Scales of oblique photographs updated

Eleonora Maset[a], Andrea Fusiello[a,*]

[a] *Dipartimento Politecnico di Ingegneria e Architettura, Universita degli Studi di Udine, Via Delle Scienze, 208, Udine, 33100, Italy*

Abstract

In this short note, we review the definition of photographic scale in the case of nadir and oblique images, and derive exact formulae for calculating the scale of oblique images in general and special cases.

Keywords: Photographic scale, Ground Sampling Distance, Oblique images

---

1. Introduction

Scale plays a versatile role in various aspects of photogrammetry and geospatial data processing.

In particular, scale variations in oblique imagery have gained increasing relevance in recent times, driven by the widespread availability of multi-perspective cameras like the Leica CityMapper and Vexcel Osprey systems. These variations can be substantial and need to be addressed during every phase of the photogrammetric workflow.

Image scale directly affects the accuracy of measurements made from oblique images. When determining the size or position of objects on the ground, it's essential to account for the scale factor. Accurate measurement requires the correct scaling of features in the images to their real-world dimensions (Höhle, 2008; Verykokou, 2020).

Traditional bundle adjustment methods assign equal weights to image point observations from different directions, but in aerial oblique images, varying ground sample distances and non-standard error distributions arise due to scale differences. This diminishes bundle adjustment accuracy. To overcome this challenge, Xie et al. (2016) incorporates pixel scale to calculate a re-weighting factor, reflecting the relative importance of individual measurements.

Another crucial aspect that demands careful scale consideration is motion compensation. This process corrects distortions caused by the movement of the imaging platform, like an aircraft or drone, during image capture. Cutting-edge methods like Adaptive Motion Compensation (AMC) (Dohr et al., 2022), that

*Corresponding author

Email addresses: `eleonora.maset@uniud.it` (Eleonora Maset), `andrea.fusiello@uniud.it` (Andrea Fusiello)

work in conjunction with CMOS sensors, rely on estimating the point-spread function behind image blurring, enabling image deconvolution. The size of this function directly relates to the extent of motion blur, mainly estimated from aircraft speed, image scale, and exposure time. It's noteworthy that different image scales result in varying degrees of motion blur, further underscoring the need to address scale variations.

2. General formulae for the photographic scale

The Ground Sampling Distance (GSD), is the 3D Euclidean distance between two adjacent pixel centres (whose distance in the image is 1 pixel) measured on the ground. Since "adjacent pixels" can be above and to the side, dependency on direction is implicit in this definition. It is also implied that the GSD depends on the position of the pixels considered.

The *photographic scale* is the ratio – denoted by $1 : m$ – between a distance measured in a photograph and the corresponding distance on the ground. The value m represents the so-called scale number. More precisely, we define two scale numbers as the ratio of the length of infinitesimal segments at a given point, in the two directions corresponding to the u and v axis in the image:

$$\begin{pmatrix} m_u \\ m_v \end{pmatrix} = \begin{pmatrix} \frac{\mathsf{d}\,S}{\mathsf{d}\,u} \\ \frac{\mathsf{d}\,S}{\mathsf{d}\,v} \end{pmatrix} \tag{1}$$

where $\mathsf{d}\,u$ is the length of the infinitesimal image segment in the u direction, $\mathsf{d}\,v$ is the length of the infinitesimal image segment in the v direction, $\mathsf{d}\,S$ is the length of the footprint (or back-projection) of the image segment on the ground. The differential is computed at the given point. With this definition, photographic scale is a 2-dimensional vector that varies from point to point in the image.

The relationship between the scale numbers and the GSD involves the pixels dimensions Δu and Δv (Kraus, 2007):

$$GSD = \begin{pmatrix} \Delta u & 0 \\ 0 & \Delta v \end{pmatrix} \begin{pmatrix} m_u \\ m_v \end{pmatrix} \tag{2}$$

where the u and v axis corresponds to columns and rows of the photosensors array, respectively.

