A BAYESIAN FORMULATION FOR SUB-PIXEL REFINEMENT IN STEREO ORBITAL IMAGERY

Ara V. Nefian[1], *Kyle Husmann*[2], *Michael Broxton*[1], *Vinh To*[1], *Michael Lundy*[1] *and Mattew D. Hancher*[1]

[1]NASA Ames Research Center, Moffett Field, CA, USA
[2]California Polytechnic State University, San Luis Obispo, CA, USA
ara.nefian@sv.cmu.edu, khusmann@calpoly.edu,
{michael.broxton, vinh.to, michael.lundy, mattew.d.hancher}@nasa.gov

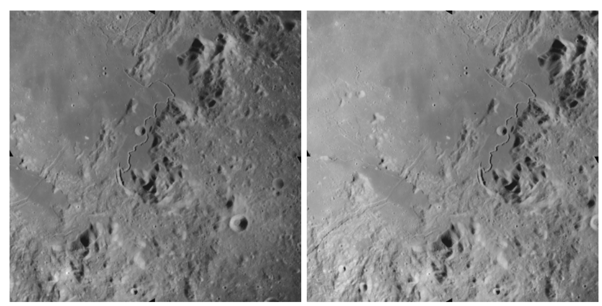
ABSTRACT

Generating accurate three dimensional planetary models is becoming increasingly more important as NASA plans manned missions to return to the moon in the next decade. This paper describes a stereo correspondence system for orbital images and focuses on a novel approach for the sub-pixel refinement of the disparity maps. Our method uses a Bayesian formulation that generalizes the Lucas-Kanade method for optimal matching between stereo pair images. This approach reduces significantly the pixel locking effect of the earlier methods and reduces the influence of image noise. The method is demonstrated on a set of high resolution scanned images from the Apollo era missions.

1. INTRODUCTION

The work [1] in this paper is motivated by the need for accurate, high resolution Lunar 3D maps that play a central role in NASA's future manned and unmanned missions to the moon. These maps support astronaut training, landing site planning, computer assisted landing, robotic exploration and provide very valuable information for lunar scientists and geologists. Although the resolution and coverage of these maps will be enhanced by data from upcoming missions, digital stereo pair scans from the Apollo era lunar missions (Figure 1) provide some of the best lunar imagery available today [3]. These images, despite their high quality, are affected by two types of noise inherent to the scanning process: the presence of film grain and dust and lint particles. The central focus of this paper is the attenuation of the effect of these scanning artifacts and improving the accuracy of the sub-pixel disparity maps.

A common technique in sub-pixel refinement is to fit a parabola to the correlation cost surface in the 8-connected neighborhood around the integer disparity estimate, and then use the parabola's minimum as the sub-pixel disparity value. This method is easy to implement and fast to compute, but exhibits a problem known as pixel-locking: the sub-pixel disparities tend toward their integer estimates and can create noticeable "stair steps" on surfaces that should be smooth [8], [10]. One way of attenuating the pixel-locking effect is through the use of a symmetric cost function [5] for matching the "left" and "right" image blocks. To avoid the high computational complexity of these methods another class of approaches relying on Lucas Kanade algorithm [1] proposes an asymmetric score where the disparity map is computed using the best matching score between the left image block and an optimally affine transformed block in the right image. Recently, several statistical approaches [2] have emerged to show encouraging results. Our sub-pixel refinement approach generalizes the earlier work by Stein et al. [8] to a Bayesian framework that models both the data and image noise. Iteratively estimating the model parameters determines the optimal disparity map that reduces the effects of image noise and attenuates the sub-pixel locking effect. The next sections describe the discrete stereo correlation (Section 2), and the subsequent sub-pixel refinement approach (Section 3) of the integer disparity map. In Section 4 we present a set of experimental results and a comparison with other sub-pixel refinement techniques.



(a) (b)

Fig. 1. Apollo Metric Camera stereo pair.

