Insights into the PAS Pana Stitching Algorithm

Richard Ladstädter, Karlheinz Gutjahr, Erich Riegler, Roland Perko

JOANNEUM RESEARCH DIGITAL, Remote Sensing and Geoinformation, Steyrergasse 17, 8010 Graz, Austria
{richard.ladstaedter,karlheinz.gutjahr,erich.riegler,roland.perko}@joanneum.at

Keywords: Phase One PAS Pana, Wide-Area Camera, Stitching Algorithm.

Abstract

In this paper, we describe a modern, efficient, accurate and reliable stitching algorithm that JOANNEUM RESEARCH (JR) has developed for the Phase One PAS Pana multi-camera system. The goal is to generate one large output image (four-channels including RGB and NIR) in nadir geometry. Based on the experience from a workflow developed in 2024, an improved stitching algorithm has been designed in 2025 which uses so called “controlled” projective transformations and simultaneous adjustment of all available image overlaps. We give insight into the role of projective transformations for fusing the images of calibrated cameras in a virtual image plane and how we moved on from the classical eight parameter transformation to a “controlled” transformation using only six, physically meaningful parameters. Experiments with real image data from a flight mission over Denver, USA demonstrate the accuracy and stability of the new stitching algorithm.

1. Introduction

1.1 The PAS Pana camera system

The PAS Pana camera system (Phase One, 2024) consists of seven camera heads based on the Phase One iXM-RS150 medium format camera equipped with a 3.76µm CMOS sensor and a 150mm (RGB) or 70mm (NIR) lens, see Table 1.

Camera	Abbrev.	c [mm]	Tilt [deg]	Orientation
RGB 1 (Left)	L2	146	+27	Portrait
RGB 2 (Left)	L1	146	+13	Portrait
RGB 3 (Nadir)	ND	146	0	Portrait
RGB 4 (Right)	R1	146	-13	Portrait
RGB 5 (Right)	R2	146	-27	Portrait
NIR 1 (Left)	IL	70	+14	Landscape
NIR 2 (Right)	IR	70	-14	Landscape

Table 1. PAS Pana camera arrangement.

Only the middle RGB camera ND is looking into nadir direction, all other cameras (L2, L1, R1, R2, IL, IR) are oblique (see Figure 1). Together, they have a field of view (FOV) of 70 degree across track and 20 degree along track for RGB. The RGB to NIR ratio is 2.1 due to the different focal lengths used.

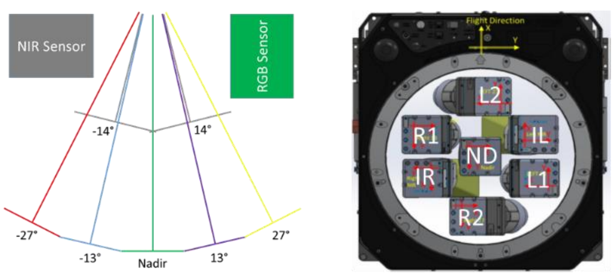

Figure 1. PAS Pana camera oblique viewing angles (left) and placement in the PAS Pana cone (right).

This camera arrangement produces a large NIR footprint, which is covered at full width by the five high-resolution RGB footprints, see Figures 2 & 3. Therefore, the PAS Pana system represents a “wide-area” camera, which allows for very efficient mapping of larger areas. At a flying height of 8000 m above ground, a ground sampling distance (GSD) of 21 cm is achieved within a 11 km swath width.

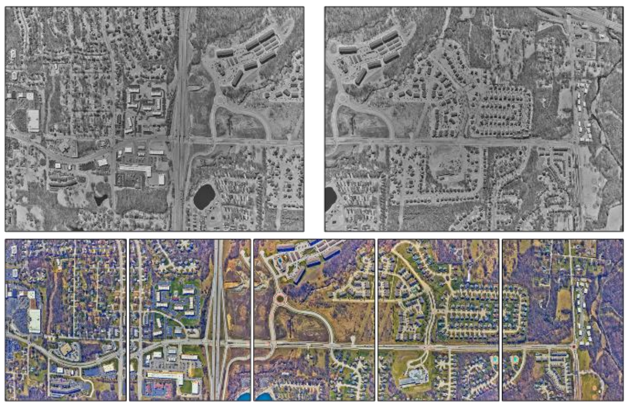

Figure 2. NIR (top) and RGB (bottom) images for a PAS Pana shot, © 2024 Surdex, a Bowman Company.

Figure 3. Coverage of the RGB and NIR images for the sample PAS Pana shot.

Note that the single NIR/NIR image overlap (~33%) is large compared to the four RGB/RGB overlaps (~15%) and that the RGB 3 image overlaps both NIR 1 and NIR 2 images.

1.2 Motivation

A first beta version of a PAS Pana stitching software has already been implemented in early 2024. PAS Pana shots of a test mission over Denver, USA have been processed by JR. The resulting fused images have been fed into an aerial triangulation workflow and finally evaluated for geometric accuracy (Jacobsen et al., 2025).

With this beta version we wanted to demonstrate the following advantages of such fused images in nadir geometry:

- One large four channel image instead of seven images reduces complexity in the automated aerial triangulation (AAT) workflow.
- Atmospheric refraction can be modelled by standard radial parameters which is not the case for the original, oblique images.
- Boresight angles calculated for the nadir RGB can directly be used for the fused image.

On the other hand, we have noticed that this beta workflow can be substantially improved by implementing these new features:

- Make use of all available image overlaps (include also RGB/RGB overlaps), see also Figure 4.
- Perform a simultaneous tie point adjustment of all overlaps (instead of sequential NIR/NIR fusion and RGB to NIR registration steps).
- Single resampling step from the sensor images to the fused output images using feathering/alpha blending for smooth transition between sensors.
- Avoid over-parameterization of the (full) projective transformations (FPT) with eight parameters by reducing them to only six parameters in a so called "controlled" projective transformation (CPT).

The last item is very important as we want to get maximum stability for the geometry of the fused image. For the PAS Pana, we do not have one single sensor covering the whole fused image format but limited image overlaps. From the analysis of the stitching parameters of the beta workflow, we can observe large variations between neighbouring images. Those variations are caused by strong correlations between some of the eight projective parameters (e.g. between scale and translation in across flight direction) and have introduced systematic errors in the fused image in the magnitude of up to 5 pixels.

1.3 General description of the new stitching algorithm

The new stitching algorithm has been developed in 2025, and the software has been released in March 2026 to Phase One. The stitching process is composed of several steps:

1) Distortion removal (based on lab or field calibration)
2) Tie point matching (coarse and fine)
3) Parameter adjustment using all sensor overlaps
4) Resampling of the fused image in the virtual image plane (VIP)

The workflow starts with the removal of the image distortion by using the calibration data given for each Phase One cameras as a Brown-Conradi parameter set (named "Australis"). Distortion removal is straightforward, but the orientation of the input image has to be considered correctly. More general, the quality of the calibration data is essential for the geometric accuracy of the fused output image.

Tie point matching will then be done in the distortion free images and can be done on a "coarse" or "fine" level, depending on the quality of the approximate transformation parameters. For a new PAS Pana system, only default transformation parameters are known, but those parameters can deviate significantly (several tens of pixels) from the actual ones. Using the "coarse" matching mode, images are downscaled and a large search radius is applied. The resulting transformation is therefore accurate only up to a few pixels.

Knowing the "coarse" transformation allows to apply a second, "fine" matching step in the full resolution image. Using smaller reference and search windows allows reaching sub-pixel accuracy. Outliers are rejected automatically by estimating residual vectors for a projective transformation derived from the set of corresponding points. Typically, coarse matching is only done once for a selected reference shot, coarse transformation parameters are stored and will then be re-used to initialize fine matching for all other images of a mission. Fine transformation parameters can also be stored and re-used, e.g. to bridge low textured areas where the tie point number is too low.

In this paper we will focus on step 3) being the core algorithm of the stitching process. In this step, the projective transformation parameters from each sensor to the VIP are determined for all seven cameras. In contrast to the previous version, the new PAS Pana stitching workflow performs simultaneous adjustment using all image overlaps (see Figure 4). This approach ensures the highest possible accuracy by using all available, highly redundant measurements in the adjustment procedure. Note, that such an advanced stitching algorithm developed for the Vexcel UltraCam can already be found in (Ladstädter et al., 2010) using tie point measurements between the low-resolution PAN channel and high resolution RGB channel.

