Underwater Computer Vision:
Two Decades of Traditional Approaches and a New Paradigm in 3-D Reconstruction based on Opti-Acoustic Stereo Imaging

Shahriar Negahdaripour

Underwater Vision & Imaging Lab
Electrical & Computer Engineering Dept.
University of Miami

Overview

- Why underwater vision is difficult, and progress has been slow;

- What has been done in favorable undersea conditions;

- Some challenging problems and how they have been overcome;

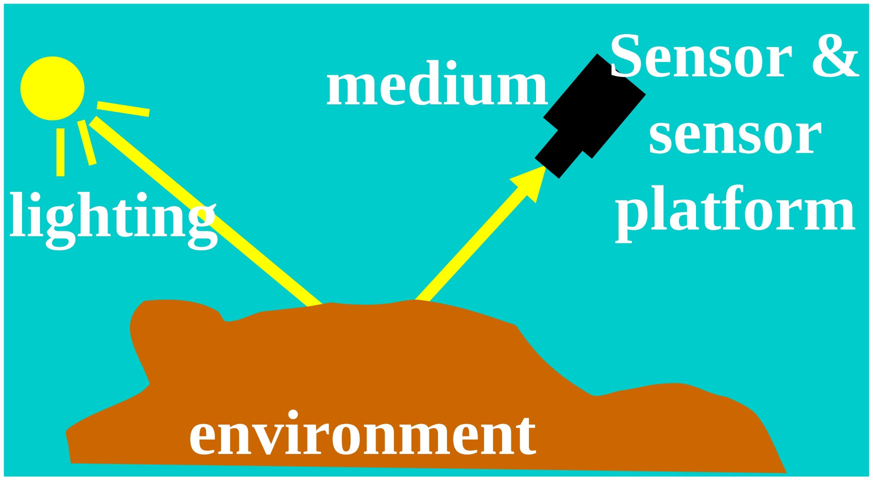
- Integration of sonar and optical imaging as a new paradigm in 3-D target mapping.

Why slow Progress?

1. More difficult problems;
2. Very small research community;
3. Need high levels of investment, resources (platform, sensors, ship time) and interdisciplinary expertise;
4. Science disciplines vary in needs (large effort goes into developing technology for specific application);
5. Fundamental problems highly depend on environmental conditions that vary with location, depth, time of day, season, etc. (e.g., sea floor –vs- water column – vs- near sea surface);
6. Many marine science needs require development of tools, NOT always challenging problems that intrigue CV researchers;
7. Scaaaaaaaaarcity of Funding sources

   -- NSF Marine Sciences: Do not typically fund CV work
   -- NSF CISE Program: Terrestrial domain priorities
   -- ONR, NOAA, USCG: Very limited funding for basic research, only for specific applications (mine detection, instrumentation, data collection, HLS) and short timelines
   -- Industrial support: Non-existent (small market, high develop. costs)
8. Very limited job opportunities: Discouraging to Graduate students

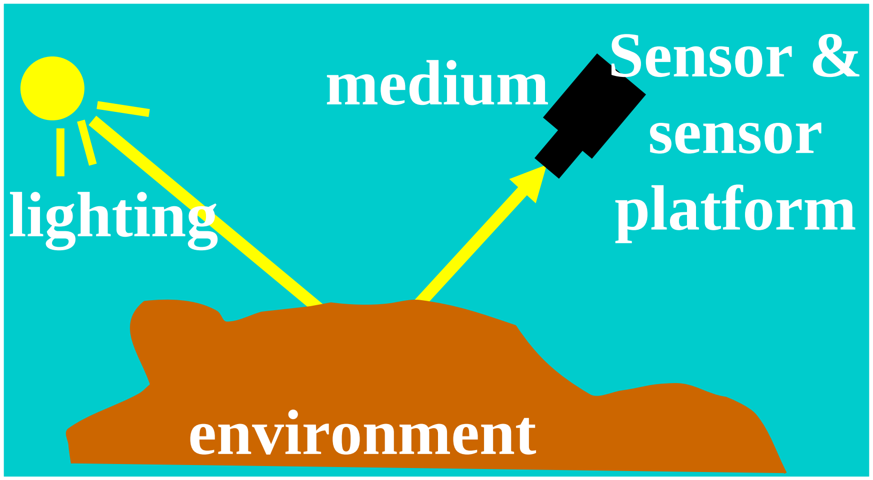
Why is vision difficult in underwater?



1. Medium

- **Poor signal-to-noise ratio (**Turbidity, absorption and scatter)
- **Limited visibility/range:** Can only see
  - short distance from camera
  - (small) parts of objects/scene in each image
- **Need to navigate and collect more data** to image the same region of the scene compared to terrestrial applications: more image registration, drift error, etc.
- **Loss/ Distortion of color cues** (wavelength-dependent attenuation)

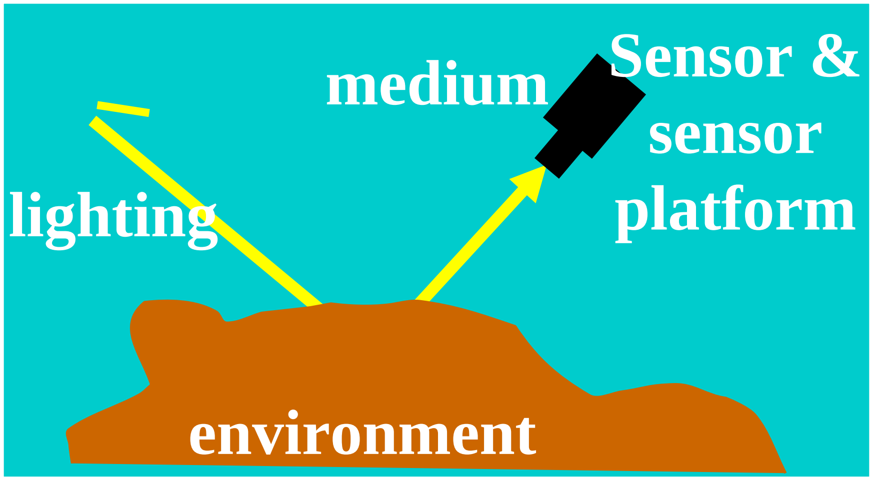
Why is vision difficult in underwater?



2. Environment: Complex & variable (unknown, unstructured, cluttered) : Vision system works

- in one place but fails 50 meters further away;
- in the morning but not a few hours later;
- in deep but not shallow waters;
- one depth but not other (e.g., in the deep, but not shallow waters).

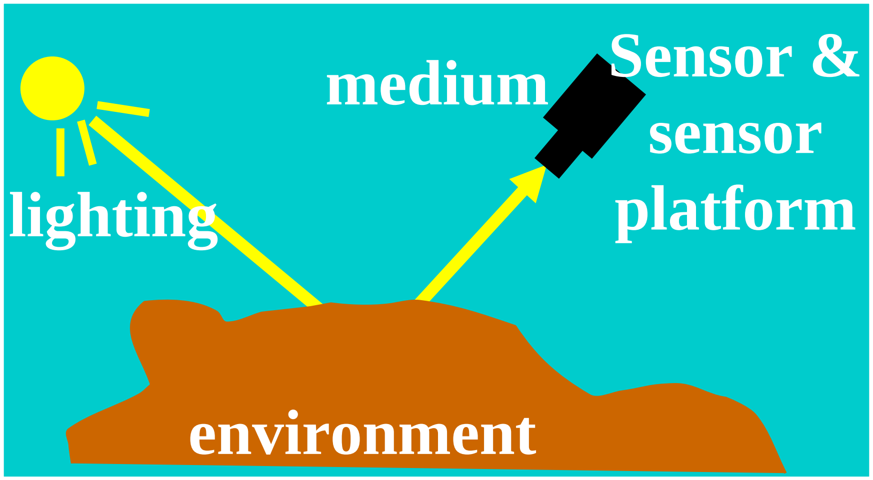
Why is vision difficult in underwater?



3. Illumination

- **Artificial lighting** (deep sea)
- **Complex cast shadows** (deep and shallow waters)

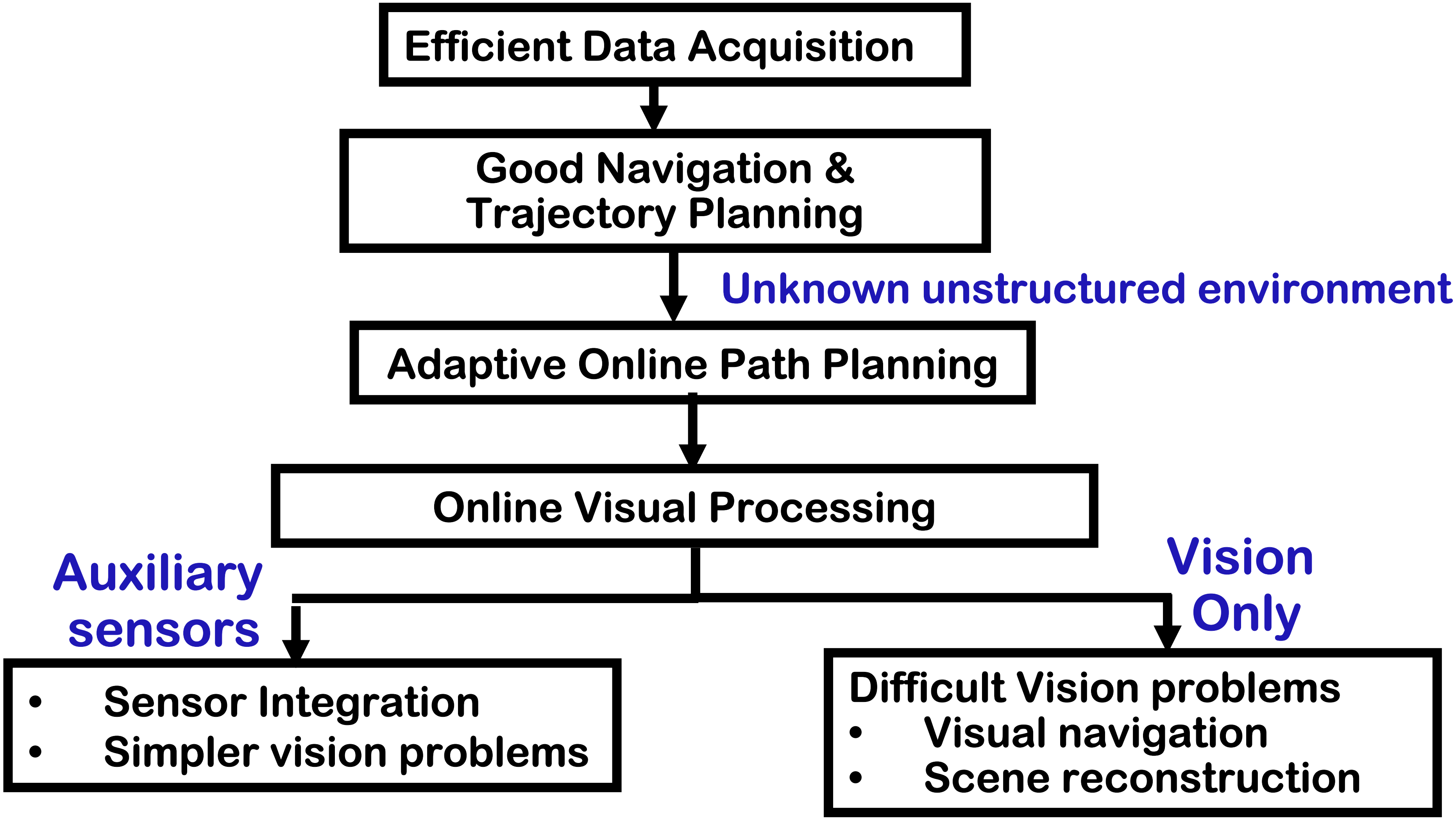
Why is vision difficult in underwater?



4. Poor positioning:

- 6 d.o.f motion in response to wave and current disturbances
- No GPS; drift from various (inexpensive) sensors
- Easier to carry out target-based rather than absolute positioning
- Auxiliary positioning sensors are very useful, but integration with camera(s) requires accurate calibration
- Much easier to carry our offline tasks, e.g. mosaicing, object modeling (potentially with manual guidance)

Operating in Unstructured Scattering Environment With Short-Range Visibility

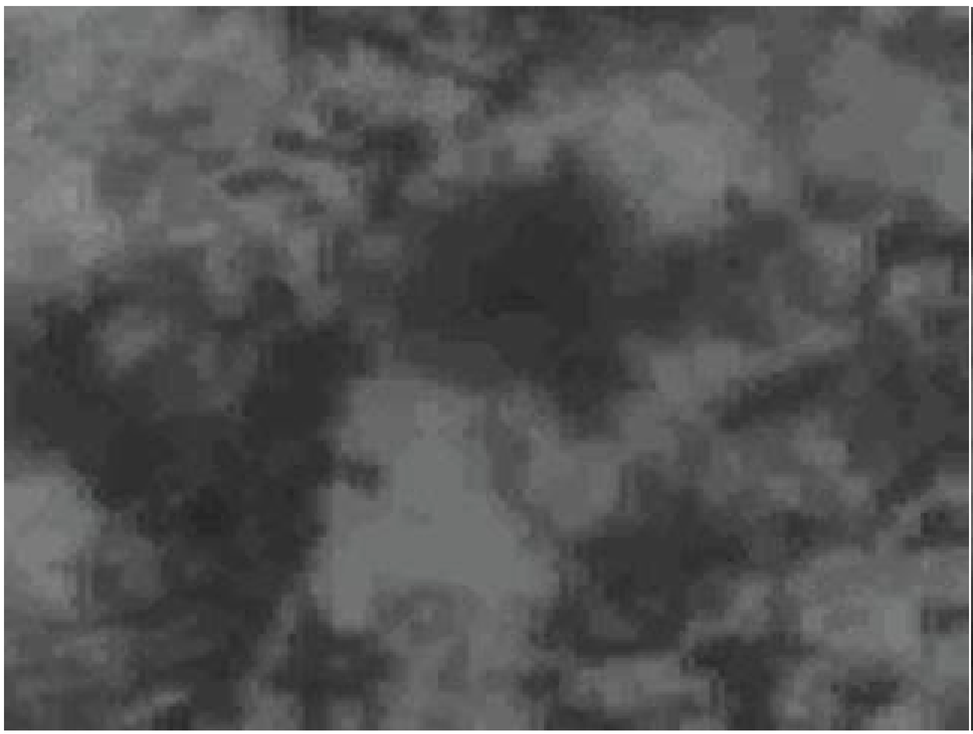


***Real-time vision-based capabilities* are critical for effective data acquisition, path planning & execution, target tracking & mapping:**

- **Optical Station Keeping (1989 – 1996)**
  Maintaining position under current disturbances
- **Video Mosaicing (1997)**
  Navigating on a 2-D map
- **Automatic Navigation (1997 – present)**
  Video-based ROV control
- **3D Reconstruction (2002 – present)**
  Benthic object mapping
  Target ID in search and inspection

Automatic Optical Station Keeping (1989-1996)

First ever ocean trial on commercial ROV
Miami Beach, FL (June 1999)



Manual ROV Control

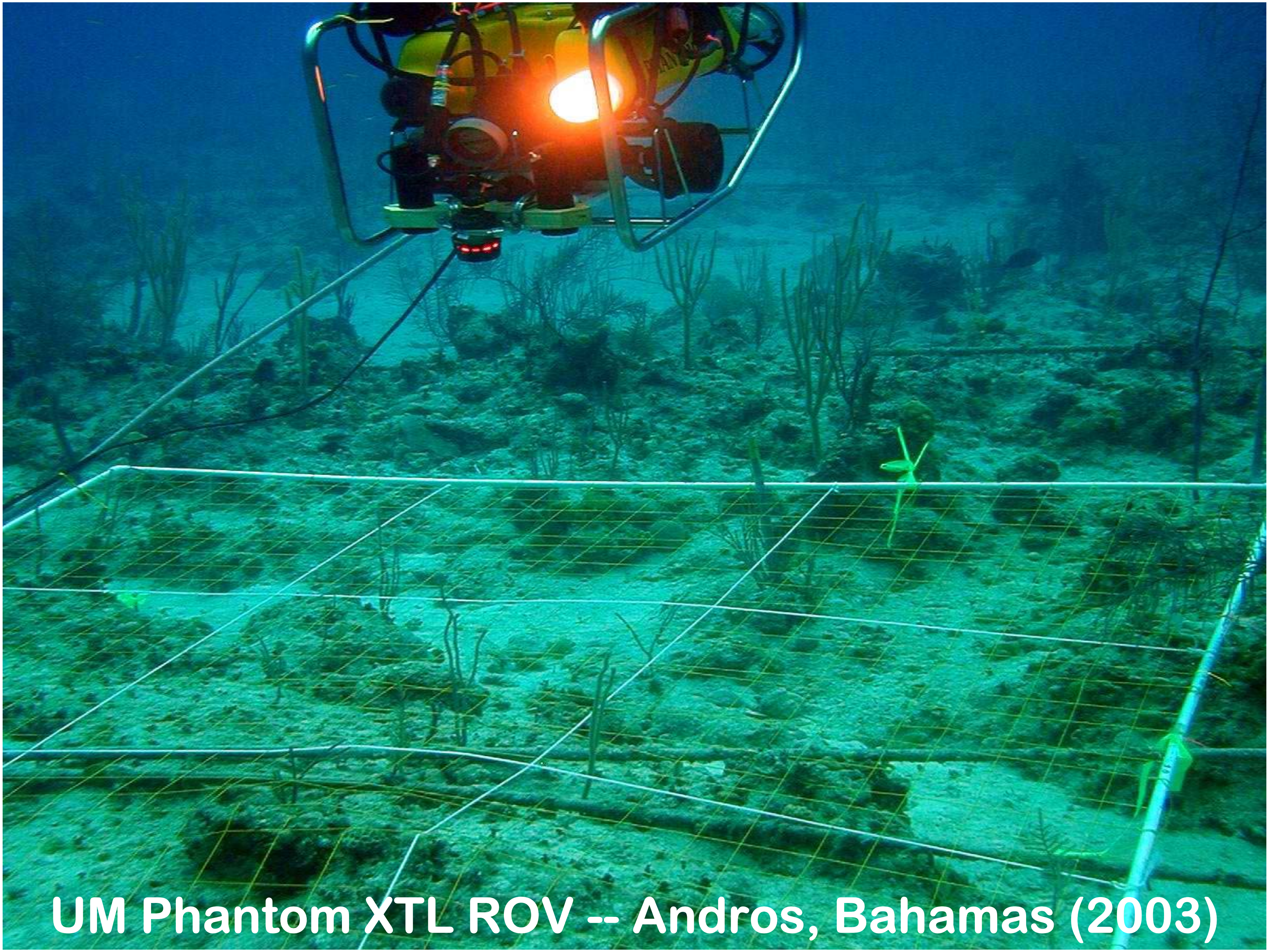
Automatic ROV Control

Strategic Environmental Research and Development Program

College of Engineering
Electrical and Computer Department

Example of Technological Developments under Favorable Visibility Conditions

Application of ROV Video Technology to complement Coral Reef Resource Mapping and Monitoring



UM Phantom XTL ROV -- Andros, Bahamas (2003)

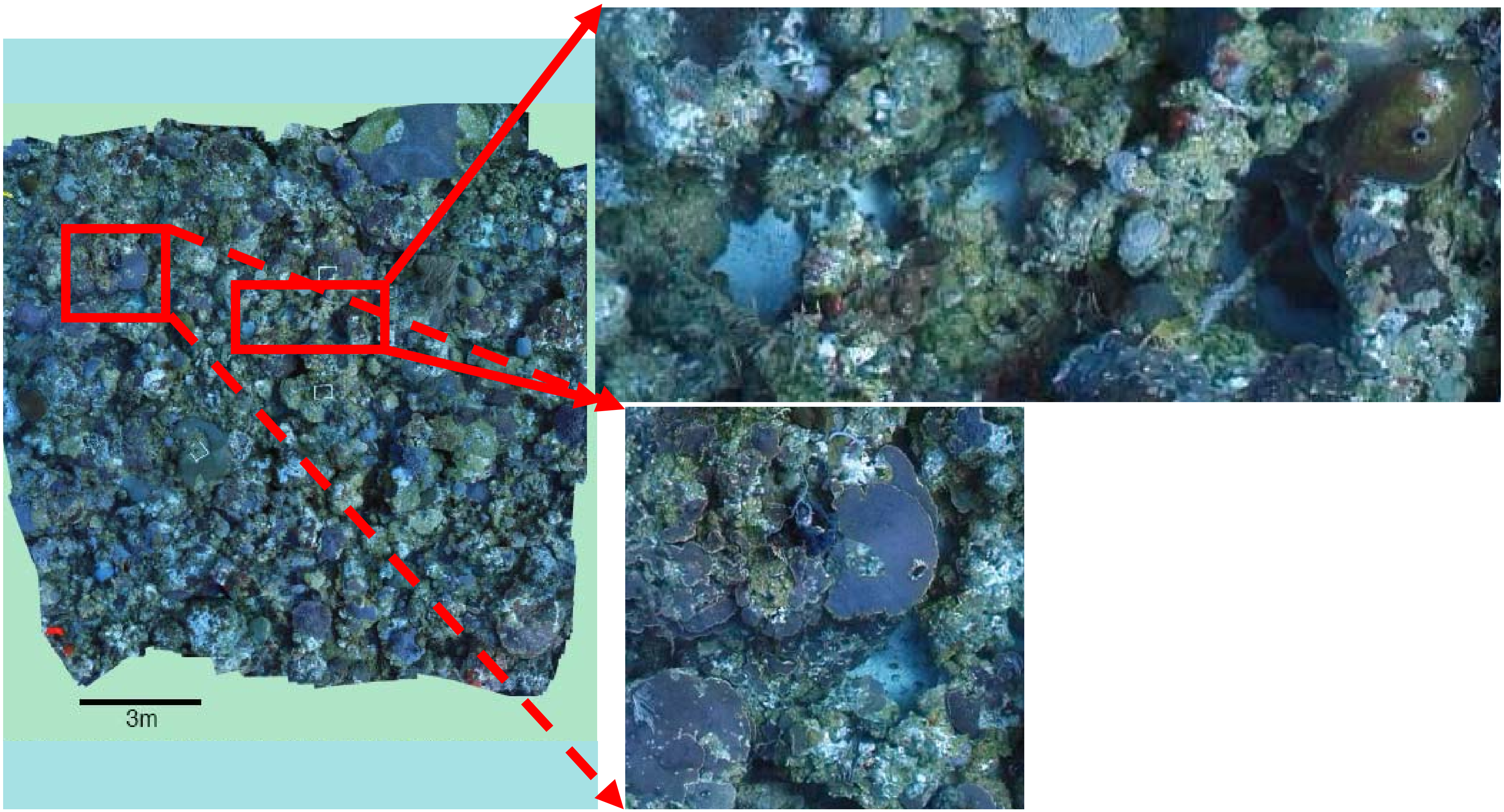
SERDP
Strategic Environmental Research and Development Program

Application of 2D Mosaicing

Deep Water Reef Community

College of Engineering
Electrical and Computer Department

Site: Dry Tortugas, FL
(stony Montastrea spp at 30[m])



