Vision-Based Trail Following

Dynamic Vision Laboratory
Dept. Computer & Information Sciences
University of Delaware, USA

The Appearance Variability of Trails

Cues for Trail Finding

- Treat the trail region as an *object*, like a person or car, that we are trying to detect in the image
  - This is a classic computer vision problem
  - Shape here means position, scale, orientation, curvature—actually fewer parameters than many other classes of object
- A machine learning approach would be to train on trail examples
- Full range of gestalt cues are available, but which are most valuable?

Visual Appearance Cues

- **Light/dark: Known *a priori*, or just based on local contrast?**

Visual Appearance Cues

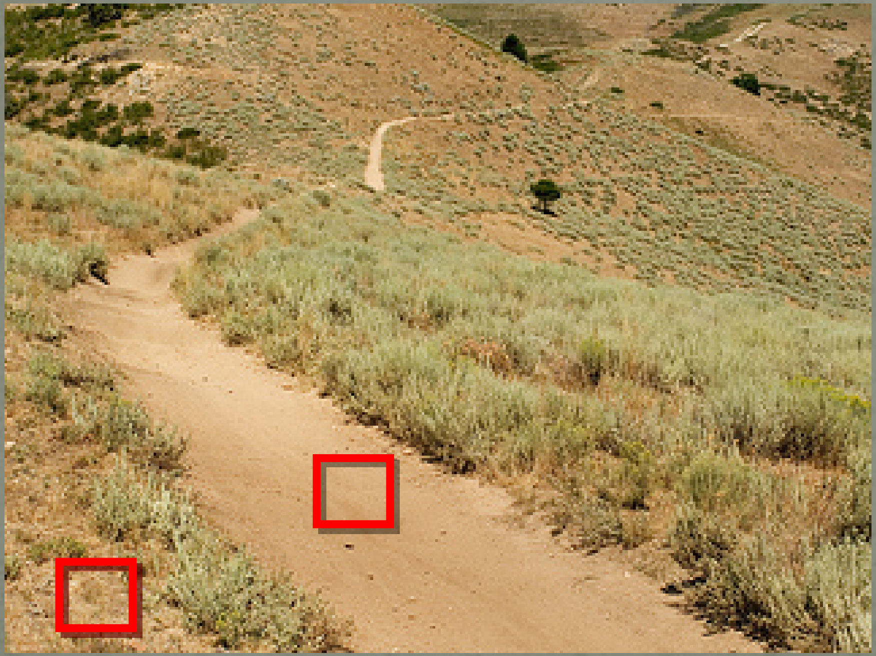
- **Color: Helps with discrimination, but more complicated to define similarity, especially with variable illumination (e.g., shadows)**

Visual Appearance Cues

- **Color: Single color may not adequately describe trail region—how to compare *mixtures* of colors?**

Visual Appearance Cues

- **Texture: Homogeneity vs. heterogeneity, isotropy vs. anisotropy**



A problematic case…

Visual Appearance Cues

- **Gross shape: Trails taper from bottom to top, nearby sides are nearly straight → *triangular* under perspective**

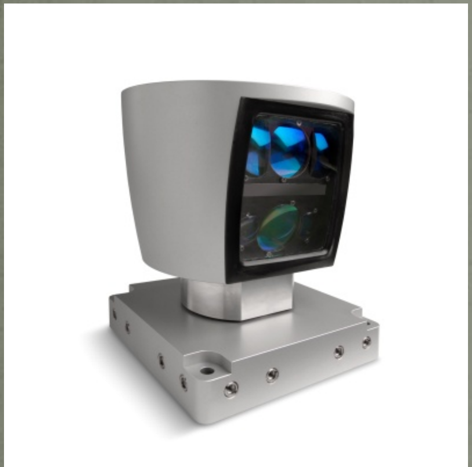
Structural Cues

- **Regardless of scene appearance, bottom line is that we don't want to run into *obstacles***
- **If we're lucky, obstacles will actually delineate the trail**
  - **Look for height contrast or variance as trail's distinguishing feature?**



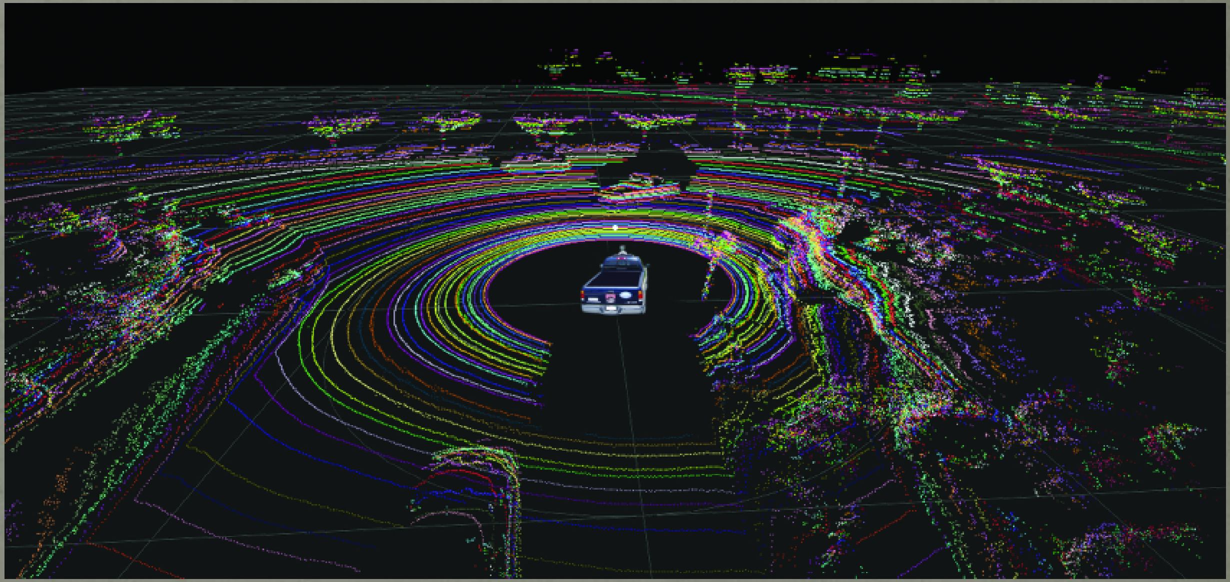
Structural Cues: From Where?