In order to derive a general expression for the photographic scale in oblique images, we will take an analytical approach. The camera station is represented by the centre of projection (COP) that we set to

$$\mathbf{X}_0 = \begin{pmatrix} 0 \\ 0 \\ H \end{pmatrix} \tag{3}$$

in the ground coordinates system (Fig. 1), so that H is the height above the ground of the camera. The camera angular attitude is given in general by three Euler angles: ω, ϕ, κ, that determine a rotation matrix $\mathtt{R} = \mathtt{R}_X(\omega)\mathtt{R}_Y(\phi)\mathtt{R}_Z(\kappa)$. Let us assume for the time being that only ϕ (the rotation angle around the Y-axis), is non-zero. The corresponding rotation matrix is

$$\mathtt{R} = \begin{pmatrix} \cos(\phi) & 0 & \sin(\phi) \\ 0 & 1 & 0 \\ -\sin(\phi) & 0 & \cos(\phi) \end{pmatrix}. \tag{4}$$

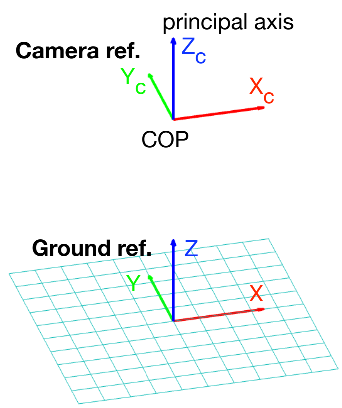

Figure 1: Ground and camera reference systems. By default the camera looks upward, so to make it look downward one needs to rotate it by π around X_c or Y_c axis.

The camera is represented in homogeneous coordinates by a 3×4 matrix:

$$\mathtt{P} = \mathtt{K}[\mathtt{R}, -\mathtt{R}\,\mathbf{X}_0] \tag{5}$$

where $\mathtt{K}$ is 3×3 diagonal matrix that contains on the diagonal the focal length (or principal distance) f:

$$\mathtt{K} = \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{6}$$

We are assuming that the image reference system has its origin in the principal point. With these settings, $\mathtt{P}$ writes:

$$\mathtt{P} = \begin{pmatrix} f\,\cos(\phi) & 0 & f\,\sin(\phi) & -H\,f\,\sin(\phi) \\ 0 & f & 0 & 0 \\ -\sin(\phi) & 0 & \cos(\phi) & -H\,\cos(\phi) \end{pmatrix}. \tag{7}$$

The action of the camera matrix (in homogeneous coordinates) is to take a 3D point $\mathbf{X} = (X, Y, Z, 1)$ in ground coordinates system and project it onto a point $\mathbf{x}$ in the image plane (See Appendix A):

$$W\,\mathbf{x} = \mathtt{P}\,\mathbf{X} \tag{8}$$

where the scalar W is the *depth* of the ground point, which is defined as the distance (in meters) of the point from the plane through the COP parallel to the image plane. The depth W is simply the third coordinate of the ground point represented in the camera reference frame.

Since the ground plane has equation $Z = 0$, the homography that maps the ground plane to the image plane is obtained by deleting the third column of $\mathtt{P}$:

$$\mathtt{P}_3 = \begin{pmatrix} f\,\cos(\phi) & 0 & -H\,f\,\sin(\phi) \\ 0 & f & 0 \\ -\sin(\phi) & 0 & -H\,\cos(\phi) \end{pmatrix}. \tag{9}$$

Its inverse, $\mathtt{P}_3^{-1}$ maps from the image plane to the ground plane:

$$\mathtt{P}_3^{-1} = \begin{pmatrix} \frac{\cos(\phi)}{f} & 0 & -\sin(\phi) \\ 0 & \frac{1}{f} & 0 \\ -\frac{\sin(\phi)}{H\,f} & 0 & -\frac{\cos(\phi)}{H} \end{pmatrix}. \tag{10}$$

Considering a generic image point of homogeneous coordinates

$$\mathbf{x} = \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \tag{11}$$

its back-projection onto the ground plane can be computed with $\mathtt{P}_3^{-1}\mathbf{x}$. After perspective division it has planimetric coordinates:

$$\bar{\mathbf{X}} = \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \dfrac{H\,\left(f\,\sin(\phi) - u\,\cos(\phi)\right)}{f\,\cos(\phi) + u\,\sin(\phi)} \\ -\dfrac{H\,v}{f\,\cos(\phi) + u\,\sin(\phi)} \end{pmatrix} \tag{12}$$

while $Z = 0$ by definition.