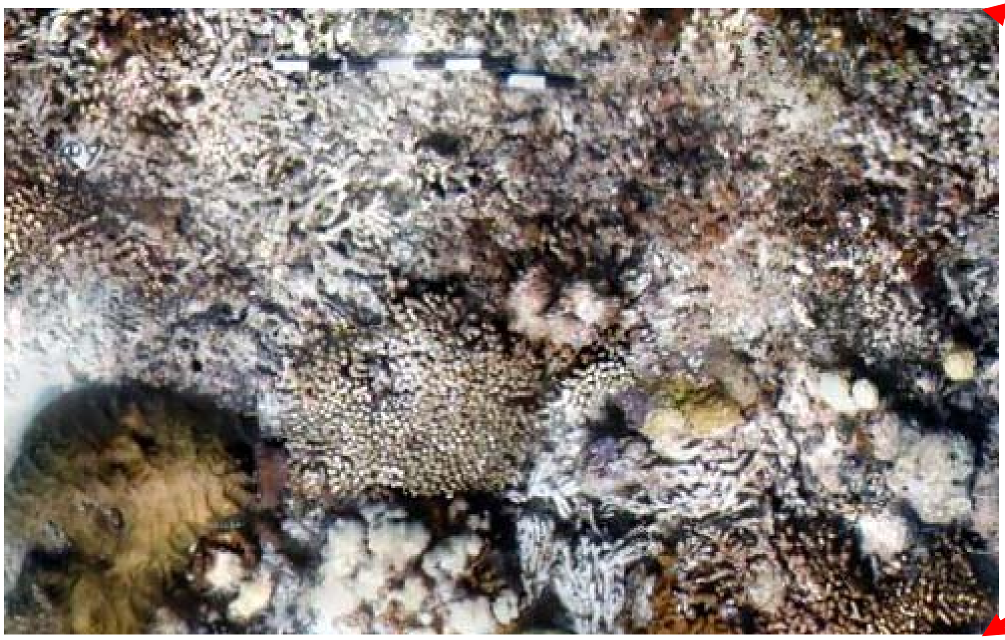
Application of 2D Mosaicing

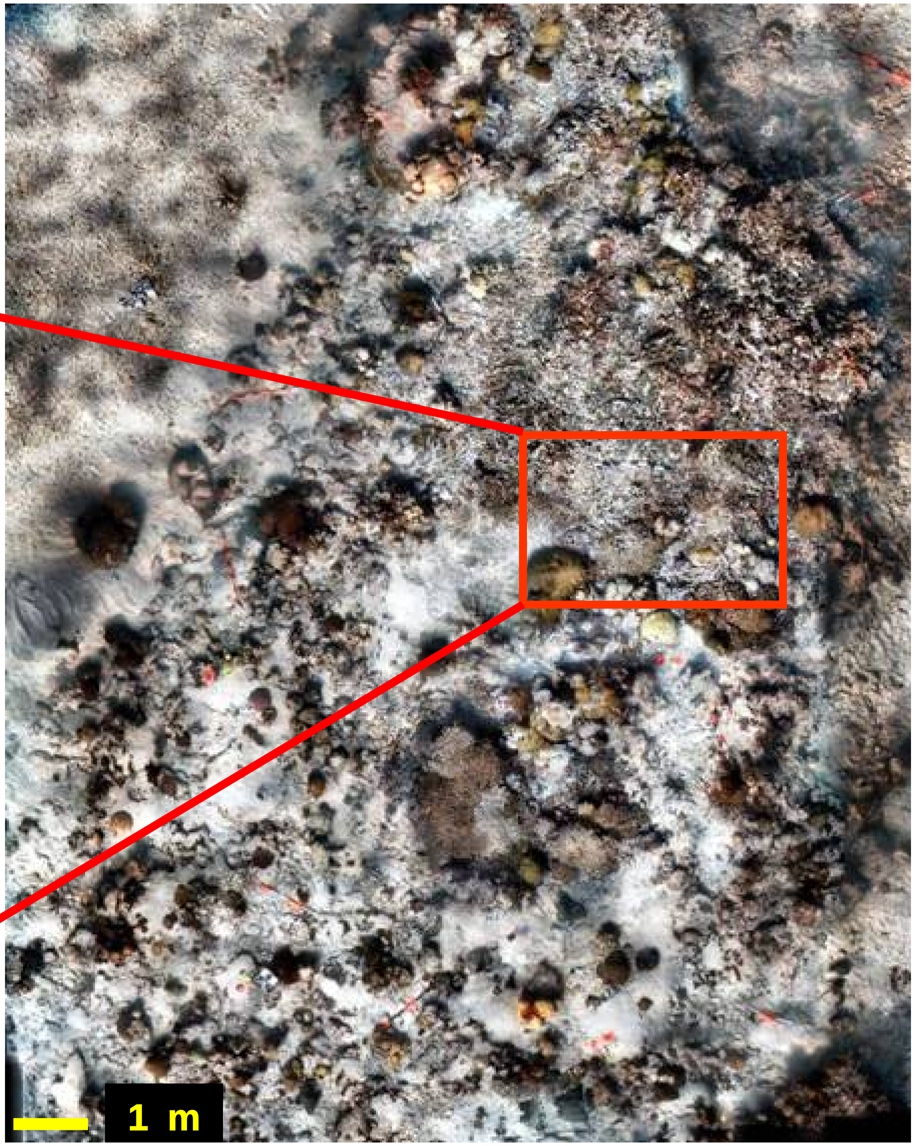
Coral Bleaching

SERDP Strategic Environmental Research and Development Program
DoD EPA DOE
RSMAS University of Miami
College of Engineering
Electrical and Computer Department

Site: Andros, Bahamas

December 2005)





DoD EPA DOE **SERDP** Strategic Environmental Research and Development Program

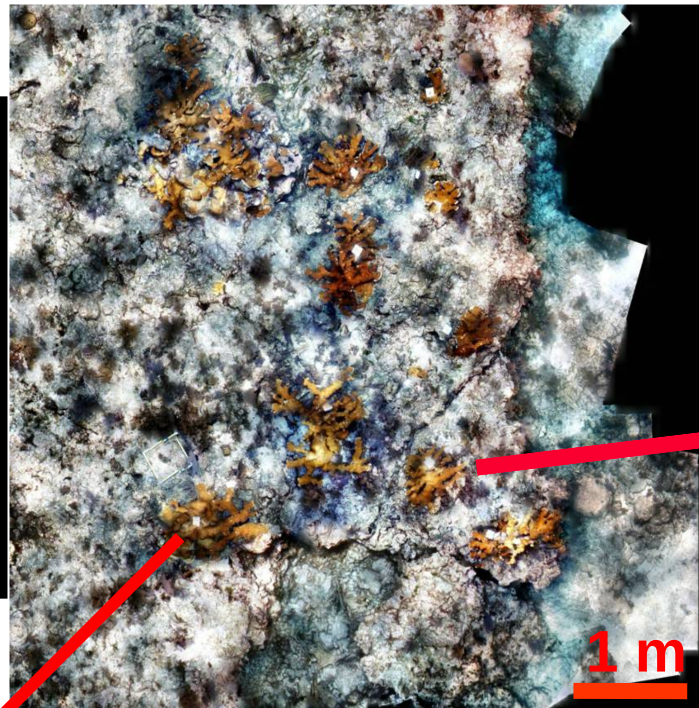
Application of 2D Mosaicing

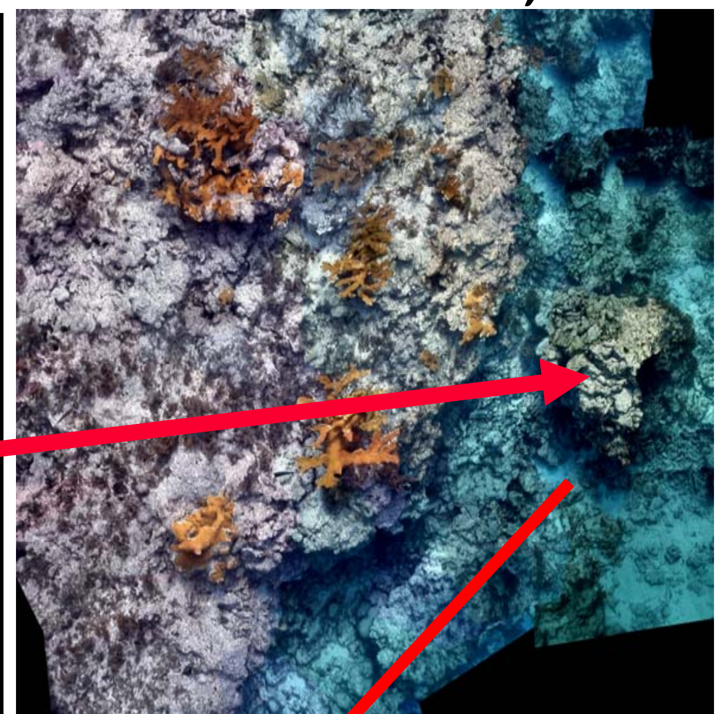
Hurricane Damage Assessment

RSMAS University of Miami
College of Engineering
Electrical and Computer Department

(Site: NOAA South East Fisheries Center – 2005/06)

Left (May’05). Right (Feb’06) after ***Hurricanes Dennis, Katrina, Rita, and Wilma*** passed over the area.





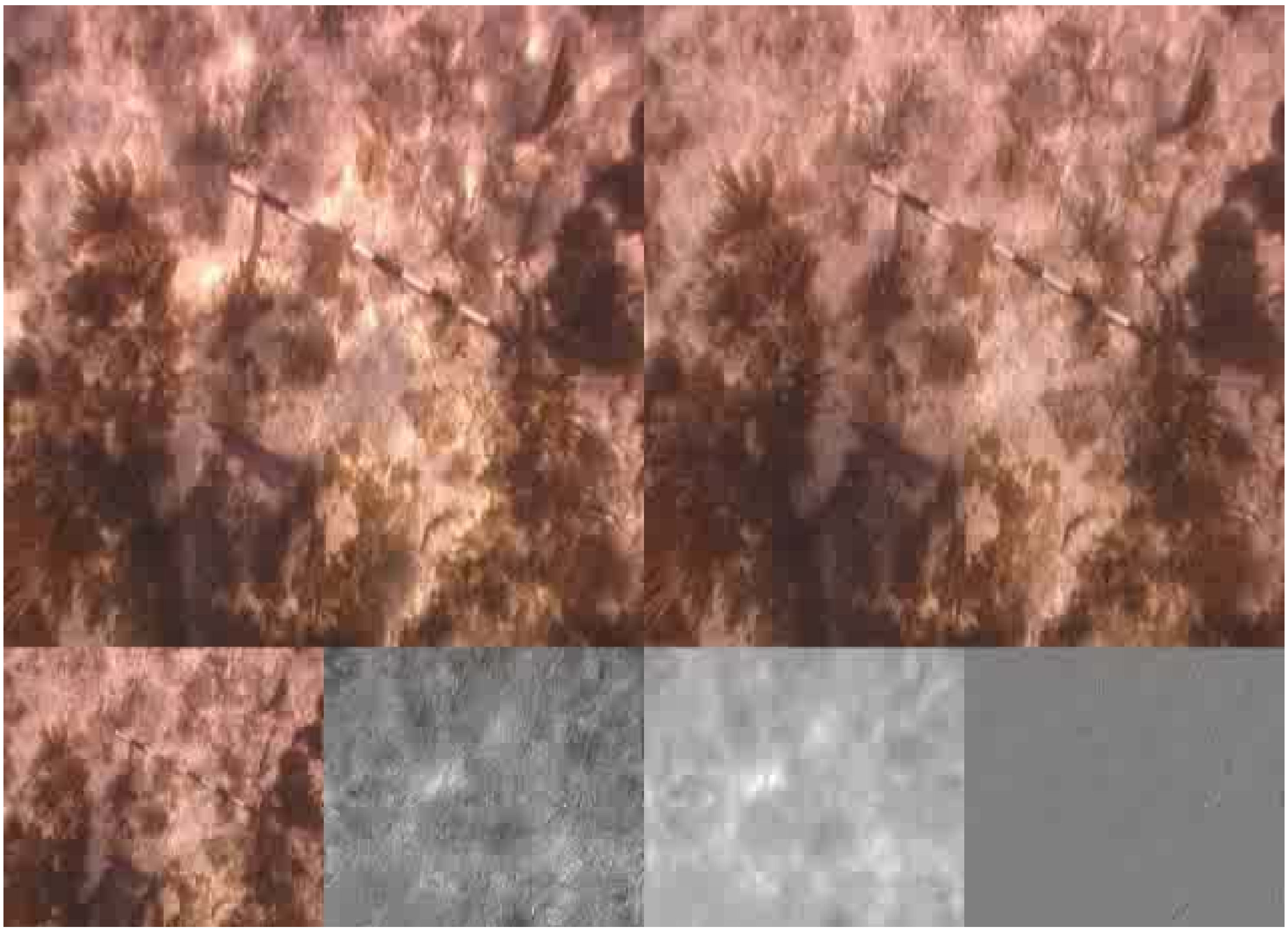
SERDP
Strategic Environmental Research and Development Program

Application of 2D Mosaicing

Ship Grounding Restoration

RSMAS
University of Miami
College of Engineering
Electrical and Computer Department

Site: Florida Keys

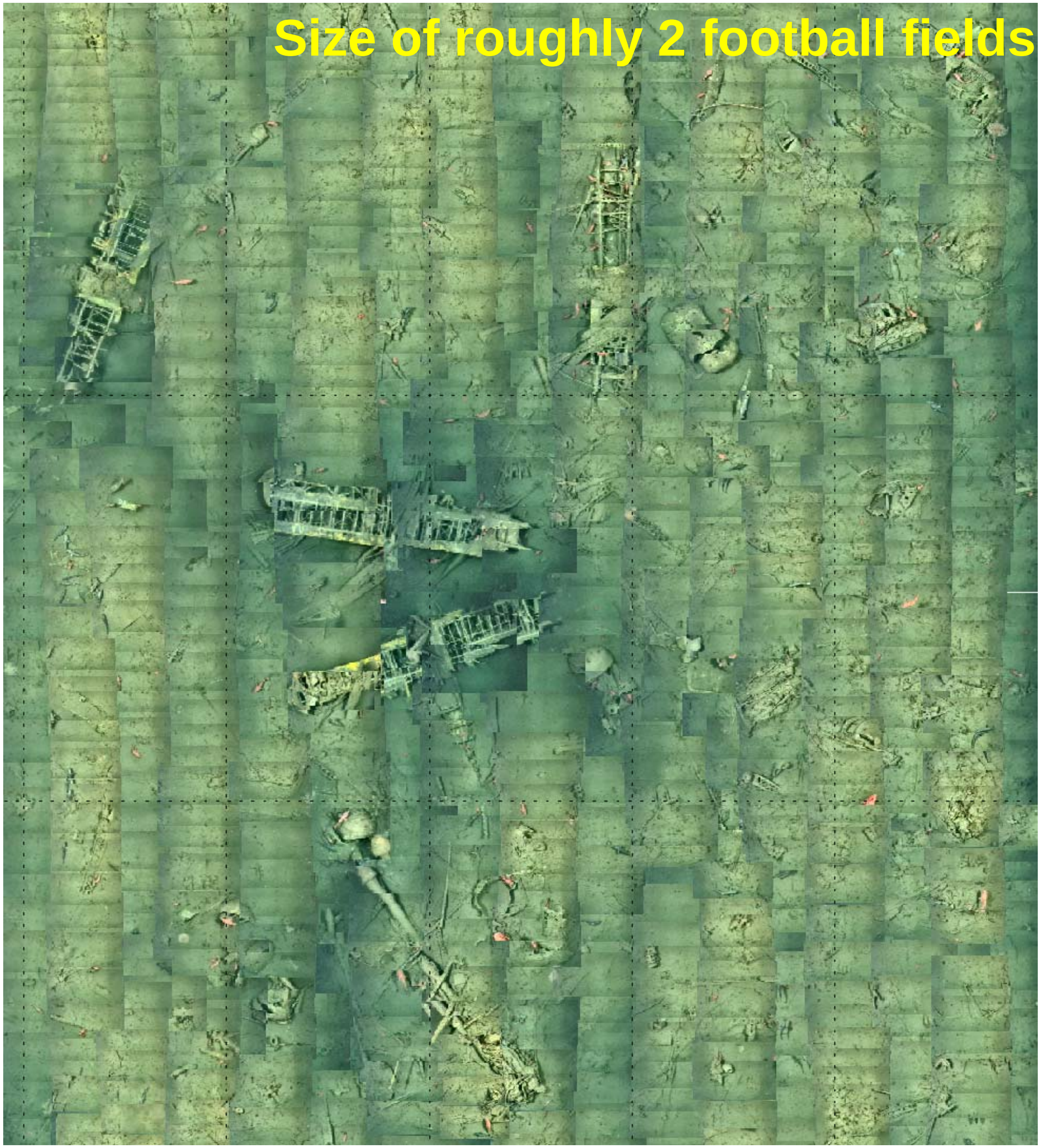


Archeology & wreck sites:

Navy ship USS Macon (Stanford/MBARI)

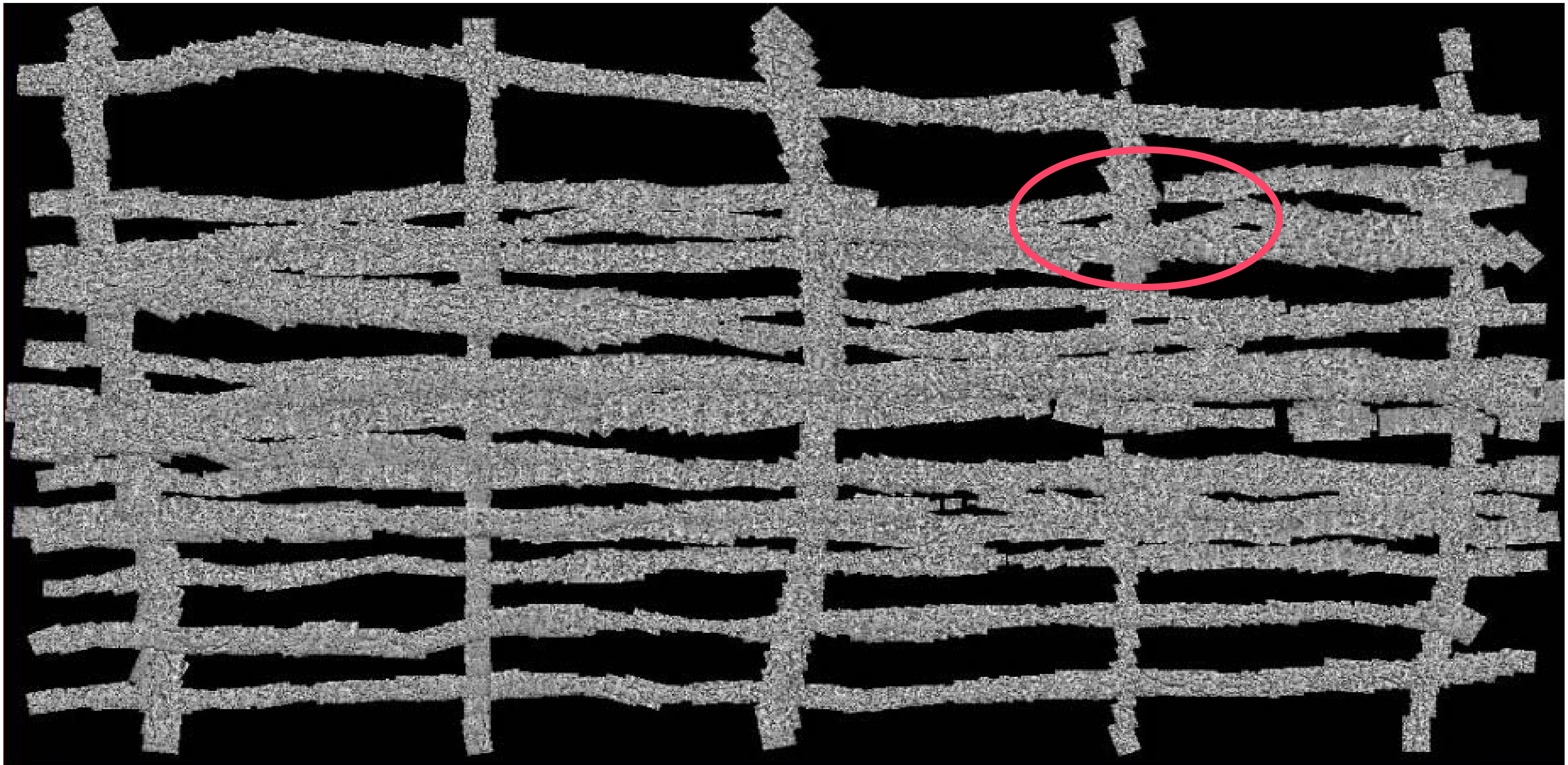
Positioning by

- Doppler Velocity Logger (DVL)
- Orientation sensors (pitch, roll, compass)
- Altitude sensor



Courtesy of Prof. Steve Rock, Stanford Univ.

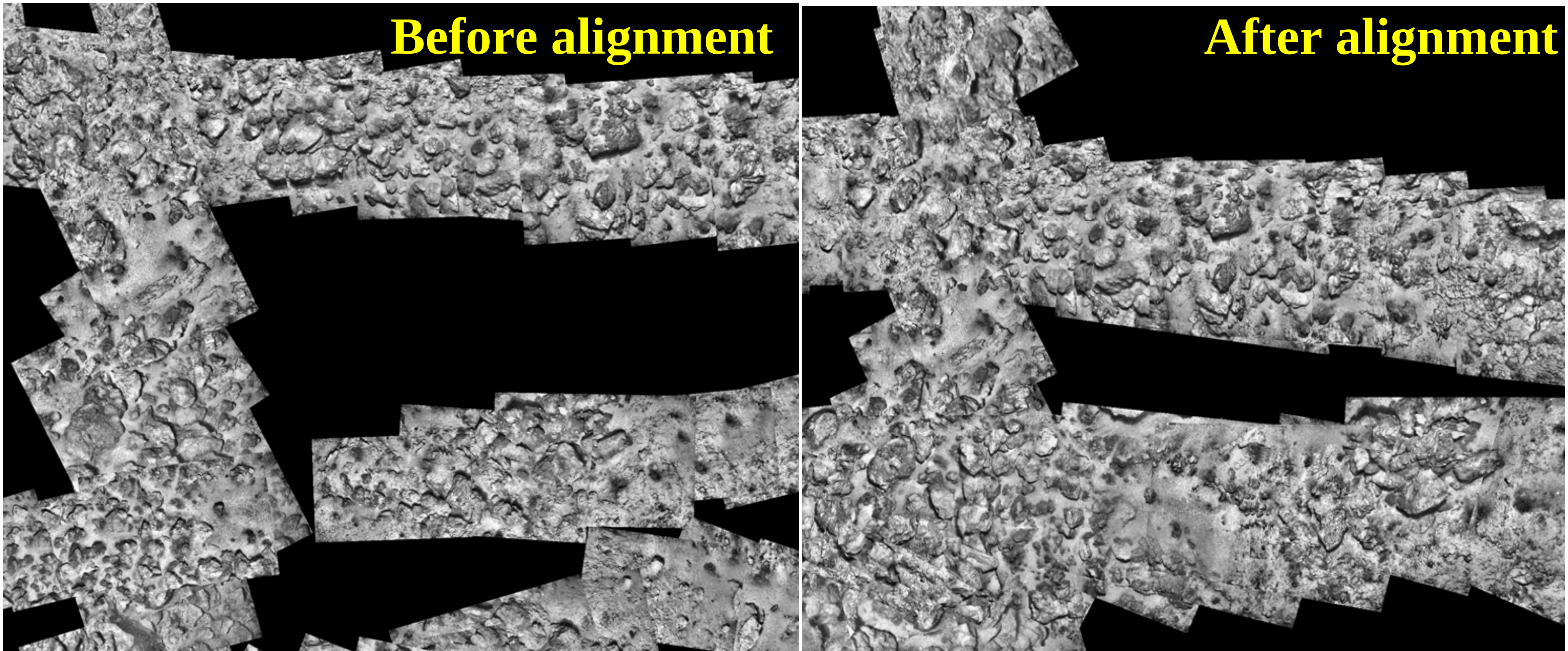
Marine Geology



Mosaic based on LBL positioning and compass information

Positioning: Long Baseline Acoustic (LBL) and compass
Courtesy of: Prof. Rafael Garcia

Marine Geology



Alignment by image registration

Positioning: Long Baseline Acoustic (LBL) and compass

Courtesy of: Prof. Rafael Garcia

Monocular Vision
Dense 3-D Reconstruction
(Plot Scale)

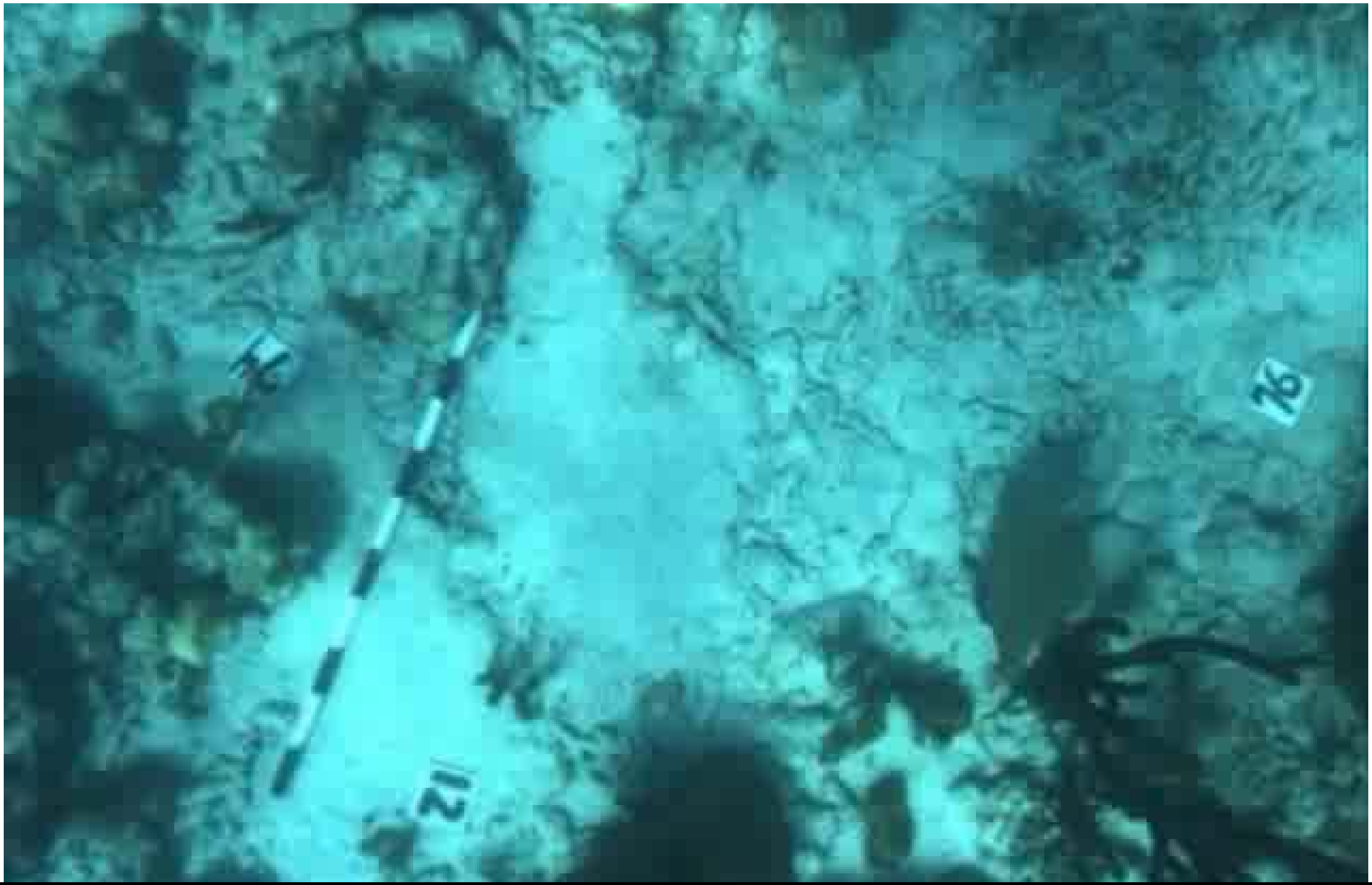
Issues:

- 3-D mapping and navigation within relatively large high-relief areas ;
- Processing large data effectively while dealing with data association problem (e.g., SLAM type formulations).