- **Laser range-finder (aka *ladar*/*lidar*)**



Velodyne

$60K

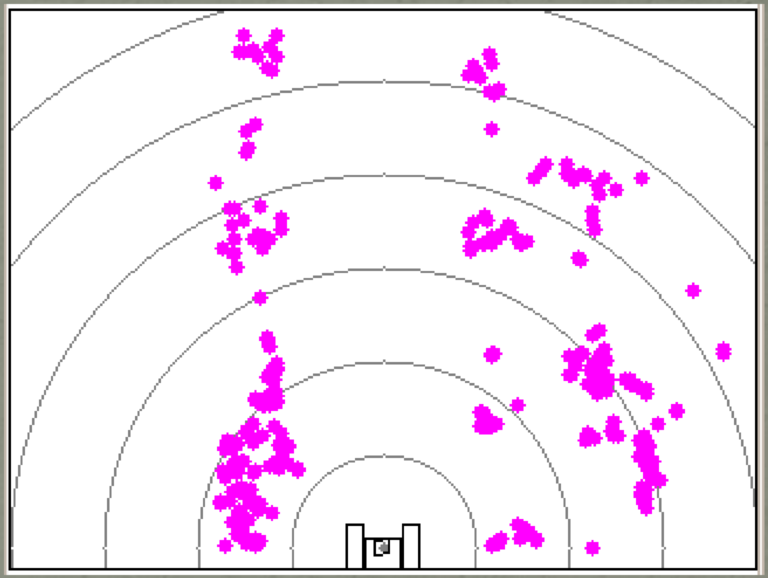


Structural Cues: From Where?

- **Laser range-finder (aka *ladar*/*lidar*)**

SICK LMS

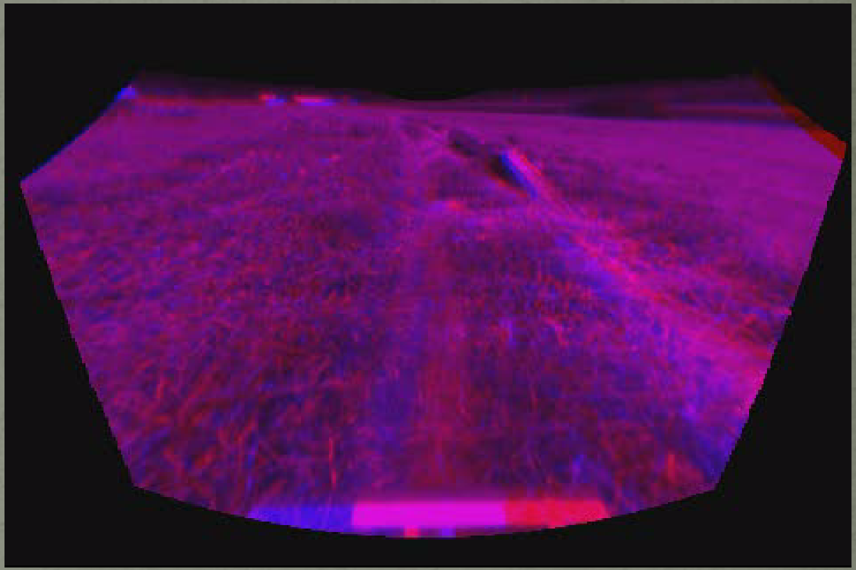
$5K



Structural Cues: From Where?

- **Stereopsis (static or motion-based)**

Left image (undistorted)



Left and right images overlaid as red & blue channels

Field

Structural Cues: From Where?

- **Stereopsis (static or motion-based)**

Left image (undistorted)

Estimated depth

Field

Structural Cues: From Where?

- **Stereopsis (static or motion-based)**

Left image (undistorted)

Estimated depth

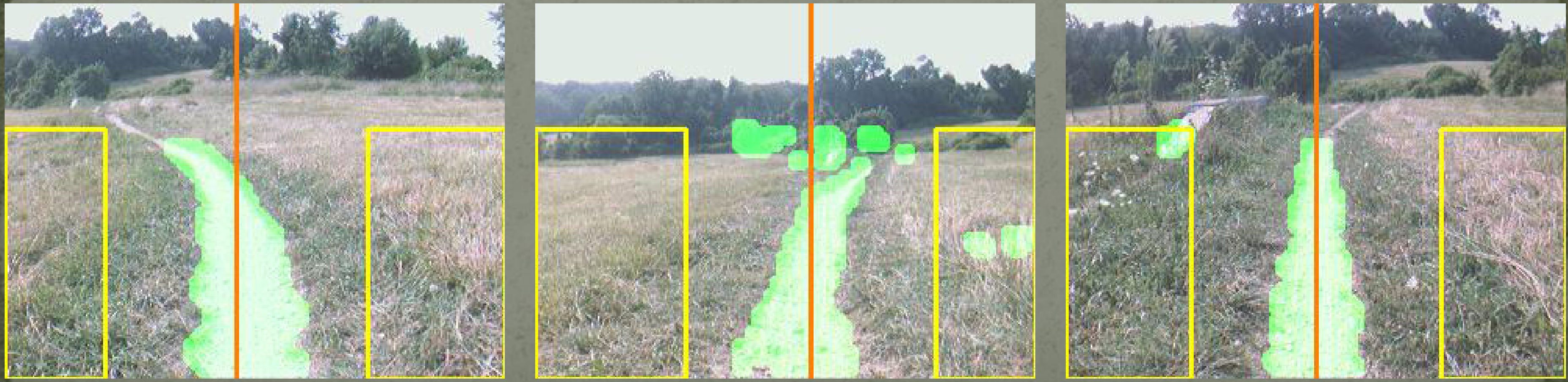
Forest

Cues for Trail Finding

- **Treat the trail region as an *object*, like a person or car, that we are trying to detect in the image**
  - **This is a classic computer vision problem**
  - **Shape here means position, scale, orientation, curvature—actually fewer parameters than many other classes of object**
- **A machine learning approach would be to train on trail examples**
- **Full range of gestalt cues are available, but which are most valuable?**
- **What about the *top-down* vs. *bottom-up* question?**

Histogram-based Trail Following

- **Very bottom-up approach:**
  1. **Assume sides of image are off-trail**
  2. **Build histogram of colors of off-trail pixels (yellow boxes)**
  3. **Classify remaining image pixels as trail/non-trail based on likelihood given by histogram**
  4. **Median *x* coordinate of trail pixels is trail center**
  5. **Adjust off-trail boxes**