For the PAS Pana, the seven cameras have a total of eleven possible overlaps (see Figure 4). Tie points are matched in one overlap between the NIR sensors, four overlaps (2-5) between the RGB sensors and six overlaps (6-11) between the NIR and RGB cameras. Parameters of only six projective transformations (CPT 1, …, CPT 7) have to be determined as one of them (for the RGB 3 camera) needs to be fixed to a unit transformation to avoid singularity of the normal equations.

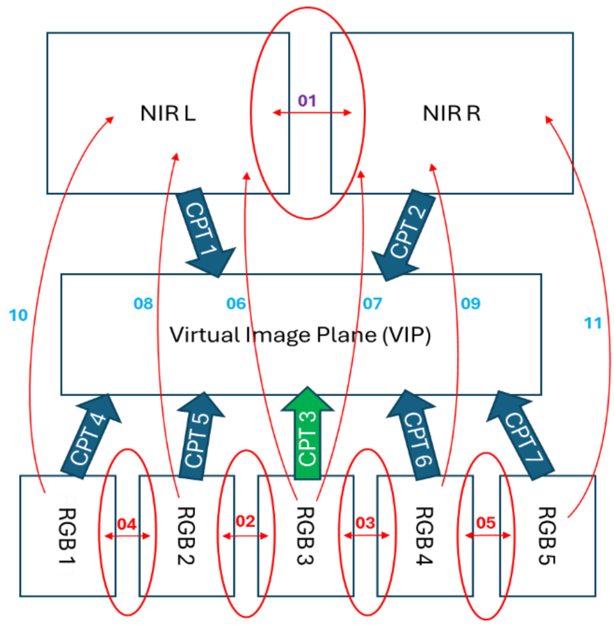

Figure 4. PAS Pana stitching scheme.

Note that this stitching scheme can easily be applied to any other multi-camera system by simply adapting to available sensors and overlap regions.

The adjustment starts with an initial solution for all transformation parameters and iterates until the residual vectors between tie point measurements in all overlaps are minimized. After convergence, adjusted parameters have been determined describing the transformation between each sensor to the VIP.

In Figure 5, tie point residuals are shown for a shot taken in the Denver 20cm mission. RGB/RGB residuals are shown in blue, RGB/NIR residuals in green and NIR/NIR residuals in magenta. Some small but systematic residuals remain which indicates some imperfections in the field calibration data used for that PAS Pana shot.

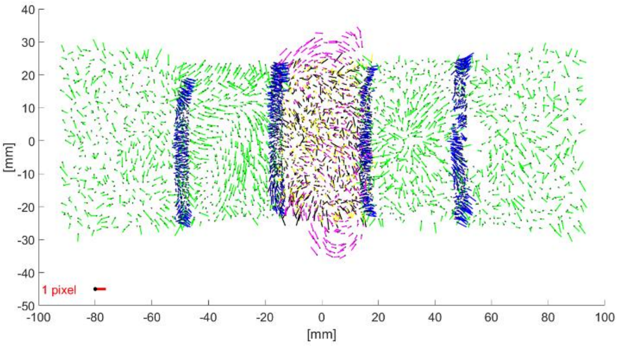

Figure 5. Residuals in the VIP after adjustment (Denver 20cm).

For the resampling of the fused image (step 4), a predefined mapping is used to determine which pixel in the VIP is mapped to which sensor and in the overlapping areas alpha blending and feathering is applied for smooth transition between the images. This approach allows for a single resampling step to generate the fused RGBI output image. This ensures not only an optimized image quality but also guarantees fastest possible image generation.

After this general workflow description, we want to focus on the theory of projective transformations in the next section.

2. Theory of 2D projective transformations

The concept of projective transformations between images forms the basic mathematical model for the stitching algorithm and shall be reviewed in more detail in this section. Some drawbacks in the standard, "full" projective transformation (FPT) will later be handled in the next section about "controlled projective transformations" (CPT).

2.1 Basic formulation of projective transformations

Equations (1) and (2) describe a projective transformation between two images (or 2D point sets) in homogeneous coordinates:

$$\bar{x}_2 \cdot w = \boldsymbol{P}_8 \cdot \bar{x}_1 \tag{1}$$

$$\bar{x}_1 = \begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix};\ \bar{x}_2 = \begin{pmatrix} x_2/w \\ y_2/w \\ 1 \end{pmatrix};\ \boldsymbol{P_8} = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & 1 \end{bmatrix} \tag{2}$$

The 3x3 matrix $\boldsymbol{P}_8$ contains only eight independent parameters as the transformation is only defined up to scale and the vector $\bar{x}_2$ has to be reconstructed by a division with the scalar w. The projective matrix $\boldsymbol{P}_8$ can be estimated from a set of tie points matched between two images. It can be shown that such a 2D projective transformation allows the transfer of points between the images of two (even uncalibrated) cameras separated by a baseline b but only for 3D points situated on a common plane (i.e. a flat 3D object). In computer vision this is called a 2D-homography (see Hartley and Zisserman, 2004).

Using a QR-decomposition, the matrix $\boldsymbol{P}_8$ can be decomposed into an orthogonal (rotational) matrix $\boldsymbol{R}$ and an upper triangular matrix $\boldsymbol{M}$ representing an affine transformation (composed of two translation parameters, two scale parameters m_x, m_y and one shear parameter s).

$$\boldsymbol{P}_8 = \boldsymbol{T} \cdot \boldsymbol{S} \cdot \boldsymbol{R} = \boldsymbol{M} \cdot \boldsymbol{R} \tag{3}$$

$$\boldsymbol{T} = \begin{bmatrix} 1 & s & dx \\ 0 & 1 & dy \\ 0 & 0 & 1 \end{bmatrix};\ \boldsymbol{S} = \begin{bmatrix} m_x & 0 & 0 \\ 0 & m_y & 0 \\ 0 & 0 & 1 \end{bmatrix};\ \boldsymbol{R} = \boldsymbol{R}(\omega, \varphi, \kappa) \tag{4}$$

2.2 2D homography for calibrated cameras

Geometrically, a 2D homography can also be interpreted as the transformation between cameras separated by a base vector $\bar{b}_{12}$ and can easily be extended to fuse images in a virtual plane (see Figure 6). As we can assume only calibrated cameras to be used for aerial mapping systems, we will now analyse projective transformation for that special case.

A photogrammetric bundle vector pointing to a 3D point in the scene is related to the bundle vector of a second camera pointing to the same point O. Both vectors as well as the base vector are co-planar. Parallaxes will be zero for objects O^∞ at infinity and they will be constant for objects O^H in a constant distance. This is the case for a nadir virtual image plane and a flight mission at constant height over flat terrain and thus the 2D homography will strictly hold for all points of the scene.

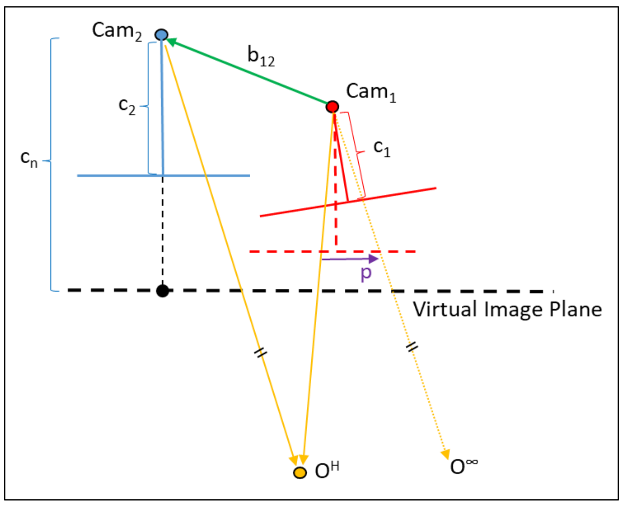

Figure 6. Fusion of calibrated cameras in a virtual image plane.