Approach

- Surface representation with minimum number of features while maintaining model accuracy;
- Identification of geometrically significance features that can be robustly tracked;
- Effective integration of optical-flow and feature-based approaches for dense surface reconstruction and camera trajectory estimation.

Monocular Vision
Dense 3-D Reconstruction
(Plot Scale)



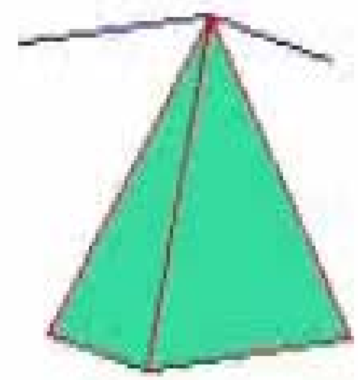
Data: Bahamas

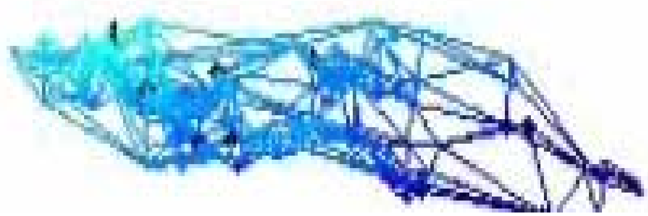
500 frames @ 512x384 resolution with total of 400 features @ 29.5 fr lifespan used for surface reconstruction

Monocular Vision
Dense 3-D Reconstruction
(Plot Scale)

Data: Bahamas

Reconstructed trajectory and piecewise planar terrain from 500 frames @ 512x384 resolution with total of 400 features @ 29.5 fr lifespan used for surface reconstruction





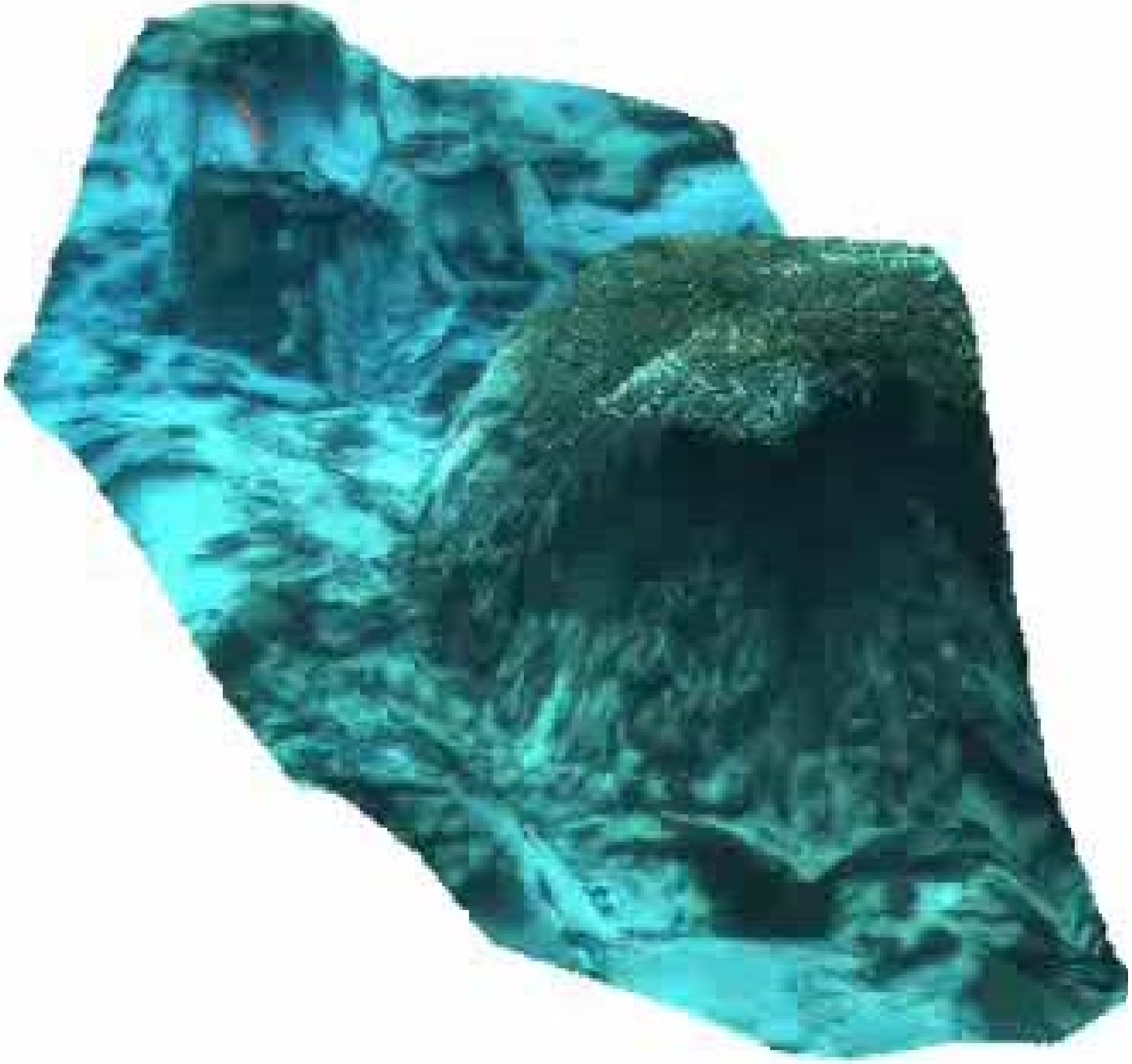
Monocular Vision
Dense 3-D Reconstruction
(Plot Scale)

Universitat de Girona

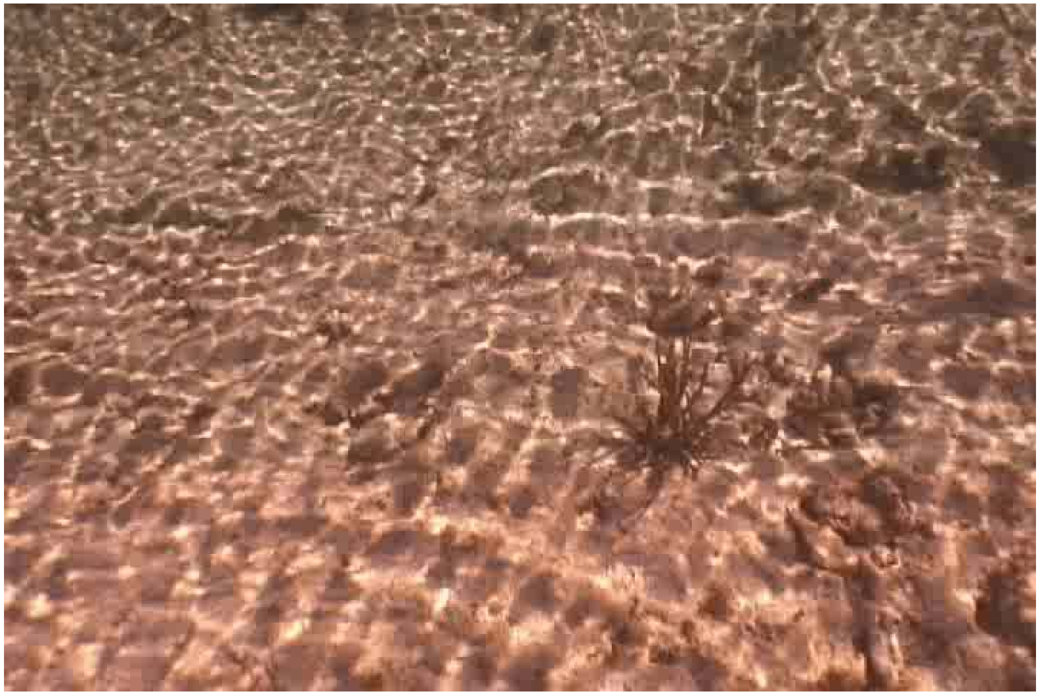
UNIVERSITY OF Miami

Data: Bahamas

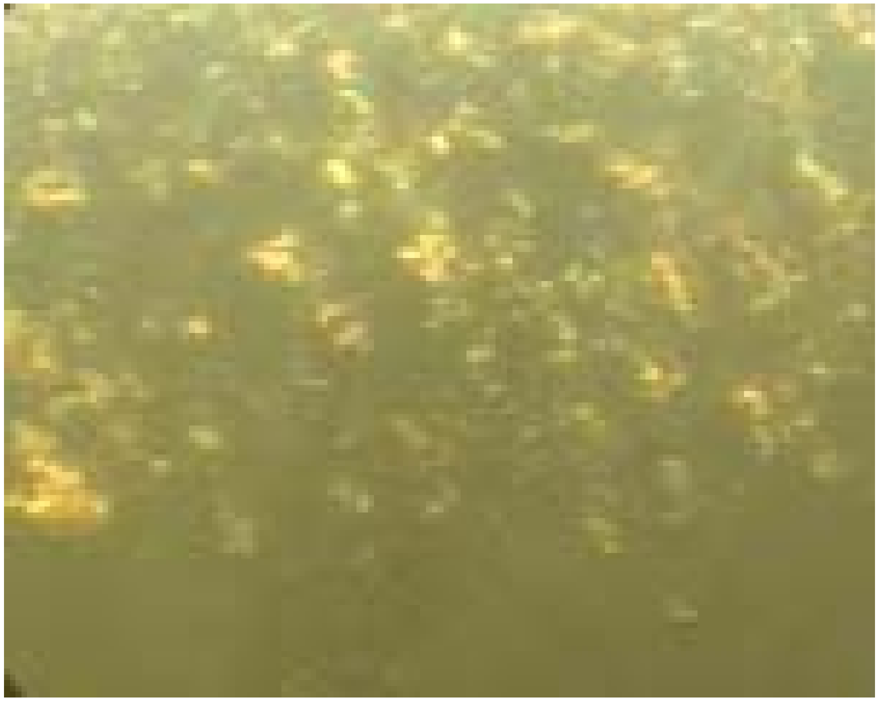
Texture-mapped reconstructed terrain from 500 frames @ 512x384 resolution with total of 400 features @ 29.5 fr lifespan used for surface reconstruction



Sun flicker and cast shadow can be quite significant



Shallow Water – Sea floor
(Florida Keys)

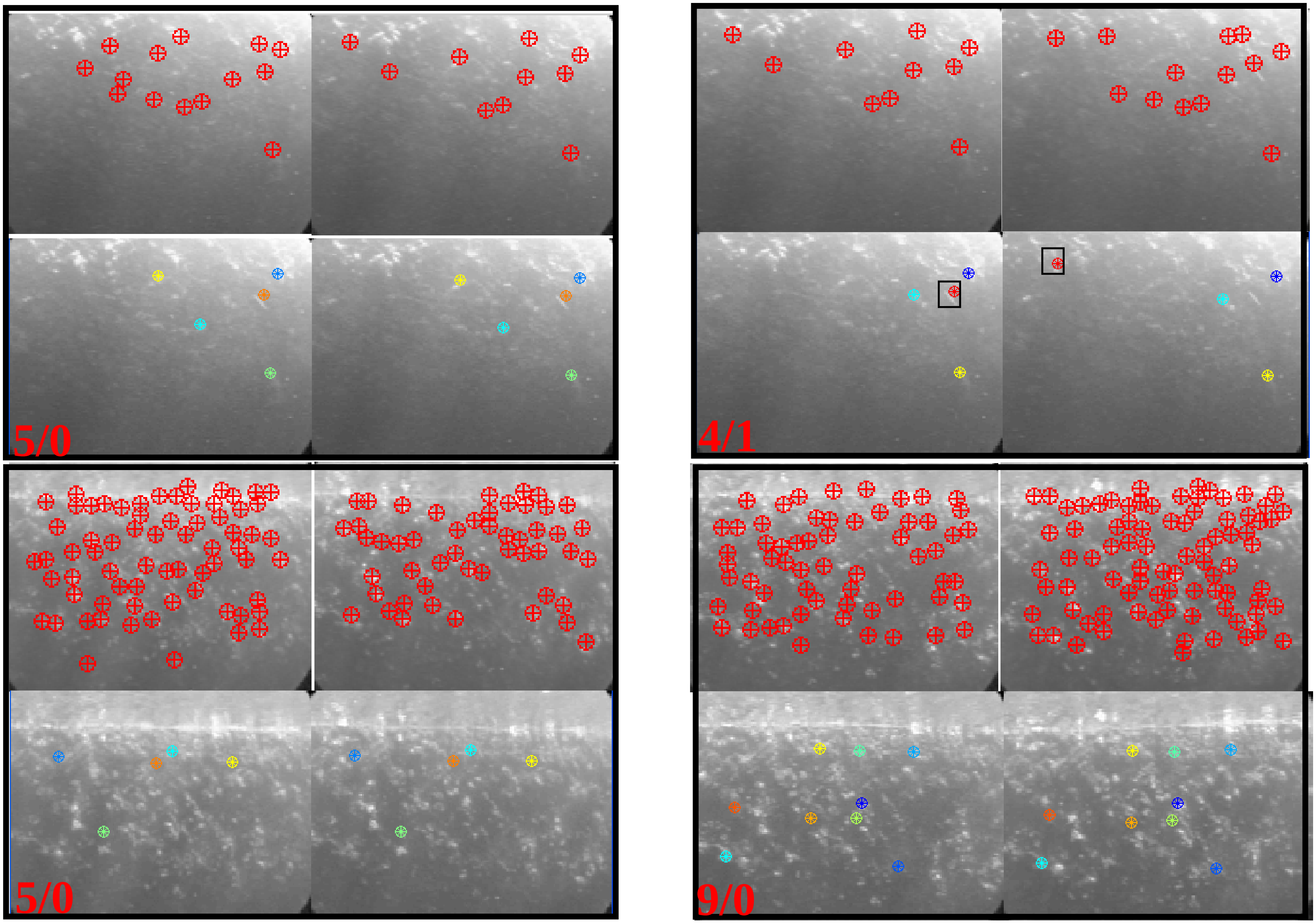


Sea Surface – Port Structure
(Panama City Naval Center)

Need a fundamental approach to deal with the problem, particularly for real-time operations

Sample Frames from Panama City Data

Feature Detection and Matching by SIFT



Stereovision?

Cast Shadows: Become signal not noise source;

Improved accuracy: Redundant spatio-temporal data;

Instantaneous Target-Based Positioning

Improved robustness (for 6 d.o.f motion):

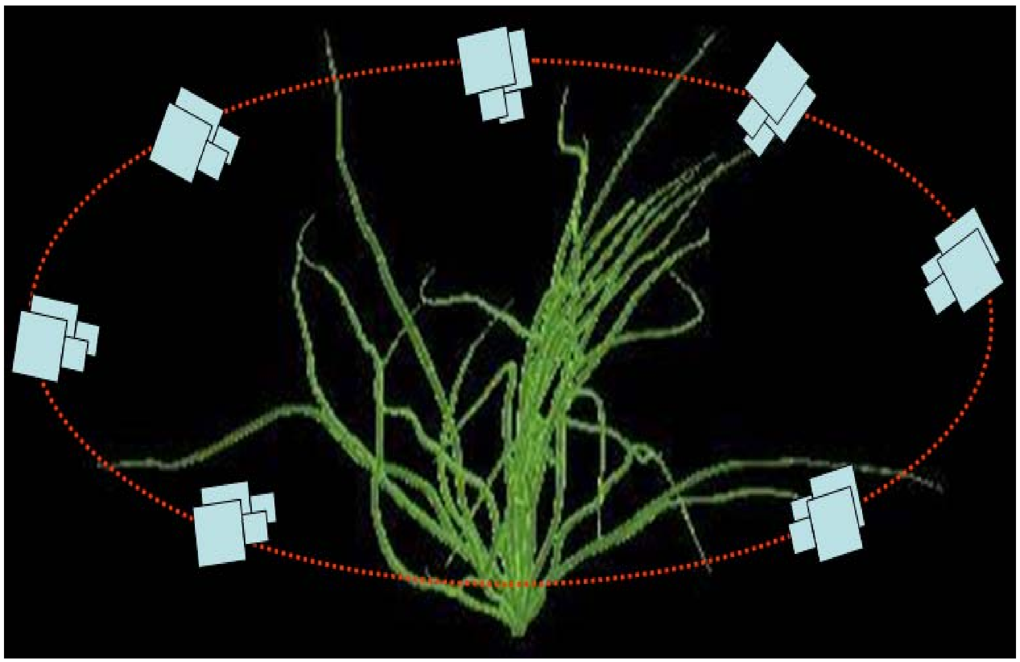
Require less features and less views;
Avoid degenerate cases;

Improved efficiency: Closed-form solution of absolute orientation in place of methods based on fundamental matrix or trifocal tensor;

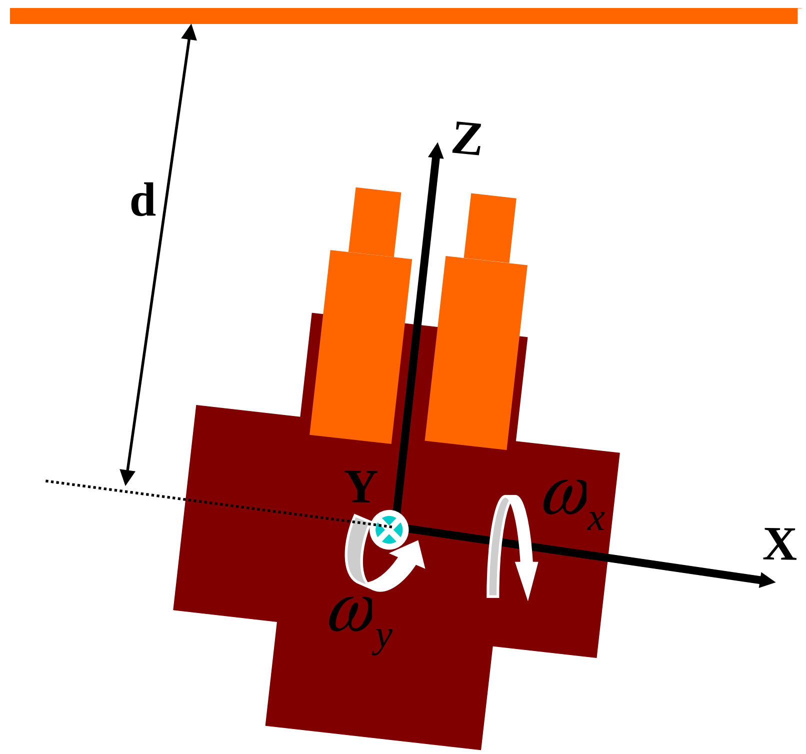
No scale drift (inherent in monocular motion vision);

More effective in dealing with dynamic feature sets (feature appearance/disappearance)

Stereovision: Has proven crucial in a range of underwater problems, handled by monocular vision in terrestrial domain

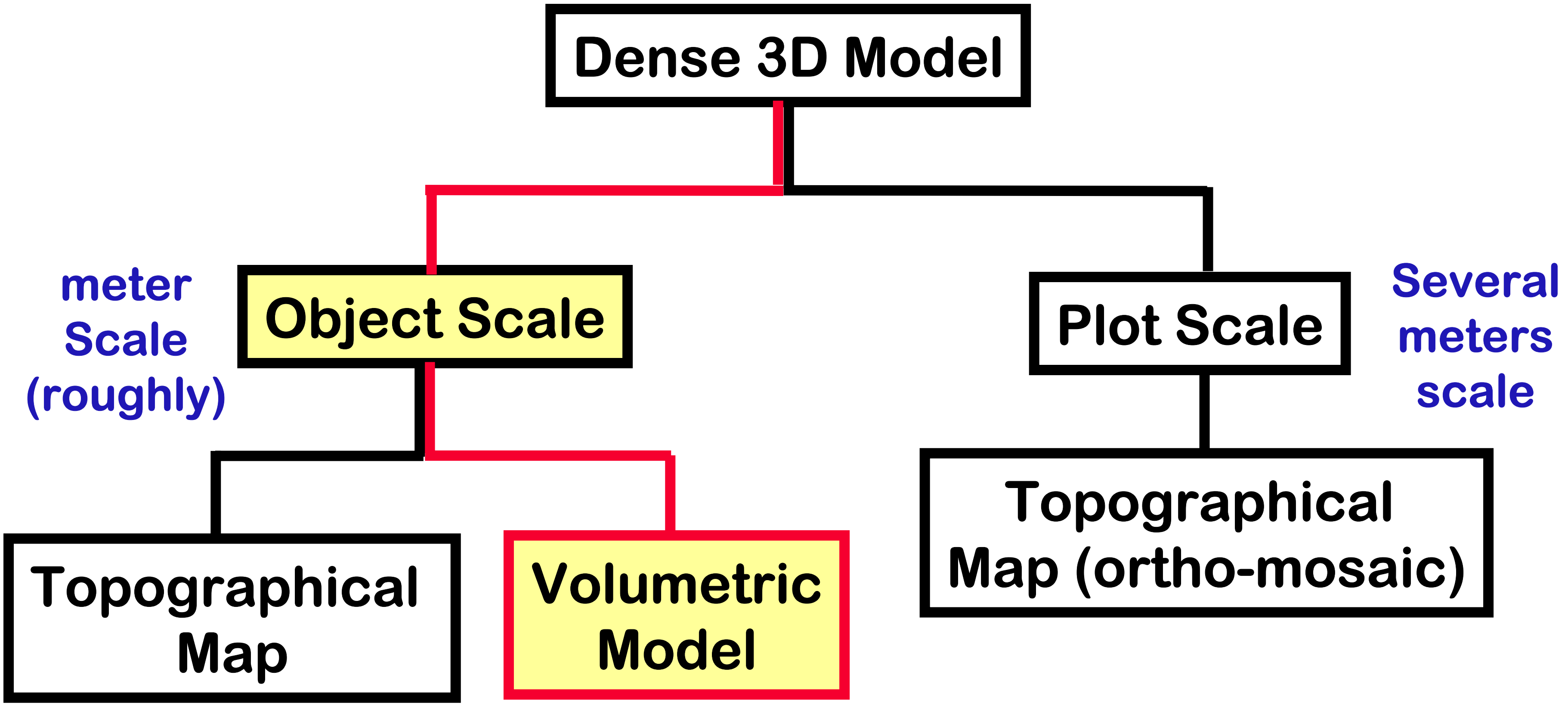


Example 1: Volumetric Modeling



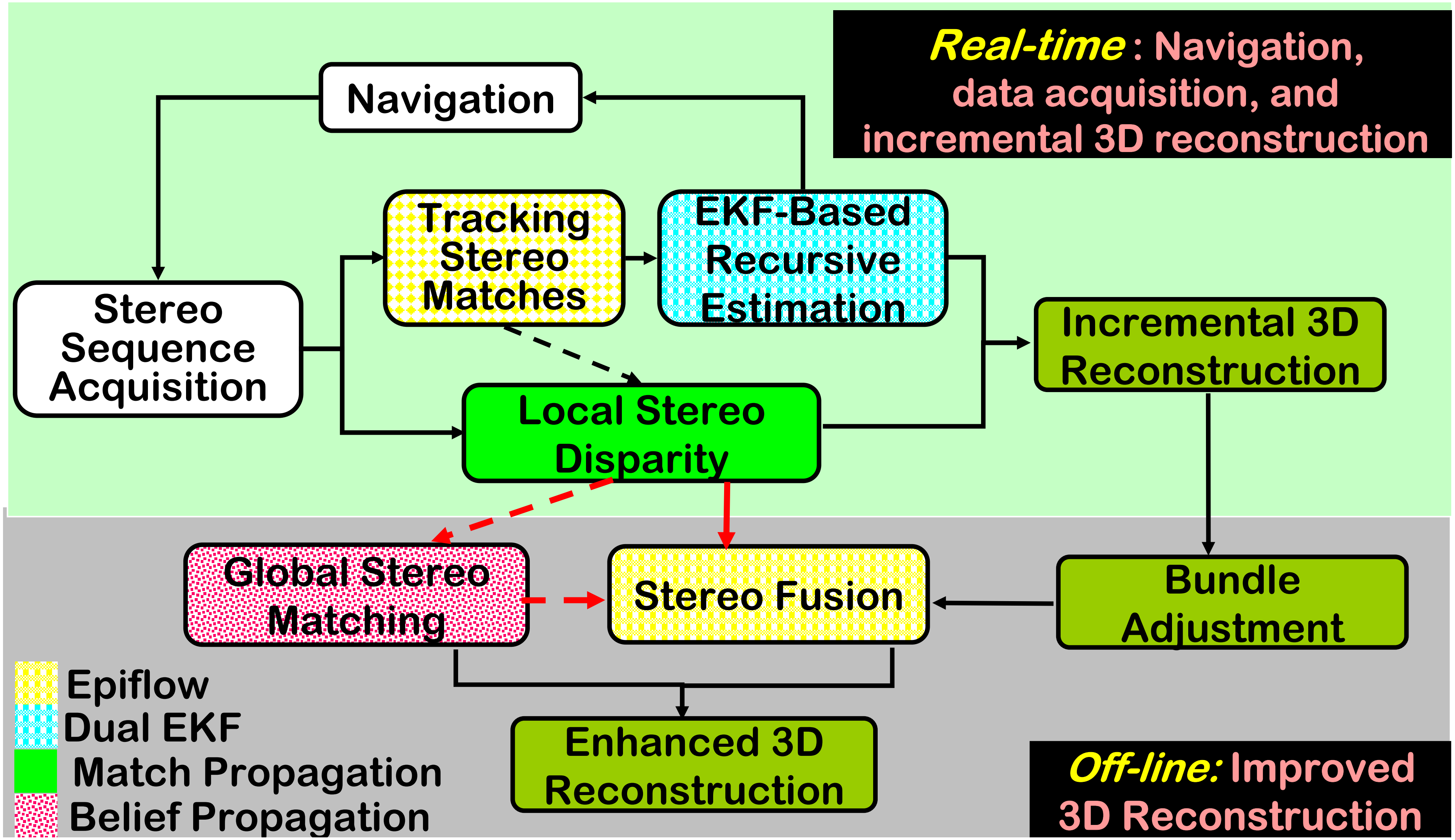
Example 2: Visual Odometry

Example 1:
Dense 3-D Scene Reconstruction



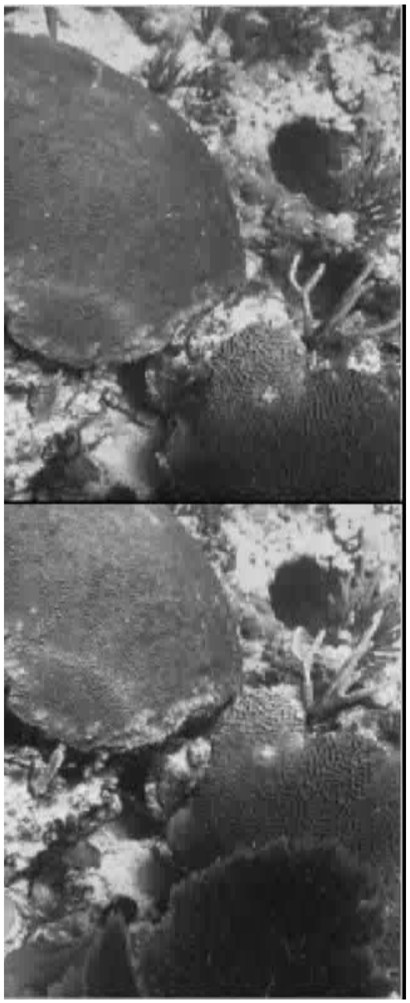
Example 1:
(Object Volumetric Modeling)