(RobVis 2008)

Shape-Guided Superpixel Grouping (IROS 2008)

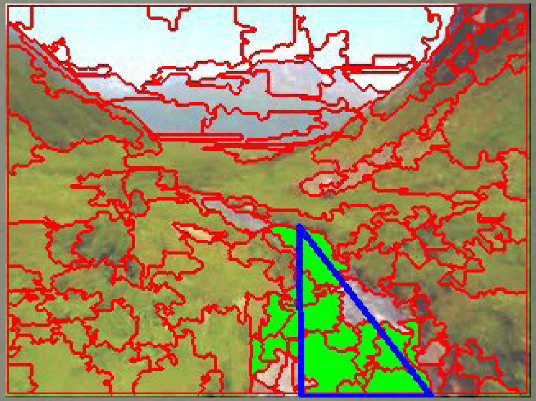
- **Another bottom-up method, but at higher level**
- ***Superpixels* (Felzenszwalb, 2004; Malik, 2001) are pixels clustered by proximity and color similarity**

Felzenszwalb

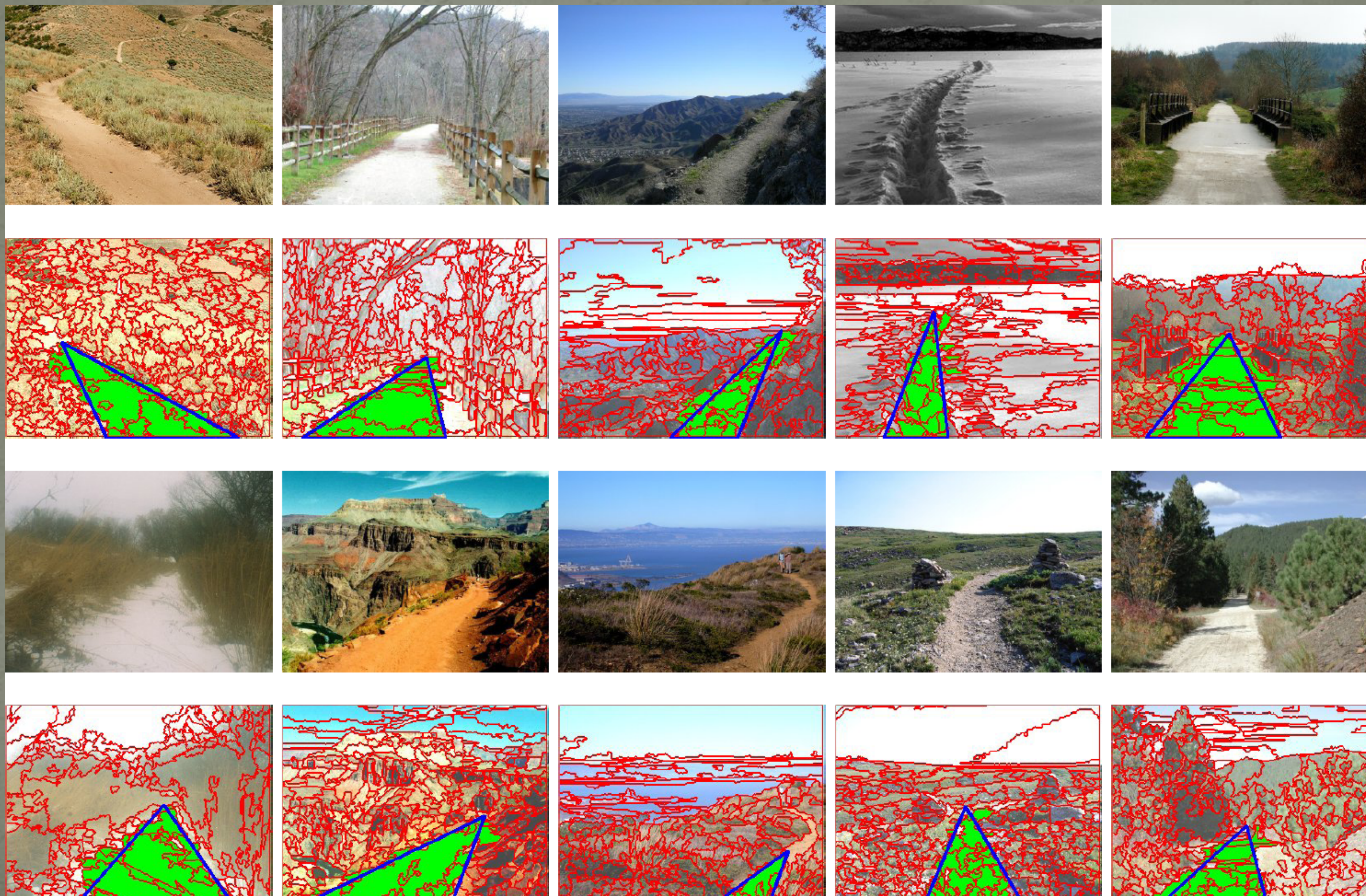
Malik

Shape-Guided Superpixel Grouping (IROS 2008)

- **Preprocess image into superpixels**
- **Repeatedly generate randomized *groupings* of superpixels as trail hypotheses**
- **Choose mostly likely grouping based on weighted combination of**
  - **Shape likelihood: How "triangular" is grouping?**
  - **Appearance likelihood: How strongly does color inside grouping contrast with colors of neighboring superpixels?**
  - **Deformation likelihood: Are overall size, width-to-height ratio, etc. of fitted triangle in expected ranges?**

