Wide-Angle Image Acquisition, Analysis and Visualisation

Reinhard Klette, Georgy Gimel'farb
CITR Tamaki, The University of Auckland
Tamaki Campus, Building 731
Auckland, New Zealand
{r.klette,g.gimelfarb}@auckland.ac.nz

Ralf Reulke
Institute of Space Sensor Technology and Planetary Exploration
German Aerospace Center, Rutherfordstr. 2
D-12489 Berlin, Germany
ralf.reulke@dlr.de

Abstract

Recent camera technology provides new solutions for wide-angle image acquisition. Multi- or single-line cameras have been designed for spaceborne and airborne scanners to provide high resolution imagery. Line cameras may also work as panorama scanners, and models of these have already been studied in computer vision for a few years. These cameras or models require studies in calibration, registration and epipolar geometry to ensure accurate imaging and stereo analysis. The resulting images or depth maps also allow new approaches in 3D scene visualisation. The paper informs about line camera models and camera hardware, calibration of line cameras, registration of captured images, epipolar geometry for along-track and panoramic stereo, stereo matching with a focus on dynamic programming, and visualisation.

1 Introduction

This paper discusses (CCD) *line cameras* with respect to theoretical models, existing hardware and applications. Related theoretical issues are in the fields of calibration, registration, epipolar geometry, stereo matching and visualisation.

Wide-angle image acquisition has its historic roots in photogrammetry. Related studies on epipolar geometry or 3D reconstruction are well established for stereo pairs of planar images, see, e.g. textbooks [6, 13, 19, 21, 39]. Compared to these studies, the computer vision literature still lacks work on images captured by line cameras which allow a new direction in photogrammetry, namely *digital photogrammetry* which replaces conventional (analog) photogrammetric images by 'streams' of digital image data. New camera architectures and different image acquisition strategies have led to a broad diversity of imagery [38].

1.1 Multi-Line Cameras

Figure 1 illustrates the architecture of a *symmetric three-line camera* where it is assumed that all three *sensor lines* are parallel and adjusted in height (see front view of focal plate), as well as collinear and symmetric (see top view of focal plate) with respect to an equal distance a between the left sensor line (or the right sensor line) to the central sensor line, and with respect to angles ω or $-\omega$ between rays emerging from the left and central sensor line, or from the right and central sensor line. In the case of rotation, the whole camera rotates around a rotation center in a distance R to the optical lens (along the optical axis of the camera), where the rotation axis is parallel to the orientation of the sensor lines in the front view. The focal length f is defined by the distance between the central sensor line and the optical lens, and the distance a between sensor lines satisfies $f = a \cdot tan(\omega)$.

In the case of an airborne three-line camera, the camera will follow motions of the aircraft which may be recorded (for example) using a differential GPS. In such an application, the left sensor line, the central sensor line, and the right sensor line capture a *backward view*, the *nadir view*, and a *forward view*, respectively.

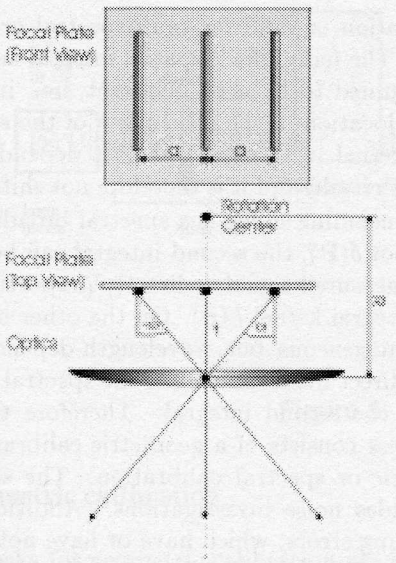

Figure 1. Model of a three-line camera.

In case of a rotated camera on a tripod, the three-line camera captures three different panoramas with respect to a fixed rotation axis, one for each of the three sensor lines.

For example, the *Wide-Angle Airborne Camera* (WAAC) from the *German Aerospace Center* (DLR), developed as a follow-up of their miniaturized *Wide-Angle Optoelectronic Stereo Scanner* (WAOSS), is a CCD-line stereo scanner working in *along-track*, or *in-track* mode, which is an example of a *pushbroom* strategy: three parallel sensor lines are assumed to be perpendicular to the aircrafts main orientation, see Fig. 2 on the right. All three lines are in one image plane with a single optical system. Knowledge of the aircraft's attitude data allows stereo processing.

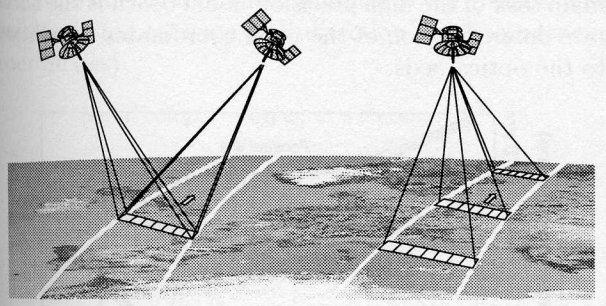

Figure 2. SPOT's across-track principle (left) where a second crossing is assumed to capture exactly the same area, and WAAC's along-track principle (right).

1.2 Single-line Camera

Recently also multi-perspective panoramic images received increasing attention for applications such as 3D scene visualization or reconstruction, see, e.g, [16, 27, 28, 35, 36, 37]. Some examples of stereo reconstructions and 3D scene visualizations based on a given set of single-center panoramic images applications can be found in [3, 20]. Originally, a small number of images taken in different directions had to be stitched together for producing a panoramic image [3, 4]. To avoid the well-known problems in image stitching it was then studied that just one column of a 'normal' CCD matrix camera is used during camera rotation for producing a panoramic image. As a logical consequence, the camera may consist of a single CCD line only. These type of camera models have been studied in [15], and a first single-line panoramic camera was discussed in [32].

Figure 3 shows a single-line panorama camera (built at DLR) which may rotate part of, or full 360°. The figure also provides parameters showing that this color camera produces images having 10,200 pixels in one column, and a full 360° image has 55,000 columns for $f = 60\ mm$ which results in a single image of size 3.3 gigabytes.

The theoretical model of a single-line camera is shown in Fig. 4. A panoramic image is acquired by rotating a line camera with respect to a fixed rotation axis and taking images consecutively at equidistant angles. A panoramic image acquisition model for single-line cameras has been formally discussed in [38]. There are three essential parameters f, R, and ω in this image acquisition model: f is the effective focal length, R is the distance between the camera's focal point and the rotation axis, and ω specifies the viewing direction of the sensor line. *Polycentric panoramas* are a collection of panoramic images acquired with respect to different camera positions (i.e. rotation axes in 3D scene space), and parameters f, R, ω may be specific for every image in a set of polycentric panoramas.