2. FAST DISCRETE STEREO CORRELATION

The integer disparity map is computed in the following steps. First, the left and right images are aligned using automatically generated tie points and geometric constraints from the camera models. Next, the images are filtered using a Laplacian of the Gaussian filter to reduce the lighting variations effect [6]. The integer disparity map $DSI(i, j, d_x, d_y)$ stores the matching cost between a left image block centered around pixel (i, j) and a right image block centered at position $(i - d_x, j - d_y)$. One efficient way

[1]This work was funded by the NASA Lunar Advanced Science and Exploration Research (LASER) program grant #07-LASER07-0148, NASA Advanced Information Systems Research (AISR) program grant #06-AISRP06-0142, and by the NASA ESMD Lunar Mapping and Modeling Program.

to compute the block matching cost is the sum of absolute differences of all block pixels. The normalized cross correlation [4], used in our system, is a computationally more complex approach for the block matching cost that provides more robust results with respect to variation in illumination conditions. A box filter [9] is applied to reduce duplicate operations in the calculation of DSI. To further reduce the computational complexity our system uses a pyramid-based approach where disparities are estimated using low resolution images, and successively refined at higher resolutions. At each level of the pyramid, we partition the image into rectangular sub-regions with similar values of the disparity determined in the previous lower resolution level of the pyramid [9].

3. SUB-PIXEL REFINEMENT

After the integer disparity estimates are computed in the fast discrete correlation step, the sub-pixel correlator refines these estimates to sub-pixel accuracy. Let $I_R(m,n)$ and $I_L(i,j)$ be two corresponding pixels in the right and left image respectively, where $i = m + d_x$, $j = n + d_y$ and d_x, d_y are the integer disparities. Using a linear approximation of the Taylor extensions around pixel (i,j) in the left image

$$I_L(i+\delta_x, j+\delta_y) \approx I_L(i,j) + \delta_x \frac{dI_L}{d_x}(i,j) + \delta_y \frac{dI_L}{d_y}(i,j) \quad (1)$$

where δ_x and δ_y are the local sub-pixel displacements. Let $e(x,y) = I_R(x,y) - I_L(i+\delta_x, j+\delta_y)$ and W be an image window centered around pixel (m,n). The local displacements are not constant across W and they vary according to:

$$\begin{aligned} \delta_x(i,j) &= a_1 i + b_1 j + c_1 \\ \delta_y(i,j) &= a_2 i + b_2 j + c_2. \end{aligned} \quad (2)$$

The goal of the method presented in [8] is to find the parameters $a_1, b_1, c_1, a_2, b_2, c_2$ that minimize the cost function

$$\mathbf{E}(m,n) = \sum_{(x,y)\in W} (e(x,y)w(x,y))^2 \quad (3)$$

where $w(x,y)$ are a set of weights used to reject outliers. Note that the local displacements $\delta_x(i,j)$ and $\delta_y(i,j)$ depend on the pixel positions within the window W. In fact, the values $a_1, b_1, c_1, a_2, b_2, c_2$ that minimize $\mathbf{E}$ can be seen as the parameters of an affine transformation that best transform the right image window to match the reference (left) image window.

In the original Lucas-Kanade method the weights are set $w(x,y) = 1$. In [8] the values the weights $w(x,y)$ determined heuristically to reject the noise and emphasize the pixel closer to the center of the window. An alternative solution is to use a set of weights derived from the Cauchy distribution [4] given by:

$$w(x,y) = \frac{\sqrt{b^2 \log(1 + \frac{I_e^2(x,y)}{b^2})}}{|I_e(x,y)|} \quad (4)$$

where b is some fixed threshold (in our experiments $b = 10^{-4}$) and $I_e(x,y) = I_R(x,y) - I_L(i,j)$.

The steps of this method are given below

- **Step 1:** Compute $\frac{dI_L}{d_x}(i,j)$, $\frac{dI_L}{d_y}(i,j)$ and the $I_R(x,y)$ values using bilinear interpolation. Initialize the parameters $a_1, b_1, c_1, a_2, b_2, c_2$.
- **Step 2:** Determine $a_1, b_1, c_1, a_2, b_2, c_2$ to minimize $\mathbf{E}$.
- **Step 3:** Compute $\delta_x(i,j)$ and $\delta_y(i,j)$ using Equation 2.
- **Step 4:** Compute a new point $(x', y') = (x,y) + (\delta_x, \delta_y)$ and the $I_R(x', y')$ values using bilinear interpolation.
- **Step 5:** Check for convergence. If norm of (δ_x, δ_y) vector falls below a fixed threshold the iterations converged. Otherwise, go to step 1.

