Seeing in Three Dimensions: Correlation and Triangulation of Mars Exploration Rover Imagery
IEEE SMC 2005

Bob Deen
Jean Lorre

Multimission Image Processing Laboratory
Jet Propulsion Laboratory
California Institute of Technology
Overview

• Correlation Process
• XYZ Derivation Process
• Results
• Future Work
Goal
Terrain Derivation

- Principal Steps
  - Acquire the stereo images [Maki]
  - Correlate stereo pairs to determine matching features and disparity
  - Triangulation of disparity and camera models to determine spatial XYZ coordinates of each pixel
  - Conversion of XYZ coordinates to integrated terrain mesh [Wright]
  - Creation of ancillary products (reachability, slope) [Leger]
- Managed by MIPL pipeline [Alexander/Zamani]
What is Correlation?

- For each pixel in one image (reference, or left)
  - Find matching pixel in the other (right) image
  - Uses an area-based search
- Difference in coordinates is called disparity
  - Directly relates to distance from the viewer
Correlation Coefficient

- Standard correlation coefficient:
  \[ \rho = \frac{\sigma_{ts}}{\sigma_t \sigma_s} \]
  - Covariance / product of standard deviations
  - Measure of similarity of areas
    - \( \rho = 1 \) for perfect match, 0 for none, -1 for inverse match
  - More computationally efficient form:
    \[ \rho = \frac{n^2 \sum ts - \sum t \sum s}{\sqrt{(n^2 \sum t^2 - \sum t)(n^2 \sum s^2 - \sum s^2)}} \]
- Other measures are possible
Geometric Warping

- Template does not exactly match right area
  - Geometry of scene affects the mapping

Top view
(world is flat plane)

Template

Images

Area matching template in right image
Geometric Warping (cont.)

- Bilinear transform warps right image to match left

  \[ x' = ax + by + c + gxy \]
  \[ y' = dx + ey + f + hxy \]

- Only a subset of parameters used
  - \( b, c, f, g \) for most cameras
    - Translation, shear, trapezoid in \( x \)
    - Translation in \( y \)
    - Models aligned cameras looking out at a flat plane
  - \( a \) added for front hazcams
    - Scale in \( x \)
    - Models aligned cameras looking at a more general scene
Function Minimization

• Goal: find best match of template
  – Maximize correlation coefficient $\rho$
    • Actually, minimize $2.0 - \rho |\rho|$
  – Warping terms $a-h$ are the free parameters

• “Amoeba” minimization used
  – Downhill simplex method
  – No partial derivatives required
  – Somewhat slow
  – Requires reasonably close starting point to avoid local minima

• Translation terms $(c,f)$ represent disparity
  – Other terms discarded after match found
Seed Points

• Starting points for amoeba
• Two methods to gather them
  – Use 1-D correlator from flight software
    • Epipolar alignment required (features on same line)
    • Must use downsampled images
      – 4x, 8x, or 16x, depending on type
    • Used operationally
  – Estimate disparities from camera and surface models
    • Epipolar alignment not required
    • Also uses downsampled images
    • Used for photometry cubes (science processing)
Remaining algorithm

- Progressive image pyramids
  - Increase size 2x at each pass until full res reached
  - Each pass provides seed points for next pass
- 3x3 box filter used
  - Reduces noise at expense of resolution
- Filling gores
  - Postprocessing step after each pass
  - Finds uncorrelated pixels, uses neighbors for seeds
  - Fills in small gaps
Image Triangulation

• Converts disparities to XYZ locations
• Requires camera models
  – Describe relationship between line/sample coordinates in image and the corresponding ray in 3-D space
  – CAHV family of camera models
    • CAHV: linear
    • CAHVOR: radial optical distortion
    • CAHVORE: generalized fisheye lens
• Steps
  – Find matching pixels via disparity map
  – Project rays through camera models to 3-D space
  – Intersection point of rays is XYZ point
    • Actually, midway between the rays’ closest approach
    • Miss distance is called the error
Rejecting Invalid Points