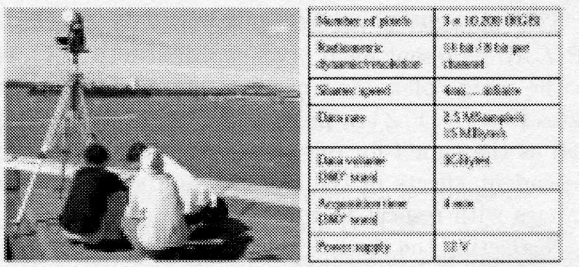

Figure 3. EYESCAN camera and [32] its basic parameters.

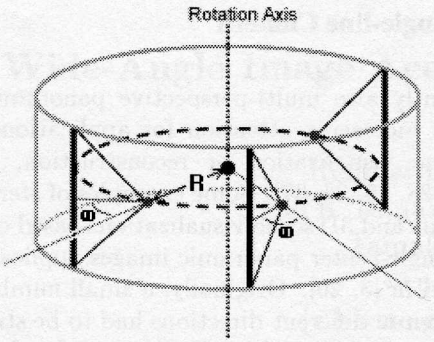

Figure 4. [15] Single-line panorama camera.

Below the new technology of line cameras is discussed in the context of photogrammetry and panoramic images.

2 Calibration

High resolution CCD-line cameras have both: a large *field of view* (FOV) and a small IFOV for any pixel. These extremely large images require an extensive calibration procedure. A calibration process [33] consists of a geometric, and a radiometric or spectral subprocess. We briefly describe the general task and sketch solutions for line cameras.

2.1 General Calibration Model

Any calibration approach can be based on the following general input-output relation which describes the measurement process,

$$E(\mathbf{a}) = \int c(\mathbf{a}, \mathbf{a}_w) \cdot h(\mathbf{a}, \mathbf{a}_w) \cdot I(\mathbf{a}_w)\, d\mathbf{a}_w + \varsigma(\mathbf{a}) \;,$$

where $I(\mathbf{a})$ are the incoming signals, $E(\mathbf{a})$ the measured values, $h(\mathbf{a}, \mathbf{a}_w)$ a system response function (in the simplest case the *point spread function* PSF), $\varsigma(\mathbf{a})$ the system noise with expected value 0 and variance σ^2, and function c specifies absolute radiometric calibration factors.

The parameter vectors $\mathbf{a} = (\mathbf{p}, t, \lambda)$ and $\mathbf{a}_w = (\mathbf{P}, t, \lambda)$, represent ideal pixel coordinates $\mathbf{p} = (x, y, f)$ in the image plane and camera-centred world coordinates $\mathbf{P} = (X, Y, Z)$ in the object space, respectively, as well as the time t for one-dimensional signals or time-dependent effects in the camera, and spectral dependencies with respect to wavelength λ.

Neglecting the time dependencies, spatial and spectral effects can be separated

$$E(\mathbf{p}, \lambda) = \int d\mathbf{P}\; h(\mathbf{p}-\mathbf{P}) \int d\lambda_w c(\lambda, \lambda_w) \cdot I(\mathbf{P}, \lambda) + \varsigma(\mathbf{P})$$

This equation is valid for multispectral or true color systems. The (simplified) spatial integral kernel $h(\mathbf{p}-\mathbf{P})$ is assumed to be shift-invariant, i.e. independent from the location in the focal plate of the sensor. The integral kernel in the spectral part depends from the wavelength itself, and it is therefore not shift-invariant.

If the incoming signal is a spectral broadband and a spatial spot $\delta(\mathbf{P})$, the second integral can be neglected and the measured signal is directly proportional to the spatial integral kernel $h(\mathbf{p})$. On the other hand a spatially homogeneous but wavelength-dependent signal allows a direct measurement of the spectral dependent kernel in the second integral. Therefore the calibration process consists of a geometric calibration, and a radiometric or spectral calibration. The second task also includes noise investigations. Additional effects and imaging errors, which have or have not to be corrected, are time dependencies and polarization.

2.2 Example of a Calibration Facility

Figure 5 shows the calibration facility in the DLR Institute of Space Sensor Technology and Planetary Exploration. The upper right part is related to radiometric calibration, and the lower part in Fig. 5 illustrates geometric calibration processes. The most important part of the calibration facility is a large collimator with integrated autocollimation device which allows camera alignment. With the help of two high accuracy micro-stages, several targets (pinholes, slits or resolution masks) can be moved into the focal plane of the collimator. This allows PSF or LSF (*line-spread function*) measurements for a static line camera.

The nodal bench carries the camera and is able to rotate and tilt the camera precisely. In this way PSF measurements can be done with fixed targets. The main task of the high precision nodal bench is the accurate determination of the pixel coordinates in relation to the optical axis.

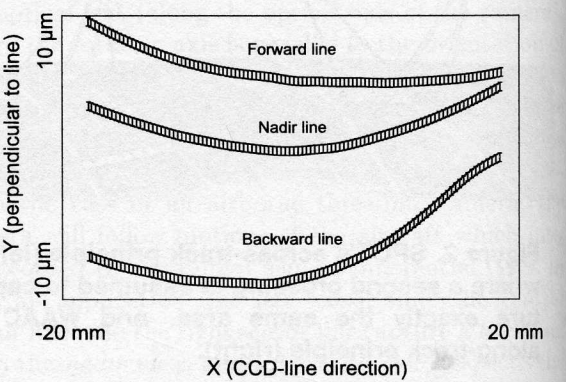

Figure 6. Result of geometric calibration.

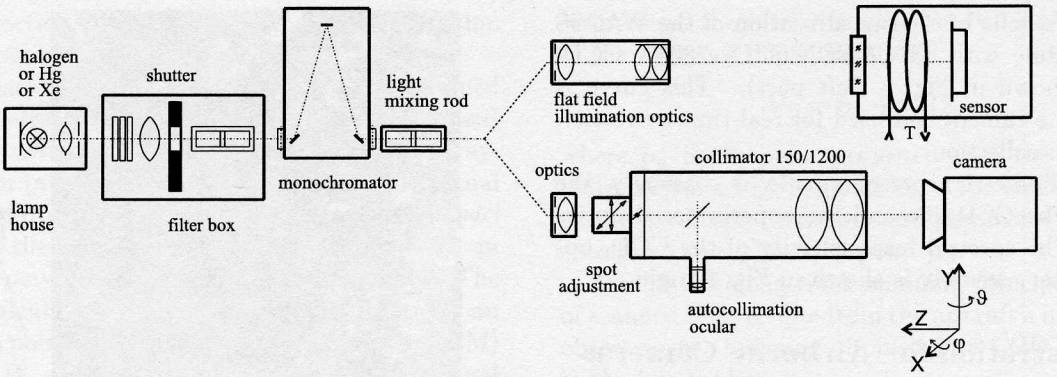

Figure 5. Example of a calibration facility.

2.3 Geometric calibration

Main tasks for geometric calibration are the highly accurate determination of the interior camera orientation and the measurements of the PSF or its Fourier transform, the *modulation transfer function* (MTF).