The derivatives of $\bar{\mathbf{X}}$ are readily computed:

$$\frac{\mathsf{d}\,\bar{\mathbf{X}}}{\mathsf{d}\,u} = \begin{pmatrix} -\dfrac{H\,f}{\left(f\,\cos(\phi) + u\,\sin(\phi)\right)^2} \\ \dfrac{H\,v\,\sin(\phi)}{\left(f\,\cos(\phi) + u\,\sin(\phi)\right)^2} \end{pmatrix} \tag{13}$$

$$\frac{\mathsf{d}\,\bar{\mathbf{X}}}{\mathsf{d}\,v} = \begin{pmatrix} 0 \\ -\dfrac{H}{f\,\cos(\phi) + u\,\sin(\phi)} \end{pmatrix}. \tag{14}$$

Since the scale is the ratio of lengths, we take the norm ($S = \|\mathsf{d}\,\bar{\mathbf{X}}\|$) thus obtaining:

$$m_u = \frac{\|\mathsf{d}\,\bar{\mathbf{X}}\|}{\mathsf{d}\,u} = \frac{H\,\sqrt{f^2 + \sin^2(\phi)\,v^2}}{\left(f\,\cos(\phi) + u\,\sin(\phi)\right)^2} \tag{15}$$

$$m_v = \frac{\|\mathsf{d}\,\bar{\mathbf{X}}\|}{\mathsf{d}\,v} = \frac{H}{|f\,\cos(\phi) + u\,\sin(\phi)|}. \tag{16}$$

These are the general formulae for the scale numbers when the camera is tilted by ϕ around the Y axis (which is parallel to the v axis in the image plane) as in Fig. 2.

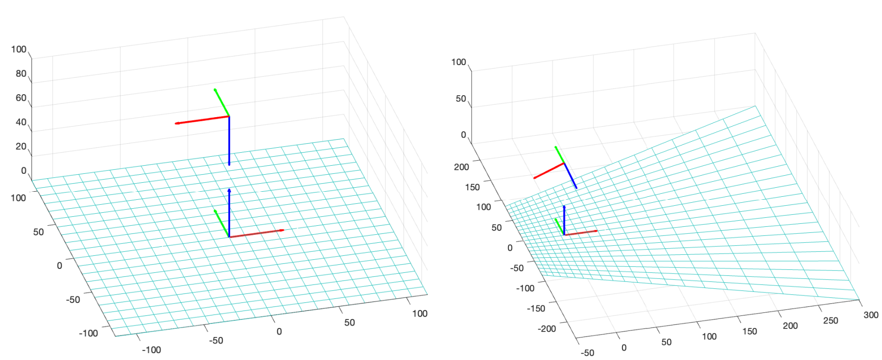

Figure 2: A simulation of nadir $\phi = \pi$ and oblique $\phi = \pi + \pi/8$ imaging. The ground and camera reference frames are shown (R=X, G=Y, B=Z). A grid of 100×100 pixels cells is projected onto the ground plane to visualise the image footprint and the GSD variation.

In the principal point ($v = 0, u = 0$), Formulae (15) and (16) reduces to

$$m_u = \frac{H}{f\ \cos^2(\phi)} \tag{17}$$

$$m_v = \frac{H}{f\ \cos(\phi)}. \tag{18}$$

Expressions (17) and (18) characterise the scale number for oblique images as a function of H/f, which is the scale number for a nadir image with the same height above ground H. Please note that *such expressions are valid only at the principal point.*

Simpler formulae are obtained using the depth of the ground point W.