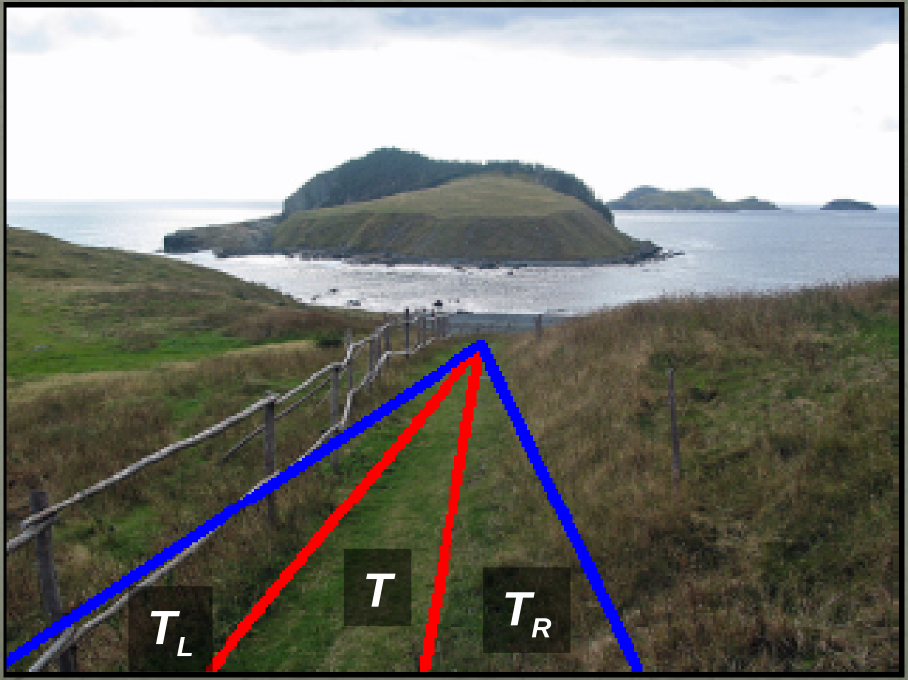


Shape-Guided Grouping Results



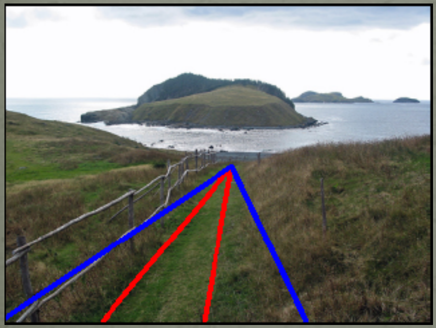
Triangular Trail Regions (IROS 2009)

- **Approximate trail boundary viewed under perspective as triangle *T* with bottom side defined by image bottom**
- **To measure contrast, look at equal-width triangular neighbor regions T_L and T_R**

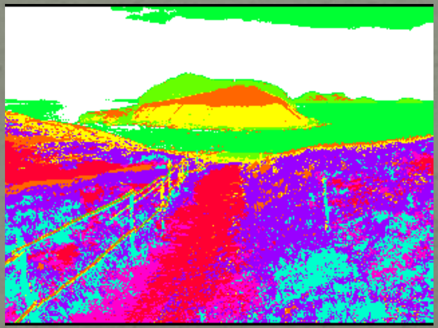


Trail region appearance characterization

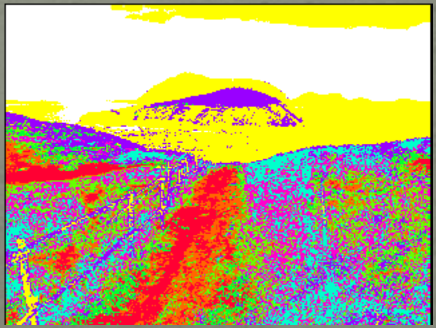
- **Compute color features (aka *textons*) via *k*-means clustering in CIE-LAB space (following Blas, 2008)**
- **In a sense this is like superpixels without proximity**
- **Clustering done over 3 different feature sets (these are used for *feature switching*)**
  - **AB (chromaticity only)**
  - **L (brightness only)**
  - **LAB (full color space)**
- **Model trail region *T*'s color distribution via texton histogram H_T**



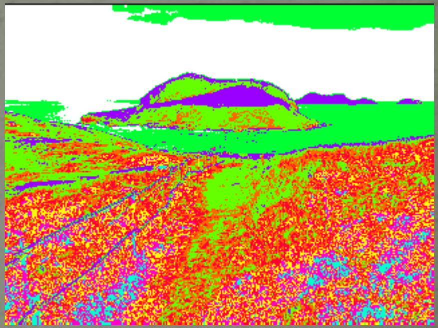
Input image



LAB textons (*k* = 8)



AB (*k* = 8)



L (*k* = 8)

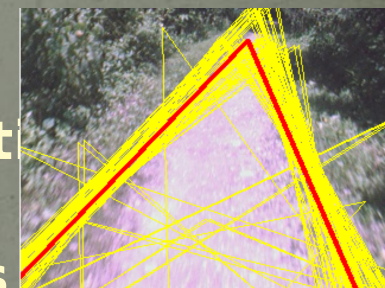
Trail likelihood function

- **Weighted sum of measures of:**
  - **Color/brightness contrast of center trail region with neighboring regions**
    - **Quantify similarity using standard histogram metric of chi-squared distance χ^2**
  - **Homogeneity of trail region—the fewer colors, the more likely**
    - **Quantify heterogeneity with entropy of histogram**