A measurement procedure for geometric calibration can be straightforward: a single pixel in one sensor line is illuminated by a pinhole spot from the collimator focus. The direction of the illuminated pixel to the optical axis is determined by the angle φ (angle around X-axis, perpendicular to line direction) and ϑ (angle around Y-axis, in line direction). It is not necessary to determine all the coordinates of all pixels because of very smooth changes of the geometric characteristic of the optics and the CCD-sensor.

A coordinate transformation from an arbitrary but fixed world coordinate system $\mathbf{P}_w = (X_w, Y_w, Z_w)$ to the camera-centered world coordinate system $\mathbf{P} = (X, Y, Z)$ can be calibrated as follows: the rotation of the nodal bench can be described by a rotation matrix around the X-axis $\mathbf{R_X}(\varphi)$ and Y-axis $\mathbf{R_Y}(\vartheta)$. The illumination direction or the collimator is the Z-axis. The result of this rotation is a direction (in homogenous coordinates)

$$[\sin\vartheta\cos\varphi,\ -\sin\varphi,\ \cos\vartheta\cos\varphi,\ 0]^T\ ,$$

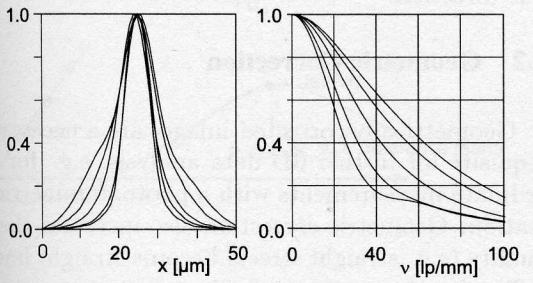

Figure 7. PSF (left) and MTF results.

The mapping of an incoming lightspot direction into a pixel location in the focal plane (for focal length f) can be described by the following transformation matrix. Ideal paraxial optics is assumed:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -f \\ 0 & 0 & f^{-1} & 0 \end{bmatrix},$$

Homogeneous coordinates $[X_1, X_2, X_3, X_4]^T$ define inhomogeneous coordinates $X = X_1/X_4$, $Y = X_2/X_4$, and $Z = X_3/X_4$. The specified relation between homogeneous coordinates of the incoming spot and focal plate coordinates leads to

$$X = f \cdot \tan\vartheta, \quad Y = -f \cdot \frac{\tan\varphi}{\cos\vartheta}, \quad Z = 0\ .$$

An example of a result of such a measurement is shown in Fig. 6, where the difference between real sensor geometry and calculated positions is less than three pixels at any end of a WAAC sensor line having 5184 pixels.

Furthermore a pixel related PSF is determined to predict the system resolution. Figure 7 shows a few single-pixel related PSFs and MTFs for one sensor line. The figure shows, that the PSF becomes broader with the distance from the center point of the camera system. The assumption of a shift-invariant spatial kernel is not strictly valid.

2.4 Radiometric calibration

Main parameters which have to be measured are *Photoresponse Non-Uniformity* (PRNU), *Spectral Response* (SR), and noise. Additional parameters are photoresponse, dark signal, *dark signal non-uniformity* (DSNU), saturation voltage, linearity, and dynamic range.

Typical results of the precalibration of the WAOSS sensor module with THOMSON THX 7808A CCD-lines are shown in Fig. 8 (left part). This curve is stored in the camera and used for real-time correction during data collection.

The radiometric correction file is basically the PRNU of the CCD-line and the superimposed optics shading. The spectral responsitivity of the CCD, optics and filter assembly is shown in Fig. 8, right.

3 Registration for Airborne Cameras

Most of the recent aerial cameras can (still) be classified as being film-based frame cameras. The film is held stationary in the focal plane during exposure, or is moved to compensate for image motion. Parameters of relative or absolute exterior orientation can be determined by control points which have to be identified in an image. For digital airborne cameras we only discuss line sensors because this seems to be the only way to compete with the very high resolution of film-based cameras.

3.1 Stereo Line Cameras

The use of three-line cameras on board of aircrafts requires special evaluation procedures because of attitude instabilities of the aircraft (roll, yaw, pitch, ground speed and altitude variation) which directly result in image distortions.

The main difference to classical film or matrix cameras is that the data evaluation of an image strip recorded from an airborne line camera requires attitude parameters for each measured image line in the strip. Figure 9 shows an image of the WAOSS camera recorded during a flight over Berlin. The influence of the aircraft's motion is better visible in the left of the image. The three curves on the right hand side of the figure show the changes in roll, pitch and yaw angles

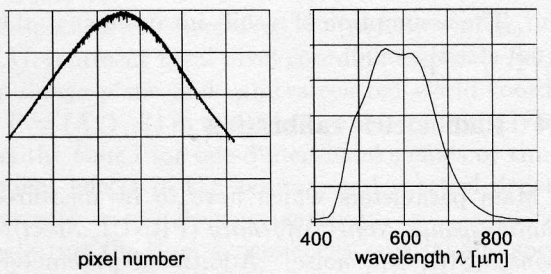

Figure 8. Radiometric correction file (left), normalized spectral sensitivity (right).

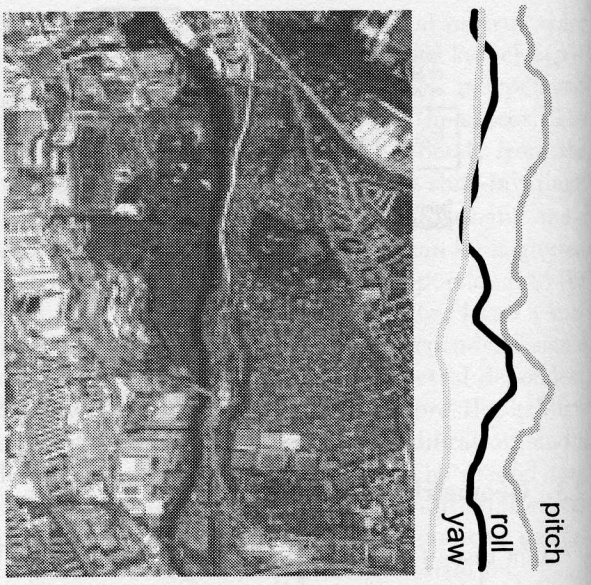

Figure 9. Detail of a captured image strip and related attitude data.

where the interval for the roll angle is $[-0.5°, 2.4°]$, for the pitch angle is $[4.0°, 4.4°]$, and for the yaw $[86°, 96°]$.

Due to the inaccuracy of a typical *inertial navigation system* (INS) conventional photogrammetric methods use additional control points on the Earth's surface to improve the attitude parameters. To improve the accuracy of attitude information an applanix-system was used. This system provides a combined hardware and software solution for the attitude parameters, which allows an absolute point accuracy in the range of one decimeter.

Using attitude and camera calibration data the image strips of the nadir, the backward and the forward looking line can be geometrically corrected assuming a flat underlying surface (*rectification on a plane*). The resulting image is equivalent to an undisturbed flight over a planar surface. Captured strips (using one of the lines of a multi-line camera) are processed with respect to the same flight path. This approach ensures a stereo image geometry equivalent to the epipolar geometry of ideal binocular vision, allowing parallax shifts in image line direction.