According to (8), $W = \mathbf{p}_3\mathbf{X}$, where $\mathbf{p}_3$ is the last row of P (see Appendix A). In our specific case $\mathbf{X}$ belongs to the ground plane ($Z = 0$), so: $\mathbf{X} = [\bar{\mathbf{X}}; 0; 1]$, thus obtaining:

$$W = -\frac{H\,f}{f\ \cos(\phi) + u\ \sin(\phi)}. \tag{19}$$

We solve for H in the resulting expression, thus getting:

$$H = -W\ \cos(\phi) - \frac{W\,u\ \sin(\phi)}{f} \tag{20}$$

and substitute it in the previous relationships, obtaining

$$m_u = \frac{\|\mathsf{d}\,\bar{\mathbf{X}}\|}{\mathsf{d}\,u} = \frac{W\ |\cos(\phi)|}{f} \tag{21}$$

$$m_v = \frac{\|\mathsf{d}\,\bar{\mathbf{X}}\|}{\mathsf{d}\,v} = \frac{W}{f}. \tag{22}$$

Although these two formulae are equivalent to (15)-(16), they are more convenient to compute as they involve a camera constant (f) and W, which is simply the third coordinate of the ground point represented in the camera reference frame. The dependency on u, v and partially the tilt angle are already accounted for by W. Nonetheless, we could not find any evidence of this definition in the literature.

The case we just illustrated was a special one. The most general case would entail a full rotation matrix. Following the widespread convention (Sigle and Heuchel, 2001) ~~used~~, e.g., in PATB (Klein, 2009) and Pix4D (Pix4D, 2021) – which corresponds to the transpose of the standard recommended by Rosenfield (1959) in expression (2.7) – a full rotation matrix where all the three angles ω, ϕ, κ are non-zero writes:

$$\mathtt{R} = \begin{pmatrix} \cos(\kappa)\cos(\phi) & -\cos(\phi)\sin(\kappa) & \sin(\phi) \\ \cos(\omega)\sin(\kappa) + \cos(\kappa)\sin(\omega)\sin(\phi) & \cos(\kappa)\cos(\omega) - \sin(\kappa)\sin(\omega)\sin(\phi) & -\cos(\phi)\sin(\omega) \\ \sin(\kappa)\sin(\omega) - \cos(\kappa)\cos(\omega)\sin(\phi) & \cos(\kappa)\sin(\omega) + \cos(\omega)\sin(\kappa)\sin(\phi) & \cos(\omega)\cos(\phi) \end{pmatrix}. \tag{23}$$

These settings lead to the following most general formulae (Appendix B):

$$m_u = \frac{H\sqrt{f^2+v^2+\cos^2(\phi)\left(-f^2+f^2\cos^2(\omega)-v^2\cos^2(\omega)-f\,v\,\sin(2\omega)\right)}}{\left(u\,\sin(\phi)+f\,\cos(\omega)\,\cos(\phi)-v\,\cos(\phi)\,\sin(\omega)\right)^2} \tag{24}$$

$$m_v = \frac{H\sqrt{f^2\cos^2(\phi)+u^2-u^2\cos^2(\omega)\cos^2(\phi)+f\,u\,\cos(\omega)\,\sin(2\phi)}}{\left(u\,\sin(\phi)+f\,\cos(\omega)\,\cos(\phi)-v\,\cos(\phi)\,\sin(\omega)\right)^2} \tag{25}$$

Again, it is to be noted that using W leads to a *much simpler formulation*:

$$m_u = \frac{W\,|\cos(\phi)|}{f} \tag{26}$$

$$m_v = \frac{W\sqrt{1-\cos^2(\phi)\sin^2(\omega)}}{f} \tag{27}$$

where the expression for m_u is formally identical to (21) and the expression for m_v contains an additional term that vanishes for $\omega = 0$.

3. Comparison with known formulae

We will start by examining the classic case in aerial photogrammetry depicted in Fig. 3 where i) the ground is flat (i.e., it has constant elevation) and ii) the image plane and ground plane are parallel (this corresponds to a *nadir* aerial photograph). It is immediate to verify that

$$m_u = m_v = \frac{H}{f} = \frac{W}{f} \tag{28}$$

from (15) - (16) and (21) - (22) respectively with $\sin(\phi) = 0$.

