. *gsize*11

A SEQUENCE OF STEREO IMAGE DATA OF A MOVING VEHICLE IN AN OUTDOOR SCENE

Yuncai Liu and Thomas S. Huang

Beckman Institute
University of Illinois at Urbana-Champaign
405 North Mathews Street
Urbana, IL 61801, U.S.A.

August 10, 1990

A SEQUENCE OF STEREO IMAGE DATA FROM A MOVING VEHICLE IN AN OUTDOOR SCENE

Yuncai Liu and Thomas S. Huang

Beckman Institute
University of Illinois at Urbana-Champaign
405 North Mathews Street
Urbana, IL 61801, U.S.A.

I. Introduction

A digital stereo image sequence data base has been created under the joint effort of the University of Illinois at Urbana-Champaign, Purdue University (Photogrammetry Group), and the U. S. Army Engineer Topographic Laboratory (AI Center). Part of this data base, specifically three pairs of stereo images taken at consecutive time instants (time instants 13, 14, and 15 from a long sequence) is now made available to the research community. More will be available later. We hope that different algorithms will be applied to these images and the results compared.

Two stationary cameras were used to take images of a moving vehicle in an outdoor scene. The problem is to determine the 3D motion of the vehicle. The image date can be used to test motion estimation algorithms based on either stereo or monocular image sequences. In the latter case, only one side of the sequence (from the left or the right camera only) is to be used. Approximate ground truth for the vehicle motion has been established; however, we choose not to reveal it to the public at large at this time.

Each of the six images has a size of 750 × 750 pixels with 8 bits/pixel. The image data are available on magnetic tape, and can be extracted by the command *tar xvf/dev/rst8 stereo*. In this report, we give formulas for coverting the image point coordinates from pixels to the actual metric distances on the image plane in the camera; and formulas for correcting geometric distortions due to the camera lens.

II. Camera Setup

The original stereo images were taken by two fixed focal length A.M.I./Bronica SQ-AM metric 70 mm camera with 40 mm nominal focal length. The cameras are parallelly mounted on the ground with a base line about three meters. The optical axes of the cameras are in the **negative z-direction** of the ground coordinate system, as shown in Figure 1. If (X', Y') are the image

Figure 1. Relationship between the ground coordinate system and the cameras

coordinates of a point in the image plane of a camera after geometric distortion correction, and (x, y, z) are the coordinates of the point in the ground coordinate system, the two are related by

$$\begin{bmatrix} X' - X_p \\ Y' - Y_p \\ -f \end{bmatrix} = s\mathbf{M} \begin{bmatrix} x - x_c \\ y - y_c \\ z - z_c \end{bmatrix}$$

where (X_p, Y_p) is the principal point of the camera and f is the focal length, whose values for the left and the right cameras can be found in Appendices A and B; s is a scale factor; $\mathbf{M}$ is a rotation matrix with

$$\mathbf{M} = \begin{bmatrix} \cos\kappa & \sin\kappa & 0 \\ -\sin\kappa & \cos\kappa & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\phi & 0 & -\sin\phi \\ 0 & 1 & 0 \\ \sin\phi & 0 & \cos\phi \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\omega & \sin\omega \\ 0 & -\sin\omega & \cos\omega \end{bmatrix}$$

where ω, ϕ and κ are rotation angles; (x_c, y_c, z_c) is the camera center position in the ground coordinate system. The values of (x_c, y_c, z_c) and ω, ϕ and κ are given in Appendix C.

III. Image Data Reduction

Each image was digitized on a Perkin-Elmer scanning microdensitometer to 4096 × 4096 pixels with 16 bits/pixel. The images are first subsampled by a factor of 2, resulting in images of size 2048 × 2048 pixels. Then, subimages are cut out by a window of size 750 × 750 pixels. See Figure 2. The coordinates of the upper-left corner of the window are (675, 415) in the 2048 × 2048 images. If (i, j) are the row and column numbers of a point counting from top and left, respectively, in a 750 × 750 image, the coordinates of the point in the image plane of the camera should be

$$X = [2\,(Y_{cut} + j) - Y_0]\,h$$

Figure 2. A subimage is cut out by a 750 × 750 window

$$Y = [X_0 - 2\ (X_{cut} + i)]\ w$$

where (X_{cut}, Y_{cut}) are the coordinates of the upper-left corner of the 750 × 750 window with the values (675, 415) in the 2048 × 2048 images; (X_0, Y_0) are the coordinates of the image center of the corresponding 4096 × 4096 images, whose value will be given in Appendix D for each image; h and w are the height and the width of a pixel in the original 4096 × 4096 images: $h = w = 13 \times 10^{-3}$ mm. The intensity range of the images is reduced to 8 bits.

IV. Image Data Correction

The original images have inherent geometrical distortions. Two major ones are the lens distortion due to the lens of a camera and the film distortion due to the instability of film bases. Of the two distortions, lens distortion has a dominant effect. The film distortion is negligible. Therefore, only the lens distortion is considered for geometric distortion correction. The formulas used for lens distortion correction are

$$X' = X + (X - X_p)\,(k_1 r^2 + k_2 r^4 + k_3 r^6) + 2p_1\,[r^2 + 2(X - X_p)^2] + p_2(X - X_p)(Y - Y_p)$$

$$Y' = Y + (Y - Y_p)\,(k_1 r^2 + k_2 r^4 + k_3 r^6) + 2p_1(X - X_p)(Y - Y_p) + p_2\,[r^2 + 2(Y - Y_p)^2]$$

where

(X', Y') are the corrected image coordinates

(X, Y) are the image coordinates before correction.