3.2 Geometric Correction

Geometrically corrected images are a necessary prerequisite for further 3D data analysis, e.g. for 3D coordinate measurements with a photogrammetric workstation. Geometric correction also increases the image quality (e.g. straight streets become straight lines) and influences the outcome of subsequent stereo matching. The knowledge of attitude parameters allows a geo-

metric correction of an image strip. The algorithm proceeds in two steps.

(1) For each CCD-line pixel it has to be determined which geometric distortion in object space was caused by aircraft movements. This task is equivalent to ray-tracing from (a yet unknown, to be calculated) actual position and direction, into the *digital elevation model* (DEM). As the original DEM is unknown as well, an ideal DEM may be assumed as a reference plane. The following calculation is based on this assumption: an intersection point with the unknown surface (the DEM) is given as $\mathbf{P}_r = \mathbf{P}_a + t \cdot \mathbf{D}_a$, where $\mathbf{P}_a$ is the actual camera location and $\mathbf{P}_r$ is the intersection point with the assumed reference plane $Z = Z_r$, and $\mathbf{D}_a$ is the actual direction vector. The latter is related to the undisturbed direction $\mathbf{D}_u$ by a rotation matrix $\mathbf{R}$,

$$\mathbf{D}_a = \mathbf{R} \cdot \mathbf{D}_u \ ,$$

where $\mathbf{R}$ contains the disturbance of the flight path: roll, pitch and yaw, as described before. In Fig. 10, angle β corresponds to the pitch of the aircraft, and the angle ω is the stereo angle.

(2) The pixel value is then backprojected from the calculated intersection point $\mathbf{P}_r$ with the reference plane into $\mathbf{p}_u$ in ideal image plane of a hypothetic camera assumed to be moving on an ideal (undisturbed) linear flight trajectory, see Fig. 10. The number of parameters for describing the external orientation is reduced this way for the whole measured swath. It suffices to specify velocity v and height h: the aircraft is assumed to have linear uniform motion with constant velocity v on a given track at a constant altitude $Z = h$.

Altogether, the task consists in determining which pixel (x, y), i.e. point $\mathbf{p}_u = (x, y, h)$ in an ideal image strip, 'sees' the intersection point $\mathbf{P}_r = (X_r, Y_r, Z_r)^T$ in the reference plane, satisfying

$$\mathbf{P}_r = \mathbf{p}_u + t \cdot \mathbf{D}_u \ ,$$

where $\mathbf{p}_u$ is the point projected onto the ideal flight trajectory (see Fig. 10), and $\mathbf{D}_u$ is the undisturbed orientation vector. Unknowns are parameter t and pixel coordinates (x, y) for the point $\mathbf{p}_u$.

Due to the along-track scanning principle the order of scanned lines results from the aircraft's movement in object space. Indexing of pixels in a CCD-line is done in the image plane.

The result of a simple approximation is $x = x_0/(v \cdot \tau)$ where τ is the scan time for a single line, and

$$x_0 = \begin{cases} X_r \ , & \text{if nadir line} \\ X_r \pm (h - Z_r) \cdot \tan\omega \ , & \text{if back- or forward line} \end{cases}$$

and

$$y = \frac{n+1}{2} + \frac{Y_r \cdot f}{(h - Z_r) \cdot \triangle_{\text{IFOV}}} \ ,$$

where n is the number of pixels per sensor line, $\triangle_{\text{IFOV}}$ is the linear size of a sensing element, and f is the focal distance. The pixel index y is independent of the viewing angle of the CCD-line (forward, backward, or nadir). Therefore the geometrical correction creates images which correspond to the stereo geometry of binocular vision.

Figure 11 shows a corrected subwindow of the image in Fig. 9. The figure shows that blurring effects in the uncorrected image can be compensated by the described pixel reordering procedure.

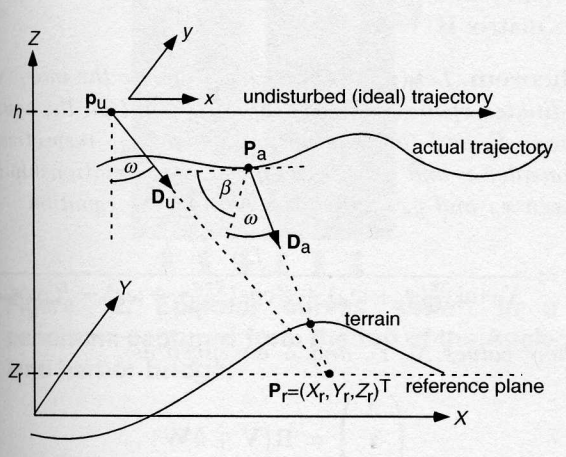

Figure 10. Aircraft attitude instabilities.

4 Epipolar Geometry

In stereo camera systems corresponding points are restricted to be on epipolar curves (in image planes) defined by a 3D surface point (or one of its projections) and the origins of the camera's coordinate systems. This allows that stereo matching only has to proceed along an epipolar line in one image of a given stereo pair of images. This is of crucial importance to ensure time-efficient search routines for corresponding points. We discuss two possible stereo approaches using line cameras.

4.1 Along-track Stereo

The epipolar geometry of cameras tilted towards each other are discussed in [21, 34]. However, a multi-line camera used for image acquisition following the along-track stereo approach is related to an even simpler stereo geometry as discussed before. After rectification, all captured images are aligned, and epipolar

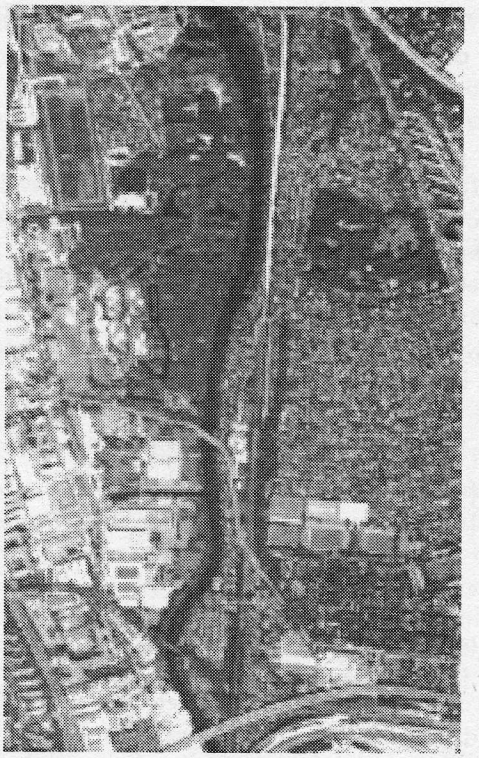

Figure 11. Detail of the corrected image.

lines of one surface point are simply identical straight lines in all channels captured by a multi-line camera with an architecture as shown in Fig. 1.