Computational Framework

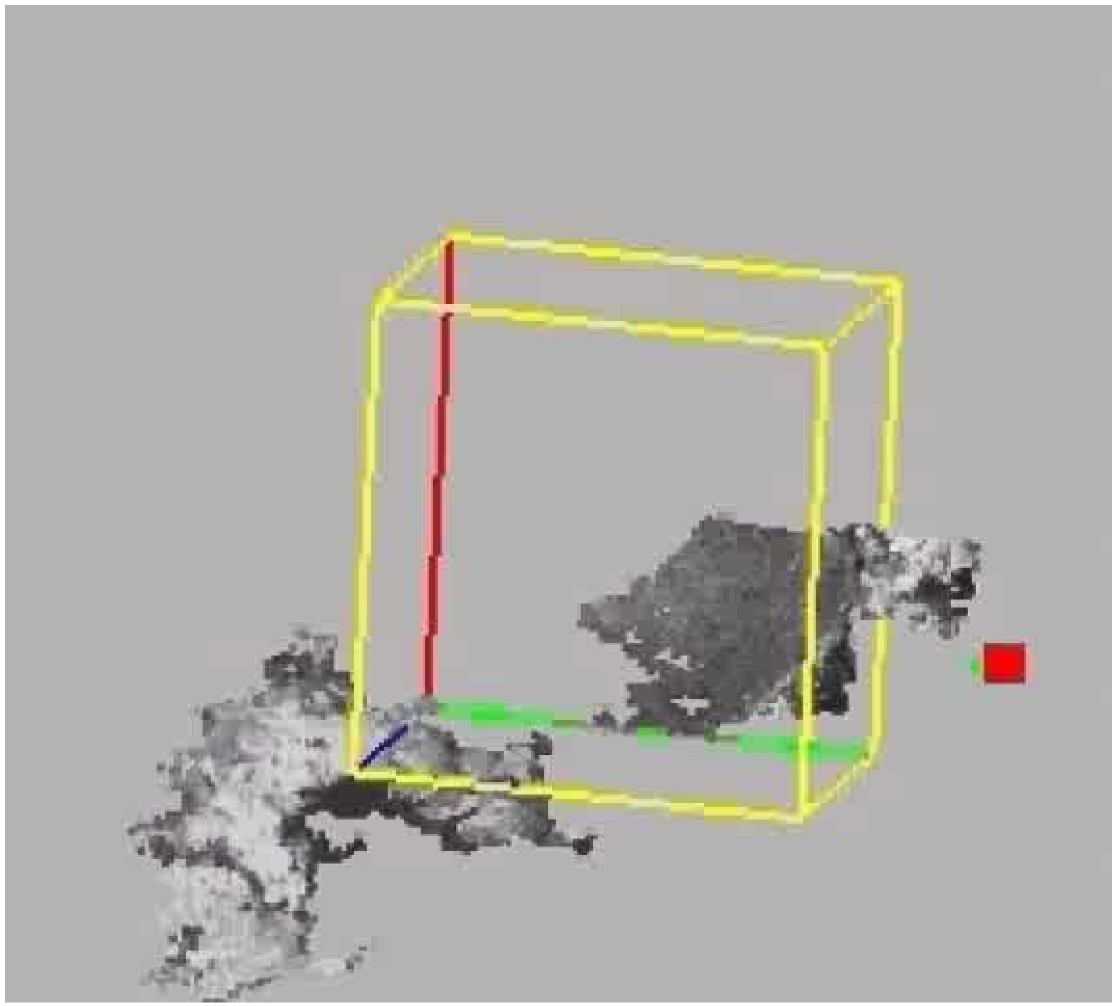


Example 1:
(Object Volumetric Modeling)

Data Collection

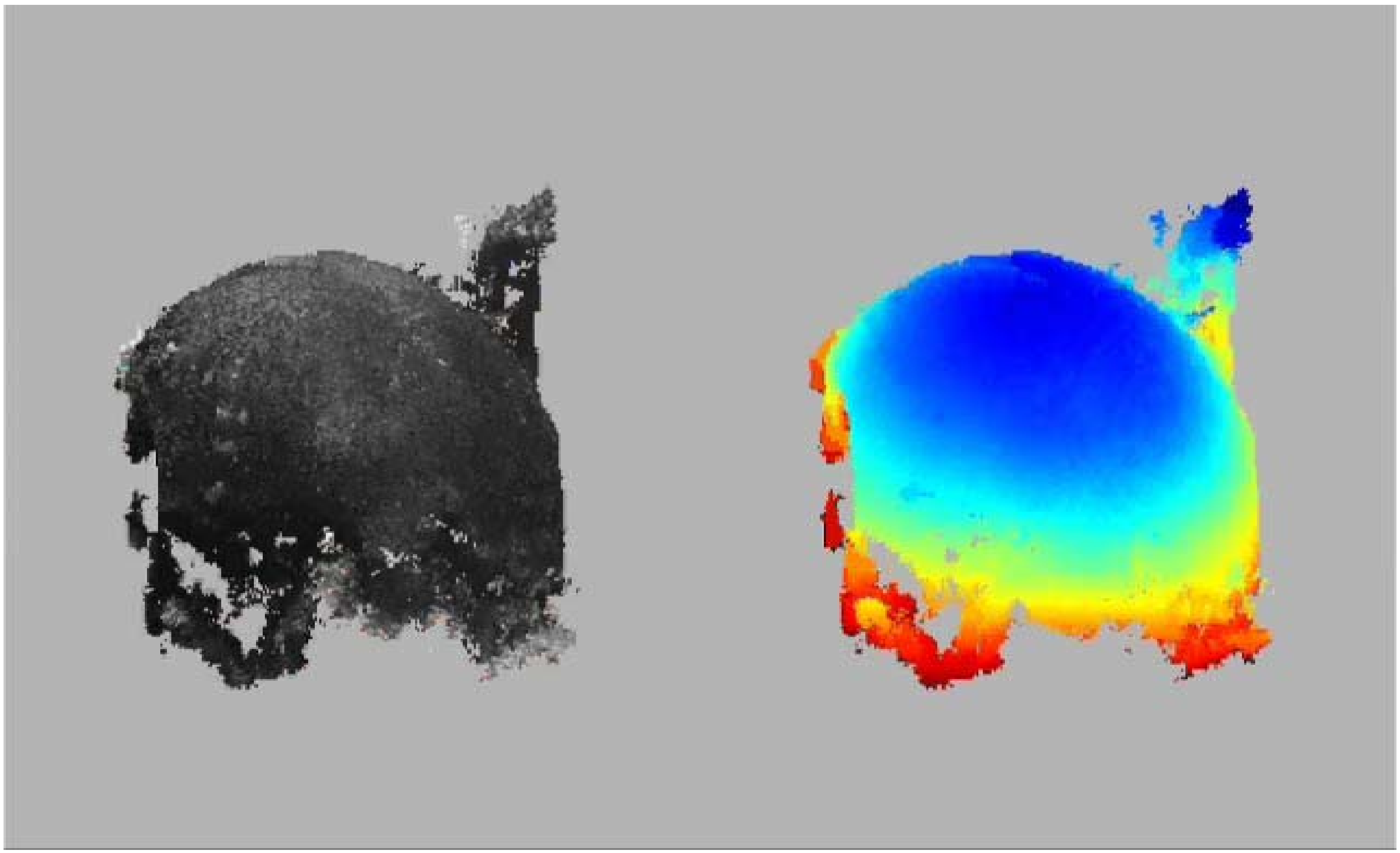


Trajectory & Object Reconstruction



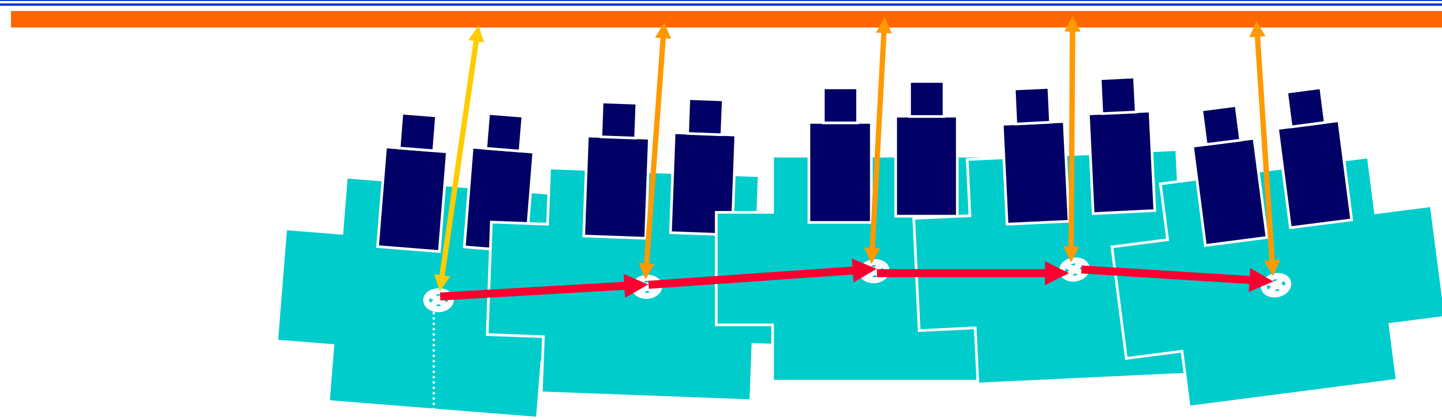
Example 1:
(Object Volumetric Modeling)

3D Model



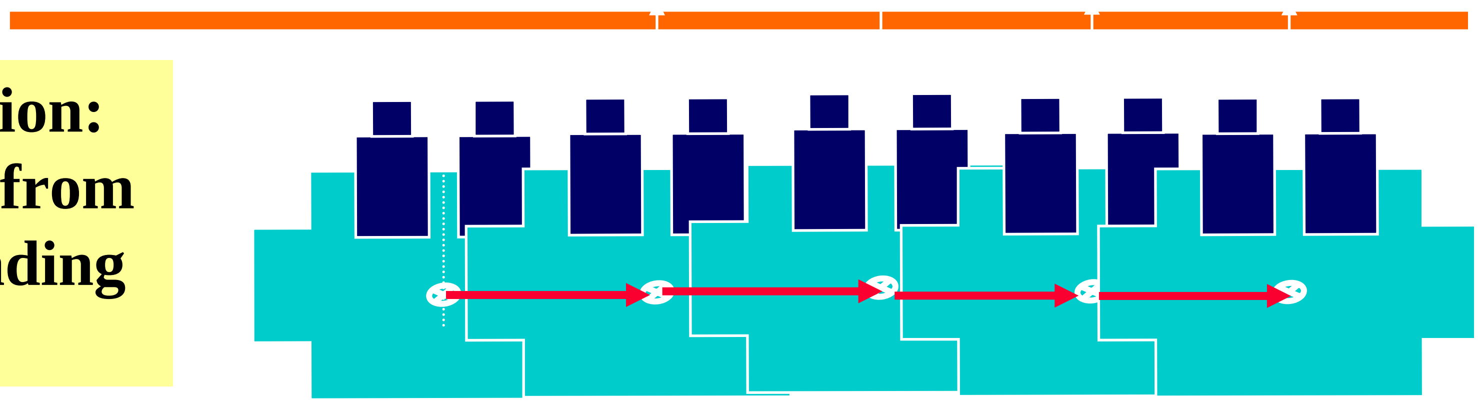
Example 2:

Shallow water Inspection and Target Mapping (Ship Hull Inspection 2003)



Actual vehicle motion

Video stabilization: Remove effects from roll motion, heading & scale change.



Computer generated video for mosaicing with 3 d.o.f in motion

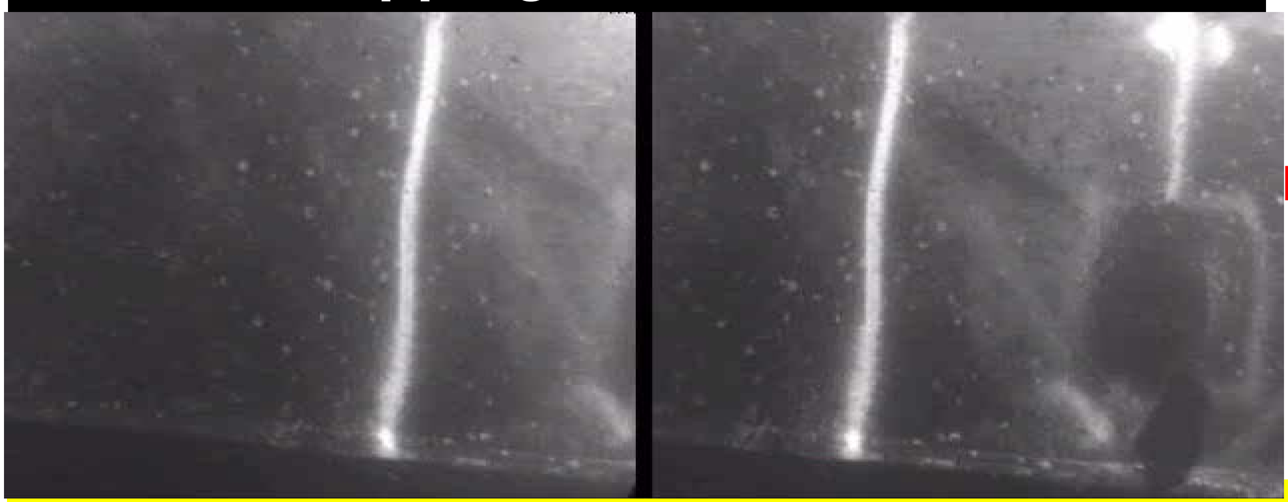
Shallow Water Inspection and Target Mapping (Ship Hull Inspection)

Ocean Demo; July 2003

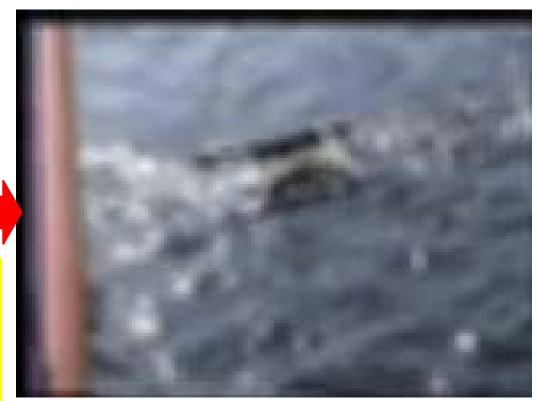
FAU SeaTech Center, Dania Beach, FL

Stereovision-based real-time system for

- **Automatic navigation**
- **2-D mosaicing**
- **3-D mapping**



Dock view



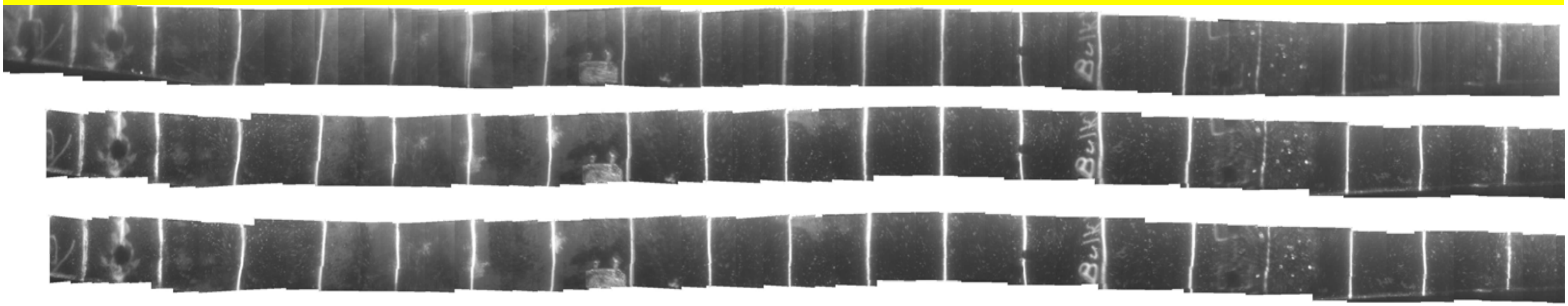
Deck view

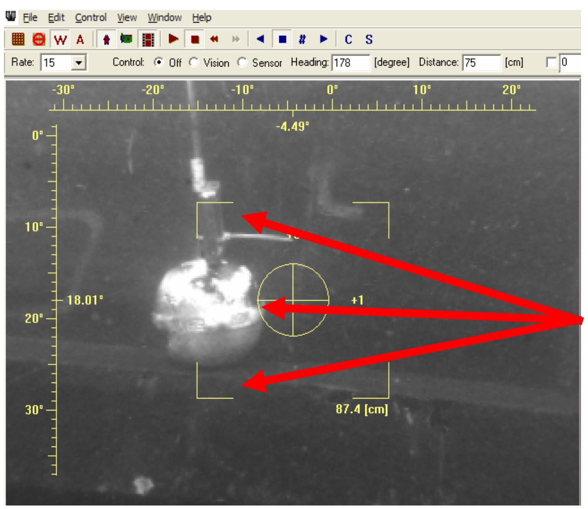
Automatic Navigation: **ROV position controlled by continuous acquisition & processing of stereo video images.**

Shallow water Inspection and Target Mapping

Ocean Demo; July 2003
FAU SeaTech Center, Dania Beach, FL

Real-time Mosaics from left-camera video
(*3 different runs verify consistency*)

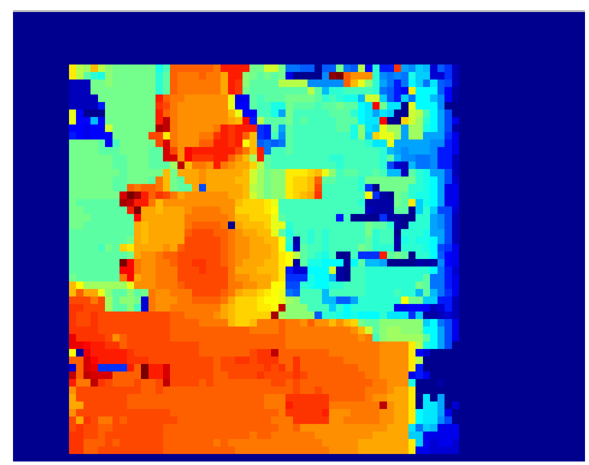




3-D Reconstruction

Sample left image and Color-coded 3-D map; (Red: close, Blue: far)

Hull, step & spherical ball placed on diving boat step detectable in 3-D map.



Application of Motion Video Processing to Sonar Imagery

- Unfortunately, visibility range limits optical imaging;
- Sonar penetrate through turbidity, and high-frequency sonar imaging systems produce images with high details
- Can we develop *same vision-based capabilities with 2-D sonar cameras* (DIDSON, BlueView) ? *How effectively*?

  -- Constructing 2-D map of targets (Photo-mosaicing);

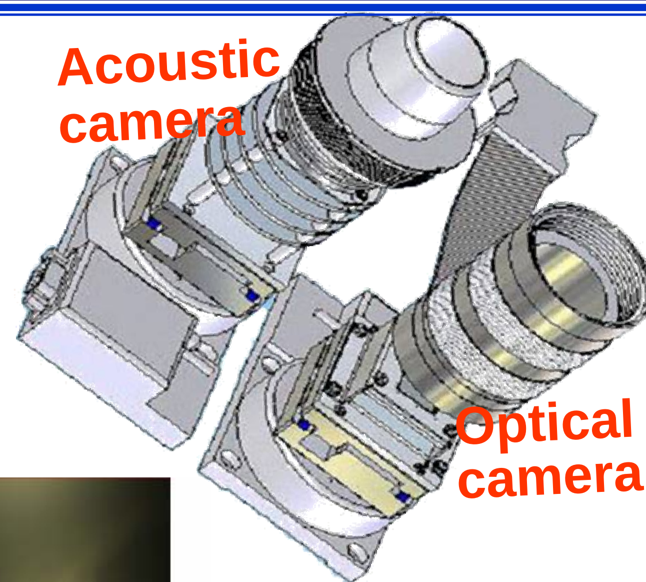
  -- Determining motion of sensor platform relative to target (Target-based positioning);

  -- Determining 3-D position of points on target surfaces (3-D reconstruction from stereo imagery).

Answer: **Need to integrate with auxiliary positioning systems**

Opti-Acoustic Stereo Imaging; New Paradigm for Underwater 3-D Target Mapping

New Paradigm: Can we achieve improved performance by integrating information from optical and sonar cameras?