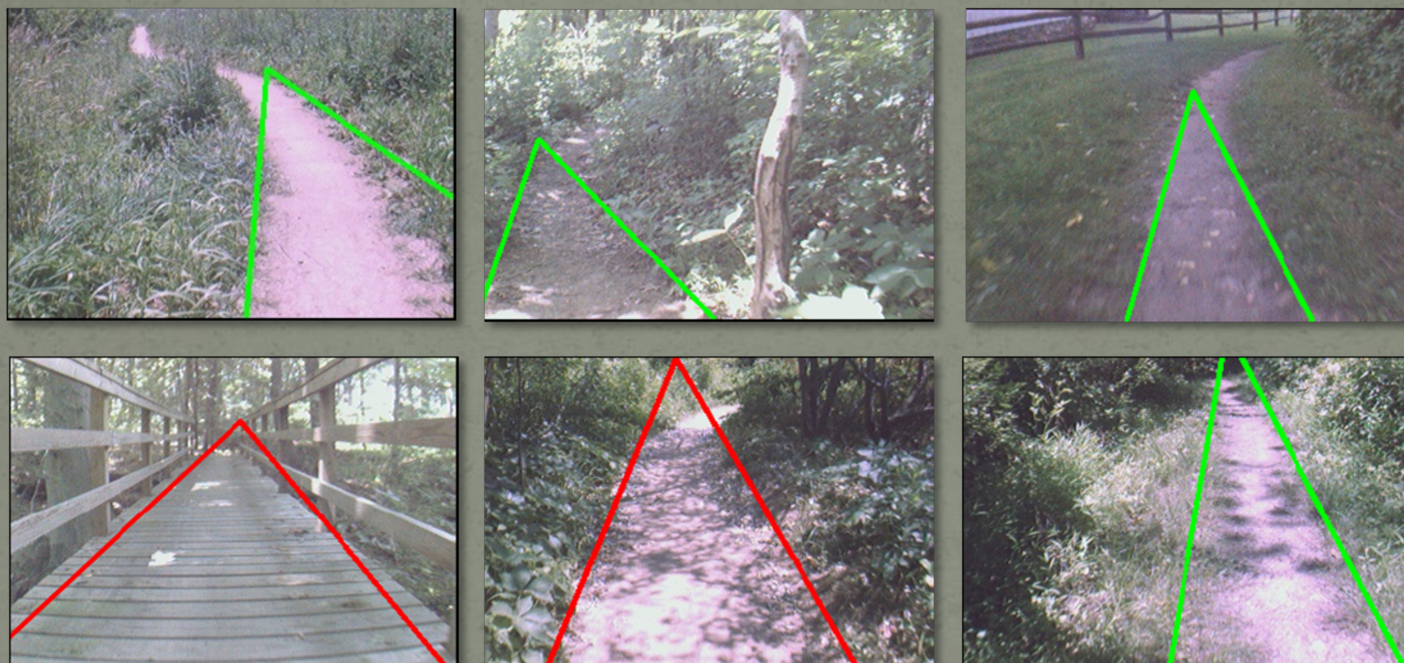
$$L_{appear}(T) = \alpha[\chi^2(h, h_L) + \chi^2(h, h_R)] + \beta(1 - H(h))$$

Likelihood maximization and tracking

- **Find and track good trail candidates via MAP estimation using *particle filtering***
  - **For static images, trail estimate is likelihood particle found after *t* iterati**
  - **For image sequences, state is sum of particles weighted by their likelihoods**
  - **Small fraction of particles are sampled from image-wide prior (rather than near previous state)**



Experimental Results – Sample Images

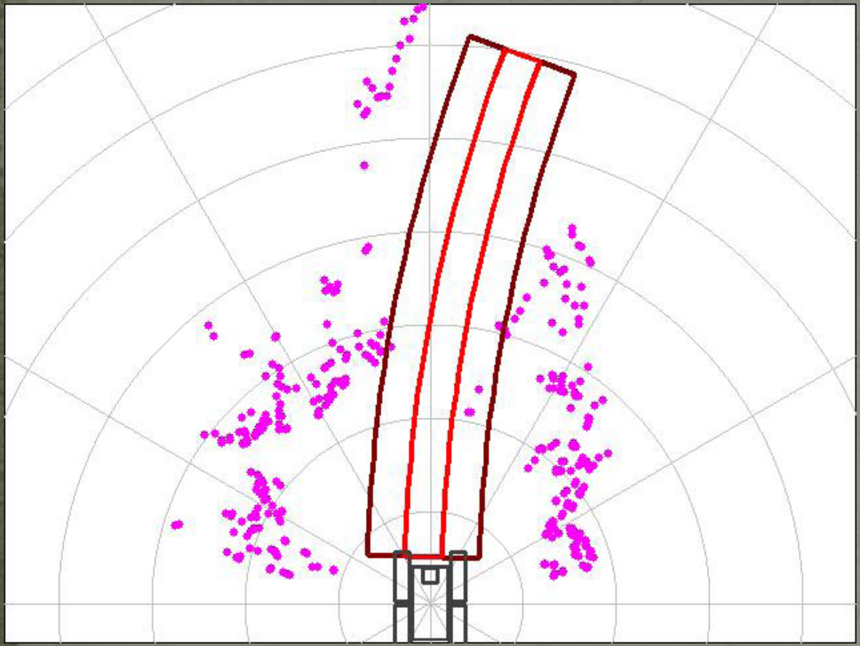


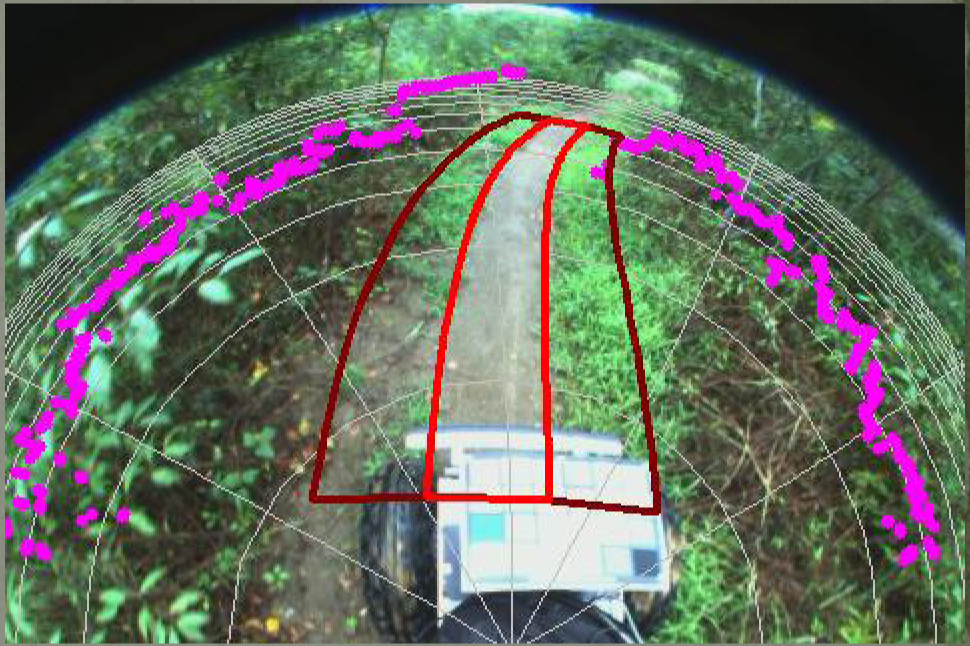
- **For display, feature set selected is indicated by color of fitted triangle: LAB = red, AB = green, L = blue**