This justifies that this stereo geometry of a multi-line camera supports the use of a stereo matching procedure which proceeds along a given straight line. A dynamic programming approach as discussed later is very suitable for such a situation.

4.2 Panoramic Stereo

As mentioned above, polycentric panoramas denote a wide range of multi-perspective panoramic images [15]. A set of panoramic images all acquired with respect to the same rotation axis is referred to as a set of *concentric panoramic images*. In [37] it is shown that epipolar geometry consists of horizontal lines if two height-aligned concentric panoramic images are *symmetric*, i.e. when the associated angular parameters of these two panoramic images are ω and $-\omega$. This matches our model in Fig. 1 and the architecture of the WAAC camera.

A panoramic image acquired with a single focal point, i.e. $R = 0$, is referred to as a *single-center panoramic image* [3, 24, 25]. This matches the EYESCAN architecture. A study about epipolar curves in a pair of single-center panoramic images can be found in [24]. Images captured by EYESCAN at the same tripod location but just with different heights are also examples of concentric panoramic images (see Fig. 17).

We briefly inform about the derivation of an epipolar curve equation for a pair of polycentric panoramic images as given in [15]. The epipolar curve equation provides a unified approach for epipolar geometry studies in any more specific class of panoramic images.

Let us consider a pair of polycentric panoramic images, a *source image* E and a *destination image* E_d, having a width (i.e. number of columns) of W and W_d, respectively. Given an image point $\mathbf{p}$ with image coordinates (x, y) on E. Possible locations of a corresponding point $\mathbf{p}_d$, with image coordinates (x_d, y_d), on E_d may be constrained by an epipolar curve which is specified in the theorem below. In this theorem we use a number k and vectors $\mathbf{V}$ and $\mathbf{W}$ which are defined as follows:

$$k = \frac{R_d \sin\omega_d + \cos(\frac{2\pi x_d}{W_d} + \omega_d)\mathbf{r}_1^T \cdot \mathbf{V} - \sin(\frac{2\pi x_d}{W_d} + \omega_d)\mathbf{r}_3^T \cdot \mathbf{V}}{\sin(\frac{2\pi x_d}{W_d} + \omega_d)\mathbf{r}_3^T \cdot \mathbf{W} - \cos(\frac{2\pi x_d}{W_d} + \omega_d)\mathbf{r}_1^T \cdot \mathbf{W}},$$

$$\mathbf{V} = \begin{pmatrix} R\sin(\frac{2\pi x}{W}) \\ 0 \\ R\cos(\frac{2\pi x}{W}) \end{pmatrix} - \mathbf{T},$$

and

$$\mathbf{W} = \begin{pmatrix} \sin(\frac{2\pi x}{W} + \omega)\cos(\tan^{-1}(\frac{y}{f})) \\ \sin(\tan^{-1}(\frac{y}{f})) \\ \cos(\frac{2\pi x}{W} + \omega)\cos(\tan^{-1}(\frac{y}{f})) \end{pmatrix}.$$

The 3×3 rotation matrix $\mathbf{R}$, and the 3×1 translation vector $\mathbf{T}$ specify the orientation and the location of the destination turning-rig coordinate system with respect to the source turning-rig coordinate system. Furthermore, $\mathbf{r}_1^T$ and $\mathbf{r}_3^T$ are the first and third row vectors of the matrix $\mathbf{R}$, respectively.

Theorem. *Let (x, y) and (x_d, y_d) denote the image coordinates of the projection of a 3D point on the source image E and the destination image E_d, respectively. Consider x and y as being given. The relationship between x_d and y_d can be described by the equation*

$$y_d = \frac{f_d Y}{X\sin(\frac{2\pi x_d}{W_d} + \omega_d) + Z\cos(\frac{2\pi x_d}{W_d} + \omega_d) - R_d\cos\omega_d},$$

where values X, Y, and Z are given as

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \mathbf{R}(\mathbf{V} + k\mathbf{W}),$$

and a corresponding point (x_d, y_d) is only valid if the value of the denominator of y_d is greater than zero.

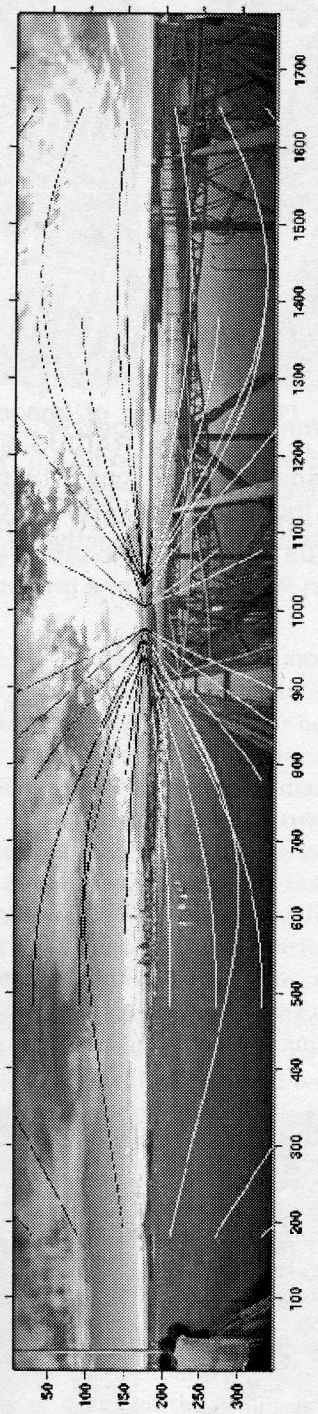

Figure 12. Epipolar curves shown in a panorama captured from the top of the Auckland harbor bridge.

This theorem from [15] provides a general epipolar curve equation for arbitrary pairs of polycentric panoramas, which describes the relationship between

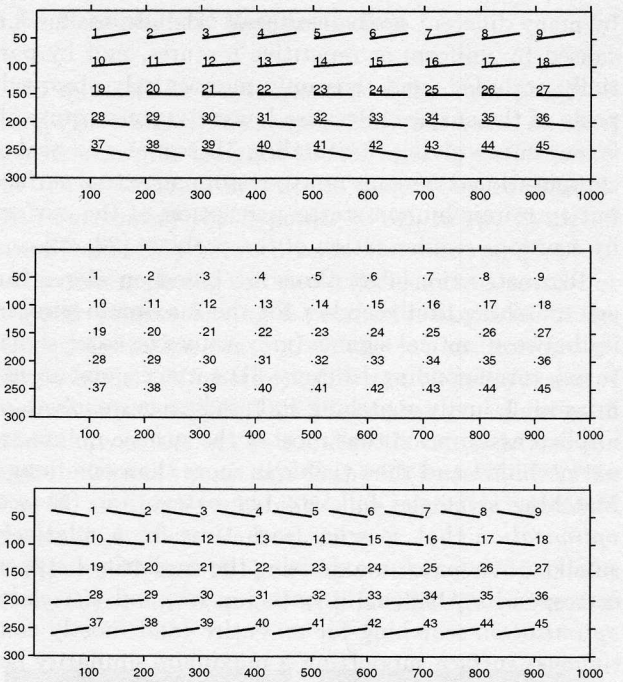

Figure 13. Epipolar curves for WAAC.

the coordinates x_d and y_d of $\mathbf{p}_d$. Figure 12 shows curves for a pair of horizontally aligned panoramas with identical values of R but different values of ω.