In the following we will show that such a 2D homography for calibrated cameras has only six independent parameters as we use a common scale parameter $m = m_x = m_y$ and there is no need for a shear parameter. For that we slightly modify Equation (4) to get Equation (5). Multiplication of $\boldsymbol{T'}$, $\boldsymbol{S'}$ and $\boldsymbol{R}$ then gives the total transformation matrix $\boldsymbol{P}_6$ (see Equation 6).

https://doi.org/10.5194/isprs-archives-XLIX-B1-2026-103-2026

$$\boldsymbol{T}' = \begin{bmatrix} 1 & 0 & dx \\ 0 & 1 & dy \\ 0 & 0 & 1 \end{bmatrix}; \boldsymbol{S}' = \begin{bmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & 1 \end{bmatrix}; \boldsymbol{R} = \boldsymbol{R}(\omega, \varphi, \kappa) \quad (5)$$

$$\boldsymbol{P}_6 = \boldsymbol{T}' \cdot \boldsymbol{S}' \cdot \boldsymbol{R} = \boldsymbol{M}' \cdot \boldsymbol{R} = \boldsymbol{P}(\omega, \varphi, \kappa, m, dx, dy) \quad (6)$$

$\boldsymbol{P_6}$ is composed of six physically meaningful parameters: Omega, Phi, Kappa rotational angles for $\boldsymbol{R}$, scale factor m for $\boldsymbol{S'}$ and translation parameters d_x, d_y for $\boldsymbol{T'}$. The geometric interpretation is quite intuitive (see again Figure 6). In order to fuse the image of the second camera with the one of the first camera in a virtual image plane (VIP), the following, consecutive steps can be done (in this order):

1) Rotation of camera #1 to be parallel to camera #2
2) Scaling of the image to compensate for the z component of the base vector b_{12}
3) Translate the image to compensate for the x- and y-parallaxes (vector p)

Note, that both the scaling for a different focal length c_2 or a defined nominal focal length c_n for the virtual camera are handled directly in the de-homogenization step and will not contribute to the matrix $\boldsymbol{S'}$. This is because we use photogrammetric bundle vectors including calibrated focal length c and principal point x_0, y_0 as given in Equations (7) and (8).

The concept of a VIP is very powerful can be used for the fusion of images from multiple, synchronized cameras with even small image overlaps. This concept has already been used for the first digital aerial camera DMC (see Hinz et al., 2000 and Tang et al., 2000) but also for combining medium cameras to systems with wider field of view. A very recent application is described by (Luo et al., 2025) and it will also be used for the PAS Pana.

To get more insight into the $\boldsymbol{P_6}$ transformation and to prove that it fully describes the relative orientation (RO) of two calibrated cameras, we can also derive it from classical photogrammetric notation. As shown also in (Luo et al., 2025), the image coordinates of a 3D point projected into two cameras Cam$_1$ and Cam$_2$ can be directly related to each other by the following equations:

$$\bar{x}_1 = \begin{pmatrix} x_1 - x_{0,1} \\ y_1 - y_{0,1} \\ -c_1 \end{pmatrix} = \frac{1}{\lambda_1} \cdot \boldsymbol{R}_1^T \cdot (\bar{S} - \bar{X}_{01}) \quad (7)$$

$$\bar{x}_2 = \begin{pmatrix} x_2 - x_{0,2} \\ y_2 - y_{0,2} \\ -c_2 \end{pmatrix} = \frac{1}{\lambda_2} \cdot \boldsymbol{R}_2^T \cdot (\bar{S} - \bar{X}_{02}) \quad (8)$$

$$\bar{b}_{12} = \bar{X}_{02} - \bar{X}_{01}; \; d\boldsymbol{R}_{12} = \boldsymbol{R}_1^T \cdot \boldsymbol{R}_2 \quad (9)$$

$$\bar{x}_2 = \frac{\lambda_1}{\lambda_2} \cdot d\boldsymbol{R}_{12}^T \cdot \bar{x}_1 - \frac{1}{\lambda_2} \cdot \boldsymbol{R}_2^T \cdot \bar{b}_{12} = \frac{\lambda_1}{\lambda_2} \cdot d\boldsymbol{R}_{12}^T \cdot \bar{x}_1 + d\bar{p} \quad (10)$$

We get a constant parallax vector $d\bar{p}$ for constant height H_2 and nadir camera Cam$_2$:

$$\boldsymbol{R}_2 = \boldsymbol{E}; \; \lambda_1 = -\frac{H_1}{d r_3^T \bar{x}_1}; \; \lambda_2 = \frac{H_2}{c_2}; \; d\bar{p} = -\frac{c_2}{H_2} \cdot \bar{b}_{12}; \quad (11)$$

The components of Equation (10) are now given explicitly:

$$x_2 = -c_2 \cdot \frac{H_1}{H_2} \cdot \frac{r_1^T \bar{x}_1}{r_3^T \bar{x}_1} - \frac{c_2}{H_2} \cdot \bar{b}_{12}^x \quad (12)$$

$$y_2 = -c_2 \cdot \frac{H_1}{H_2} \cdot \frac{r_2^T \bar{x}_1}{r_3^T \bar{x}_1} - \frac{c_2}{H_2} \cdot \bar{b}_{12}^y \quad (13)$$

A re-formulation is needed to include $d\bar{p}$ into a single, projective transformation (see Equations (16), (17) and compare with (1)):

$$x_2 = -\frac{c_2}{r_3^T \bar{x}_1} \cdot \left[\frac{H_1}{H_2} \cdot r_1^T \bar{x}_1 + \frac{r_3^T \bar{x}_1}{H_2} \cdot \bar{b}_{12}^x\right] \quad (14)$$

$$y_2 = -\frac{c_2}{r_3^T \bar{x}_1} \cdot \left[\frac{H_1}{H_2} \cdot r_2^T \bar{x}_1 + \frac{r_3^T \bar{x}_1}{H_2} \cdot \bar{b}_{12}^y\right] \quad (15)$$

$$\boldsymbol{P}_6 = \begin{bmatrix} \left(m\, r_{11} + \frac{\bar{b}_{12}^x}{H_2} r_{31}\right) & \left(m\, r_{12} + \frac{\bar{b}_{12}^x}{H_2} r_{32}\right) & \left(m\, r_{13} + \frac{\bar{b}_{12}^x}{H_2} r_{33}\right) \\ \left(m\, r_{21} + \frac{\bar{b}_{12}^y}{H_2} r_{31}\right) & \left(m\, r_{22} + \frac{\bar{b}_{12}^y}{H_2} r_{32}\right) & \left(m\, r_{23} + \frac{\bar{b}_{12}^y}{H_2} r_{33}\right) \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad (16)$$

$$\bar{x}_2 \cdot w = \boldsymbol{P}_6 \cdot \bar{x}_1 \quad (17)$$

From comparison of (16) with (5) we can also directly relate the scale and translation parameters to geometric properties:

$$m = \frac{H_1}{H_2} = \frac{(H_2 - \bar{b}_{12}^z)}{H_2}; \; d_x = \frac{\bar{b}_{12}^x}{H_2}; \; d_y = \frac{\bar{b}_{12}^y}{H_2} \quad (18)$$

It is also noteworthy that the matrix $\boldsymbol{P_6}$ can be set up directly from known (calibrated) relative orientation parameters (relative orientation angles ω, φ, κ, base vector b_{12} and a defined object distance H). This can be well exploited for co-registering e.g. images of a thermal camera with a RGB camera without having to match any tie points between those images or fusing images of individual cameras with almost no overlap (cp. Ladstädter et al., 2023).

From Equation (11) we can also estimate the components of the (absolute) parallax vector $d\bar{p}$ from the base vector components:

$$p_x = -c_2 \cdot \frac{\bar{b}_{12}^x}{H} = -c_2 \cdot d_x; \; p_y = -c_2 \cdot \frac{\bar{b}_{12}^y}{H} = -c_2 \cdot d_y \quad (19)$$

With dH being the relative terrain height variations, we can also estimate approximate "relative" parallaxes, describing how the absolute parallax varies for neighboring points at different heights:

$$dp_x = p_x \cdot \frac{dH}{H}; \; dp_y = p_y \cdot \frac{dH}{H}, \quad (20)$$

In Table 2, absolute parallaxes are given for a PAS Pana system operated at 6000 m over ground.

