation and selection for epipolar geometry and generic manifold fitting. Int J Comput Vis 50:35–61. doi:10.1023/A:1020224303087
Triggs B, Mclauchlan PF, Hartley RI, Andrew F (2000) Bundle adjustment – a modern synthesis. In: Triggs B, Zisserman A, Szeliski R (eds) Vision algorithms: theory and practice: proceedings of the international workshop on vision algorithms, Corfu, Greece, September 1999, vol 1883, Lecture notes in computer science. Springer, London, pp 298–372
Turpin RD, Ramey EH, Case JB, Coleman CG, Lynn WD, Michaelis OE (1966) Definitions of terms and symbols used in photogrammetry. In: Thompson MM, Eller RC, Radlinski WA, Speert JL (eds) Manual of photogrammetry, vol II, 3rd edn. American Society of Photogrammetry, Falls Church, pp 1125–1161
Tuytelaars T, Mikolajczyk K (2007) Local invariant feature detectors: a survey. Found Trends Comput Graph Vis 3:177–280. doi:10.1561/0600000017
Ullman S (1979) The interpretation of structure from motion. Proc R Soc B Biol Sci 203:405–426. doi:10.1098/rspb.1979.0006
Verhoeven G (2008a) Exploring the edges of the unseen: an attempt to digital aerial UV photography. In: Remote sensing for archaeology and cultural heritage management: proceedings of the 1st International EARSeL workshop CNR, Rome, September 30–October 4, 2008. Aracne, Rome, pp 79–83

Verhoeven G (2008b) Imaging the invisible using modified digital still cameras for straightforward and low-cost archaeological near-infrared photography. J Archaeol Sci 35:3087–3100. doi:10.1016/j.jas.2008.06.012

Verhoeven G (2009a) Beyond conventional boundaries. New technologies, methodologies, and procedures for the benefit of aerial archaeological data acquisition and analysis. PhD thesis, Nautilus Academic Books, Zelzate

Verhoeven G (2009b) Providing an archaeological bird's-eye view: an overall picture of ground-based means to execute low-altitude aerial photography (LAAP) in archaeology. Archaeol Prospect 16:233–249. doi:10.1002/arp.354

Verhoeven G (2011) Taking computer vision aloft: archaeological three-dimensional reconstructions from aerial photographs with PhotoScan. Archaeol Prospect 18:67–73. doi:10.1002/arp.399

Verhoeven G (2012a) Methods of visualisation. In: Edwards HGM, Vandenabeele PV (eds) Analytical archaeometry: selected topics. Royal Society of Chemistry, Cambridge, pp 3–48

Verhoeven G (2012b) Near-infrared aerial crop mark archaeology: from its historical use to current digital implementations. J Archaeol Method Theory 19:132–160. doi:10.1007/s10816-011-9104-5

Verhoeven G (2012c) Straightforward archeological orthophotos from oblique aerial images. SPIE Newsroom. doi:10.1117/2.1201210.004506

Verhoeven G, Schmitt KD (2010) An attempt to push back frontiers: digital near-ultraviolet aerial archaeology. J Archaeol Sci 37:833–845. doi:10.1016/j.jas.2009.11.013

Verhoeven G, Loenders J, Vermeulen F, Docter R (2009a) Helikite aerial photography: a versatile means of unmanned, radio controlled, low-altitude aerial archaeology. Archaeol Prospect 16:125–138. doi:10.1002/arp.353

Verhoeven G, Smet P, Poelman D, Vermeulen F (2009b) Spectral characterization of a digital still camera's NIR modification to enhance archaeological observation. IEEE Trans Geosci Remote Sens 47:3456–3468. doi:10.1109/TGRS.2009.2021431

Verhoeven G, Doneus M, Briese C, Vermeulen F (2012a) Mapping by matching: a computer vision-based approach to fast and accurate georeferencing of archaeological aerial photographs. J Archaeol Sci 39:2060–2070. doi:10.1016/j.jas.2012.02.022

Verhoeven G, Taelman D, Vermeulen F (2012b) Computer vision-based orthophoto mapping of complex archaeological sites: the ancient quarry of Pitaranha (Portugal-Spain). Archaeometry 54:1114–1129. doi:10.1111/j.1475-4754.2012.00667.x

Vermeulen F (2002) The potenza valley survey (Marche). In: New developments in italian landscape archaeology: theory and methodology of field survey, land evaluation and landscape perception, pottery production and distribution. Proceedings of a three-day conference held at the University of Groningen, 13–15 Apr 2000. Archaeopress, Oxford, pp 104–106

Vermeulen F (2004) Fotografia aerea finalizzata nelle Marche centrali: un progetto integrato. Archeologia Aerea Studi di Aerotopografia Archeologica 1:91–118

Vermeulen F, Taelman D (2010) From cityscape to landscape in Roman Lusitania: the Municipium of Ammaia. In: Changing landscapes: the impact of Roman towns in the Western Mediterranean. Proceedings of the International Colloquium, Castelo de Vide, Marvão. 15–17 May 2008. AnteQuem, Bologna, pp 311–324

Wells J, Wells R (2012). Kite aerial photography in the near infra-red and ultra-violet. http://www.armadale.org.uk/phototech05.htm. Accessed 11 February 2013

Whittlesey JH (1973) Balloons, 'flying mattresses' and photography. Expedition 15:30–39

Wilson DR (1975) Photographic techniques in the air. In: Wilson DR (ed) Aerial reconnaissance for archaeology, vol 12, Research report series. The Council for British Archaeology, London, pp 12–31

Wilson DR (2000) Air photo interpretation for archaeologists, 2nd edn. Tempus, Stroud

Wolf PR, Dewitt BA (2000) Elements of photogrammetry with applications in GIS, 3rd edn. McGraw-Hill, Boston

Xu G, Jun-ichi Terai, Heung-Yeung Shum (2000) A linear algorithm for camera self-calibration, motion and structure recovery for multi-planar scenes from two perspective images. In: Proceedings of the IEEE conference on computer vision and pattern recognition. IEEE Computer Society Press, Los Alamitos, 13–15 June 2000, pp 474–479. doi:10.1109/CVPR.2000.854886

Yu G, Morel J-M (2011) ASIFT: an algorithm for fully affine invariant comparison. IPOL. doi:10.5201/ipol.2011.my-asift

Zantopp R (1995) Methode und Möglichkeiten der Luftbildarchäologie im Rheinland. In: Jürgen K (ed) Luftbildarchäologie in Ost- und Mitteleuropa/Aerial archaeology in Eastern and Central Europe: Internationales Symposium, Kleinmachnow, Land Brandenburg, 26–30 Sept 1994, vol 3, Forschungen zur Archäologie im Land Brandenburg. Verlag Brandenburgisches Landesmuseum für Ur- und Frühgeschichte, Potsdam, pp 155–163