Good Optical Visibility

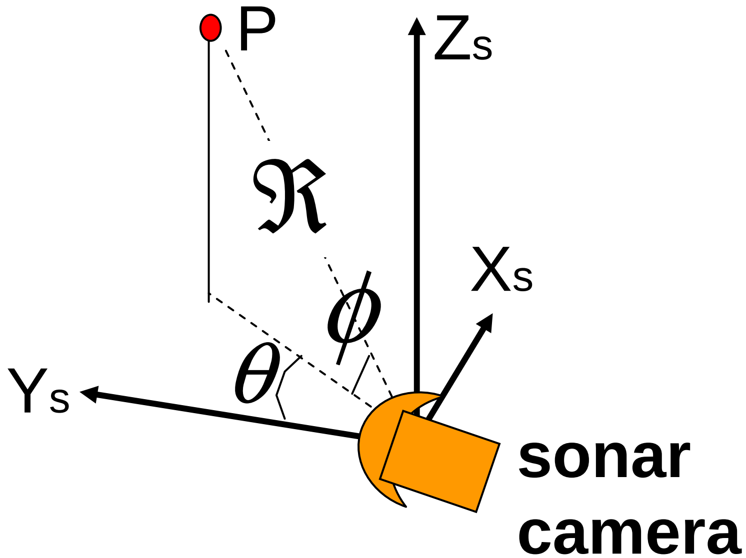
Both cameras provide valuable information

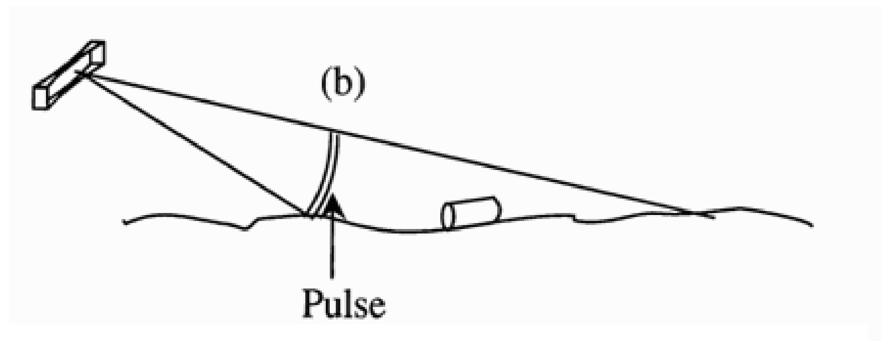
Poor Optical Visibility

Rely solely on sonar camera

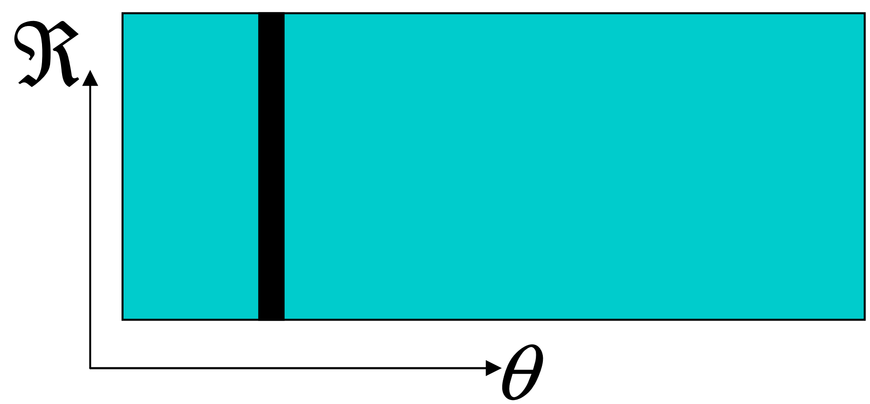
Currently exploring sensor fusion in the context of *Opti-Acoustic Stereo Imaging*

2-D Imaging Sonar

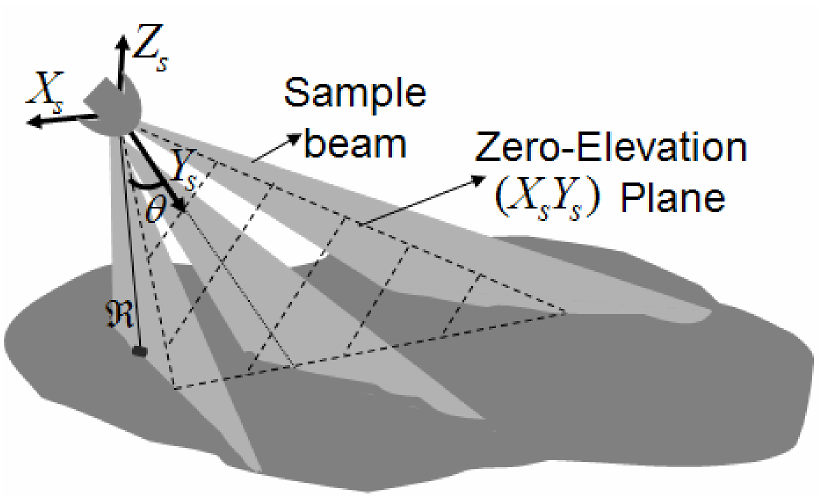


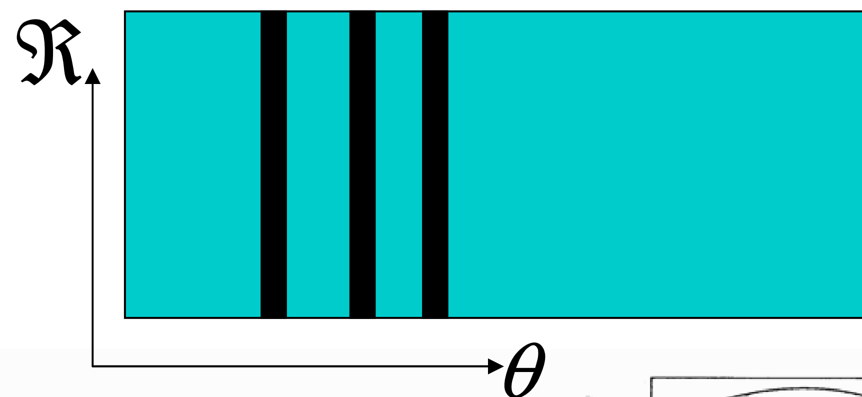


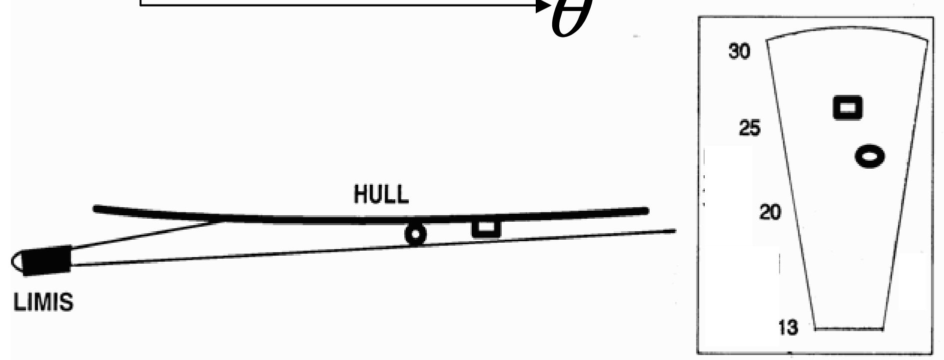
Scene "illumination" for particular (azimuth) direction



2-D Imaging Sonar





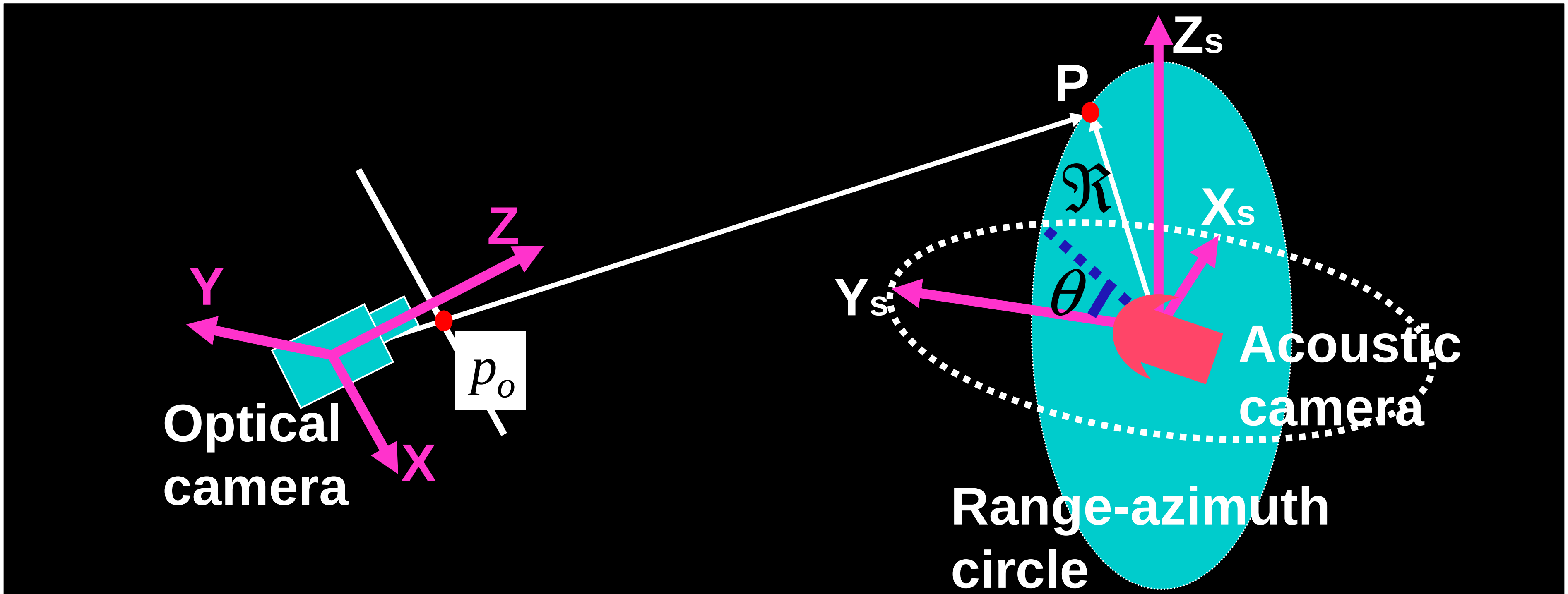


Target illumination and the image

- **Each array of sonar measurements correspond to target scene ranges** $\Re$ **for a particular bearing (azimuth direction)** θ **;**

- **2-D Image** $I(\Re, \theta)$ **is constructed from range measurements at different bearing directions.**

Opti-Acoustic Stereo Imaging

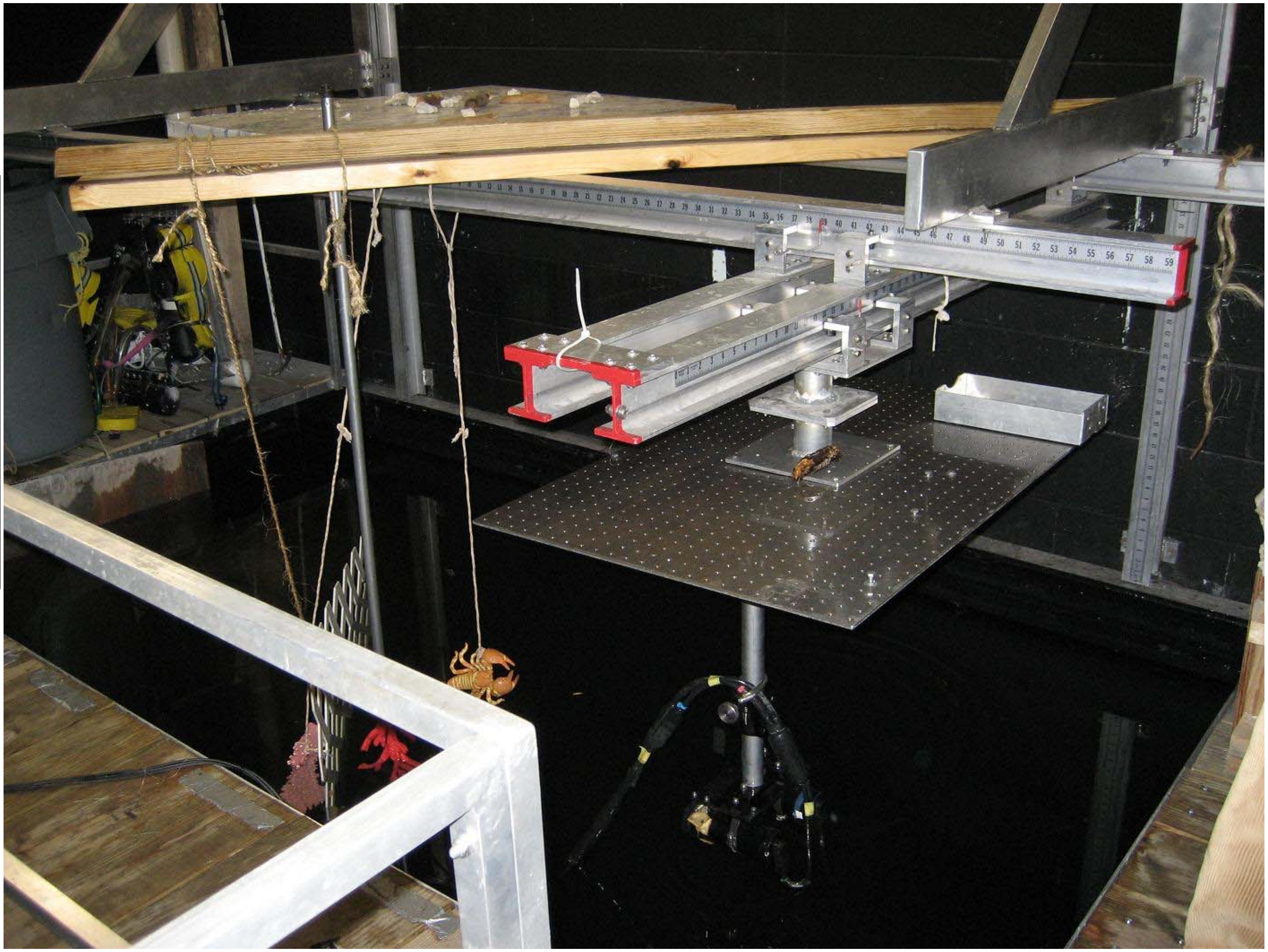


Opti-acoustic stereo pair: Projections of a scene feature onto optical and sonar views in the form of ***opti-acoustic correspondences:***

$$\begin{cases} p_o = (x, y) \\ p_s = (\Re, \theta) \end{cases} \rightarrow \begin{cases} p_o = (x, y) \\ p_s = (\Re \sin\theta, \Re \cos\theta) \end{cases}$$

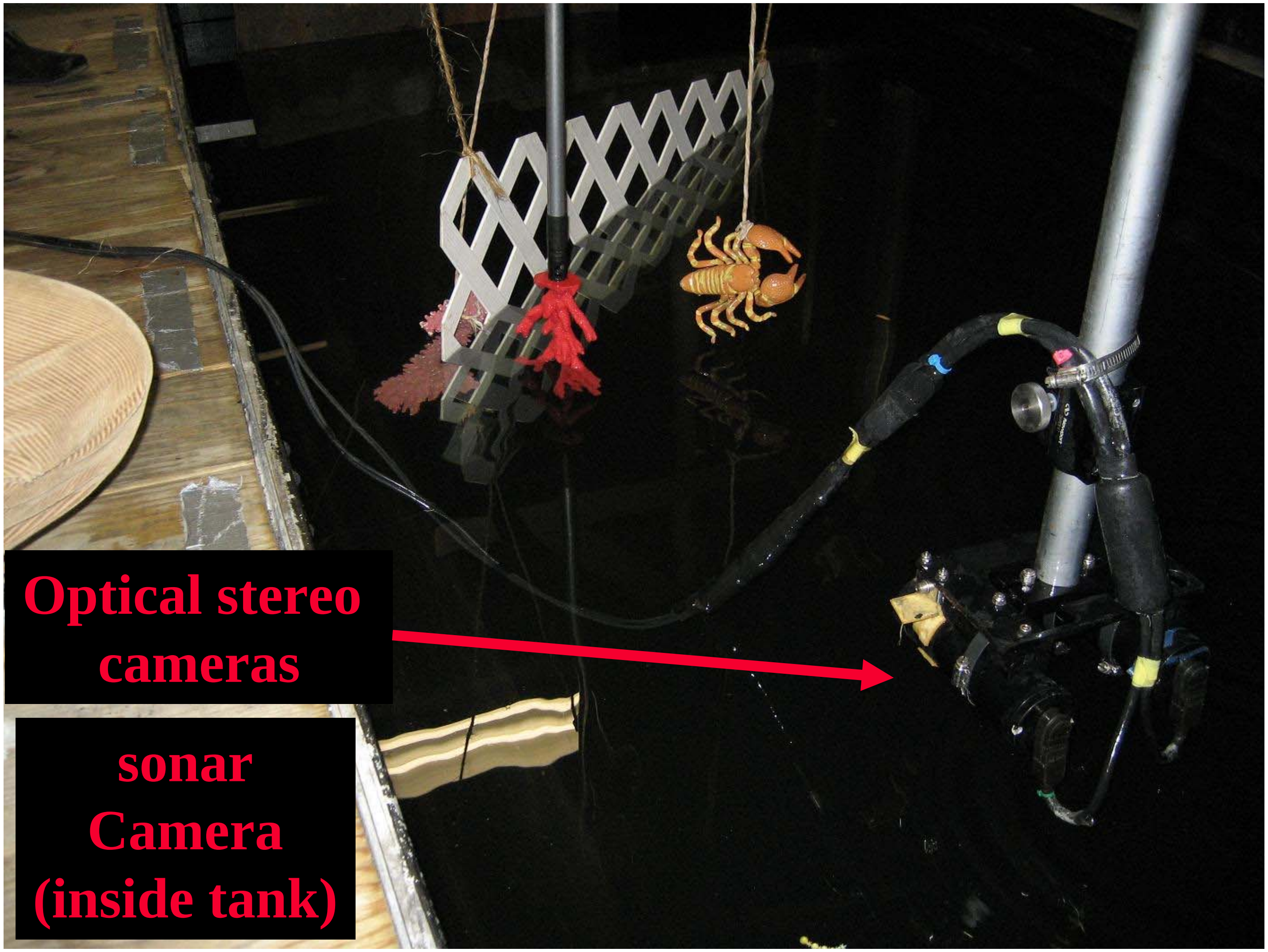
Opti-Acoustic Stereo Imaging

Water tank set for controlled underwater experiments

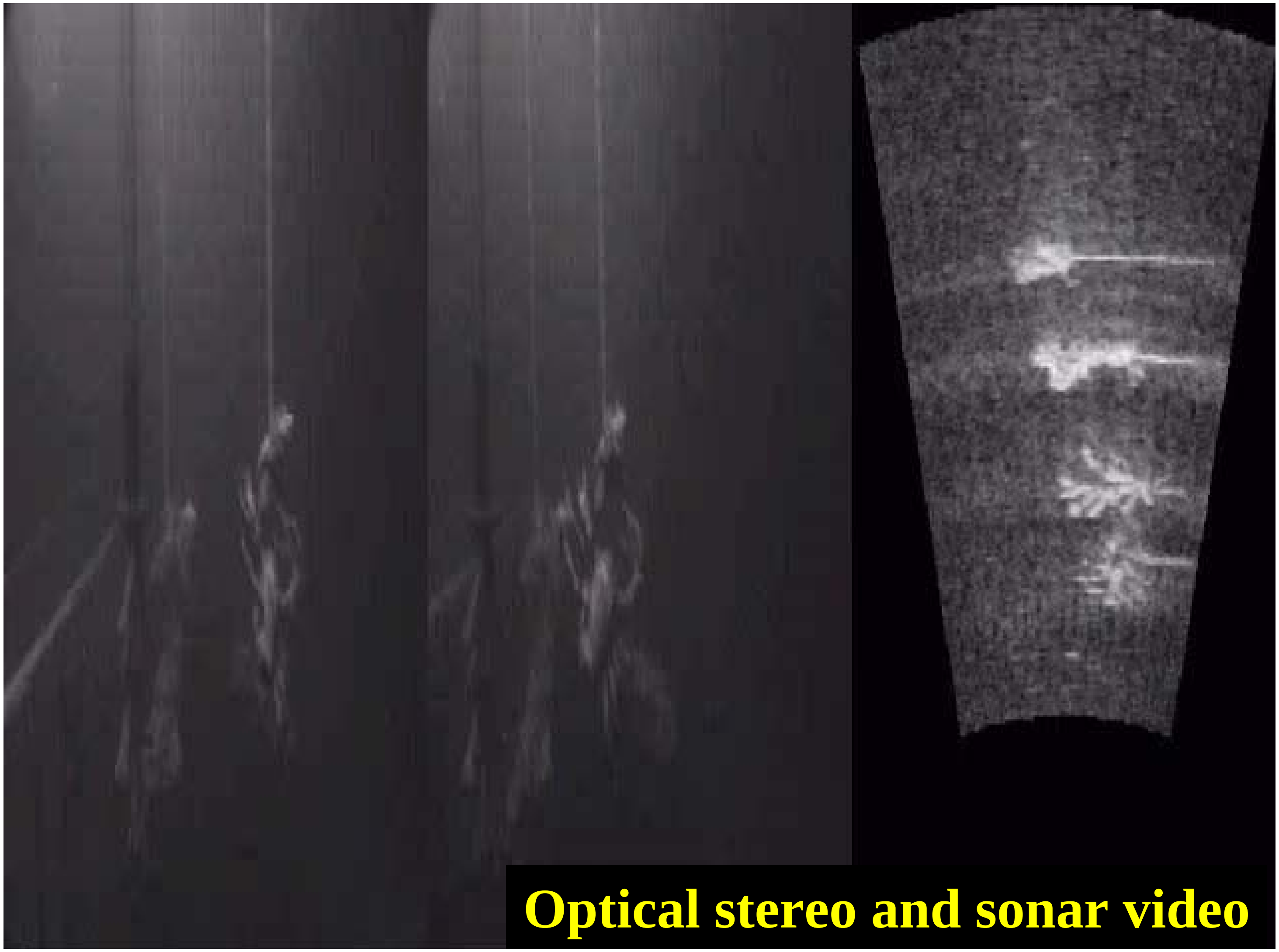


Opti-Acoustic Stereo Imaging

Water tank set up and targets

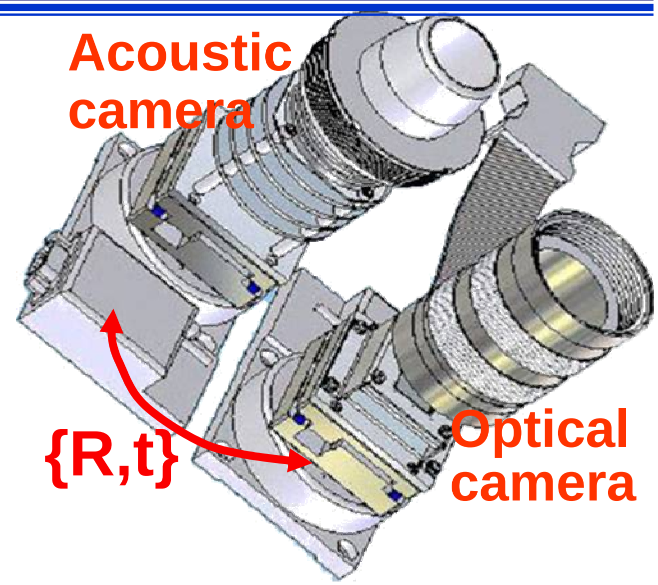


Opti-Acoustic Stereo Imaging



Optical stereo and sonar video

Opti-Acoustic Stereo Imaging; New Paradigm for Underwater 3-D Target Mapping



Epipolar Geometry of Opti-Acoustic Stereo Imaging

- For ***an acoustic-image point s***, what is epipolar contour in optical image?