One shortcoming of this method is directly related to the cost function that is minimized. The cost function $\mathbf{E}$ has a low tolerance to noise and often creates erroneous disparity information in areas affected by image noise. The method introduced in this paper tries to overcome this problem by replacing the cost function $\mathbf{E}$ with a probabilistic framework that allows to optimally estimate the parameters of the noise model.

In our approach the probability of a pixel in the right image is given by the following Bayesian model:

$$P(I_R(m,n)) = \sum_{k=0,1} P(I_R(m,n)|z=k)P(z=k) \quad (5)$$

The first mixture component $(z = 0)$ is a normal density function with mean $I_L(i+\delta_x, j+\delta_y)$ and variance σ_p

$$P(I_R(m,n)|z=0) = \mathcal{N}(I_R(m,n)|I_L(i+\delta_x, j+\delta_y), \sigma_p) \quad (6)$$

The second mixture component $(z = 1)$ in Equation 6 models the image noise using a normal density function with mean μ_n and variance σ_n.

$$P(I_R(m,n)|z=1) = \mathcal{N}(I_R(m,n)|\mu_n, \sigma_n) \quad (7)$$

Let $\mathbf{I}_R(m,n)$ be a vector of all pixels values in a window W centered in pixel (m,n) in the right image. Then,

$$P(\mathbf{I}_R(m,n)) = \prod_{(x,y)\in W} P(I_R(x,y)) \quad (8)$$

The parameters $\lambda = \{a_1, b_1, c_1, a_2, b_2, c_2, \sigma_p, \mu_n, \sigma_n\}$ that maximize the model likelihood in Equation 8 are determined using the Expectation Maximization (EM) algorithm. Maximizing the model likelihood in Equation 8 is equivalent to maximizing the auxiliary function:

$$\begin{aligned} \mathbf{Q}(\lambda) &= \sum_k P(k|\mathbf{I}_R, \lambda_t) \log P(\mathbf{I}_R, k, \underline{\delta}|\lambda) \\ &= \sum_k \sum_{x,y} P(k|I_R(x,y), \lambda_t) \log P(I_R(x,y)|k, \theta) P(k|\lambda) \end{aligned}$$

The EM algorithm for the stereo image pair is described by the following steps:

- **Initialization:** Iteration start with a set of values chosen at random or based on a priori knowledge of the data. In our system $a_1 = 1, b_1 = 0, c_1 = 0, a_2 = 0, b_2 = 1, c_2 = 0, \sigma_p = 10^{-3}, \mu_n = 0.0, \sigma_n = 10^{-2}$.
- **Estimation:** Compute the a posteriori probabilities

$$\begin{aligned} \gamma_{xy}^{(k)} &= P(k|I_R(x,y), \lambda_t) \\ &= \frac{P(I_R(x,y)|k, \lambda_t)P(k|\lambda_t)}{\sum_l P(I_R(x,y)|l, \lambda_t)P(l|\lambda_t)} \end{aligned} \quad (9)$$

for $k = 0, 1$.

- **Maximization:** Determine the parameters λ s.t. $\frac{d\mathbf{Q}}{d\underline{\lambda}} = 0$ with the constraint $\sum_k P(k|\lambda) = 1$.

$$\begin{bmatrix} A_{00} & A_{01} \\ A_{10} & A_{11} \end{bmatrix} \begin{bmatrix} a_1 \\ b_1 \\ c_1 \\ a_2 \\ b_2 \\ c_2 \end{bmatrix} = - \begin{bmatrix} \sum \gamma_{xy}^{(0)} x I_e I_x \\ \sum \gamma_{xy}^{(0)} y I_e I_x \\ \sum \gamma_{xy}^{(0)} I_e I_x \\ \sum \gamma_{xy}^{(0)} x I_e I_y \\ \sum \gamma_{xy}^{(0)} y I_e I_y \\ \sum \gamma_{xy}^{(0)} I_e I_y \end{bmatrix} \tag{10}$$