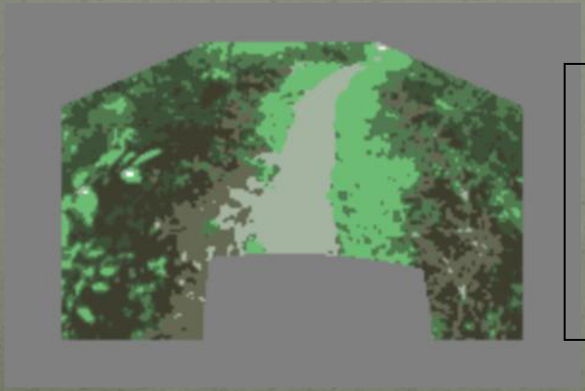
Omnidirectional Trail Following

- **Triangle approach works visually, but results cannot easily be translated into robot coordinates**
  - **IGVC 2008 showed that camera with narrow field of view was very limiting**
- **As with IGVC 2009, calibrated *omnidirectional* camera allows for trail shape hypothesis to be expressed in robot rather than image coordinates**
- **Don't have to process whole image—just "look" where you need to**

4-D Omnidirectional Trail State

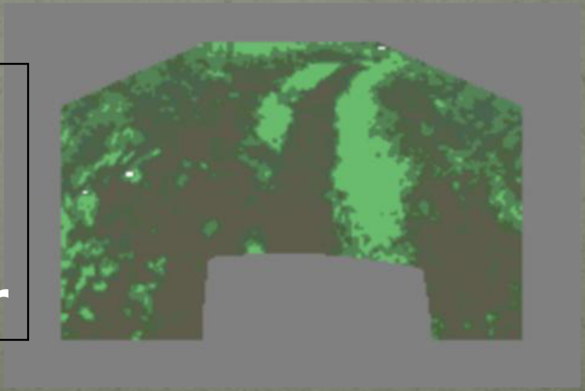






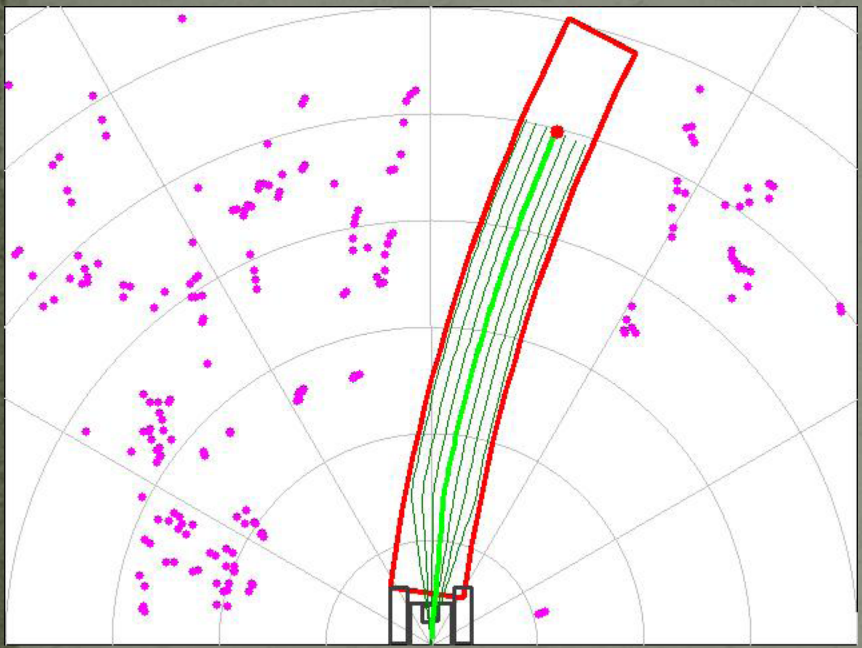
LAB textons

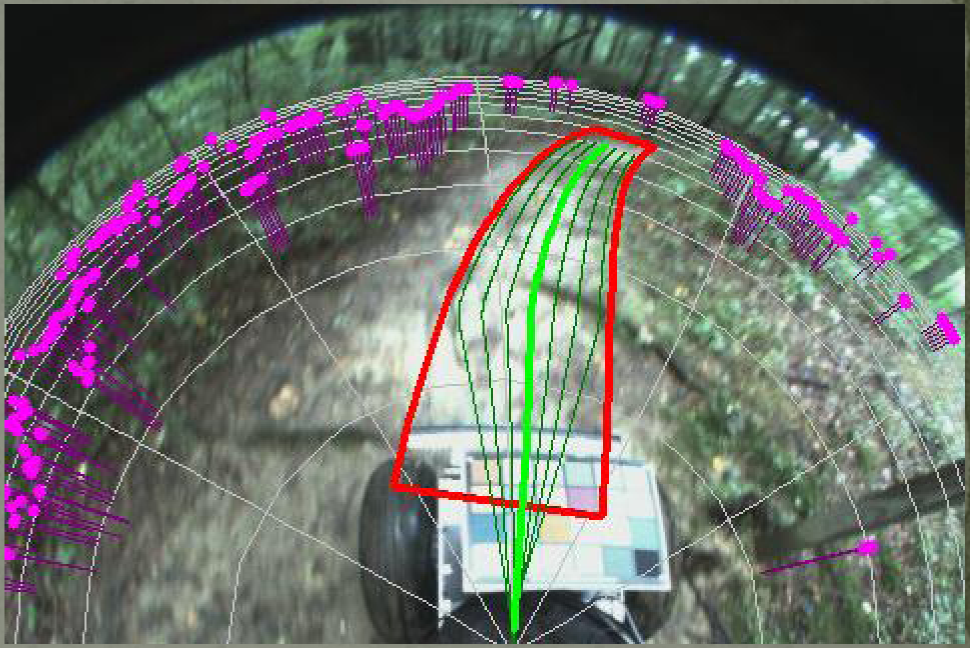
1. **Width**
2. **Curvature**
3. **Lateral offset**
4. **Heading error**