In this particular case, the scale is constant because W is constant (and equal to H), the image plane and the ground plane being parallel. The scale does not depend on the point (nor on the direction) and it is constant over the image.

It also turns out that, in this special case:

$$\frac{H}{f} = \frac{L}{l} = \frac{S}{s} \tag{29}$$

by triangle similarity in Fig. 3.

If either condition i) or ii) fails, the scale is no longer constant. Failure of i) corresponds to a topographic variation of the surface (relief) and it is not the focus of this short note. Failure of condition ii) corresponds to oblique images.

All previous known formulae consider a special case of oblique image (see Fig. 4), where the camera has only one inclination angle γ that specify a rotation around an axis orthogonal to the *principal plane*, i.e., the vertical plane containing the principal axis. Let us assume that this is also the Y-axis of the camera. The camera Z-axis is contained in the principal plane (it coincides with the principal axis), while the Y-axis is perpendicular to the screen. The segments s and S and the angles α and γ lie in the principal plane.

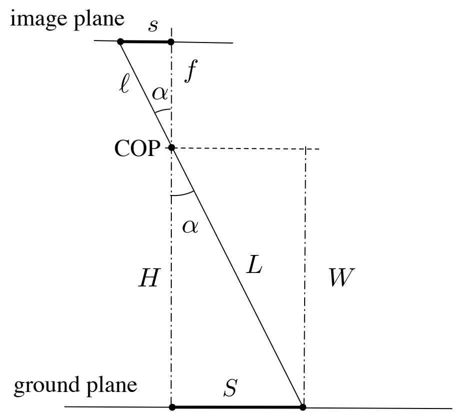

Figure 3: Geometric construction for defining the GSD in the case of nadir images (with flat ground).

3.1. Digital Photogrammetry approaches

In ~~this~~ the special case depicted in Fig. 4, Höhle (2008) defines the scale number as:

$$m_L = \frac{L}{\ell} = \frac{H\cos(\alpha)}{f\cos(\alpha+\gamma)} \tag{30}$$

since

$$\ell = \frac{f}{\cos(\alpha)} \tag{31}$$

$$L = \frac{H}{\cos(\alpha+\gamma)} \tag{32}$$

where γ is the tilt angle, i.e., the angle formed by the camera's principal axis and the vertical direction ($\gamma = 0$ for nadir views), and α is the angle between the ray of the given image point and the principal axis (see Fig. 4). It is easy to see that with this definition the photographic scale varies across oblique images.

From the similarity of the triangles in Fig. 4 it follows that

$$m_L = \frac{L}{\ell} = \frac{W}{f}. \tag{33}$$

In the formula with W the scale variation is due to W itself, since the image plane and the ground plane are not parallel.

It is easy to see that the definition of m_L in (30) corresponds to our m_v both in (16) and (22). The latter derives from (33); as for the former, one needs to substitute $u = f\tan(\alpha)$, thus:

$$m_v = \frac{H\ \cos(\alpha)}{|f\ \cos(-\phi+\alpha)|} \tag{34}$$

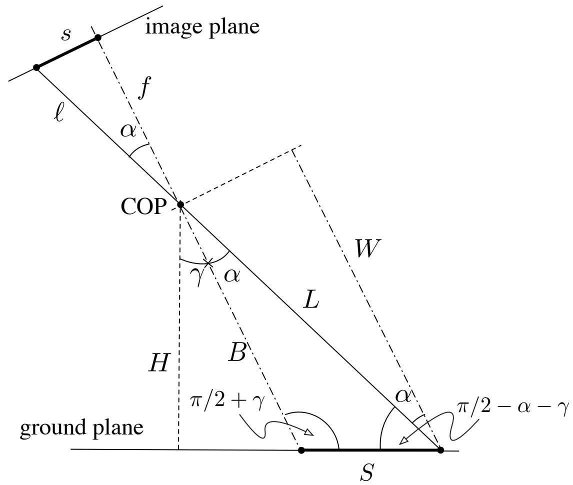

Figure 4: Geometric construction for defining the GSD in the oblique case. This drawing is a view of the principal plane.

where $\phi = \pi - \gamma$. In this special case ($\omega = 0, \kappa = 0$), the v direction is orthogonal to the principal plane.