Cameras	b_x [m]	b_y [m]	p_x [µm]	p_y [µm]
NIR 1 / NIR 2	-0.102	-0.180	-1.2	-2.2
RGB 1 / RGB 3	0.140	-0.050	3.5	-1.2
RGB 2 / RGB 3	-0.051	0.090	-1.3	2.3
RGB 4 / RGB 3	0.051	-0.090	1.3	-2.3
RGB 5 / RGB 3	-0.140	0.050	-3.5	1.2

Table 2. Absolute parallaxes for H = 6000m.

While absolute parallaxes are compensated for in the stitching process, relative parallaxes are not and will introduce geometric errors. Fortunately, at that height, even large height variations dH of +/-250m (a factor of 0.04 compared to H) is resulting in a maximum relative parallax dp_x of only +/-0.15µm which can be neglected. We have shown, up to now, that using projective transformations based on six parameters can be used to generate fused images with a very high geometric quality and return to the task of deriving them from tie point measurements.

2.3 Estimation of projective transformations from tie points

The $\boldsymbol{P}_8$ transformation is linear in the eight parameters and estimation of those parameters from tie point measurements is straightforward. However, this is not the case for the $\boldsymbol{P}_6$ transformation (see Equation 16). For getting to the CPT algorithm, linearization of the problem is needed and will be given in section 3.1. Note that (Luo et al., 2025) have used the same approach (without giving the linearized equations) but restricted the estimated parameters to three rotational angles only. Their idea is to "calibrate" the relative orientation (RO) from all tie point measurements of a production flight and use them for creating high precision fused images in the VIP. Moreover, a photogrammetric bundle adjustment which estimates RO parameters from tie point measurements can also be used and should deliver the same result. This method has already been used for the DMC, where only the rotational parameters of the four cones and the focal lengths of three of the cones have been adjusted while projection centers have been fixed (Dörstel et al., 2002).

3. Estimation of a "controlled projective transformation"

We use the term "controlled perspective transformation" (CPT) to emphasize that we want to estimate $\boldsymbol{P}_6$ based on six physically meaningful parameters. Moreover, the authors want to control the initial values of those parameters and decide which of them will be treated as free parameters and which of them can be fixed to known values. We also want to be able to introduce known parameters (e.g. calibrated relative orientation angles) with their sigma values to be used as "directly observed parameters" in the adjustment. The basic idea is to initialize a $\boldsymbol{P}_6$ matrix with an approximated parameter set and then estimate improved CPT parameters in an iterative adjustment such that tie point residuals are minimized. We have developed a CPT prototype software already in early 2025 and have now tested it against the results of the stitching workflow based on full projective transformations.

3.1 Derivation of the CPT formulas

In contrast the FPT, the elements of $\boldsymbol{P}_6$ are built by non-linear functions of six parameters, therefore Equation (6) needs to be linearized to be used in an iterative adjustment:

$$\boldsymbol{P}_6 = (\boldsymbol{P}_6) + d\boldsymbol{P}_6 = [(\mathbf{M}') + \mathrm{d}\mathbf{M}')] \cdot [(\boldsymbol{R}) + (\boldsymbol{R})d\boldsymbol{A}] \quad (21)$$

$$(\boldsymbol{P}_6) + d\boldsymbol{P}_6 = (\boldsymbol{P}_6) + (\boldsymbol{P}_6)d\boldsymbol{A} + d\boldsymbol{M}'(\boldsymbol{R}) + O(2) \quad (22)$$

For the projection of an image point directly into the virtual image plane (VIP) we need to set up a 3D photogrammetric bundle vector by applying inner orientation (IO) parameters also for the virtual camera:

$$\bar{x}_i = \begin{pmatrix} x_i - x_{0,i} \\ y_i - y_{0,i} \\ -c_i \end{pmatrix}; \;\; \bar{x}^i_{VIP} = \begin{pmatrix} x^i_{VIP} \\ y^i_{VIP} \\ -c_{VIP} \end{pmatrix} \quad (23)$$

$$\bar{x}^i_{VIP} = \boldsymbol{P}_6 \cdot \bar{x}_i = (\boldsymbol{P}_6) \cdot \bar{x}_i + d\boldsymbol{P}_6 \cdot \bar{x}_i \quad (24)$$

We now explicitly write the three components of the homogeneous transformation:

$$x^i_{VIP} = \frac{-c_i}{w} \cdot [(\bar{p}_1) + d\bar{p}_1] \cdot \bar{x}_i = \frac{-c_i}{w} \cdot (Z_x + \Delta Z_x) \quad (25)$$

$$y^i_{VIP} = \frac{-c_i}{w} \cdot [(\bar{p}_2) + d\bar{p}_2] \cdot \bar{x}_i = \frac{-c_i}{w} \cdot \left(Z_y + \Delta Z_y\right) \quad (26)$$

$$w = [(\bar{p}_3) + d\bar{p}_3] \cdot \bar{x}_i = (w) + \Delta w \quad (27)$$

Small changes of the CPT parameters will cause small changes of the matrix P, so we sort for them and re-write functions ΔZ_x and ΔZ_y:

$$\Delta Z_x = [A_x \cdot \Delta\omega + B_x \cdot \Delta\varphi + C_x \cdot \Delta\kappa + D_x \cdot \Delta \mathrm{m} + E_x \cdot \Delta d_x] \quad (28)$$

$$\Delta Z_y = \left[A_y \cdot \Delta\omega + B_y \cdot \Delta\varphi + C_y \cdot \Delta\kappa + D_y \cdot \Delta \mathrm{m} + E_y \cdot \Delta d_y\right] \quad (29)$$

$$w = (w) + A_w \cdot \Delta\omega + B_w \cdot \Delta\varphi + C_w \cdot \Delta\kappa \quad (30)$$

We can then give the partial derivatives for the x- and y-coordinate for the point projected into the VIP using the quotient rule (here only given for the Omega angle):

$$\frac{\partial x^i_{VIP}}{\partial\omega} = \frac{-c_2}{w^2} \cdot (A_x \cdot w - Z_x \cdot A_w) \quad (31)$$

$$\frac{\partial y^i_{VIP}}{\partial\omega} = \frac{-c_2}{w^2} \cdot \left(A_y \cdot w - Z_y \cdot A_w\right) \quad (32)$$

For a single camera (with index i) we can therefore find an improved projection into the VIP by adding small changes to the x- and y-component caused by small changes of the initial CPT parameters:

$$x^i_{VIP} = \left(x^i_{VIP}\right) + \frac{\partial x^i_{VIP}}{\partial\omega} \cdot d\omega + \cdots + \frac{\partial x^i_{VIP}}{\partial y} \cdot dy \quad (33)$$

$$y^i_{VIP} = \left(y^i_{VIP}\right) + \frac{\partial y^i_{VIP}}{\partial\omega} \cdot d\omega + \cdots + \frac{\partial y^i_{VIP}}{\partial y} \cdot dy \quad (34)$$

For a tie point matched in the overlap of the images of camera i and j, respectively, we formulate the constraint that they will have the same position in the VIP and therefore get the final formulas to be used in the simultaneous adjustment of all tie points matched in all overlaps:

$$\left(x^i_{VIP}\right) - \left(x^j_{VIP}\right) = \frac{\partial x^j_{VIP}}{\partial\omega_j} \cdot d\omega_j + \cdots - \left(\frac{\partial x^i_{VIP}}{\partial\omega_i} \cdot d\omega_i + \cdots\right) \quad (35)$$

$$\left(y^i_{VIP}\right) - \left(y^j_{VIP}\right) = \frac{\partial y^j_{VIP}}{\partial\omega_j} \cdot d\omega_j + \cdots - \left(\frac{\partial y^i_{VIP}}{\partial\omega_i} \cdot d\omega_i + \cdots\right) \quad (36)$$

3.2 Simulation results

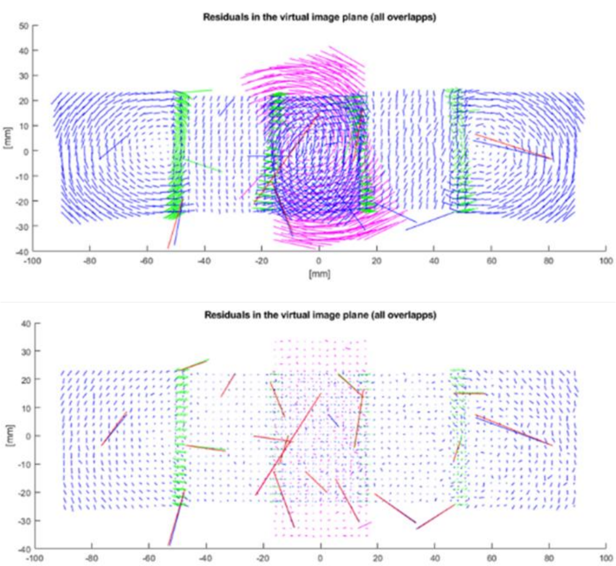

Figure 7. CPT adjustment using simulated tie points.