where $A_{01} = A_{10}$ and

$$A_{00} = \begin{bmatrix} \sum \gamma_{xy}^{(0)} x^2 I_x^2 & \sum \gamma_{xy}^{(0)} xy I_x^2 & \sum \gamma_{xy}^{(0)} x I_x^2 \\ \sum \gamma_{xy}^{(0)} xy I_x^2 & \sum \gamma_{xy}^{(0)} y^2 I_x^2 & \sum \gamma_{xy}^{(0)} y I_x^2 \\ \sum \gamma_{xy}^{(0)} x I_x^2 & \sum \gamma_{xy}^{(0)} y I_x^2 & \sum \gamma_{xy}^{(0)} I_x^2 \end{bmatrix}$$

$$A_{11} = \begin{bmatrix} \sum \gamma_{xy}^{(0)} x^2 I_y^2 & \sum \gamma_{xy}^{(0)} xy I_y^2 & \sum \gamma_{xy}^{(0)} x I_y^2 \\ \sum \gamma_{xy}^{(0)} xy I_y^2 & \sum \gamma_{xy}^{(0)} y^2 I_y^2 & \sum \gamma_{xy}^{(0)} y I_y^2 \\ \sum \gamma_{xy}^{(0)} x I_y^2 & \sum \gamma_{xy}^{(0)} y I_y^2 & \sum \gamma_{xy}^{(0)} I_y^2 \end{bmatrix}$$

$$A_{01} = \begin{bmatrix} \sum \gamma_{xy}^{(0)} x^2 I_x I_y & \sum \gamma_{xy}^{(0)} xy I_x I_y & \sum \gamma_{xy}^{(0)} x I_x I_y \\ \sum \gamma_{xy}^{(0)} xy I_x I_y & \sum \gamma_{xy}^{(0)} y^2 I_x I_y & \sum \gamma_{xy}^{(0)} y I_x I_y \\ \sum \gamma_{xy}^{(0)} x I_x I_y & \sum \gamma_{xy}^{(0)} y I_x I_y & \sum \gamma_{xy}^{(0)} I_x I_y \end{bmatrix}$$

where $I_x = \frac{dI_L}{d_x}(i,j)$, $I_y = \frac{dI_L}{d_y}(i,j)$ and $I_e = I_R(x,y) - I_L(i,j)$ and

$$\sigma_p = \frac{\sum \gamma_{xy}^{(0)} (I_R(x,y) - I_L(i+\delta_x, j+\delta_y))^2}{\sum \gamma_{xy}^{(0)}} \tag{11}$$

$$\mu_n = \frac{\sum \gamma_{xy}^{(1)} I_R(x,y)}{\sum \gamma_{xy}^{(1)}} \tag{12}$$

$$\sigma_n = \frac{\sum \gamma_{xy}^{(1)} (I_R(x,y) - \mu_n)^2}{\sum \gamma_{xy}^{(1)}} \tag{13}$$

$$P(k|\theta) = \frac{\sum \gamma_{xy}^{(k)}}{\sum 1} \tag{14}$$

- **Convergence test:** The EM iteration stop when the absolute difference of the $P(\mathbf{I}_R(m,n))$ at consecutive iterations fall below a fixed threshold or after a maximum number of iterations. Otherwise, continue with the estimation step.

Note that Equation 10 is similar to the equation used to determine the parameters $a_1, b_1, c_1, a_2, b_2, c_2$ in step 2 of the the method presented in [8] and described earlier in this section. In our method the values $\gamma_{xy}^{(0)}$ are interpreted as a posteriori probabilities and are optimally re-estimated using EM to learn the noise and data parameters. In this way, our approach can be seen as a generalization of the Lucas-Kanade method. The following steps summarize our approach:

- **Step 1:** Compute $\frac{dI_L}{d_x}(i,j)$, $\frac{dI_L}{d_y}(i,j)$ and the $I_R(x,y)$ values using bilinear interpolation. Initialize the model parameters λ.
- **Step 2:** Compute iteratively the model parameters λ using the EM algorithm described by Equations 9- 14.
- **Step 3:** Compute $\delta_x(i,j)$ and $\delta_y(i,j)$ using Equation 2.
- **Step 4:** Compute a new point $(x', y') = (x, y) + (\delta_x, \delta_y)$ and the $I_R(x', y')$ values using bilinear interpolation.
- **Step 5:** If the norm of (δ_x, δ_y) vector falls below a fixed threshold the iterations converged. Otherwise, go to step 1.