Assume that a three-line camera as shown in Fig. 1 is used for capturing a triple of panoramic images E_1 (for ω), E and E_2 (for $-\omega$). The focal length f_1 for E_1, and f_2 for E_2 satisfies $f_1 = f_2 = f/cos\omega$. We assume that we know parameters W, W_1 and W_2 which are the numbers of columns in these three panoramic images. The task is that we have a point (x, y) in E, and we have to specify epipolar curves for all the possible locations of corresponding points (x_1, y_1) in E_1, and (x_2, y_2) in E_2. It follows that

$$y_1 = \frac{y}{\cos\omega} \cdot \frac{\sin(\theta + \omega) - \sin\omega}{\sin\theta}$$

with $\theta = \frac{2\pi}{W_1} x_1 - \frac{2\pi}{W} x$, and

$$y_2 = \frac{y}{\cos\omega} \cdot \frac{\sin\omega - \sin(\theta + \omega)}{\sin\theta}$$

with $\theta = \frac{2\pi}{W_2} x_2 - \frac{2\pi}{W} x$. Figure 13 illustrates these epipolar lines for focal length $f = 21.7\ mm$, rotation radius $R = 50\ cm$ and $\omega = 25°$.

5 Stereo Matching

From a mathematical point of view, stereo reconstruction is an ill-posed problem with no unique solution because a stereo pair of images might be produced

by many different optical surfaces. The ill-posedness is caused by uniform or repetitive textures, and by partially occluded (and thus only monocularly observed) parts of the surface. Because human vision is quite efficient in perceiving surrounding 3D scenes, the goal of computational stereo is not to restore an actual surface but to mimic human stereo perception of the surface by a proper regularization of the problem [22, 29].

Reconstruction algorithms are based on stereo image matching that searches for the maximum similarity between optical signals (grey values or color signatures) corresponding to every 3D surface point in the images. Usually matching is based on an explicit or implicit assumption that most of the surface points are not occluded and thus visible in more than one image. Matching strategies fall into two categories: (*i*) *local optimization* that searches each time for a relatively small surface patch maximizing the similarity between corresponding regions in both images, and (*ii*) *global optimization* searching for an entire (and mostly continuous) surface supporting a maximum similarity between both images [2, 7].

5.1 Correlation Methods

When the corresponding image signals only have random deviations, the similarity for a stereo pair can be measured in terms of Cartesian or Euclidean distance between both signals. But this assumption does not hold in practice. More general assumption of uniform contrast and brightness variations results in a cross-correlation measure of similarity [7, 12]. Due to its theoretical and practical simplicity, correlation-based stereo matching is used for many years, and still it is the most popular matching technique in photogrammetry.

To take account of projective geometric deformations of corresponding image regions, photogrammetric matching is usually performed by an exhaustive least-square search for the most appropriate terrain slope [7, 14]. To simplify the search, a planar terrain patch is assumed, and projective transformations are affinely approximated. Experiments show that in this latter case a combined directed and exhaustive search for the optimal least-square correlation allows to correctly match pairs of substantially deformed images of the same 3D scene obtained from different positions by various cameras [11].

Figures 14 and 15 show some results of a least-square correlation matching of entire images of size 83×65 depicting the same 3D model-board scene "RADIUS-M" [30]. The correct matching was obtained for many other image pairs although it failed for some pairs with

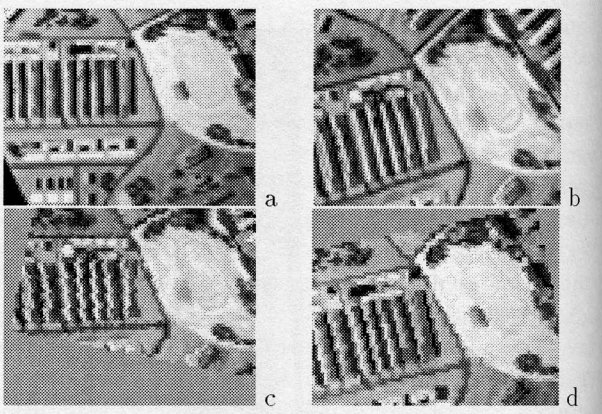

Figure 14. [11] Matching of images M11 (a) and M19 (b) from the RADIUS-M set: the affinely transformed images M19 (c) and M11 (d) yield the best least-square correlation found (0.68 and 0.82, respectively).

dominating homogeneous regions.

Independent local matching usually results in discontinuous and mutually incompatible surface patches so that a complicated post-processing has to be involved to obtain a self-consistent final surface. More reliable reconstruction is obtained if the reconstructed surfaces are assumed to be continuous and with no repetitive texture and the set of possible matches is restricted to only stable ones [31]. The stable matching means that for every stereo correspondence there are no competing variants with higher correlation values that are not over-competed by other correspondences in this matching.

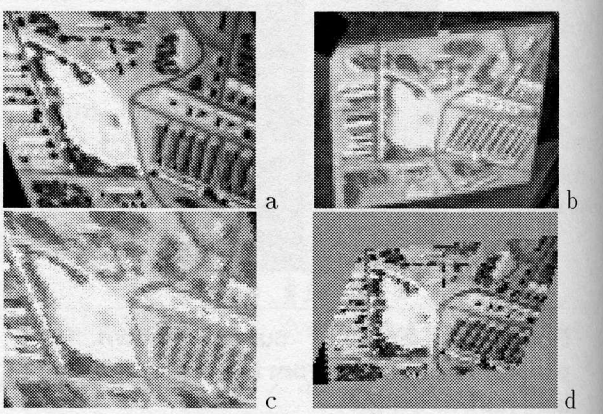

Figure 15. [11] Matching of images M12 (a) and M40 (b) from the RADIUS-M set: the affinely transformed images M40 (c) and M12 (d) yield the best least-square correlation found (0.76 and 0.66, respectively).

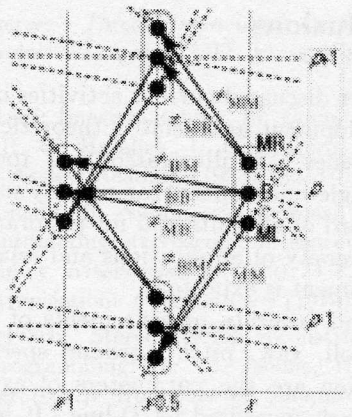

Figure 16. [8, 9] Nodes of the symmetric GVP (B, ML, MR - the binocularly visible 3D point or the point observed only monocularly by left or right stereo camera, respectively; $\pi_{s_c|s_p}$ - probability of transition from the state s_p to s_c in the preceding and current GVP nodes).