To test the CPT adjustment algorithm, simulated tie points have been generated for a PAS Pana sensor configuration. The known transformation parameters for each sensor have been modified by small deviations (up to +/- 5 deg for the rotation angles). In addition, measurement noise (and some gross errors) have been added to the tie points image coordinates. Figure 7 shows the second and fifth iteration of the CPT adjustment. Almost all systematic residuals have been removed at that stage, and outliers have been identified and excluded from further processing. After ten iterations, the adjustment has converged and was able to reproduce the known parameter values with a high accuracy.

4. Real data experiments and results

The same flight mission over Denver already used for testing the beta workflow has been selected again for testing the new PAS Pana stitching algorithm (but this time, field calibration was used). The missions have been flown on the same day (October 21st, 2023) at two different altitudes using 90% forward overlap (see Table 3). The 20cm mission has been flown just before the 15cm mission. For both missions, six strips in North/South direction and two cross strips in East/West direction have been provided.

Mission	Altitude	GSD	#Strips	#Shots
Denver / high	9790m	20cm	6+2	1174
Denver / low	7610m	15cm	6+2	1436

Table 3. Details on the PAS Pana Denver missions (Oct. 2023).

While distortion removal and tie pointing (steps 1 & 2) have been done using the very performant stitching software, the adjustment (step 3) has been done with a Matlab prototype for both the FPT and CPT variants for direct comparison. Step 4 (output image generation) has been skipped as no aerial triangulation using the fused output images was planned. Note that only every 10th shot has been processed for this evaluation.

4.1 Denver 20cm mission

Figure 8 shows the number of tie points matched separated for the NIR/NIR, RGB/RGB and RGB/NIR overlaps. The total number of tie points per shot over all overlaps is just below 5000 and is stable over all flight lines.

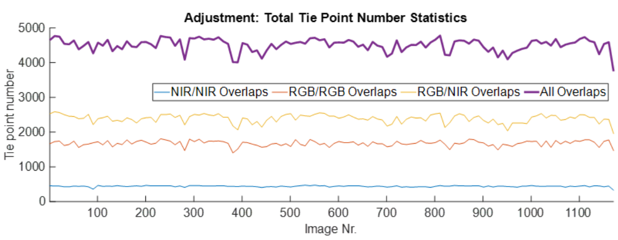

Figure 8. Tie point number used for adjustment (Denver 20cm).

Param.	Min	Max	Mean	Median	Stdev.
Omega [°]	-0.164	-0.103	-0.135	-0.134	9.1e-3
Phi [°]	-28.253	-28.119	-28.178	-28.177	1.8e-2
Kappa [°]	-0.678	-0.662	-0.669	-0.669	2.2e-3
d_x [pixel]	-57.127	-17.192	-34.418	-34.397	5.5
d_y [pixel]	-10.643	9.032	-1.396	-1.370	3.0
m_x [ ]	0.9995	1.0007	1.0001	1.0001	2.4e-4
m_y [ ]	1.0004	1.0010	1.0008	1.0008	8.2e-5

Table 4. NIR/NIR fusion (Beta workflow, Denver 20cm).

Table 4 shows a statistic of the parameters calculated in the NIR/NIR fusion step in the beta workflow to be compared later with the results of the new algorithm.

4.1.1 FPT solution

As described in section 2.1, physically meaningful parameters have been extracted from the $\boldsymbol{P}_8$ transformations calculated for all seven sensors and plotted in Figures 9 and 10. Tables 5 and 6 show a statistic of the parameters extracted from the FPT solution for the RGB 1 and NIR L sensor, respectively.

Stability of the new approach using simultaneous adjustment is improved compared to the beta workflow as it can be seen by comparing standard deviations of Table 6 with Table 4. There is also no visible trend in any of the stitching parameters of any sensor, which also proves that the new, simultaneous adjustment delivers more stable results already in the FPT variant.

Param.	Min	Max	Mean	Median	Stdev.
Omega [°]	-0.0024	0.0407	0.0205	0.0204	8.0e-3
Phi [°]	27.348	27.411	27.381	27.3800	1.2e-2
Kappa [°]	0.0948	0.1032	0.1004	0.1005	1.4e-3
d_x [mm]	-0.1984	-0.0676	-0.1350	-0.1344	2.6e-2
d_y [mm]	-0.0408	0.0604	0.0130	0.0140	1.9e-2
m_x [ ]	0.9989	0.9997	0.9993	0.99934	1.3e-4
m_y [ ]	0.9994	0.9998	0.9997	0.99967	6.5e-5

Table 5. Sensor RGB 1/ FPT solution (20cm mission).

Param.	Min	Max	Mean	Median	Stdev.
Omega [°]	-0.0354	-0.0069	-0.0208	-0.0209	5.9e-3
Phi [°]	14.151	14.198	14.172	14.1725	9.7e-3
Kappa [°]	0.4555	0.4614	0.4592	0.4592	9.7e-3
d_x [mm]	-0.1340	-0.0230	-0.0733	-0.0735	2.3e-2
d_y [mm]	-0.0357	0.0306	-0.0021	-0.0019	1.4e-2
m_x [ ]	0.9998	0.9999	0.9999	0.99986	2.7e-5
m_y [ ]	0.9999	1.0000	1.0000	0.99999	2.1e-5

Table 6. Sensor NIR L / FPT solution (20cm mission).

From Figure 9 we can see that scale parameters stay mostly below 1.0 and deviations from 1.0 increase for the outer sensors RGB 1 and RGB 5. A similar effect can be seen for the translation parameters (plotted as parallaxes), which already indicates strong correlation between those parameters.

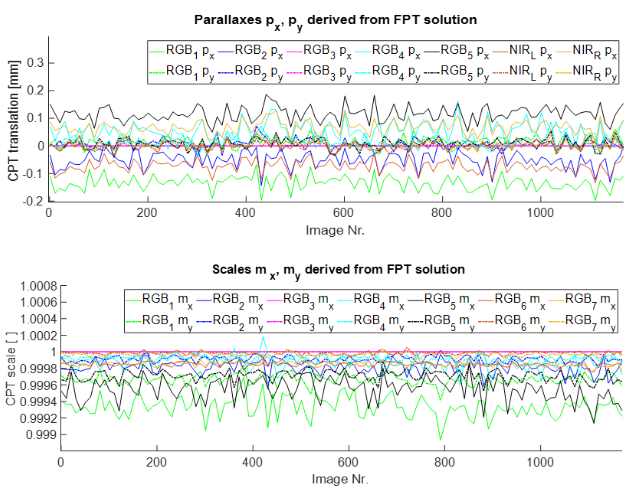

Figure 9. Translation and scale parameters (FPT, Denver 20cm).

For plotting, the rotational angles have been reduced by their mean value for better comparison. Looking at the plots, Phi has the biggest standard deviation (corresponding to the smaller overlap dimension *x*), kappa the smallest one (see Figure 10).