- Nine filters used to remove bad points from result
  - Correlation match must exist
  - Vertical disparity must be < 4 pixels
  - Vertical disparity must be within 0.75 pixels of average over 51x51 pixel area
  - XYZ intersection must be computable (not parallel)
  - Absolute miss distance must be < 0.05 meters
  - Miss distance/range must be < 0.005 meters/meter.
  - Computed Z values must be within limits (-20,+20 meters for Opportunity, -40,+20 for Spirit)
  - Rays must not diverge
  - Range must be within 1000 times the camera baseline
Speed

- Great improvement from Mars Pathfinder
  - 1 hour for 256x256 -> 2-4 minutes for 1024x1024
- Four critical parameters
  - Template size
    - Large == slower, but more accurate
  - Warping parameters $a-h$
    - All would be nice, but each adds significant speed penalty
  - Gore-filling passes
    - Inefficiently implemented
    - Diminishing returns
  - Tolerance parameter
    - Tells amoeba how hard to work;
    - Lower == more precise but much slower
Accuracy

• Very hard to measure
  – Spot (target) tests validate XYZ range
    • Spots easy to correlate
  – Generate pseudo-left image by projecting through disparity map
    • Hard to see subpixel errors
  – Manual spot checks
    • Hard to see subpixel errors
  – Artificially-generated imagery
    • Not representative of real-world performance
  – Relative comparison between correlator runs
    • Helps with speed vs. accuracy tradeoffs

• End-to-end tests show accuracy sufficient for ops
  – Can it be improved?
Compression Noise

- Primary source of error
  - Patchiness in the output
  - Correlator constrains allowable compression rates
  - 9 bpp (left) vs. 0.5bpp (right)

Seeing in Three Dimensions: Correlation and Triangulation of MER Imagery
IEEE SMC, Oct. 10-12, 2005, Waikoloa, Hawaii, USA
Anti-Integer Bias

• Subpixel results in the presence of noise are biased away from integer values, toward half-integer values
  – More noise results in more pronounced effect

• Self-correlation with noise introduced
  – Histogram plot of vertical disparity + line number
Anti-Integer Bias (cont.)

noise = 5 DN  
noise = 10 DN  
noise = 20 DN  
noise = 30 DN  
noise = 200 DN
Anti-Integer Bias (cont.)

- Believed to be due to bilinear interpolation
  - Performed as part of geometric warp
  - Interpolation smooths the noise
    - Smoother noise is easier to match
    - Maximum interpolation (and thus smoothing) at half-integer

- Possible solutions
  - More degrees of freedom in warp
    - Causes interpolation even at integer values
  - Larger windows
    - Increase signal-to-noise ratio
  - Different interpolation algorithms
    - Not yet investigated
Inappropriately Correlated Areas

• Correlation results in bad places
  – Sky
  – Horizon
  – Undifferentiated terrain
  – Noise

• Most problematic issue operationally
  – Produces artifacts in terrain mesh

• Horizon mask used to trim sky/horizon
  – Stopgap measure
Future Directions

• Avoid bad spots in sky, sand, etc.
  – Interest operator
  – Heuristics
  – Compression knowledge

• Checking via right->left correlation
  – Cost/benefit ratio not high at the moment

• Handling compression noise
  – Different interpolation
  – Recognize noise regionally via compressor metadata

• Areas with different depths (rock edges)
  – Adaptive windows
Long-Baseline Stereo

• Move rover to obtain long (~10m) baseline

• Work in progress
  – Results not calibrated yet
    • Slopes, etc. still useful despite that
  – Shows promise for long-term planning
  – Can get useful results >1km away
Long-Baseline Stereo (cont.)
Conclusion

• Results are accurate and reliable enough for Mars remote ops
  – Almost 2 years on each rover (!)

• Image compression noise is biggest problem
  – Imposes constraint on allowed compression rates
    • And thus mission data volume

• Still many avenues for improvement
Questions

- Bob.Deen@jpl.nasa.gov