(X_p, Y_p) are the coordinates of the calibrated principal point of the lens.

$r = \sqrt{(X - X_p)^2 + (Y - Y_p)^2}$, the distance from point (X, Y) to the principal point.

k_1, k_2, k_3 are the radial distortion coefficients of the lens.

p_1, p_2 are the decentering distortion coefficients of the lens.

All the calibration parameters of the lenses are given in Appendices A and B.

Acknowledgement

The presons responsible for the creation of the image sequence data base are:

University of Illinois: T. S. Huang, S. D. Blosteine, Y. C. Liu, M. K. Leung.

U. S. Army Engineer Topographic Laboratory: M. McDonald, W. Seemuller, A. Werkheiser.

Purdue University: E. M. Mikhail, F. C. Paderes.

At University of Illinois, the work has been supported by National Science Foundation Grant IRI-89-08255 and Grant IRI-89-02728, and the State of Illinois Department of Commerce and Community Affairs Grant SCCA 90-82140.

APPENDIX A

Calibration Report for Left Camera

Camera Type: A.M.I./Bronica SQ-AM	Nom. focal length: 40 mm
Lens Type: Zenzanon-S	Test focal setting: infinity

1) Principal Distance: f = 41.357 mm ± 0.007 mm

2) Calibrated Principal Point:

$X_p =$ 0.035 mm ±0.005 mm

$Y_p =$ - 0.019 mm ±0.006 mm

3) Calibrated Lens Distortion Parameters:

$k_1 = \ 5.32800E-05 \quad \pm 4.10E-07$

$k_2 = -4.12600E-08 \quad \pm 5.30E-10$

$k_3 = \ 8.58900E-12 \quad \pm 1.90E-13$

$p_1 = -4.16400E-06 \quad \pm 1.20E-06$

$p_2 = -3.87400E-06 \quad \pm 1.20E-06$

APPENDIX B

Calibration Report for Right Camera

Camera Type: A.M.I./Bronica SQ-AM

Nom. focal length: 40 mm

Lens Type: Zenzanon-S

Test focal setting: infinity

1) Principal Distance: f = 41.337 mm ± 0.007 mm

2) Calibrated Principal Point:

X_p = 0.020 mm ±0.003 mm

Y_p = 0.037 mm ±0.005 mm

3) Calibrated Lens Distortion Parameters:

$k_1 = \quad 4.99579E-05 \quad \pm 5.17E-07$

$k_2 = -3.62418E-08 \quad \pm 8.74E-10$

$k_3 = \quad 5.41550E-12 \quad \pm 4.85E-13$

$p_1 = -2.29110E-06 \quad \pm 5.38E-07$

$p_2 = -7.36123E-06 \quad \pm 6.02E-07$

APPENDIX C

CAMERA POSITIONS AND ORIENTATIONS MEASURED IN THE GROUND COORDINATE SYSTEM

1. Left Camera:

Image frame 13:

(x_c, y_c, z_c) = (35413.20, 1985.56, 68171.65)

(ω, ϕ, κ) = (– 0^0 21 ' 54.703", 9^0 3' 23.145", – 1^0 1' 27.530")

Image frame 14:

(x_c, y_c, z_c) = (35413.94, 1989.10, 68175.55)

(ω, ϕ, κ) = (– 0^0 22 ' 13.866", 9^0 3' 32.466", – 1^0 1' 19.863")

Image frame 15:

(x_c, y_c, z_c) = (35411.88, 1986.32, 68168.27)

(ω, ϕ, κ) = (– 0^0 21 ' 59.404", 9^0 3' 27.535", – 1^0 1' 36.469")

2. Right Camera:

Image frame 13:

(x_c, y_c, z_c) = (38418.24, 1979.32, 67881.57)

(ω, ϕ, κ) = (– 0^0 43 ' 58.416", 9^0 43' 52.854", – 1^0 7' 2.598")

Image frame 14:

(x_c, y_c, z_c) = (38417.79, 1977.91, 67879.94)

(ω, ϕ, κ) = (– 0^0 44 ' 2.496", 9^0 43' 54.974", – 1^0 7' 16.905")

Image frame 15:

$(x_c,\ y_c,\ z_c)$ = (38418.54, 1977.87, 67879.70)

$(\omega,\ \phi,\ \kappa)$ = ($-\ 0^0$ 43 ' 48.873", 9^0 43' 0.437", $-\ 1^0$ 7' 38.582")

APPENDIX D

PARAMETERS FOR THE IMAGE COORDINATE TRANSFORMATION FROM THE DISCRETE SYSTEM TO THE SYSTEM IN THE IMAGE PLANE OF A CAMERA

$$X = [2\,(Y_{cut} + j) - Y_0]\,h$$

$$Y = [X_0 - 2\,(X_{cut} + i)]\,w$$

where

$(X_{cut}, Y_{cut}) = (675, 415)$

$h = 13 \times 10^{-3}$ mm, $h = 13 \times 10^{-3}$ mm

$(X_0, Y_{0)}$ are different for different images.

1. Left Camera:

Image frame 13: $(X_0, Y_0) = (2075, 2055)$

Image frame 14: $(X_0, Y_0) = (2081, 2064)$

Image frame 15: $(X_0, Y_0) = (2061, 2068)$

2. Right Camera:

Image frame 13: $(X_0, Y_0) = (2071, 2055)$

Image frame 14: $(X_0, Y_0) = (2082, 2052)$

Image frame 15: $(X_0, Y_0) = (2070, 2057)$