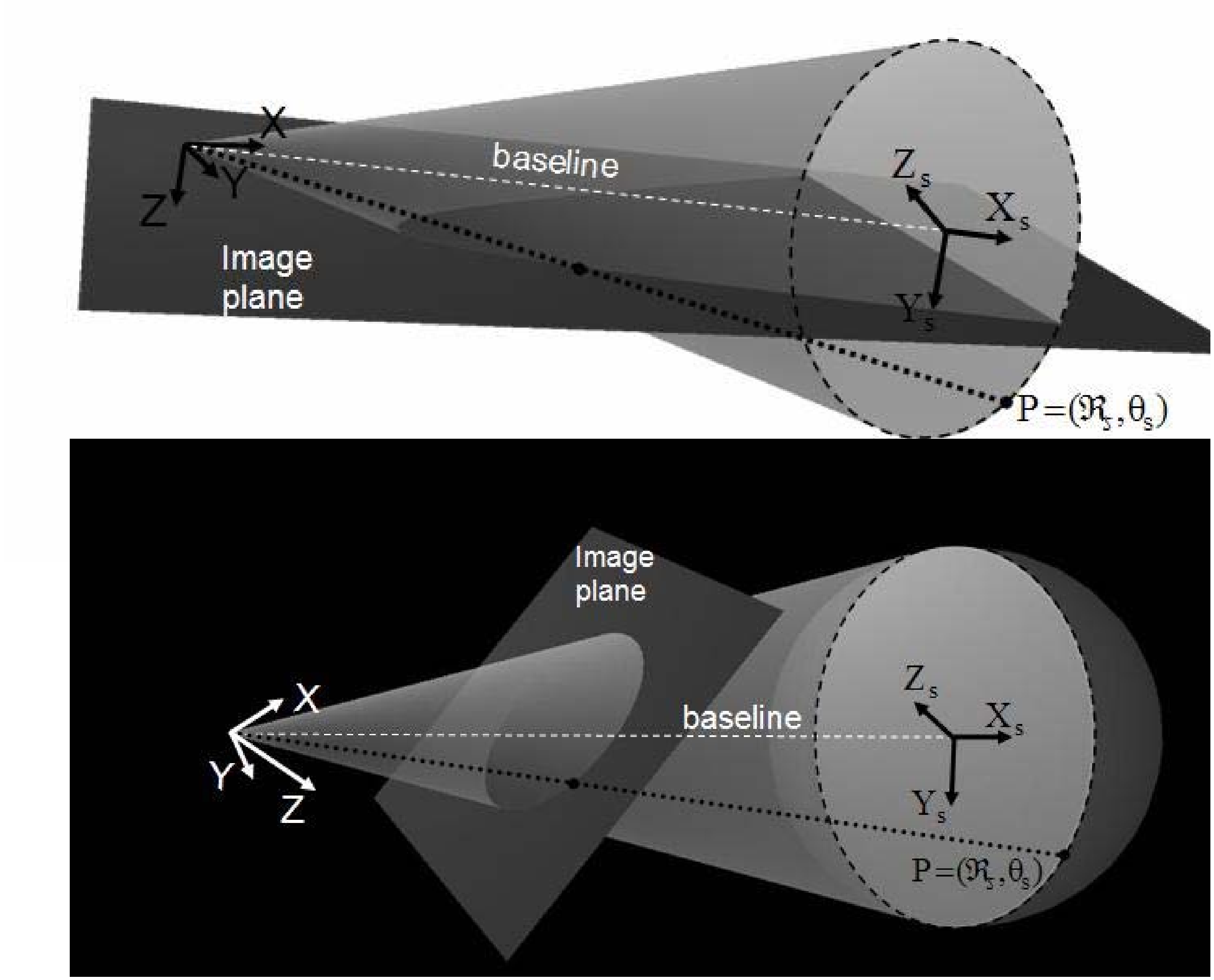
$$\mathrm{p}^{\mathrm{T}} U(s)\mathrm{p} = 0;$$

$$U(\Re,\theta) = \left(\frac{(\tan\theta\, t_y - t_x)^2}{t.t - \Re^2}\right) I + (r_1 - \tan\theta\, r_2)(r_1 - \tan\theta\, r_2)^T +$$

$$\left(\frac{\tan\theta\, t_y - t_x}{t.t - \Re^2}\right)\left((r_1 - \tan\theta\, r_2)t^T R + R^T t(r_1 - \tan\theta\, r_2)^T\right)$$

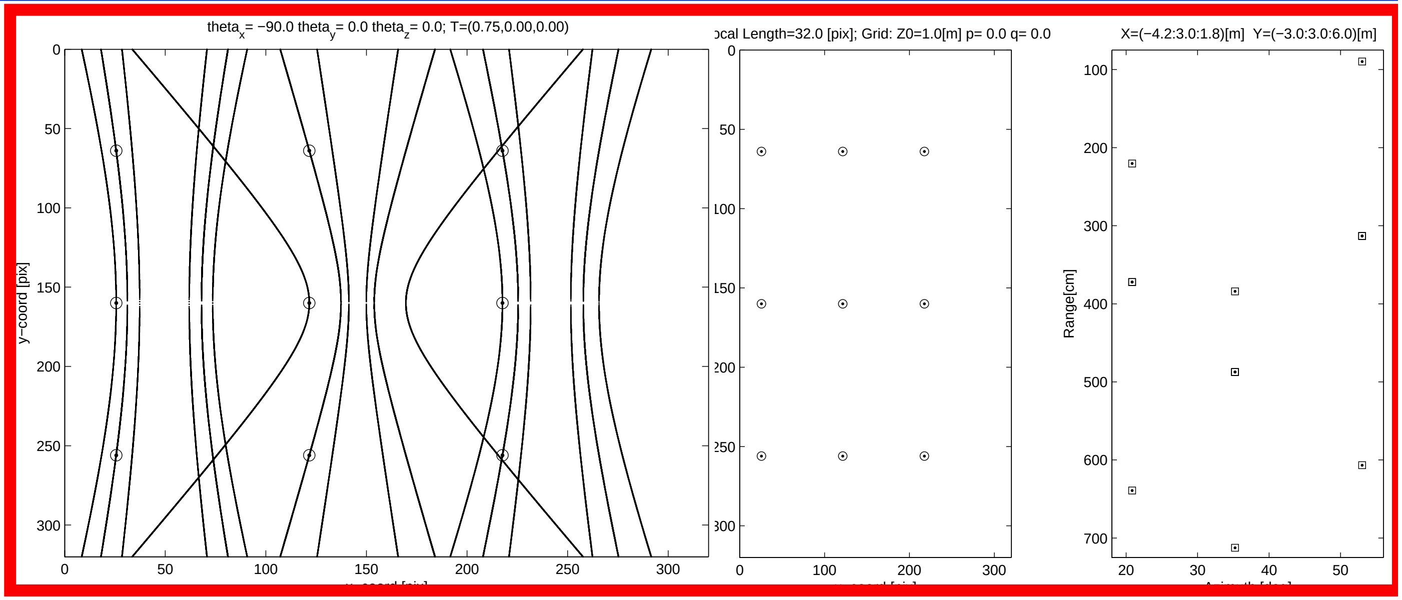
Opti-Acoustic Stereo Epipolar Geometry

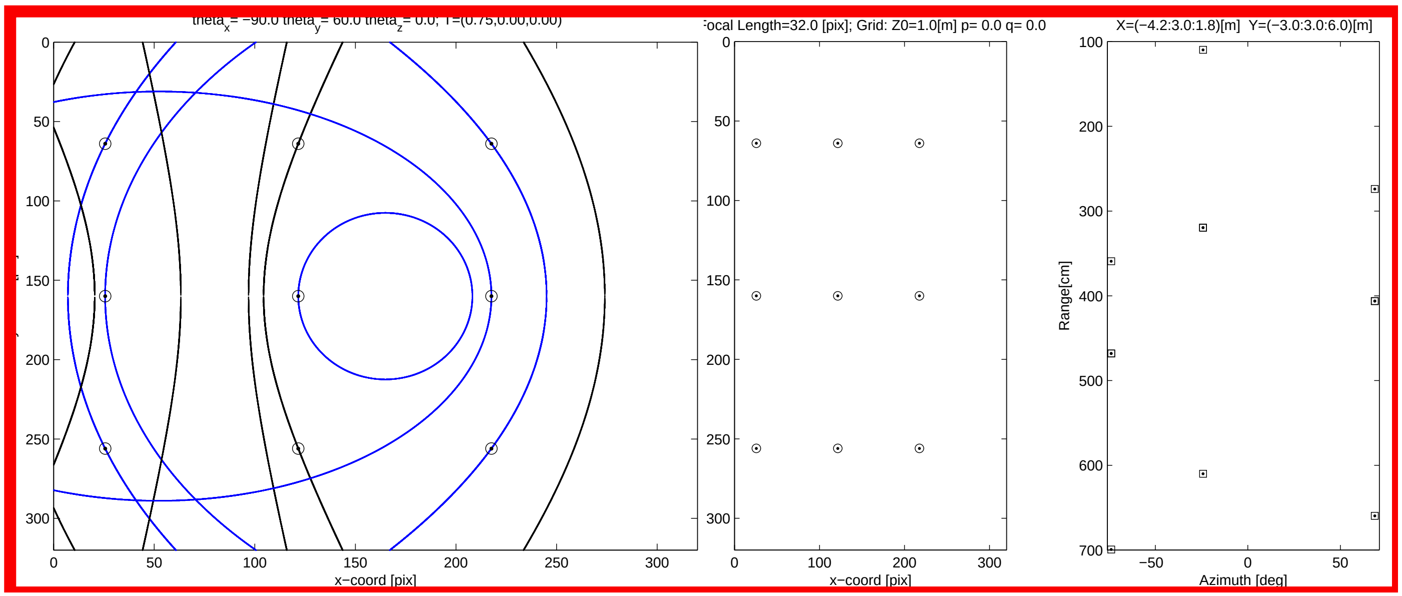
Epipolar contours in optical image are quadrics: Either hyperbolas or ellipses.



Opti-Acoustic Stereo Imaging; New Paradigm for Underwater 3-D Target Mapping

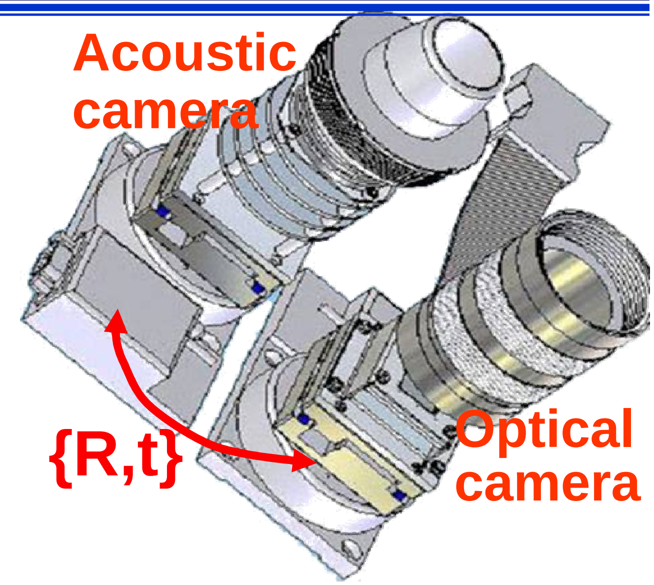
Epipolar Geometry of Opti-Acoustic Stereo Imaging



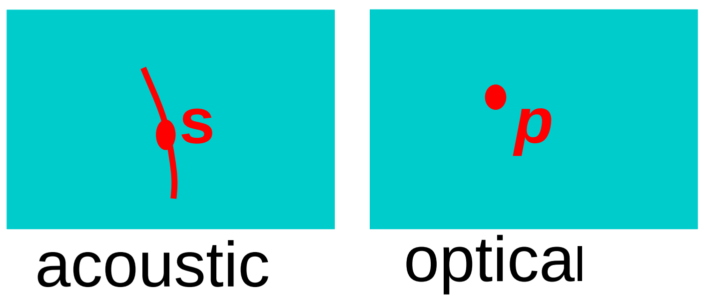


Opti-Acoustic Stereo Imaging; New Paradigm for Underwater 3-D Target Mapping

Epipolar Geometry of Opti-Acoustic Stereo Imaging



- For every **optical-image point *p***, what is epipolar contour in sonar image?



$$\Re = \sqrt{N(\theta) / D(\theta)}$$

$$N(\theta) = (u_{31}\sigma_2 - u_{32}\sigma_1)^2 + ((u_{12}\sigma_1 - u_{11}\sigma_2)\sin\theta + (u_{22}\sigma_1 - u_{21}\sigma_2)\cos\theta)^2$$

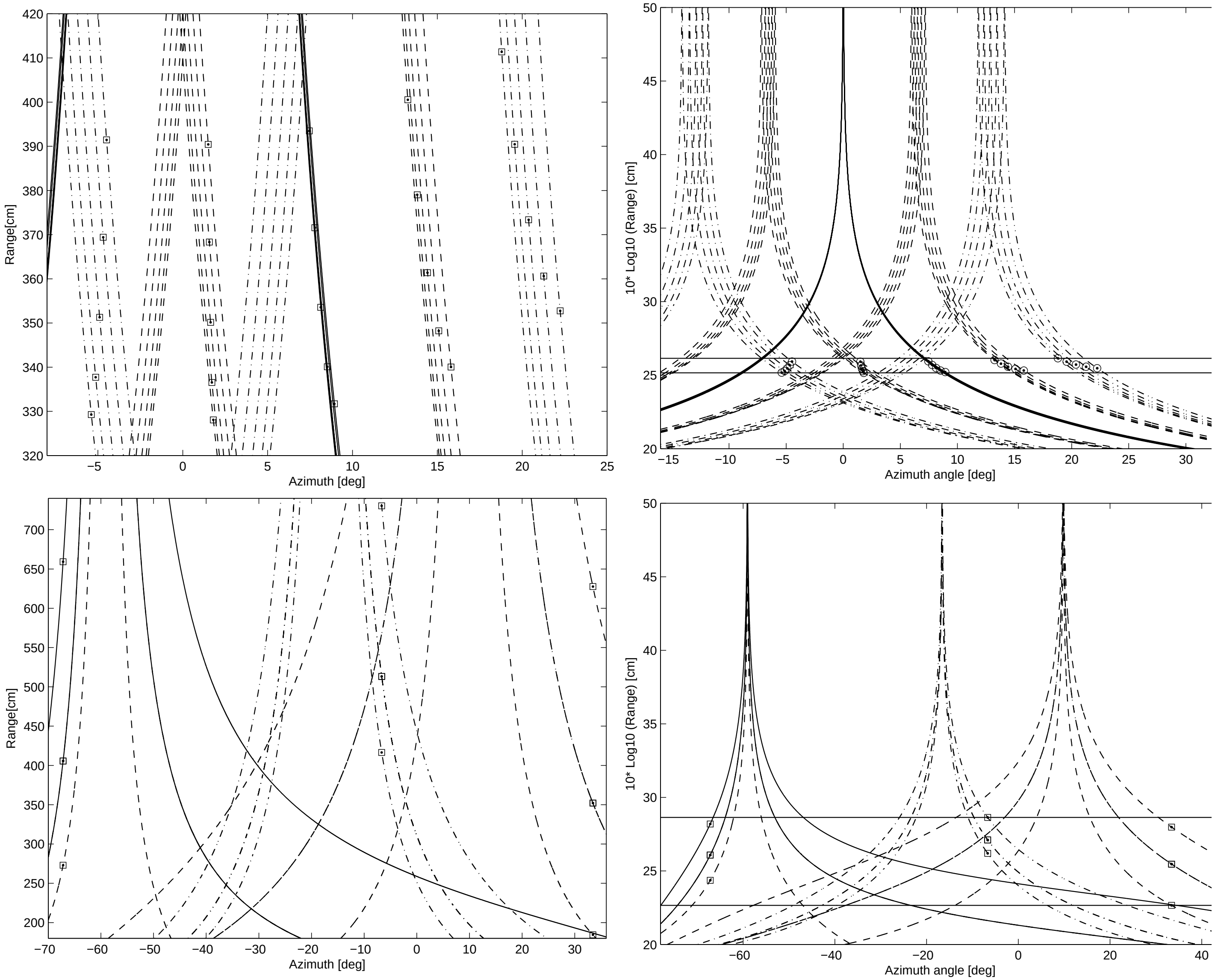
$$D(\theta) = ((u_{31}u_{12} - u_{32}u_{11})\sin\theta + (u_{31}u_{22} - u_{32}u_{21})\cos\theta)^2$$

$$u_{k1} = r_{3k}y - r_{2k} \qquad u_{k2} = r_{3k}x - r_{1k}$$

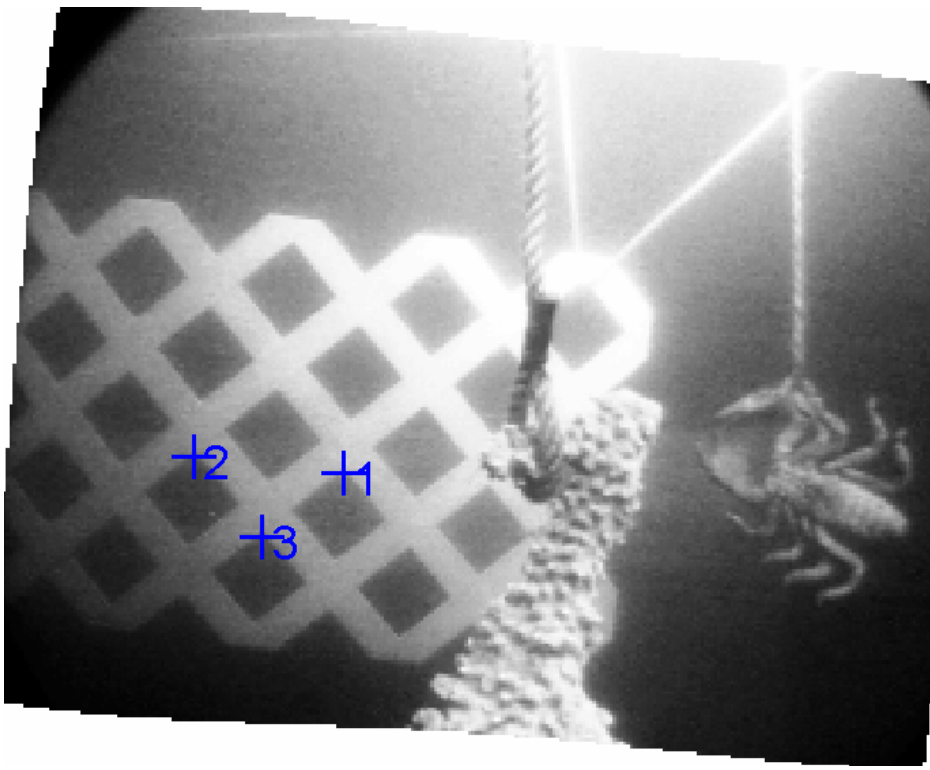
$$\sigma_i = t_x u_{1i} + t_y u_{2i} + t_z u_{3i}$$

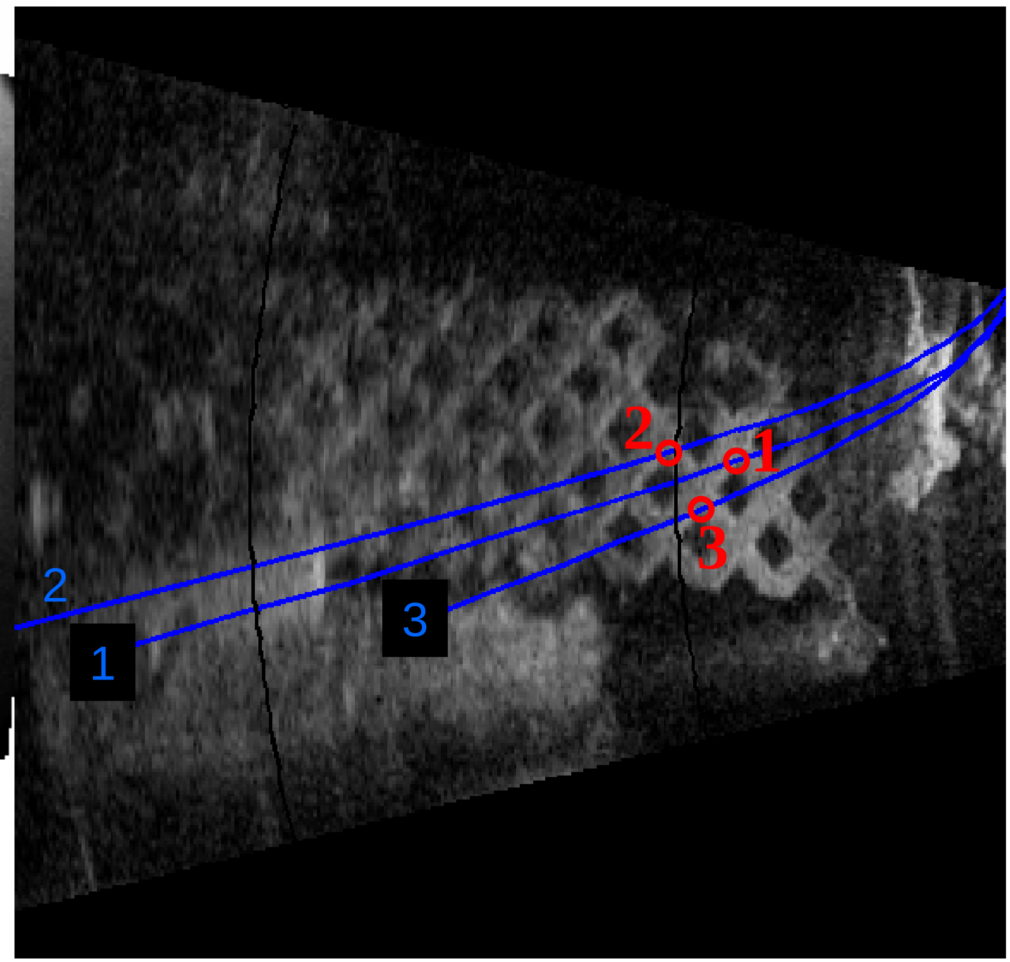
Opti-Acoustic Stereo Imaging; New Paradigm for Underwater 3-D Target Mapping

Epipolar Geometry of Opti-Acoustic Stereo Imaging



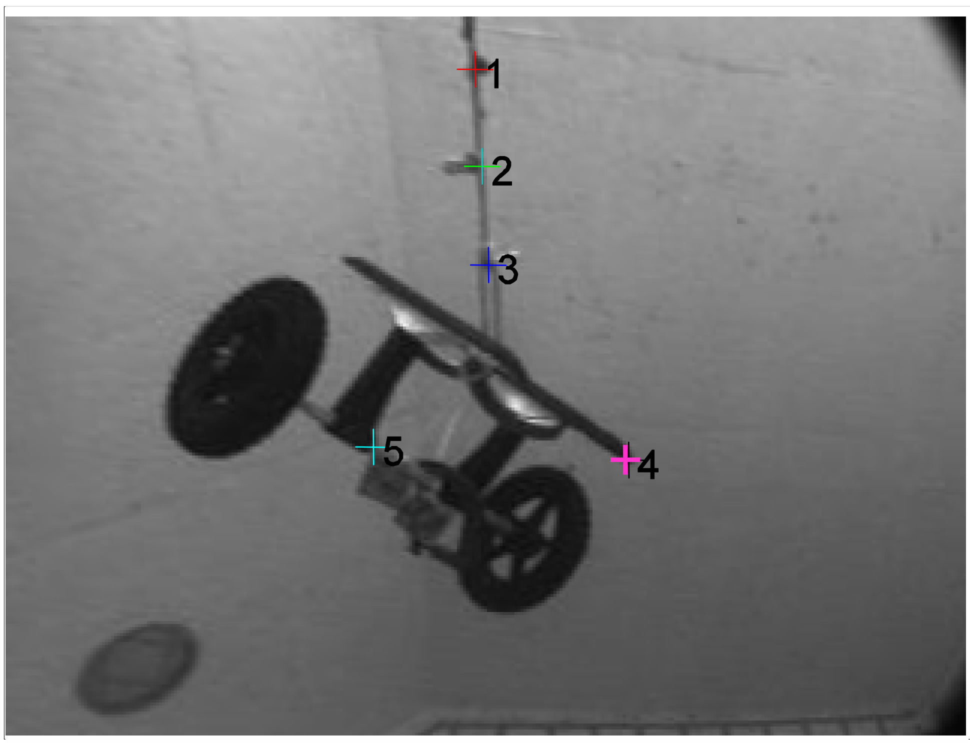
Opti-Acoustic Stereo Epipolar Geometry

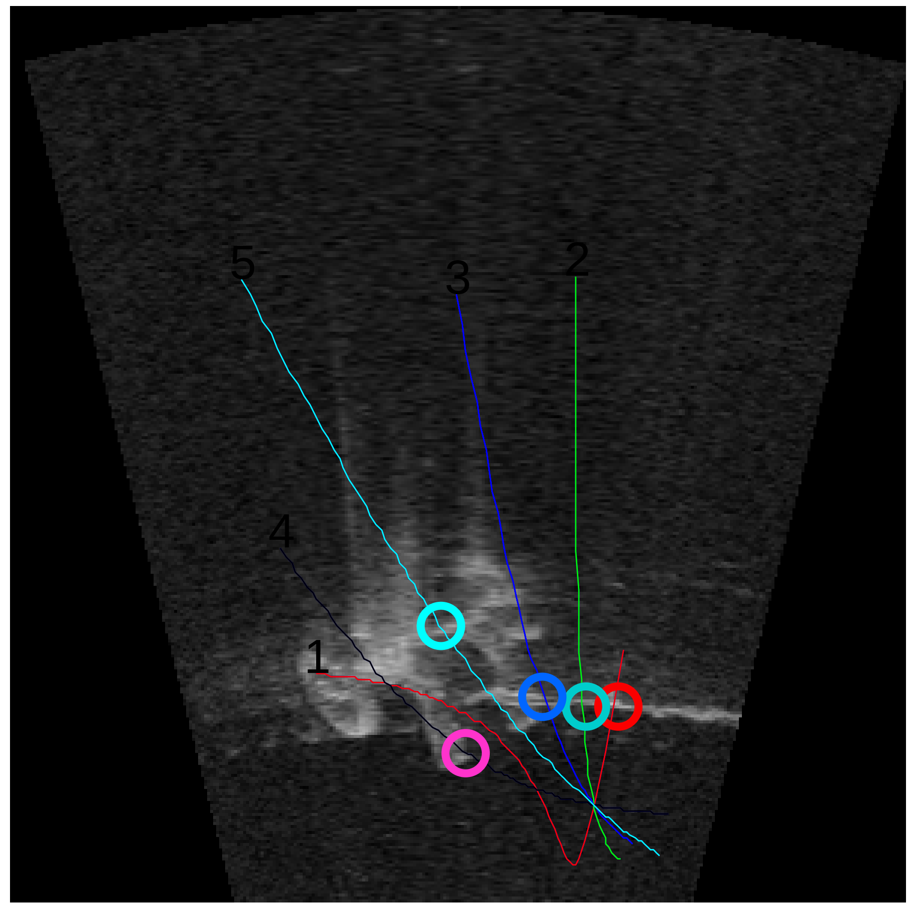




**water tank data
January 2007**

Opti-Acoustic Stereo Epipolar Geometry





Teledyne Benthos pool facility
(N. Falmouth, MA)
November'06

Epipolar Geometry of Opti-Acoustic Stereo Imaging

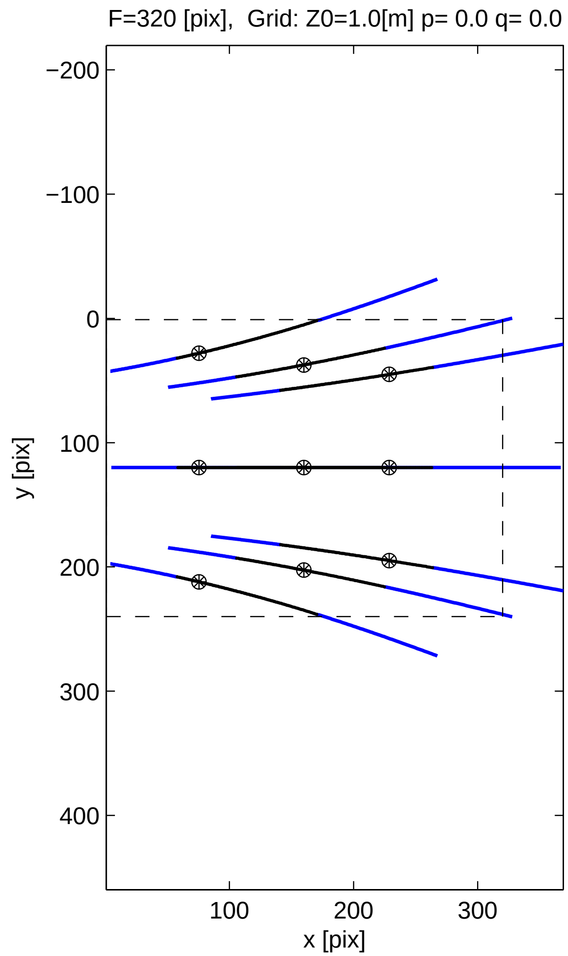
Visibility Constraint:

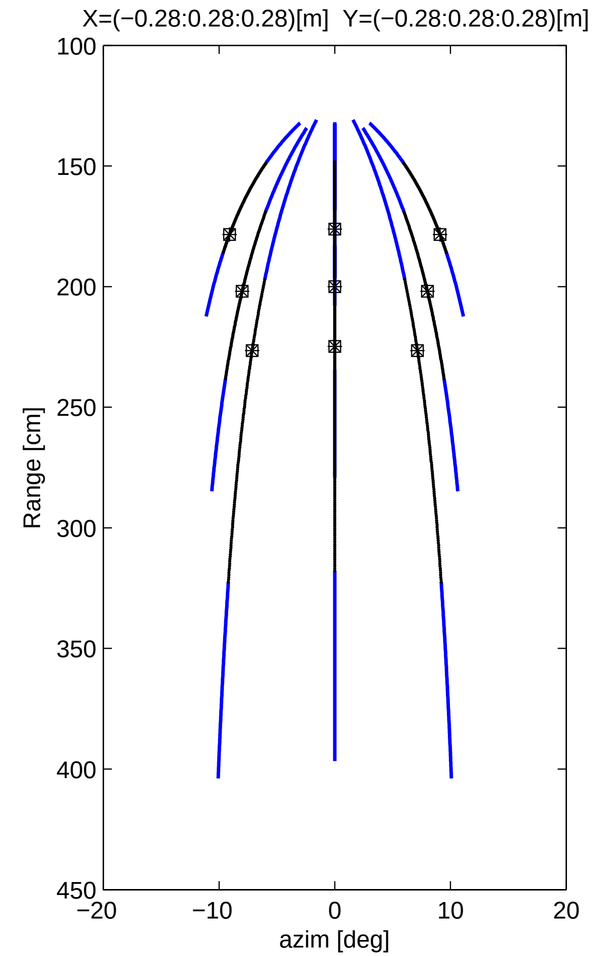
Each acoustic beam insonifies the scene within small sector in elevation direction:

DIDSON & BlueView: ±7[deg]

Black: elevation ±7[deg]

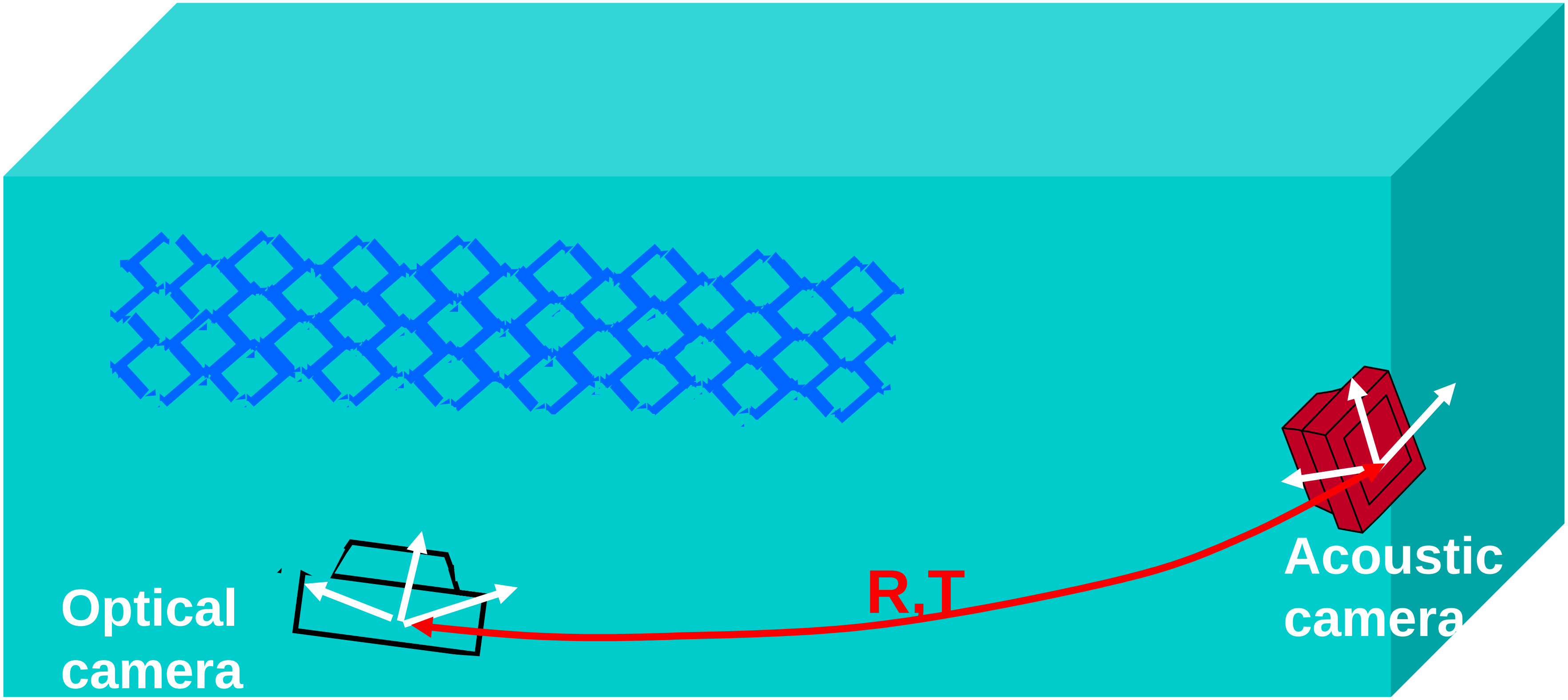
Blue: elevation ±12[deg]





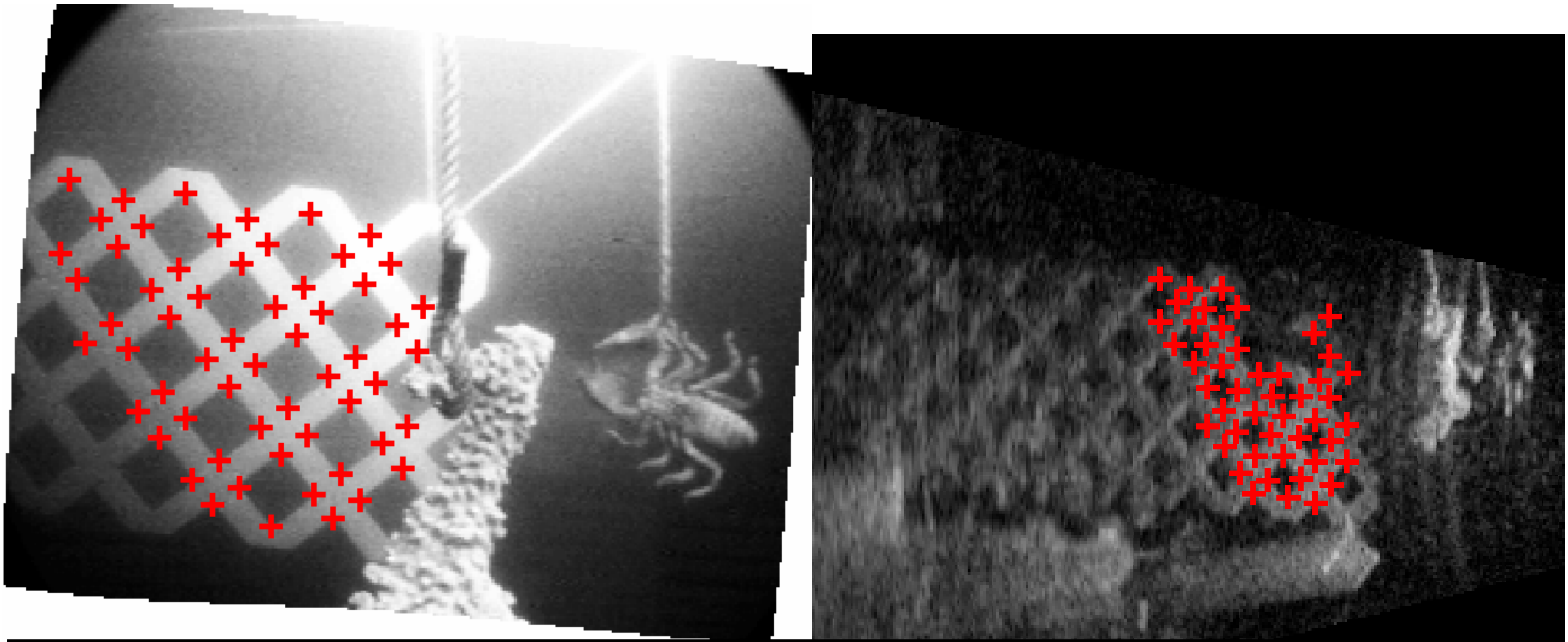
Opti-Acoustic Stereo Imaging; New Paradigm for Underwater 3-D Target Mapping

Stereo Calibration



Opti-Acoustic Stereo Imaging; New Paradigm for Underwater 3-D Target Mapping

Stereo Calibration

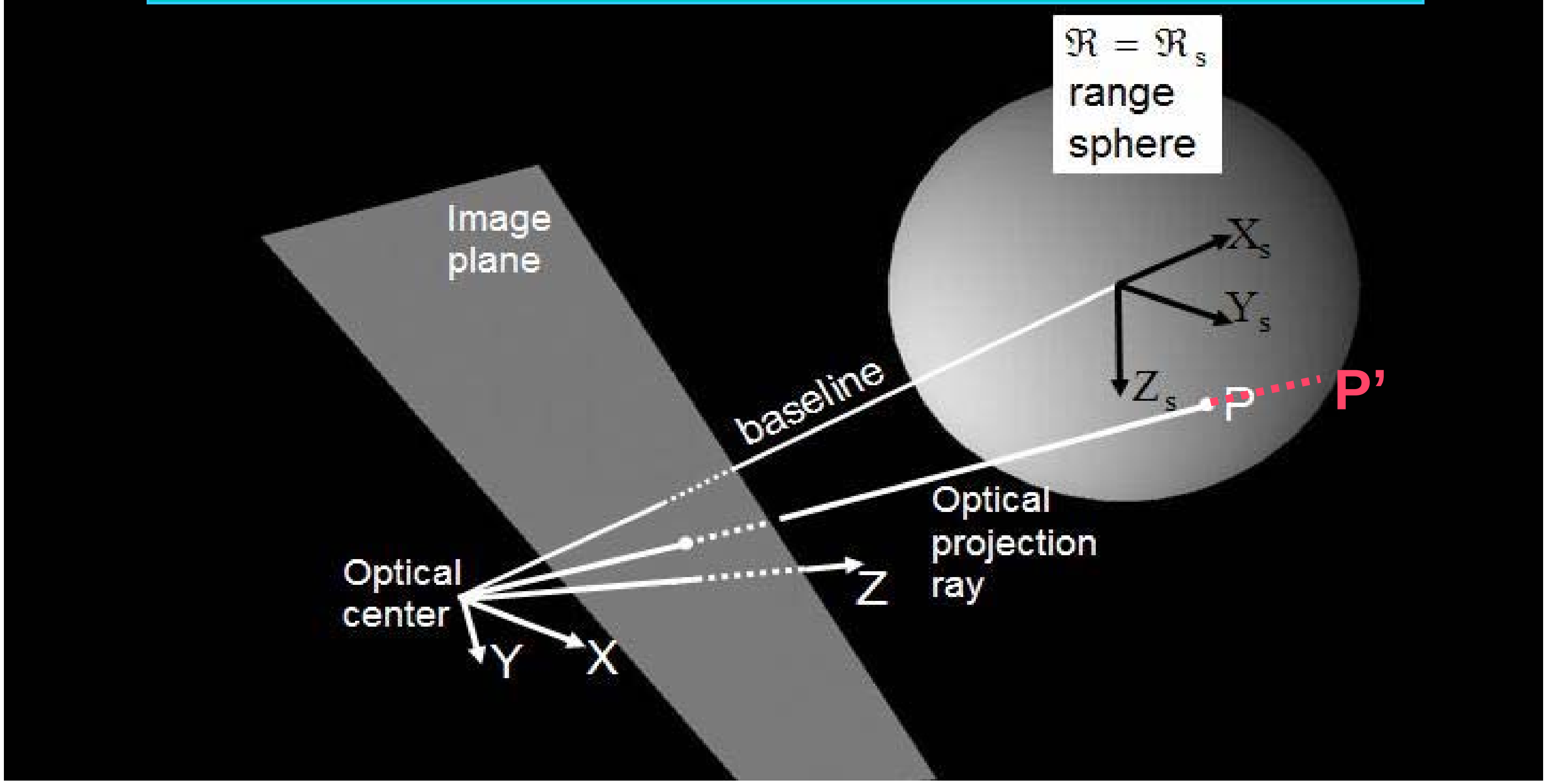


water tank data
January 2007

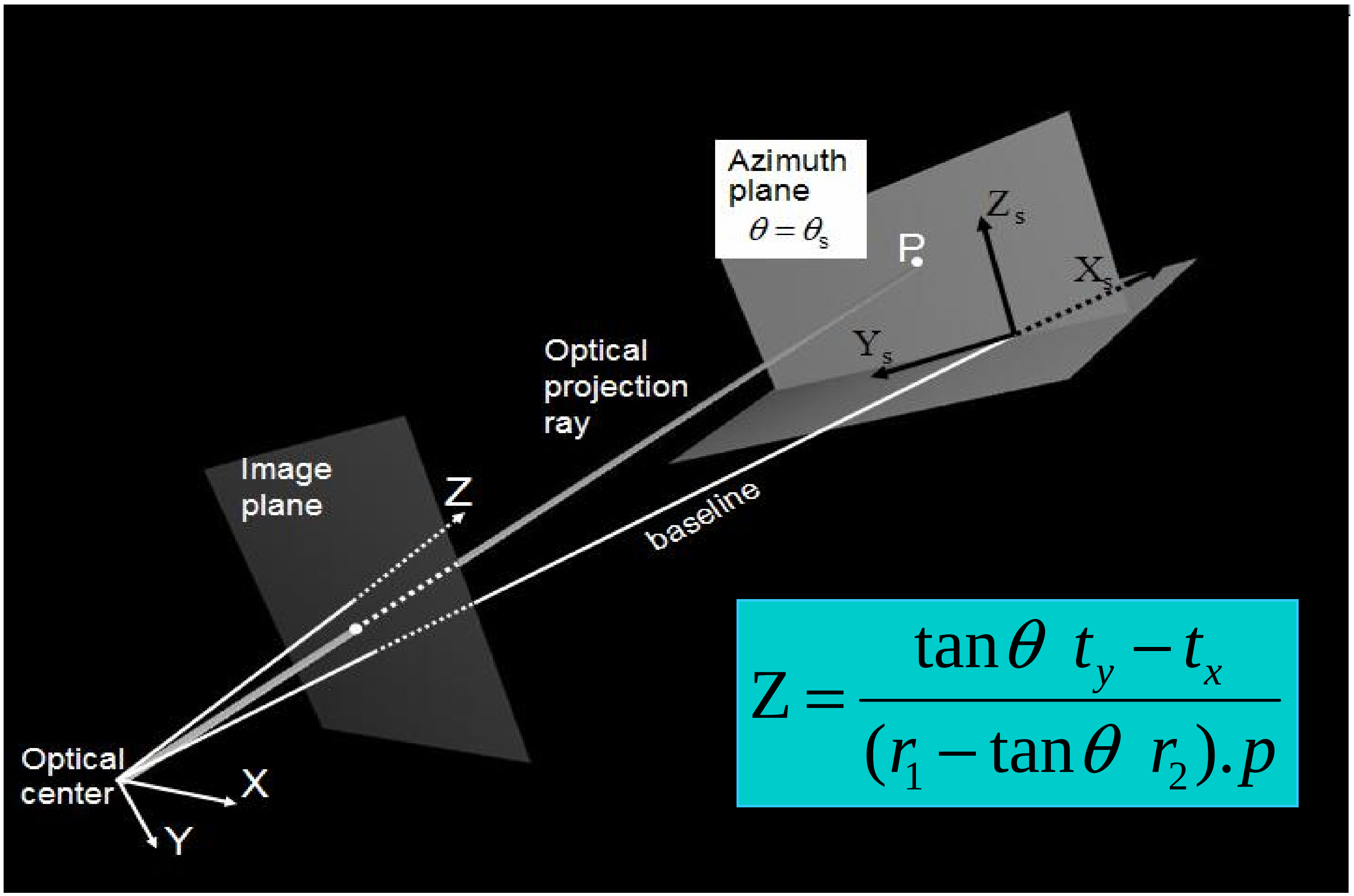
Opti-Acoustic Stereo Imaging; New Paradigm for Underwater 3-D Target Mapping

Triangulation: “optical ray” and “range sphere”

$$Z^2 + 2(t^T R p_o) Z - (\Re^2 - (t.t)) = 0$$



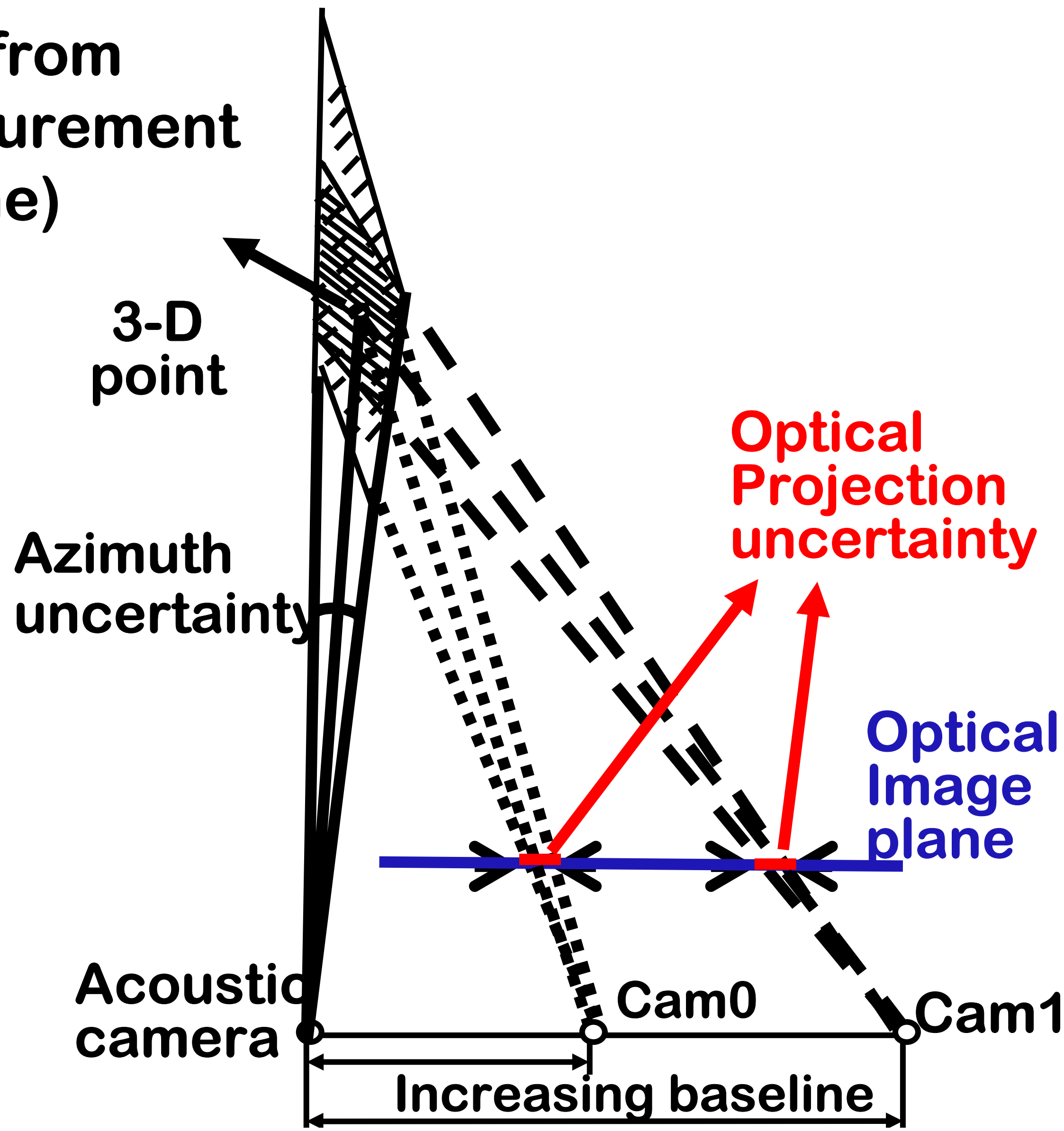
Triangulation: “optical ray” with “azimuth plane”



$$Z = \frac{\tan\theta \; t_y - t_x}{(r_1 - \tan\theta \; r_2).p}$$

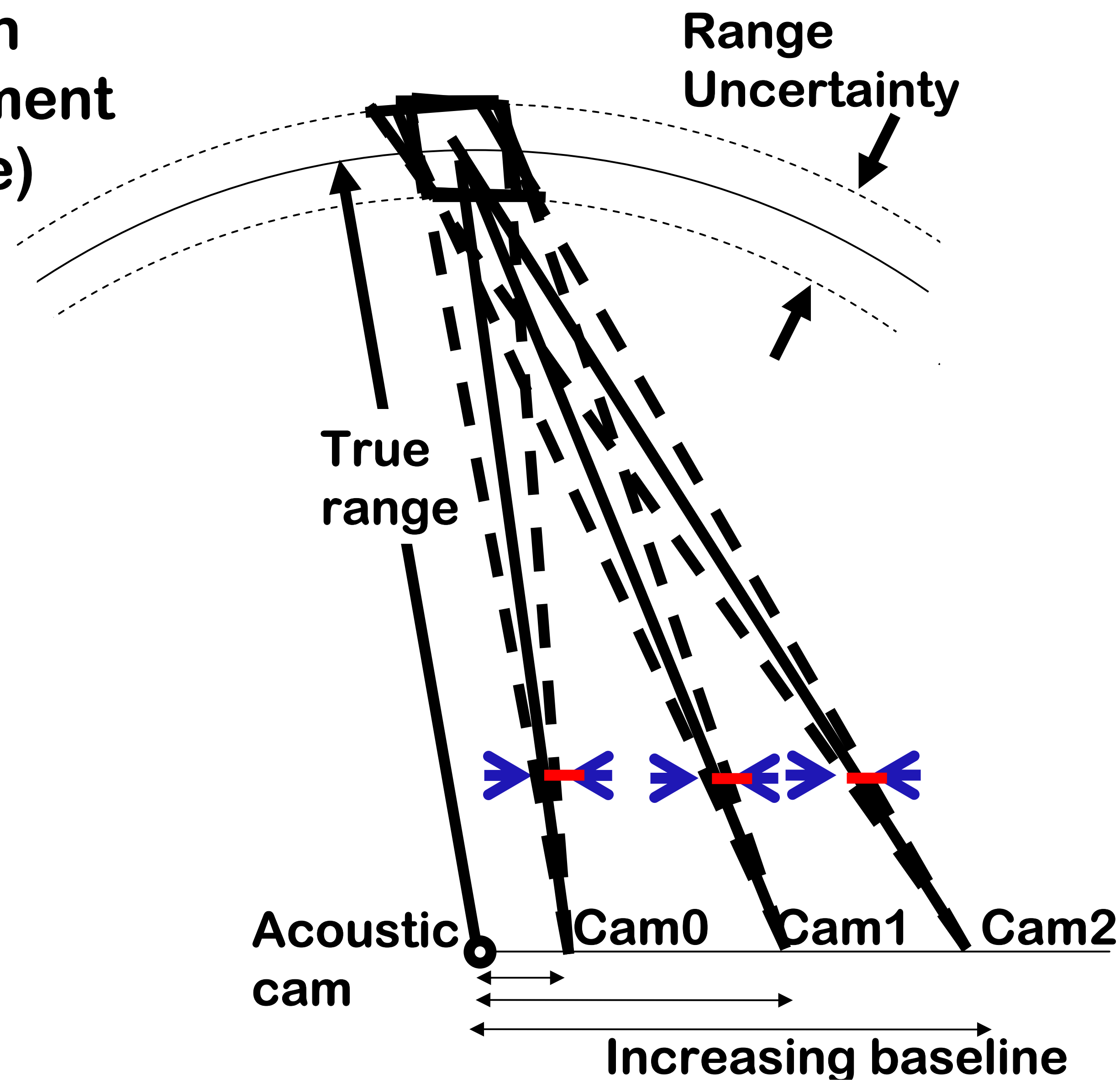
Opti-Acoustic Stereo Imaging; New Paradigm for Underwater 3-D Target Mapping