In this same situation one could also try to compute the scale by differentiation. The following relationships can be written with reference to Fig. 4:

$$s = \frac{f \, \sin(\alpha)}{\cos(\alpha)} \tag{35}$$

$$B = \frac{H}{\cos(\gamma)} \tag{36}$$

$$S = \frac{B \, \sin\alpha}{\sin(\pi/2 - \alpha - \gamma)} = \frac{H \, \sin(\alpha)}{\cos(\gamma + \alpha) \, \cos(\gamma)} \tag{37}$$

where (37) derives from the law of sines. Then:

$$\frac{\mathsf{d}\, S}{\mathsf{d}\, \alpha} = \frac{2\, H}{\cos(2\,\gamma + 2\,\alpha) + 1} \tag{38}$$

$$\frac{\mathsf{d}\, s}{\mathsf{d}\, \alpha} = -\frac{f}{\sin^2(\alpha) - 1} \tag{39}$$

and

$$m_S = \frac{\mathsf{d}\, S}{\mathsf{d}\, s} = -\frac{2\, H \, \left(\sin^2(\alpha) - 1\right)}{f \, \left(\cos(2\,\gamma + 2\,\alpha) + 1\right)} = \frac{H \cos^2(\alpha)}{f \cos^2(\gamma + \alpha)}. \tag{40}$$

It is easy to see that this definition of m_S in (40) corresponds to m_u in (15) where one needs to set $v = 0$ and $u = f \tan(\alpha)$ thus:

$$m_u = \frac{H \cos^2(\alpha)}{f \cos^2(-\phi + \alpha)} \tag{41}$$

where we identify $\phi = \pi - \gamma$.

Formula (30) was presented by Höhle (2008) as the GSD for oblique images, and later used by many authors (e.g. (Xie et al., 2016; Pepe et al., 2018; Balamuralidhar et al., 2021). In a later paper Höhle (2013) renamed it GSD_{ortho}, and gave a new formula for what he called GSD_{ground}, that matches (40).

3.2. Conventional Analytical Photogrammetry approaches

Conventional approaches to scale determination in oblique images, originating from analogue-analytical photogrammetry (Lane, 1950; Moffitt, 1967; Moffitt and Mikhail, 1980; Wolf, 1983), focus on scale along two specific directions: the x-scale, perpendicular to the principal line's direction and the y-scale, which is along the former. The principal line is defined as the intersection of the principal plane with the image plane, and its orientation depends on the camera's angular attitude.

For example, in (Verykokou, 2020), the x-scale s_x is defined as

$$s_x = \frac{T'T_1'}{TT_1} \tag{42}$$

where T' is the projection on the image plane of a world point T, T_1' is the projection of T' on the principal line and T_1 is the projection of T on the principal plane. For a camera with only one inclination angle γ (Fig. 4), Verykokou (2020) derives the following equation:

$$s_x = \frac{f \cos(\alpha + \gamma)}{(H - \Delta H) \cos(\alpha)}. \tag{43}$$

Assuming that T lies on the ground ($\Delta H = 0$), it is easy to see that (43) is the inverse of m_v expressed in (34).

The y-scale, computed along the principal line, is defined as

$$s_y = \frac{T'T_2'}{TT_2} = \frac{d^2 \cos^2(\theta)}{f(H - \Delta H)} \tag{44}$$

where T' and T_2' are the projections on the oblique image of two world points T and T_2 belonging to the principal plane and the segment $T'T_2'$ lies on the principal line. The angle formed by the camera's principal direction and the true horizon line is denoted with θ and d is the distance between T' and the horizon point. Furthermore, the following relation holds (Verykokou, 2020):

$$\frac{d \cos(\theta)}{f} = \frac{\sin(\theta - \alpha)}{\cos \alpha} \tag{45}$$

Substituting (45) in (44) and given that $\theta = \pi/2 - \gamma$ and $\sin(\pi/2 - \gamma - \alpha) = \cos(\gamma + \alpha)$, it results $s_y = m_u^{-1}$, with m_u expressed as in (41) for $v = 0$. This clarifies that in fact the expression for s_y holds only for points belonging to the principal line.