4. EXPERIMENTAL RESULTS

The stereo processing system described in the previous sections is used to generate three dimensional lunar models from the Apollo metric camera (AMC) images captured during the Apollo 15, 16 and 17 missions [7]. AMC is a calibrated wide field (75deg) of view orbital mapping camera that photographed overlapping images (80%) used as stereo pairs. The image set consisting of 8000 stereo pairs was obtained by scanning original film negatives captured by AMC. The scans (Figure 1) capture the full dynamic range and resolution of the original film resulting in digital images of size 22,000 $\times$ 22,000 pixels representing a resolution of $10m^2$/pixel.

Film grain and the dust particles are inherent to the scanning process and can significantly limit the accuracy of the stereo processing system. One example where dust particle noise occurs in one of the stereo pair images is shown in detail in Figure 2 a and b. Figure 2 c illustrates the integer disparity map obtained by running the fast discrete correlation method described in Section 2. Figure 2 d, e and f compares the horizontal sub-pixel disparity maps obtained using the parabola method, the Lucas Kanade method with Cauchy weights (Equation 4) and the Bayesian approach introduced in Section 3 respectively. The Bayesian approach reduces the "stair-steps" effect of the parabola based method apparent in the horizontal and vertical artifacts, at the same time with reducing the effect of dust noise that affect the weighted Lucas Kanade method. Figure 3 displays a 3D oblique view of Hadley Rille, produced using our stereo processing system from the image pair shown in Figure 1.

5. CONCLUSIONS AND FUTURE WORK

The method for sub-pixel disparity maps generation introduced in this paper uses a novel statistical formulation for optimally determining the stereo correspondence and reducing the effect of image noise. The method proposed here was tested on a set of very large resolution scanned images from the Apollo era missions. Our approach outperforms previous methods based on Lucas Kanade optical flow formulations at the cost of a higher computational complexity. Further research will be directed towards reducing the complexity of the current approach and generating a high resolution 3D map from the entire set of stereo image pairs captured during the Apollo missions.

6. ACKNOWLEDGMENTS

We would like to thank Mark Robinson for supplying high resolution scans of the Apollo Metric Camera images.

7. REFERENCES

[1] Simon Baker, Ralph Gross, and Iain Matthews. Lucas-Kanade 20 Years On: A Unifying Framework. *International Journal of Computer Vision*, 56:221–255, 2004.