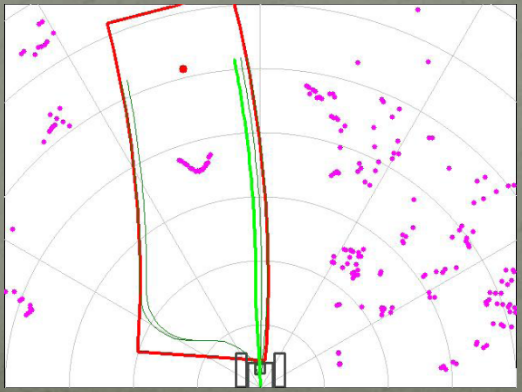


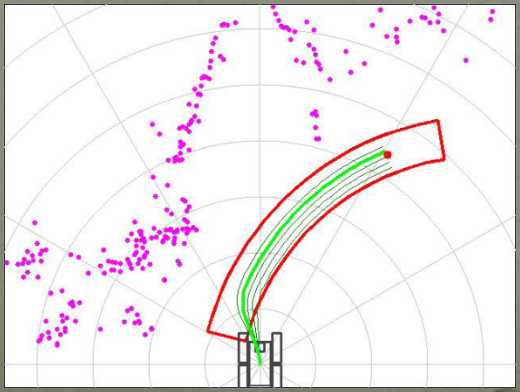
AB textons

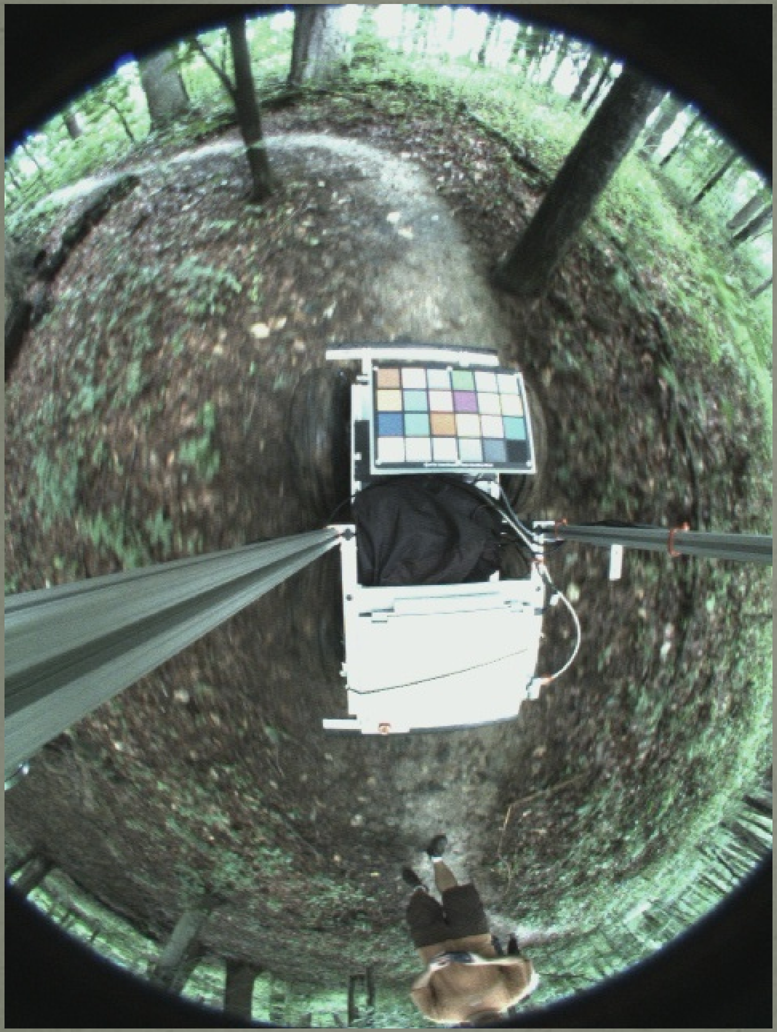
Omnidirectional Motion Planning





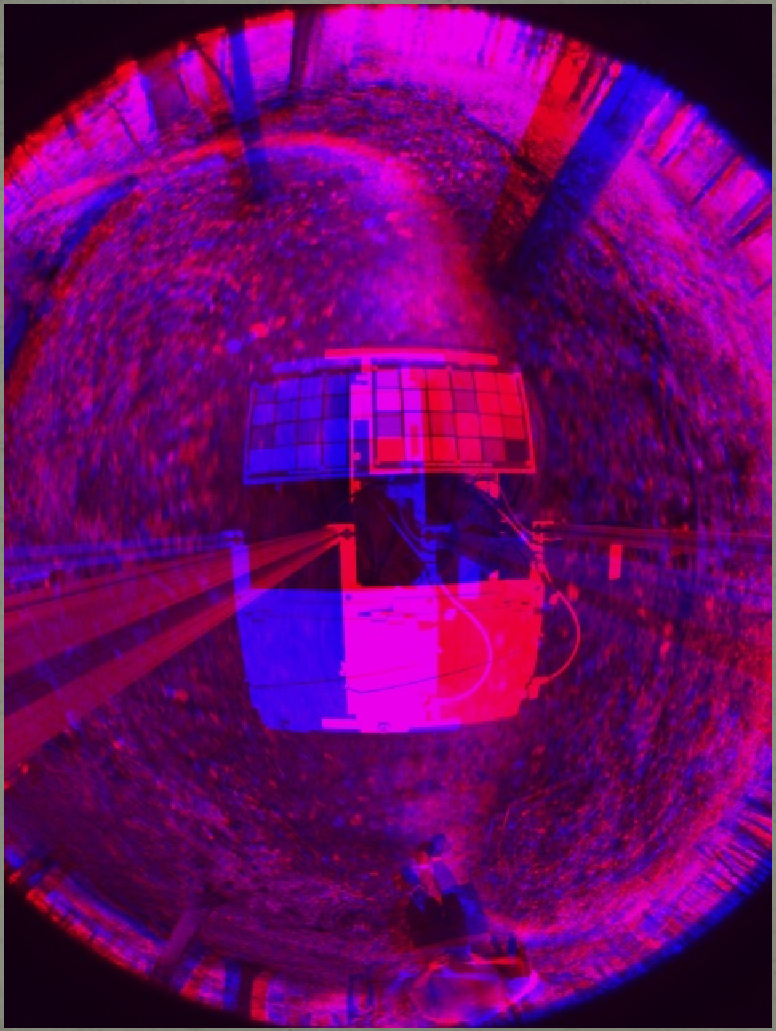




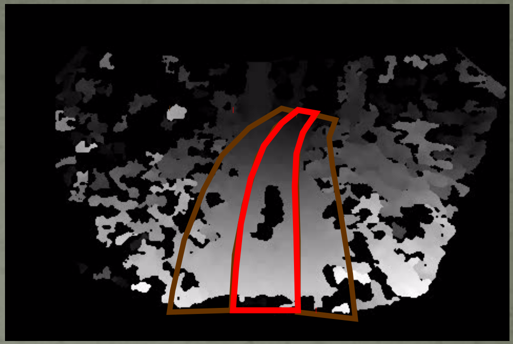


Left camera view

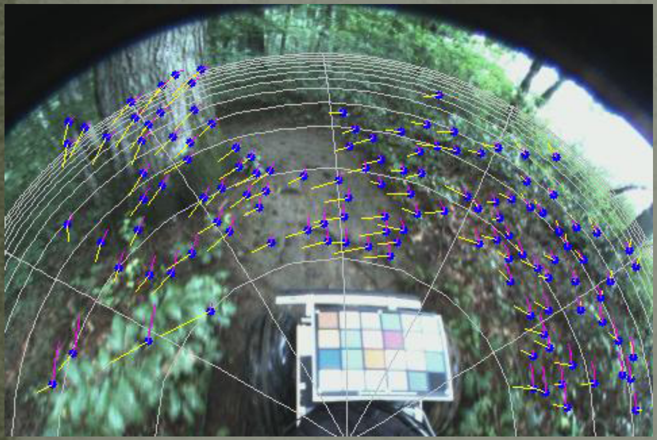
Right camera view



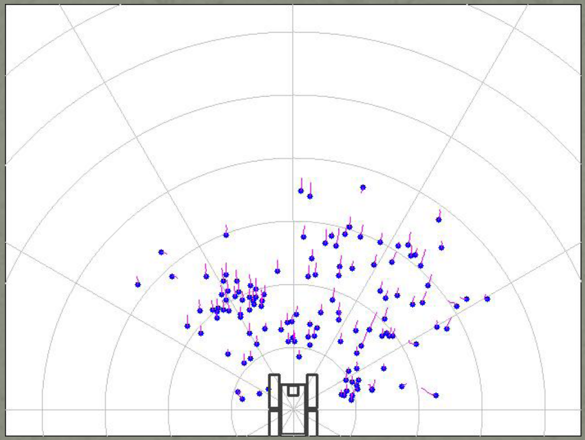
Ongoing Work



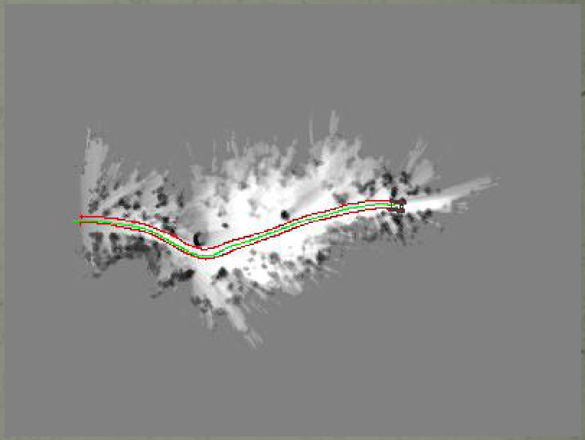
- **Combine structural information (ladar + stereo) with appearance in trail likelihood function**