The classic formulae (Lane, 1950; Verykokou, 2020) differ from our approach in two key aspects: i) they decompose scale along different *extrinsic* directions (i.e. directions depending on the external orientation of the camera), and ii) they reference elements like the principal line, the true horizon line and the horizon point, which are not readily available in our formulation. In contrast, our approach takes a different perspective by breaking down scale variation along two *intrinsic* directions, namely along the u and v axes of image coordinates.

4. Conclusions

Scale plays a versatile role in various aspects of photogrammetry and geospatial data processing. However, determining the scale of oblique images can be a somewhat intricate task in modern *digital* photogrammetric procedures. This obstacle arises because available formulae involves computing geometrical elements that are typically associated with *analytical* methods and may not be readily available. In addition, known formulae refers only to a special case where the cameras has only one inclination angle (Fig. 4).

The primary objective of this short note is to introduce a *general* alternative definition of image scale – based on image space derivatives – as outlined in equations (26) -(27). This approach seamlessly integrates with digital photogrammetric procedures (e.g., motion compensation, bundle adjustment), distinguishing itself from prior analytical methods that predominantly relied on 2D graphical interpretations for image scale determination. By presenting this unique perspective, our approach has the potential to provide fresh insights and enhance our understanding of the inherent characteristics of oblique photographs in the context of modern digital photogrammetry.

Acknowledgement

The authors are grateful to Fabio Crosilla and Isabella Toschi for their valuable comments. This study was carried out within the Interconnected Nord-Est Innovation Ecosystem (iNEST) and received funding from the European Union Next-GenerationEU (Piano Nazionale di Ripresa e Resilienza (PNRR) – Missione 4 Componente 2, Investimento 1.5 – D.D. 1058 23/06/2022, ECS00000043). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

References

Balamuralidhar, N., Tilon, S., Nex, F., 2021. Multeye: Monitoring system for real-time vehicle detection, tracking and speed estimation from uav imagery on edge-computing platforms. Remote Sensing 13. doi:10.3390/rs13040573.

Dohr, S., Muick, M., Schachinger, B., Gruber, M., 2022. Adaptive motion compensation: A modern approach. White paper - Vexcel Imaging GmbH, Graz Austria. URL: https://www.vexcel-imaging.com/wp-content/uploads/2022/04/2022-03-31_WhitePaper_AdaptiveMotionCompensation_A-modern-approach.pdf.

Faugeras, O., 1993. Three-Dimensional Computer Vision: A Geometric Viewpoint. The MIT Press, Cambridge, MA.

Höhle, J., 2008. Photogrammetric measurements in oblique aerial images. Photogrammetrie, Fernerkundung, Geoinformation , 7–14.

Höhle, J., 2013. Oblique aerial images and their use in cultural heritage documentation. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XL-5/W2, 349–354. doi:10.5194/isprsarchives-XL-5-W2-349-2013.

Klein, T., 2009. K2 - photogrammetry. http://www.k2-photogrammetry.de/products/patb.html. Accessed: October 2023.

Kraus, K., 2007. Photogrammetry - Geometry from Images and Laser Scans - 2nd edition. Walter de Gruyter.

Lane, B., 1950. Scales of oblique photographs. Photogrammetric Engineering 16, 409–414.

Moffitt, F., 1967. Photogrammetry. volume 12 of *Geodetickỳ a kartografickỳ obzor*. International textbook in civil engineering. International Textbook Company.

Moffitt, F., Mikhail, E., 1980. Photogrammetry. Series in civil engineering, Harper & Row.

Pepe, M., Fregonese, L., Scaioni, M., 2018. Planning airborne photogrammetry and remote-sensing missions with modern platforms and sensors. European Journal of Remote Sensing 51, 412–436. doi:10.1080/22797254.2018.1444945.

Pix4D, 2021. How are the internal and external camera parameters defined? https://support.pix4d.com/hc/en-us/articles/202559089-How-are-the-Internal-and-External-Camera-Parameters-defined. Accessed: October 2023.