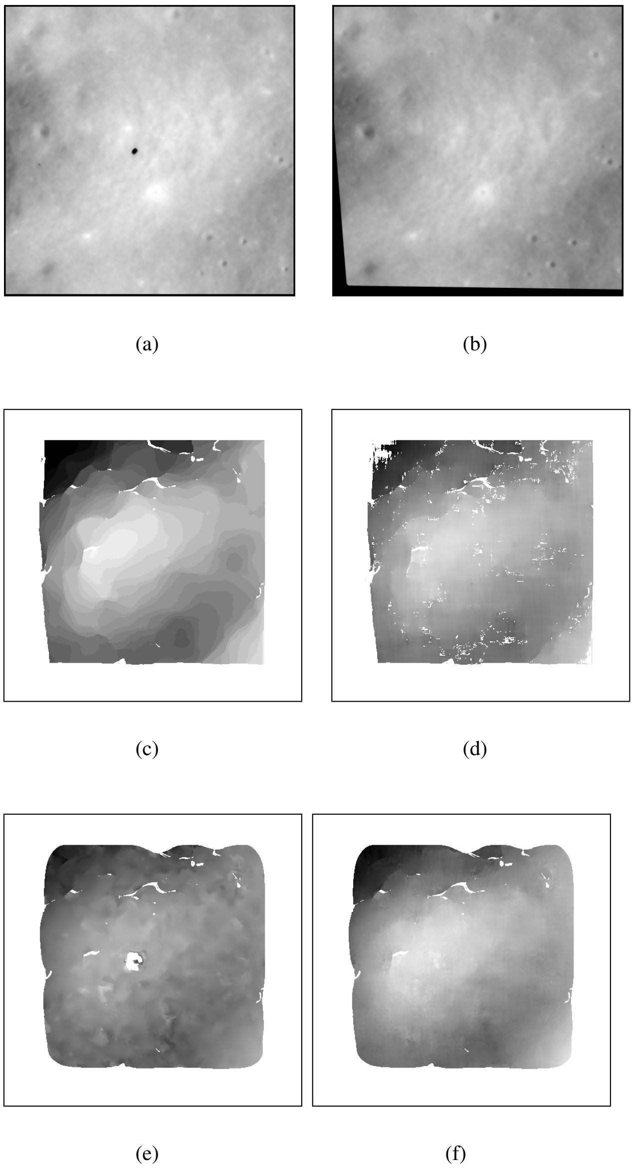


Fig. 2. (a) Left Image, (b) Right Image, (c) Horizontal integer disparity map, (d) Horizontal disparity map using the parabola method, (e) Horizontal disparity map using 2D affine transform method and Cauchy weights, (f) Horizontal disparity map using the Bayes approach.

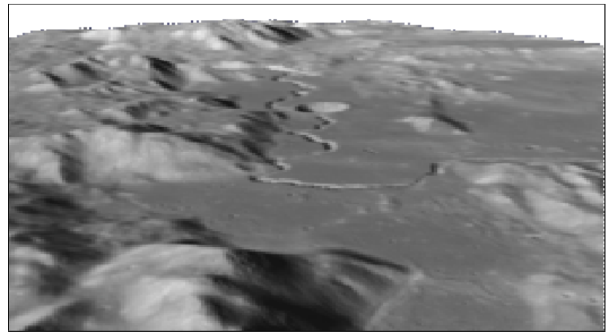


Fig. 3. Oblique view of Hadley Rille derived from high resolution scans AS15-M-1135 and AS15-M-1136.

[2] Li Cheng and T. Caelli. Bayesian stereo matching. *Computer Vision and Pattern Recognition Workshop, 2004. CVPRW '04. Conference on*, pages 192–192, June 2004.

[3] S. J. Lawrence, M. S. Robinson, M. Broxton, J. D. Stopar, W. Close, J. Grunsfeld, R. Ingram, L. Jefferson, S. Locke, R. Mitchell, T. Scarsella, M. White, M. A. Hager, T. R. Watters ad E. Bowman-Cisneros, J. Danton, and J. Garvin. The Apollo Digital Image Archive: New Research and Data Products. In *Proc of the NLSI Lunar Science Conference*, page 2066, 2008.

[4] Christian Menard. *Robust Stereo and Adaptive Matching in Correlation Scale-Space*. PhD thesis, Institute of Automation, Vienna Institute of Technology (PRIP-TR-45), January 1997.

[5] Diego Nehab, Szymon Rusinkiewicz, and James Davis. Improved sub-pixel stereo correspondences through symmetric refinement. *Computer Vision, IEEE International Conference on*, 1:557–563, 2005.

[6] H.K. Nishihara. PRISM: A Practical real-time imaging stereo matcher. *Optical Engineering*, 23(5):536–545, 1984.

[7] M.S. Robinson, E.M. Eliason, H. Hiesinger, B.L. Jolliff, A.S. McEwen, M.C. Malin, M.A. Ravine, D. Roberts, P.C. Thomas, and E.P. Turtle. LROC - Lunar Reconaissance Orbiter Camera. In *Proc of the Lunar and Planetary Science Conference (LPSC) XXXVI*, page 1576, March 2005.

[8] Andrew Stein, Andres Huertas, and Larry Matthies. Attenuating stereo pixel-locking via affine window adaptation. In *IEEE International Conference on Robotics and Automation*, pages 914 – 921, May 2006.

[9] Changming Sun. Rectangular Subregioning and 3-D Maximum-Surface Techniques for Fast Stereo Matching. *International Journal of Computer Vision*, 47(1-3), 2002.

[10] Richard Szeliski and Daniel Scharstein. Sampling the Disparity Space Image. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 26:419 – 425, 2003.