- **Visual odometry for obstacle registration and map creation**



Tracking of corresponding stereo features



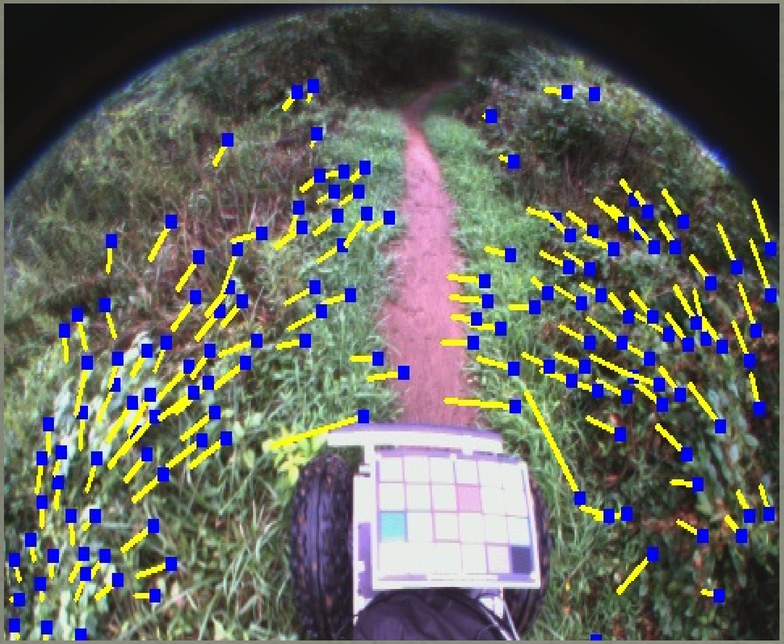
Feature triangulation and robust 3-D motion estimation



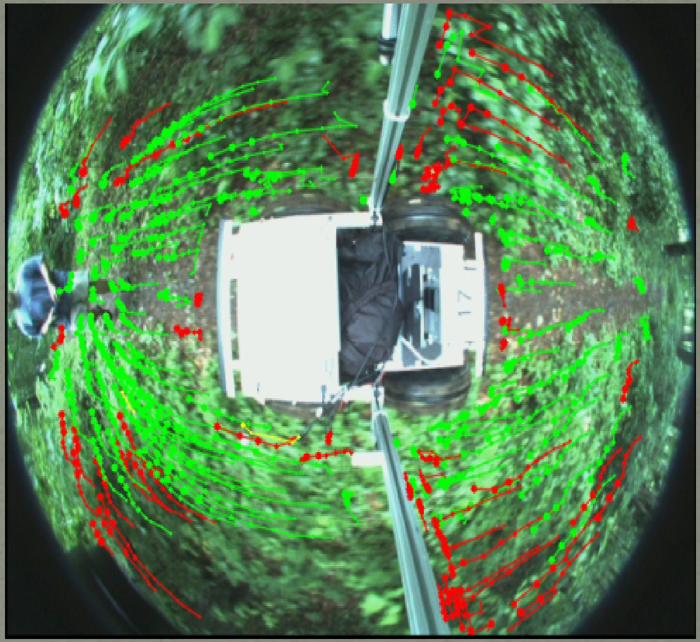
Motion integration and accumulation of obstacle observations in global map

- **Incorporate trail color model into tracked state**

Stereo + visual odometry



Stereo depth estimation for small & negative obstacles



Optical flow for visual odometry