5.2 Global Approaches

An alternative (but theoretically less innovative) approach to exclude inconsistent correlation-based matches is based on following a desired terrain by choosing at each step its most appropriate local extension. To suppress error accumulation (it can break the process far away from the actual surface), adaptive selections of regions to be correlated can be used [18].

Stereo by global optimization is based on stating and solving the statistical problem of estimating *hidden Markov models* (HMM) of epipolar terrain profiles. Every digital epipolar profile is modelled by a Markov chain of transitions between the neighboring nodes of a specific *graph of variants of the profile* (GVP). In general, every node has potential visibility states as shown in Fig. 16. Each state indicates whether the node is binocularly (B) or only monocularly visible (ML,MR) in the left or the right image of a stereopair [9].

The prior Markov–chain model of a profile is combined with a conditional model of image signals. The resulting posterior model of the profile is used to measure similarity between both images and implement Bayesian or other optimum statistical decision rules to recover the HMM. The HMM allows for taking account of possible partial occlusions, and the global optimization is usually performed by dynamic programming [1, 2, 5, 8, 23, 26]. But some other approaches such as a gradient-based global matching using similarity measures that allow to partly avoid traps of multiple local maxima were also proposed [17].

5.3 Probabilistic Regularization

In most of the above-mentioned dynamic programming stereo algorithms, regularization is based on some heuristic weights assigned to *monocularly visible points* (MVP) of a terrain in order to make them compatible with *binocularly visible points* (BVP) in specifying an overall similarity between both stereo images [5, 9].

By reformulating the stereo problem as a search for epipolar profiles having a maximum likelihood ratio, the regularization can be related to transition probabilities in the GVP between different visibility states of successive nodes [10]. This allows that probabilities of transitions to the MVPs are directly derived from the probabilities of transitions to the BVPs in the GVP in Fig. 16, and both probabilities are explicitly related to the similarity between both images.

Figures 17 and 18 show the vertical stereopair of a close-range scene acquired with EYESCAN and results

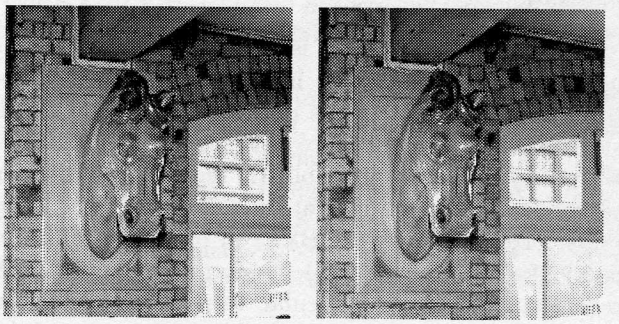

Figure 17. Upper (left) and lower (right) images of a vertical close-range stereopair "Horse" 886 × 818 (segment of a larger panoramic stereo image captured by EYESCAN at two different heights).

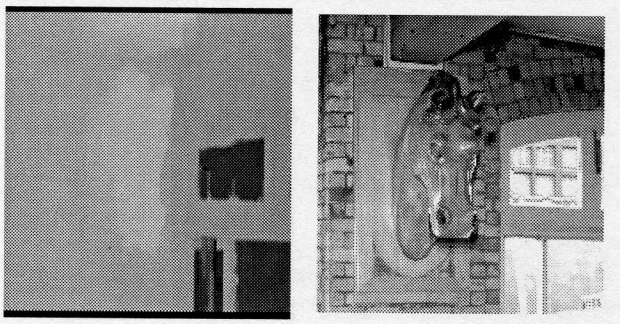

Figure 18. SDPS reconstruction of the "Horse" scene: the grey-coded range (DEM) image (left) and the corresponding cyclopean image (right) formed by fusing images of the stereo pair in line with the DEM.

Figure 19. Backward (left) and nadir (right) images obtained by the 3-line scanner WAAC to form the horizontal stereopair "Berlin" (a 2400×2400 window of a larger stereo image).

of the *symmetric dynamic programming stereo* (SDPS) reconstruction [9] of the scene. This variant of the SDPS exploits the above-mentioned probabilistic regularization.

The DEM of this scene is represented in Fig. 18 as a grayscale-encoded range image. The related cyclopean image of the scene is obtained by merging both stereo images in accordance with the DEM. These results show that the probabilistic regularization yields rather accurate stereo matching (the corresponding cross-correlation between BVPs in both stereo images is 0.92 for the above DEM). Similar results of a SDPS reconstruction of a particular terrain are shown in Figs. 19 and 20. In both examples most points of the DEM seem to be correct although there are unavoidable errors in regions with uniform textures (such as, for instance, river and lake surfaces in the "Berlin" scene or windows in the "Horse" scene).

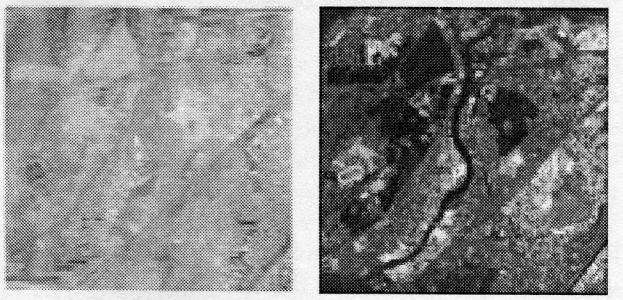

Figure 20. SDPS reconstruction of the "Berlin" scene: the grey-coded range (DEM) image (left) and the corresponding cyclopean image (right) formed by fusing images of the stereo pair in line with the DEM.

6 Conclusions

The paper discussed current activities in wide-angle sensor development and related theoretical and algorithmical issues. It illustrated that today's digital multi- or single-line cameras can already successfully be used on board of aircrafts and for panoramic imaging. A broad diversity of applications and related research and development is expected.

Direct and accurate measurements of attitude parameters (roll, yaw, pitch), ground speed, and altitude variation are the parameters of exterior orientation for each measured CCD-line. It has been discussed that attitude and camera calibration data allow that CCD-line scanner data can be geometrically corrected and successfully used for stereo analysis, shape reconstruction and visualisation. Digital resolutions recently already available in line cameras approach the optimum resolutions available with film-based cameras, time-consuming chemical photo processing and expensive large-scale hardcopy scanners. Complete processing lines connecting digital image acquisition, transmission, and stereo analysis will redefine digital photogrammetry and panoramic imaging within the next years.

Acknowledgements

The authors thank Anko Börner (DLR), Chia-Yen Chen (CITR), Fay Huang (CITR), Hartmut Korsitzky (DLR), Hao Li (CITR), Martin Scheele (DLR), Karsten Scheibe (DLR), Hans-Jörg Schönherr (KST), Reinhard Schuster (DLR), Michael Solbrig (DLR), Shou Kang Wei (CITR) and Jian Zhong (CITR) for joint work which has been discussed or cited in this paper.

References

[1] H. H. Baker, T. O. Binford: Depth from edge and intensity based stereo. In: Proc. 7th Intern. J. Conf. *Artificial Intelligence*, (1981) Vol. 2, 631–636.