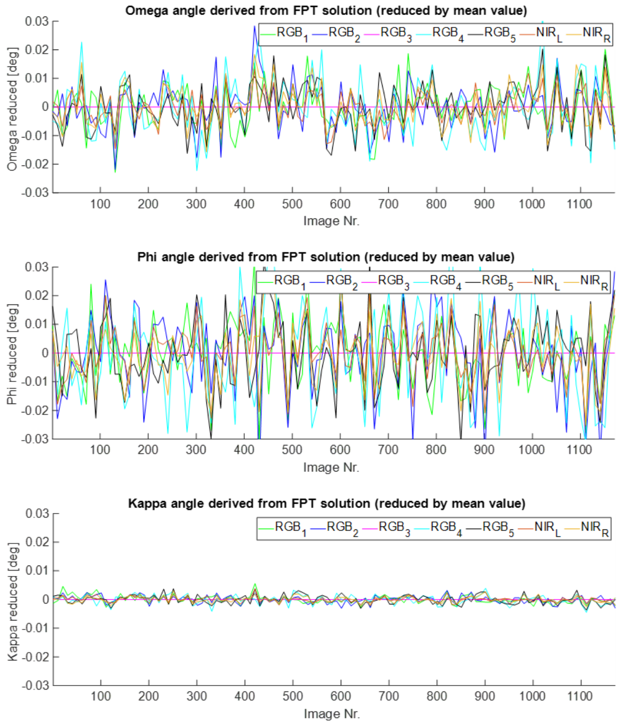

Figure 10. Rotation parameters (FPT solution, Denver 20cm).

The magnitude of the angle variations for Omega and Phi of +/- 0.02 deg is quite high (corresponding to a magnitude of about +/-10 Pixels in the VIP) and is mainly caused by the correlation with other parameters.

RGB/RGB, RGB/NIR and NIR/NIR tie point measurements have been weighted 1.0:0.5:0.3 in the adjustment to account for different resolutions and spectral properties. The components of the residual vectors in the VIP have an RMS value of about +/- 1µm and are also very stable for all images of that mission, see Figure 11.

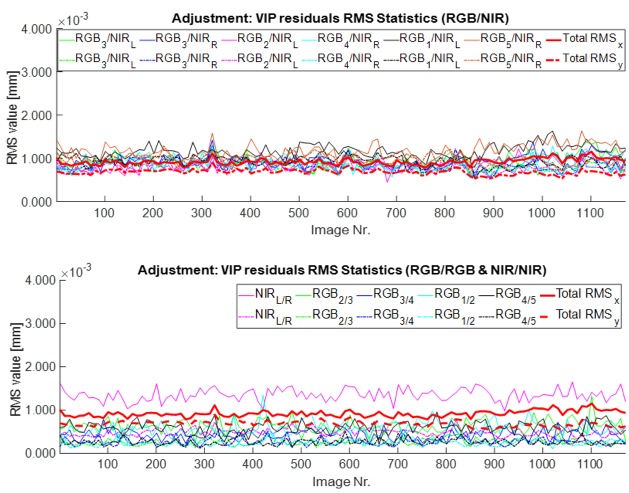

Figure 11. RMS for x/y residuals (FPT solution, Denver 20cm).

4.1.2 CPT solution

Using the same tie point measurements, the adjustment was repeated for each PAS Pana shot using the CPT equations. As the high correlation between rotational and translational parameters is also valid for the CPT, we fix the latter ones to zero (as parallaxes are very small for that height), which can easily be done in the CPT variant. The scale parameters are constrained to be identical. Statistics for the selected sensors RGB 1 and NIR L (Tables 7 and 8) can now directly be compared with the FPT solution (Tables 5 and 6).

Param.	Min	Max	Mean	Median	Stdev.
Omega [°]	0.0136	0.0165	0.0149	0.0149	5.6e-4
Phi [°]	27.314	27.318	27.316	27.3162	8.1e-4
Kappa [°]	0.0987	0.1036	0.1010	0.1009	9.2e-4
d_x [mm]	0.0	0.0	0.0	0.0	0.0
d_y [mm]	0.0	0.0	0.0	0.0	0.0
$m_x = m_y$ [ ]	0.9999	1.0000	1.0000	0.99996	2.7e-5

Table 7. Sensor RGB 1/ CPT (20cm mission).

Param.	Min	Max	Mean	Median	Stdev.
Omega [°]	-0.0207	-0.0191	-0.0200	-0.0200	3.4e-4
Phi [°]	14.139	14.142	14.141	14.1407	5.1e-4
Kappa [°]	0.4573	0.4606	0.4590	0.4590	6.4e-4
d_x [mm]	0.0	0.0	0.0	0.0	0.0
d_y [mm]	0.0	0.0	0.0	0.0	0.0
$m_x = m_y$ [ ]	0.9999	1.0000	1.0000	0.99998	2.0e-5

Table 8. NIR L / CPT (20cm mission).

The rotational parameters are also plotted for the CPT solution (see Figure 12) and can be compared with Figure 10 (be aware that the plotting scale is different by one magnitude).

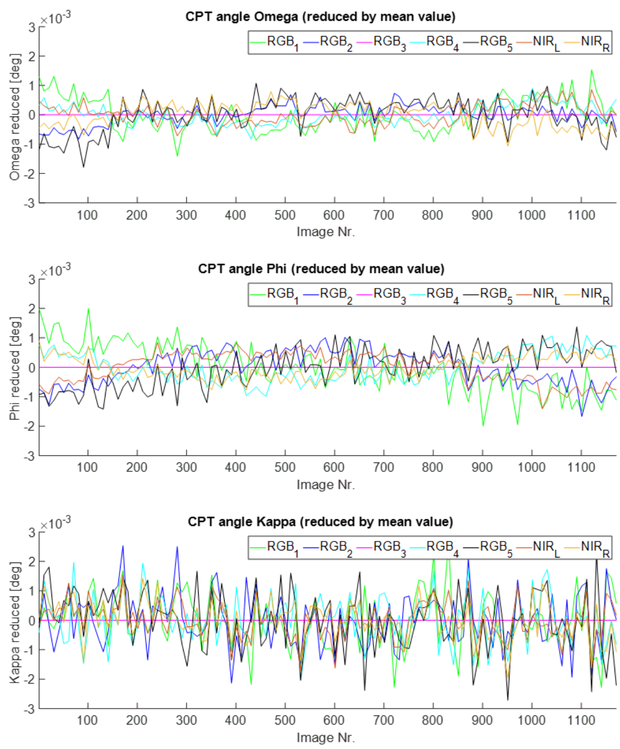

Figure 12. Rotation parameters (CPT solution, Denver 20cm).

https://doi.org/10.5194/isprs-archives-XLIX-B1-2026-103-2026 | © Author(s) 2026. CC BY 4.0 License.

Figure 13 shows the CPT scale parameters (identical for *x*- and *y*-component), being also mostly below 1.0 (i.e. the scale fixed for the RGB 3 sensor), compare also with Figure 9 for the FPT solution).

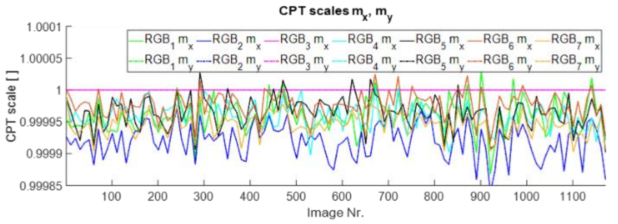

Figure 13. Scale parameters (CPT solution, Denver 20cm).

Looking at the RMS values of the residuals in the VIP (Figure 14), we see that the total RMS value is slightly higher (+/- 1.2 µm) compared to the FPT solution (Figure 11).

Especially residuals from the RGB 1/NIR L, RGB 5/NIR R and NIR/NIR overlaps show higher RMS values which indicates that the remaining systematic errors cannot be modelled by the CPT.

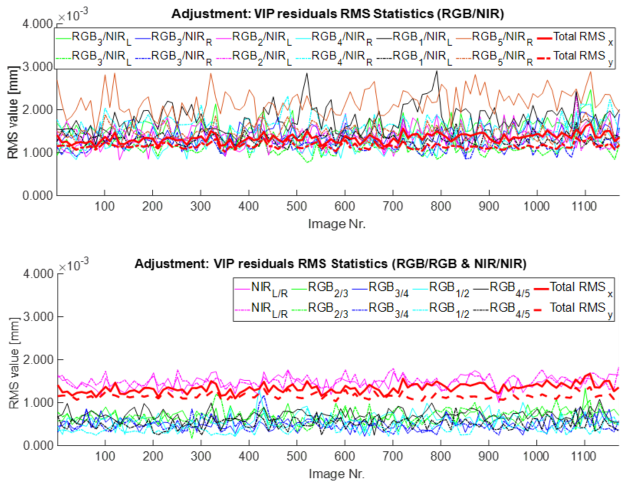

Figure 14. RMS for x/y residuals (CPT solution, Denver 20cm).