Rosenfield, G.H., 1959. The problem of exterior orientation in photogrammetry. Photogrammetric Engineering 25, 536–553.

Sigle, M., Heuchel, T., 2001. Match-AT: Recent developments and performance. Photogrammetrix Week 2001. URL: https://ifpwww.ifp.uni-stuttgart.de/publications/phowo01/Sigle.pdf.

Verykokou, S., 2020. Georeferencing procedures for oblique aerial images. Doctoral Dissertation. doi:10.26240/heal.ntua.18725.

Wolf, P., 1983. Elements of Photogrammetry. 2nd ed., Mc-Graw-Hill, NY.

Xie, L., Hu, H., Wang, J., Zhu, Q., Chen, M., 2016. An asymmetric reweighting method for the precision combined bundle adjustment of aerial oblique images. ISPRS Journal of Photogrammetry and Remote Sensing 117, 92–107. doi:10.1016/j.isprsjprs.2016.03.017.

Appendix A. Perspective projection equation

Perspective projection in homogeneous coordinates writes (Faugeras, 1993):

$$\lambda \mathbf{x} = \mathtt{P}\,\mathbf{X} \tag{A.1}$$

where the λ is a scale factor that accounts for the homogeneous representation. To derive an expression for λ let us expand the product $\mathtt{P}\,\mathbf{X}$, where $\mathtt{P} = \mathtt{K}[\mathtt{R}, \mathbf{t}]$, $\mathbf{t} = -\mathtt{R}\mathbf{X}_0$, $\mathtt{R}$ is a rotation matrix and $\mathtt{K}$ is given by (6):

$$\mathtt{P}\,\mathbf{X} = \begin{pmatrix} f\,t_1 + R_{1,1}\,X_1\,f + R_{1,2}\,X_2\,f + R_{1,3}\,X_3\,f \\ f\,t_2 + R_{2,1}\,X_1\,f + R_{2,2}\,X_2\,f + R_{2,3}\,X_3\,f \\ t_3 + R_{3,1}\,X_1 + R_{3,2}\,X_2 + R_{3,3}\,X_3 \end{pmatrix} \tag{A.2}$$

Since $\mathbf{x}_3 = 1$, it turns out that $\lambda = t_3 + R_{3,1}\,X_1 + R_{3,2}\,X_2 + R_{3,3}\,X_3$, and this is the Z coordinate of the point represented in the camera reference system. i.e., the depth of the point W.

Appendix B. Derivation of general formulae

All the derivation have been done with the help of the Matlab Symbolic Toolbox, with the following code.

```
syms omega phi kappa alpha f  H u v W  real

K = [f 0 0 ; 0 f 0; 0 0 1];

a = [0 phi 0];

Rx =  [1     0          0
       0 cos(a(1))  -sin(a(1))
       0 sin(a(1))  cos(a(1))]; % omega

Ry = [cos(a(2))  0 sin(a(2))
     0           1        0
    -sin(a(2)) 0  cos(a(2))]; % phi

Rz = [cos(a(3)) -sin(a(3))  0
      sin(a(3))   cos(a(3))  0
     0             0          1]; % kappa

R = Rx*Ry*Rz;

COP = [0 0 H]';
```

```
P = K * [R , -R*COP];

X_bar = inv(P(:,[1 2 4])) * [u v 1]';
X_bar = (simplify(X_bar(1:2)./ X_bar(3), 'Steps', 20));

% comment these lines for formulae without W
H = (simplify(solve(W == (P(3,:) * [X_bar; 0; 1]), H), 20));
X_bar = (simplify(subs(X_bar), 20));

Du = (simplify(jacobian(X_bar, u),'Steps',20));
Dv = (simplify(jacobian(X_bar, v),'Steps',20));

mu = (simplify(norm(Du),'Steps',20));
mv = (simplify(norm(Dv),'Steps',20));

% uncomment these lines for special case formulae
%    v = 0;
%    u = f * tan(alpha);
%    mu = simplify(subs(mu),'Steps',20);
%    mv = simplify(subs(mv),'Steps',20);
```