[2] H. H. Baker: Surfaces from mono and stereo images. *Photogrammetria* **39** (1984) 217–237.

[3] S. E. Chen: QuickTimeVR - an image-based approach to virtual environment navigation. *Proc. SIGGRAPH'95*, (1995) 29–38.

[4] Ch.-Y. Chen, R. Klette: Image stitching - comparisons and new techniques. In: Proc. *CAIP*, (September 1999) 615–622.

[5] I. J. Cox, S. L. Hingorani, S. B. Rao: A maximum likelihood stereo algorithm. *Computer Vision and Image Understanding* **63** (1996) 542–567.

[6] O. Faugeras: *Three-Dimensional Computer Vision: A Geometric Viewpoint.* MIT Press, Cambridge, US (1993).

[7] W. Förstner: Image matching. In: *Computer and Robot Vision* (R. W. Haralick, L. G. Shapiro), Chapter 16. Addison-Wesley, Reading (1993) 289–378.

[8] G. L. Gimel'farb: Symmetric approach to the problem of automating stereoscopic measurements in photogrammetry. *Kibernetika* **2** (1979) 73–82 [in Russian; English translation: *Cybernetics* **8** (1979) 235–247].

[9] G. Gimel'farb: Stereo terrain reconstruction by dynamic programming. In: B. Jaehne, H. Haussecker, P. Geisser (eds.): *Handbook of Computer Vision and Applications* **2**: *Signal Processing and Pattern Recognition.* San Diego, Academic Press (1999) 505–530.

[10] G. Gimel'farb, H. Li: Probabilistic regularization in symmetric dynamic programming stereo. In: Proc. *Image and Vision Computing New Zealand 2000*, Hamilton, New Zealand (2000) 144–149.

[11] G. Gimelfarb, J. Zhong: Matching multiple views by the least square correlation. *Multi-Image Analysis* (R. Klette, Th. Huang, G. Gimel'farb, eds.), **LNCS-2032** Springer, Berlin (2001) 107–117.

[12] M. J. Hannah: Digital stereo image matching techniques. *Intern. Arch. Photogrammetry and Remote Sensing* **XXVII** (1988) 280–293.

[13] R. Hartley, A. Zisserman: *Multiple View Geometry in Computer Vision.* University Press, Cambridge, UK (2000).

[14] U. V. Helava: Object-space least-square correlation. *Photogrammetric Engineering and Remote Sensing.* **54** (1988) 711–714.

[15] F. Huang, S. K. Wei, R. Klette: Epipolar geometry in polycentric panoramas *Multi-Image Analysis* (R. Klette, Th. Huang, G. Gimel'farb, eds.), **LNCS-2032** Springer, Berlin (2001) 40–51.

[16] H. Ishiguro, M. Yamamoto, S. Tsuji: Omni-directional stereo. *IEEE-PAMI*, **14** (1992) 257–262.

[17] H. Jahn: Stereo matching for pushbroom stereo cameras. *Intern. Arch. Photogrammetry and Remote Sensing* **XXXIII**:B3H, Comm. III (2000) 436–443.

[18] T. Kanade, M. Okutomi: A stereo matching algorithm with an adaptive window: theory and experiment. *IEEE-PAMI* **16** (1994) 920–932.

[19] K. Kanatani: *Geometric Computation for Machine Vision.* Clarendon, Oxford, US (1993).

[20] S.B. Kang, R. Szeliski: 3-d scene data recovery using omnidirectional multibaseline stereo. *IJCV*, **25** (1997) 167–183.

[21] R. Klette, K. Schlüns, A. Koschan: *Computer Vision - Three-Dimensional Data from Images.* Springer, Singapore, (1998).

[22] V. R. Kyreitov: *Inverse Problems of Photometry.* Computing Centre, Siberian Branch, Academy of Sciences of the USSR, Novosibirsk (1983) [In Russian].

[23] S. A. Lloyd: Stereo matching using intra- and inter-row dynamic programming. *Pattern Recognition Letters* **4** (1986) 273–277.

[24] L. McMillan, G. Bishop: Plenoptic modeling: an image-based rendering system. *Proc. SIGGRAPH'95*, (1995) 39–46.

[25] T. Matsuyama, T. Wada: Cooperative distributed vision - dynamic integration of visual perception, action, and communication. In: Proc. 2nd Internat. Workshop *Cooperative Distributed Vision*, Kyoto, (1998) 1–40.

[26] Y. Ohta, T. Kanade: Stereo by intra- and inter-scanline search using dynamic programming. *IEEE-PAMI* **7** (1985) 139–154.

[27] S. Peleg, M. Ben-Ezra: Stereo panorama with a single camera. *CVPR99*, (1999) I:395–401.

[28] S. Peleg, Y. Pritch, M. Ben-Ezra: Cameras for stereo panoramic imaging. Technical report, Computer Vision Lab, Institute of Computer Science of The Hebrew University, (2000).

[29] T. Poggio, V. Torre, C. Koch: Computational vision and regularization theory. *Nature* **317** (1985) 317–319.

[30] *RADIUS Model Board Imagery and Groundtruth.* CD-ROM, vols. 1 and 2. ISL, University of Washington, Seattle (1996).

[31] R. Šára: Stable monotonic matching for stereoscopic vision. *Robot Vision* (R. Klette, S. Peleg, G. Sommer, eds.), **LNCS-1998** Springer, Berlin (2001) 184–192.

[32] K. Scheibe, H. Korsitzky, R. Reulke, M. Scheibe, M. Sobrig: EYESCAN - a high resolution digital panoramic camera. *Robot Vision* (R. Klette, S. Peleg, G. Sommer, eds.), **LNCS-1998** Springer, Berlin (2001) 77–83.

[33] R. Schuster: Sensor calibration and geometric calibration of a three line stereo camera. *Int. Arch. of Photogrammetry and Remote Sensing*, **30**, 5W1 (1995) 265–271.

[34] Y. Shirai: *Three-Dimensional Computer Vision.* Springer, Berlin, (1986).

[35] H.-Y. Shum, L.-W. He: Rendering with concentric mosaics. *Proc. SIGGRAPH'98*, Los Angeles, (August 1999) 299–306.

[36] H. Shum, A. Kalai, S. Seitz: Omnivergent stereo. *ICCV99*, Korfu/Greece, (1999) 22–29.

[37] H. Shum, R. Szeliski: Stereo reconstruction from multiperspective panoramas. *ICCV99*, Korfu/Greece, (1999) 14–21.

[38] S. K. Wei, F. Huang, R. Klette: Classification and characterization of image acquisition for 3D scene visualization and reconstruction applications. *Multi-Image Analysis* (R. Klette, Th. Huang, G. Gimel'farb, eds.), **LNCS-2032** Springer, Berlin (2001) 83–94.

[39] G. Xu, Z. Zhang: *Epipolar Geometry in Stereo, Motion and Object Recognition - A Unified Approach.* Kluwer, Amsterdam (1996).