4.2 Denver 15cm mission

The same evaluation as for the 20cm mission has been repeated for the 15cm mission. Tie point numbers are again very stable (see Figure 15) and only slightly reduced compared with the 20cm mission.

The total number drops to about 3000 for only some shots around number 350 and 1350 due to low textured areas covering larger image parts at higher resolution. This is still by far sufficient to ensure high quality of the stitched output image.

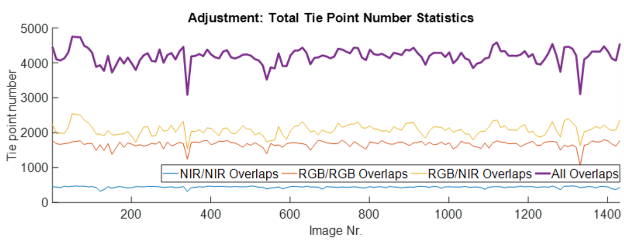

Figure 15. Total tie point number used for adjustment.

4.2.1 FPT solution

Tables 9 and 10 show the statistics of extracted CPT parameters which can directly be compared to the 20cm mission (Table 7 and 8). Median and standard deviation values show a good coincidence between the two mission heights.

Param.	Min	Max	Mean	Median	Stdev.
Omega [°]	-0.003	0.044	0.024	0.0241	9.6e-3
Phi [°]	27.340	27.421	27.373	27.3730	1.4e-2
Kappa [°]	0. 097	0.105	0.101	0.1010	1.5e-3
d_x [mm]	-0.209	-0.048	-0.120	-0.1173	3.1e-2
d_y [mm]	-0.045	0.067	0.018	0.0187	2.3e-2
m_x [ ]	0.9987	0.9999	0.9994	0.99943	1.9e-4
m_y [ ]	0.9994	0.9999	0.9997	0.99972	9.4e-5

Table 9. Sensor RGB 1/ FPT solution (15cm mission).

Param.	Min	Max	Mean	Median	Stdev.
Omega [°]	-0.042	0.001	-0.018	-0.0179	7.6e-3
Phi [°]	14.148	14.200	14.169	14.1679	9.8e-3
Kappa [°]	0.456	0.462	0.459	0.4591	9.5e-4
d_x [mm]	-0.129	-0.020	-0.069	-0.066	2.3e-2
d_y [mm]	-0.053	0.045	0.001	0.0024	1.8e-2
m_x [ ]	0.9998	0.9999	0.9998	0.99984	3.1e-5
m_y [ ]	0.9999	1.0000	1.0000	0.99996	2.3e-5

Table 10. Sensor NIR L / FPT solution (15cm mission).

Accordingly, the plots for the rotational parameters (Figure 16), translational and scale parameters (Figure 17) look very similar to the corresponding plots of the 20cm mission.

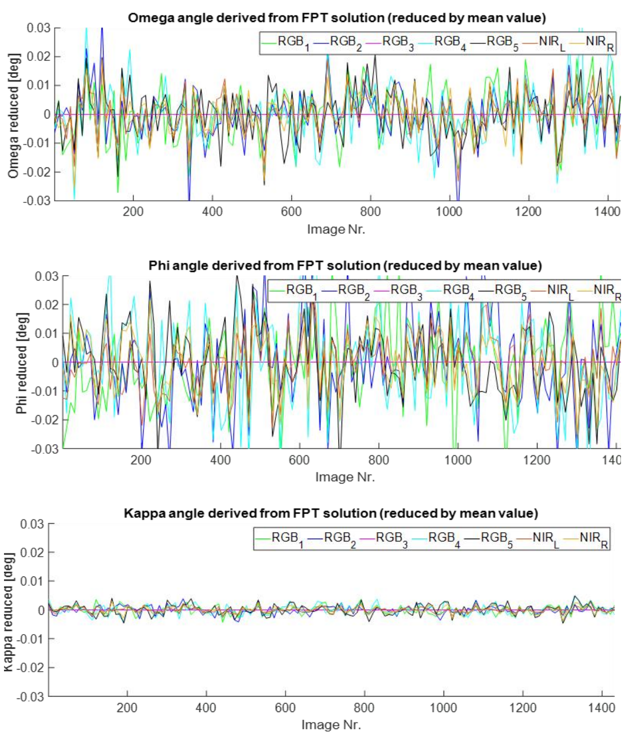

Figure 16. Rotation parameters (FPT solution, Denver 15cm).

The RMS values of the residuals remaining in the VIP after adjustment have the same mean magnitude as for the 20cm mission. However, they show a higher variability and reach up to about +/- 2µm for the RGB/NIR tie points (see Figure 18).

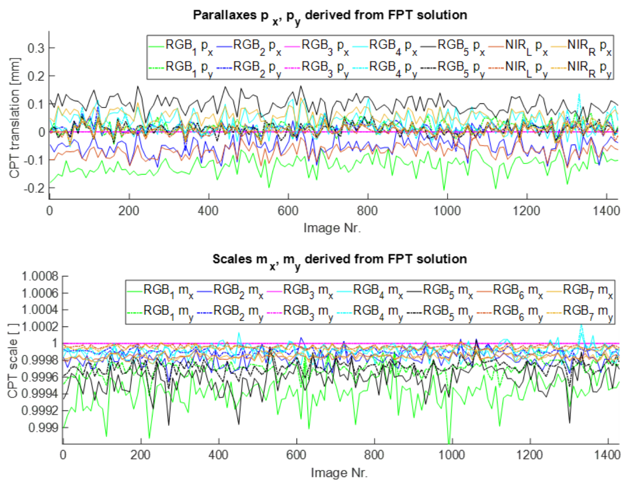

Figure 17. Translation and scale parameters (FPT, Denver 15cm).

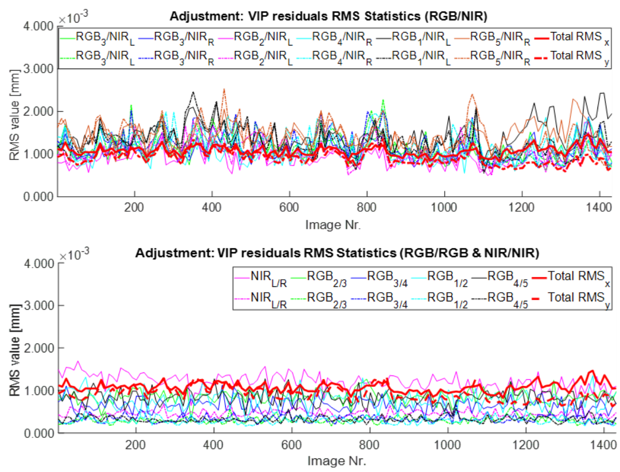

Figure 18. RMS for x/y residuals (FPT solution, Denver 15cm).

4.2.2 CPT solution

Again, for the CPT solution, the translational parameters (not plotted here) have been fixed to zero to avoid the strong correlation with the angle parameters. Transformation parameters for all sensors are plotted for direct comparison with the FPT solution and the CPT solution of the 20cm mission.

From Tables 11 & 12, it can be seen, that the CPT angles for the 15cm and 20cm mission deviate only by 2e-3 deg for Omega and are even identical for the Phi and Kappa. Scale parameters deviate only by 1e-5 (RGB 1) and 2e-5 (NIR L) between the two missions, see also Figures 20 and 13, respectively.

Param.	Min	Max	Mean	Median	Stdev.
Omega [°]	0.015	0.019	0.016	0.0165	5.3e-4
Phi [°]	27.315	27.318	27.316	27.3162	7.4e-4
Kappa [°]	0.098	0.106	0.101	0.1010	1.4e-3
d_x [mm]	0.0	0.0	0.0	0.0	0.0
d_y [mm]	0.0	0.0	0.0	0.0	0.0
$m_x = m_y$ [ ]	0.9999	1.0000	1.0000	0.99995	3.2e-5

Table 11. Sensor RGB 1/ CPT solution (15cm mission).

Param.	Min	Max	Mean	Median	Stdev.
Omega [°]	-0.020	-0.018	-0.019	-0.0188	3.9e-4
Phi [°]	14.139	14.141	14.140	14.1396	5.6e-4
Kappa [°]	0.457	0.461	0.459	0.4588	7.3e-4
d_x [mm]	0.0	0.0	0.0	0.0	0.0
d_y [mm]	0.0	0.0	0.0	0.0	0.0
$m_x = m_y$ [ ]	0.9999	1.0000	1.0000	0.99996	2.1e-5

Table 12. Sensor NIR L / CPT solution (15cm mission).

Rotational parameters are plotted for the CPT solution (Figure 19) and can be compared with the 20cm mission (see Figure 12). In contrast to the FPT solution, the standard deviation for the Kappa angle is larger than for Omega and Phi but still guarantees subpixel accuracy over the whole sensor format.

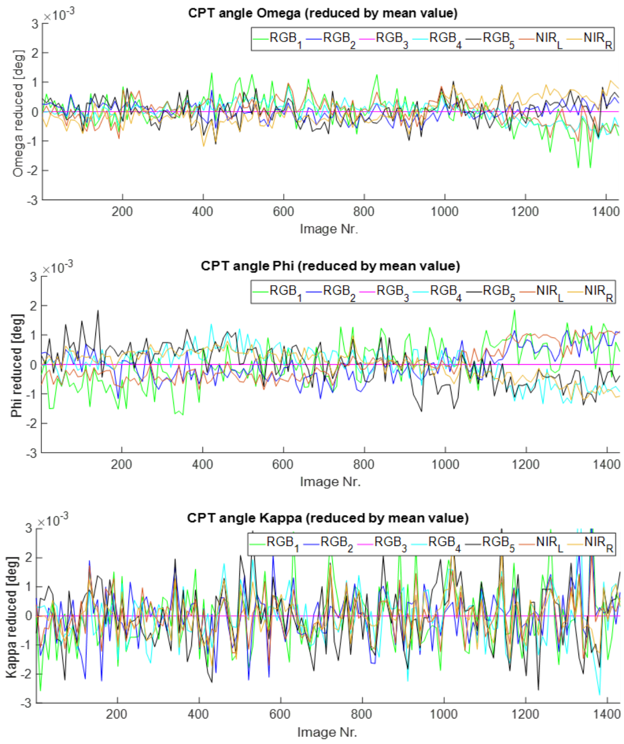

Figure 19. Rotation parameters (CPT solution, Denver 15cm).

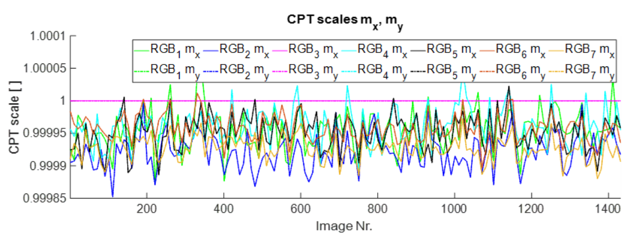

Figure 20. Scale parameters (CPT solution, Denver 15cm).

The total RMS values of the residuals remaining in the VIP after adjustment are slightly bigger (~+/-1.5µm) as for the 20cm mission (+/-1.3µm). Again, the RGB/NIR tie points show a higher variability and reach up to about +/- 2.5µm (see Figure 21).

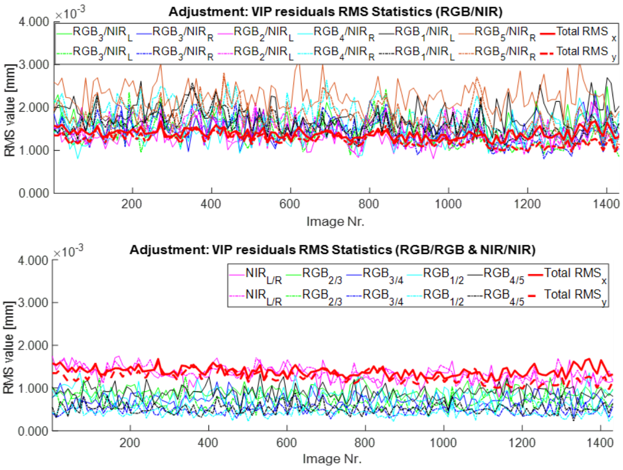

Figure 21. RMS for x/y residuals (CPT solution, Denver 15cm).

Finally, Figure 22 shows a sample output image displayed as RGB and CIR product. The image format is 48.800 x 12400 pixels and the nominal focal length is set to 132mm. The data volume for a 16bit output image is 4.6 GByte. Calculation of such an image takes about 3 minutes excluding file I/O.

Figure 22. Fused RGB & CIR image sample (20cm mission).

5. Conclusions and outlook

We have presented the theoretical background of a modern stitching algorithm for synchronized multi-camera systems and showed first results from processing a large data set of PAS Pana shots acquired in two flight heights.

The new simultaneous adjustment of all image overlaps shows much improved stitching stability compared to the beta workflow from 2024 even when using the standard, full projective transformations (FPT) build on eight parameters.

From theoretical considerations we have also seen that only six parameters of a "controlled" projective transformation (CPT) can model the relative orientation of calibrated cameras and can be used to fuse images in a virtual image plane (VIP). The strong correlation between rotation angles and translational parameters (due to long focal lengths and small image overlaps) exists also for the CPT parameter set but it can be solved by fixing the translational parameters to known, fixed values.

As a result, the variance of the CPT angles is one order of magnitude lower than for the FPT solution. On the other hand, CPT cannot model remaining systematic image residuals to the same extent as the FPT variant. This can be seen in the RMS values of the tie point residuals in the VIP after adjustment. In the opinion of the authors however, this just indicates imperfection of the camera calibration which can be better removed by the FPT but at the cost of deformation of the image geometry. To evaluate the absolute accuracy of the fused image (for both variants), an aerial triangulation and careful analysis of the bundle adjustment results would be needed, which was out of scope of this paper.

As a next step, it is planned to integrate the stitching software solution of JOANNEUM RESEARCH into the Phase One image processing SDK.

Acknowledgements

The authors would like to thank Phase One for supporting research and software development for the PAS Pana camera system and Scott Merritt from SURDEX/Bowman for providing test data.

References

Dörstel, C., Zeitler, W., Jacobsen, K., 2002: Geometric Calibration of the DMC: Method and Results, *International Archives of Photogrammetry and Remote Sensing* (34) Part 1 Com I, pp 324 – 333.

Hartley, R., Zisserman, A., 2004: Multiple View Geometry in Computer Vision, 2nd edition, Cambridge University Press.

Hinz, A., Dörstel, C., Heier, H., 2000: Digital Modular Camera: System Concept and Data Processing Workflow, Proc. IAPRS Vol. XXXIII, Amsterdam.

Jacobsen, K., Ladstädter, R. & Bosch, R., 2025: Geometric Accuracy of PhaseOne PAS Pana, ISPRS, ERASeL & DGPF Joint Istanbul Workshop, 29.-31.1.2025, in: Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLVIII-M-6-2025, 153–159.

Ladstädter, R., Gutjahr, K., Perko, R., Gregorac, A., 2023: Simultane Kalibrierung von RGB/TIR Multi-Kamerasystemen, in Proceedings: Internationale Geodätische Woche No. 22, 12.-18.2.2023, Obergurgl, Tirol, p. 150-158.

Ladstädter, R., Gruber, M. & Wiechert, A., 2010: Monolithic Stitching: One sensor geometry for multiple sensor cameras, ASPRS 2010 Annual Conference, 26.-30. April 2010, San Diego, California.

Luo, X., Zhao, H., Liu, Y., Liu, N., Chen, J., Yang, H., Pan, J., 2025. A High-Precision Virtual Central Projection Image Generation Method for an Aerial Dual-Camera. MDPI Journal, Remote Sensing, 2025, 17, 683.

Phase One, 2024: Aircraft-systems product information web page: https://www.phaseone.com/solutions/geospatial-solutions/aircraft-systems/pas-pana/ (last visit: March. 2026).

Tang L., Dörstel, C., Jacobsen. K., Heipke. C., Hinz, A., 2000: Geometric accuracy potential of the Digital Modular Camera; Proc. IAPRS, Vol. XXXIII, Amsterdam.