3-D reconstruction from bearing (azimuth) measurement (Impact of baseline)

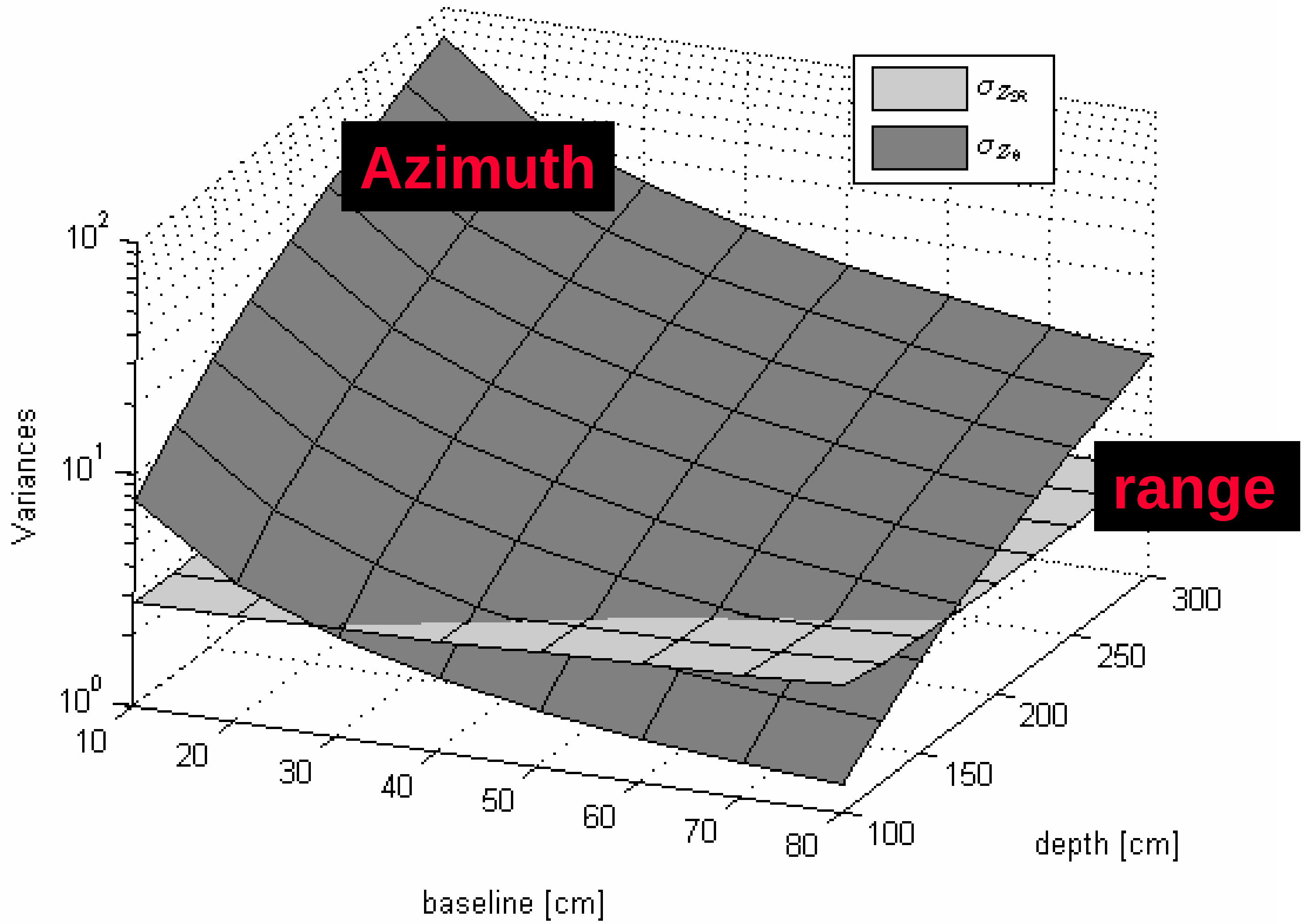


Opti-Acoustic Stereo Imaging; New Paradigm for Underwater 3-D Target Mapping

3-D reconstruction from range measurement (Impact of baseline)

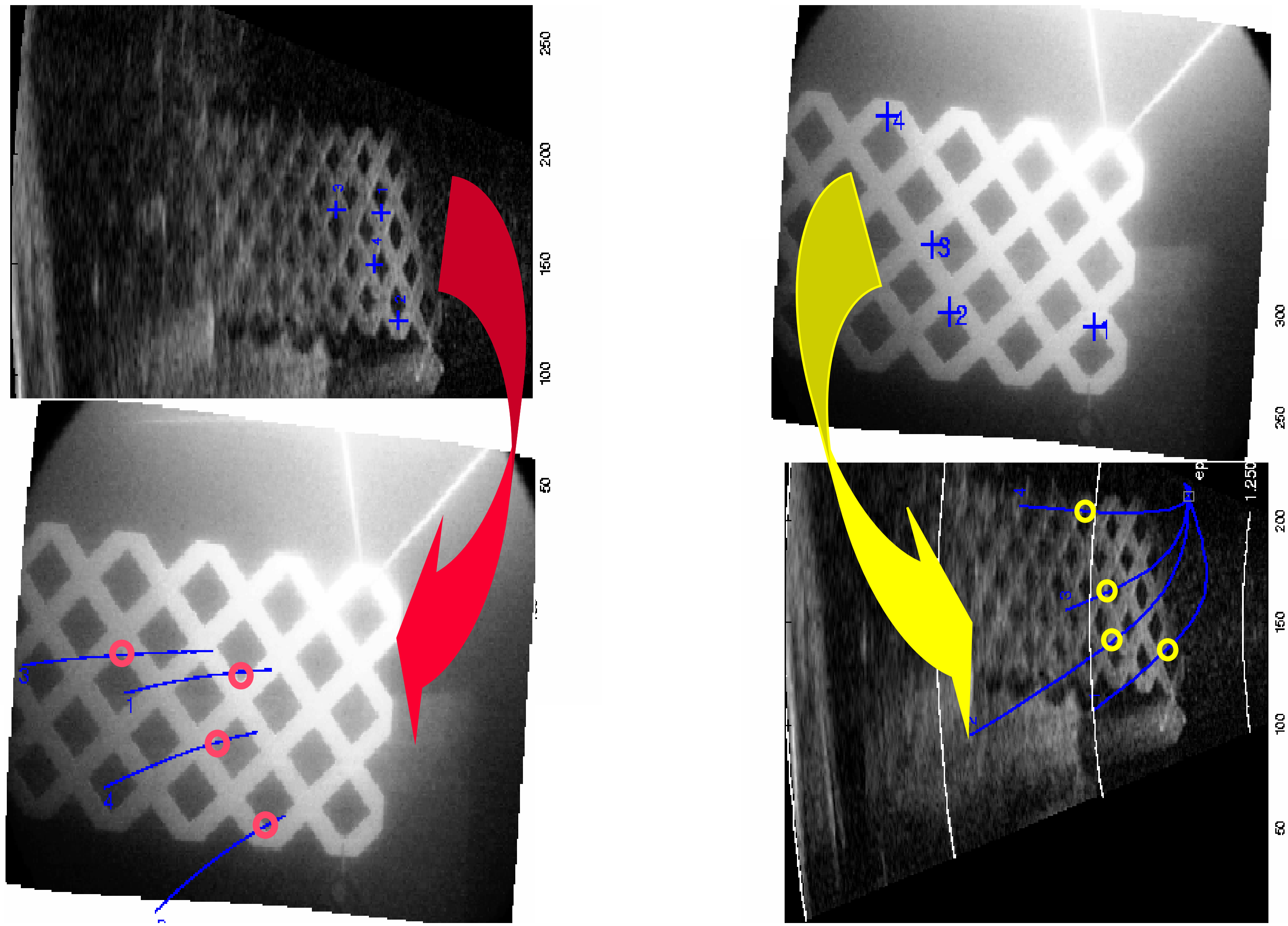


Opti-Acoustic Stereo Imaging; New Paradigm for Underwater 3-D Target Mapping



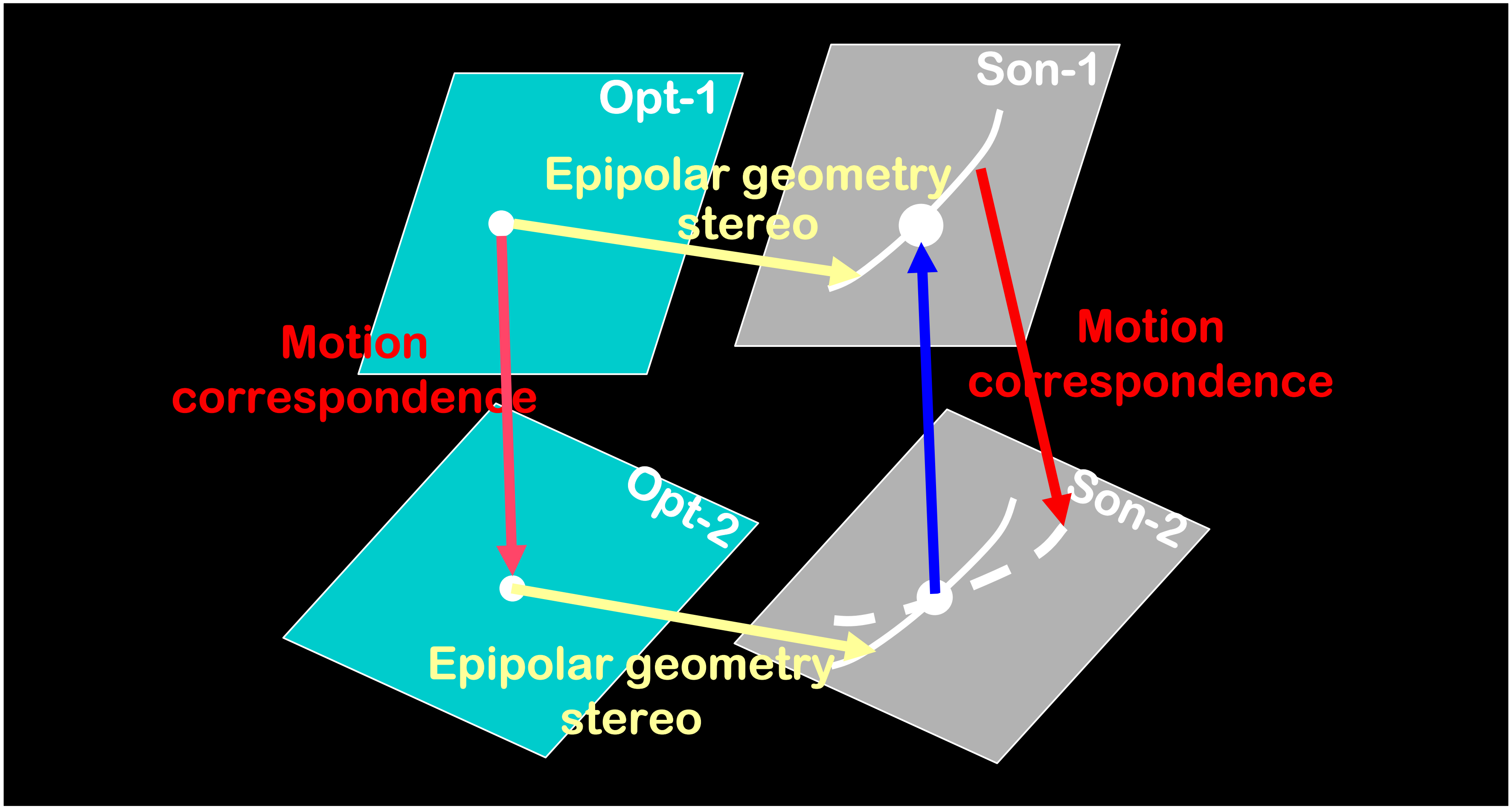
Azimuth and range measurements can be utilized more cleverly, accounting for target range and stereo baseline.

Opti-Acoustic Correspondence Problem?



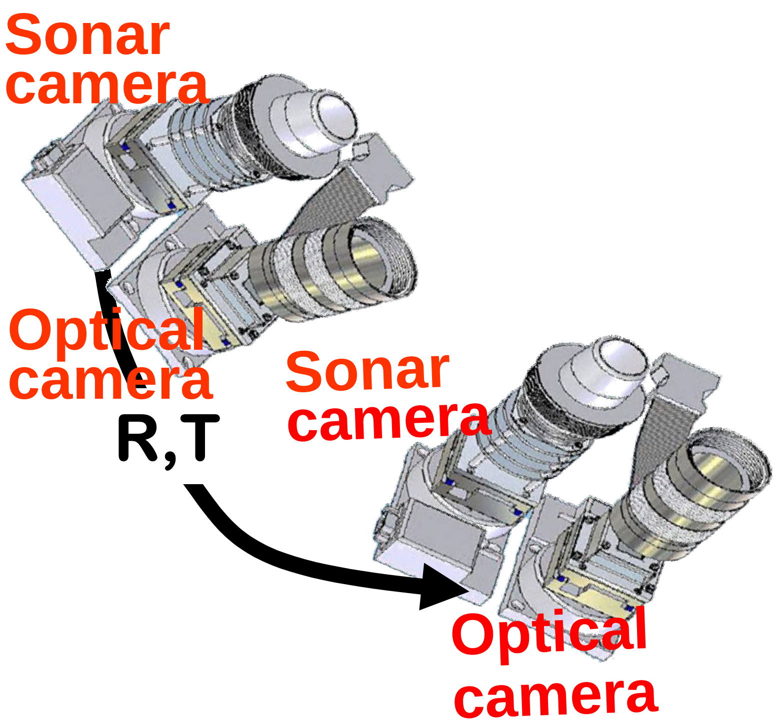
Opti-Acoustic Stereo Imaging; New Paradigm for Underwater 3-D Target Mapping

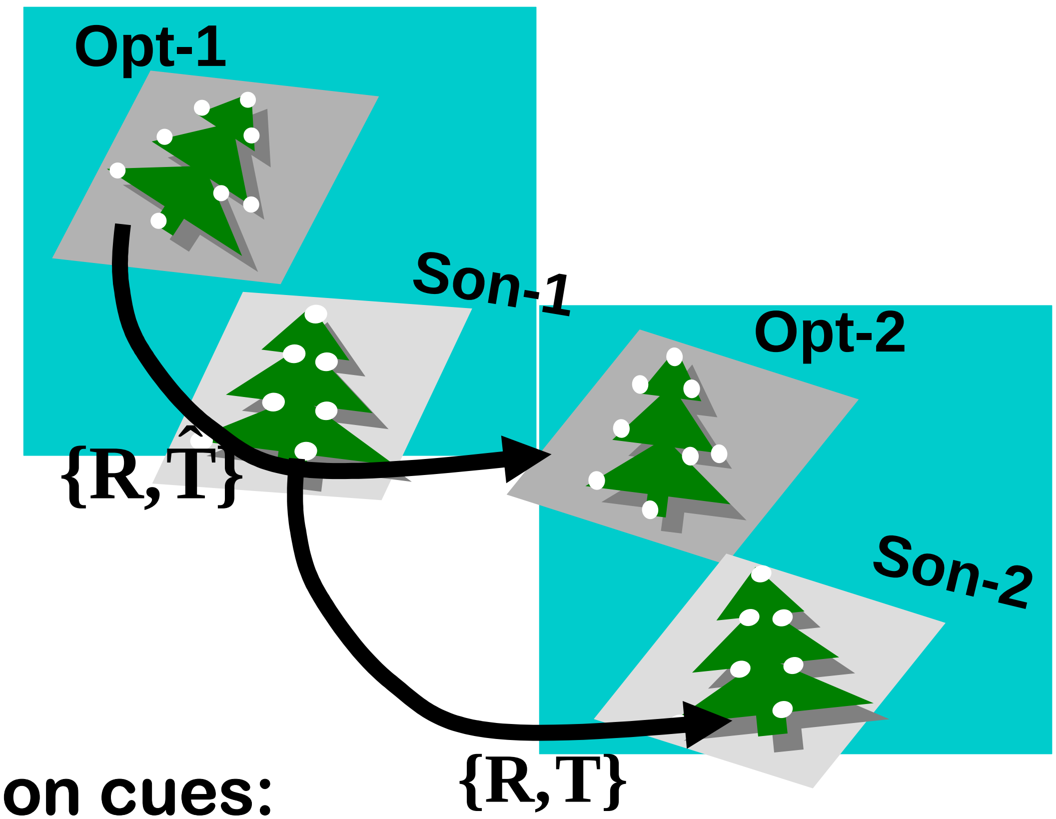
Current Approach:
Integration of motion and stereo cues



Opti-Acoustic Stereo Imaging; New Paradigm for Underwater 3-D Target Mapping

3-D Camera Motion Estimation (BenCOS'07)





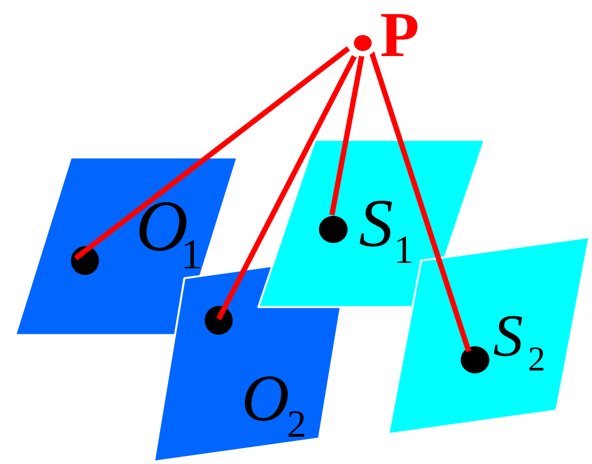
Integration of visual motion cues:

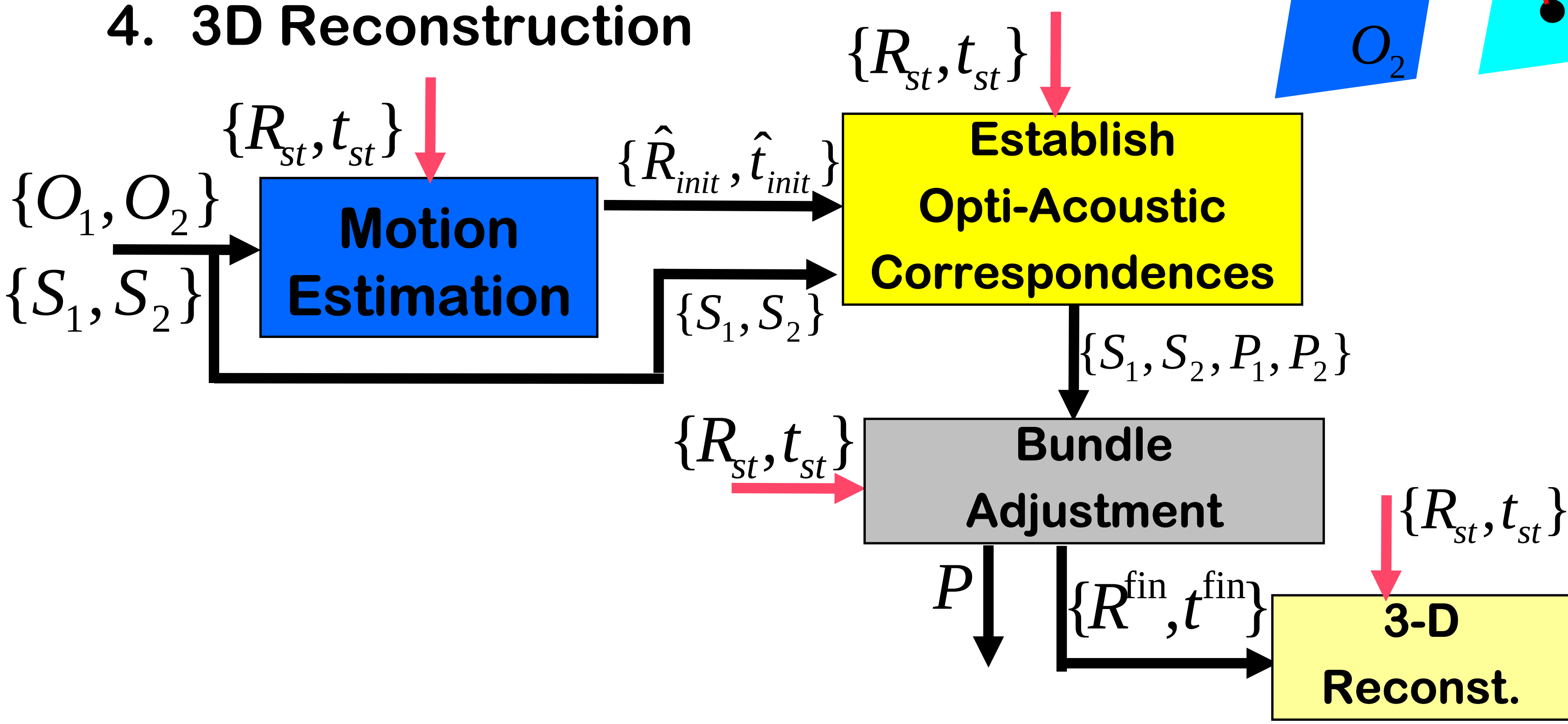
- *Resolving scale factor ambiguity* of monocular motion
- *More robust & accurate estimation* of sensor platform motion

Opti-Acoustic Stereo Imaging; New Paradigm for Underwater 3-D Target Mapping

Automated 3-D Reconstruction:

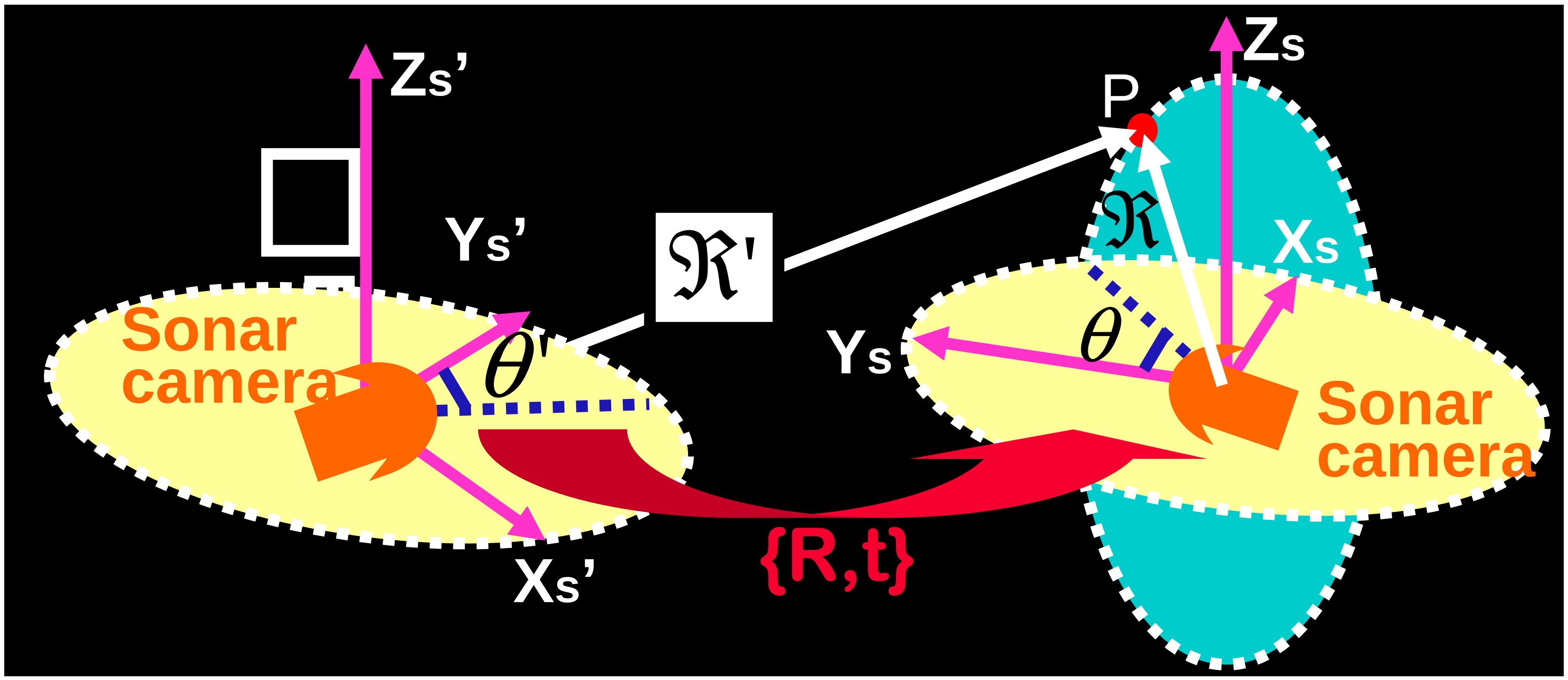
1. Motion Estimation
2. Opti-Acoustic Matching
3. Bundle Adjustment
4. 3D Reconstruction





Opti-Acoustic Stereo Imaging; New Paradigm for Underwater 3-D Target Mapping

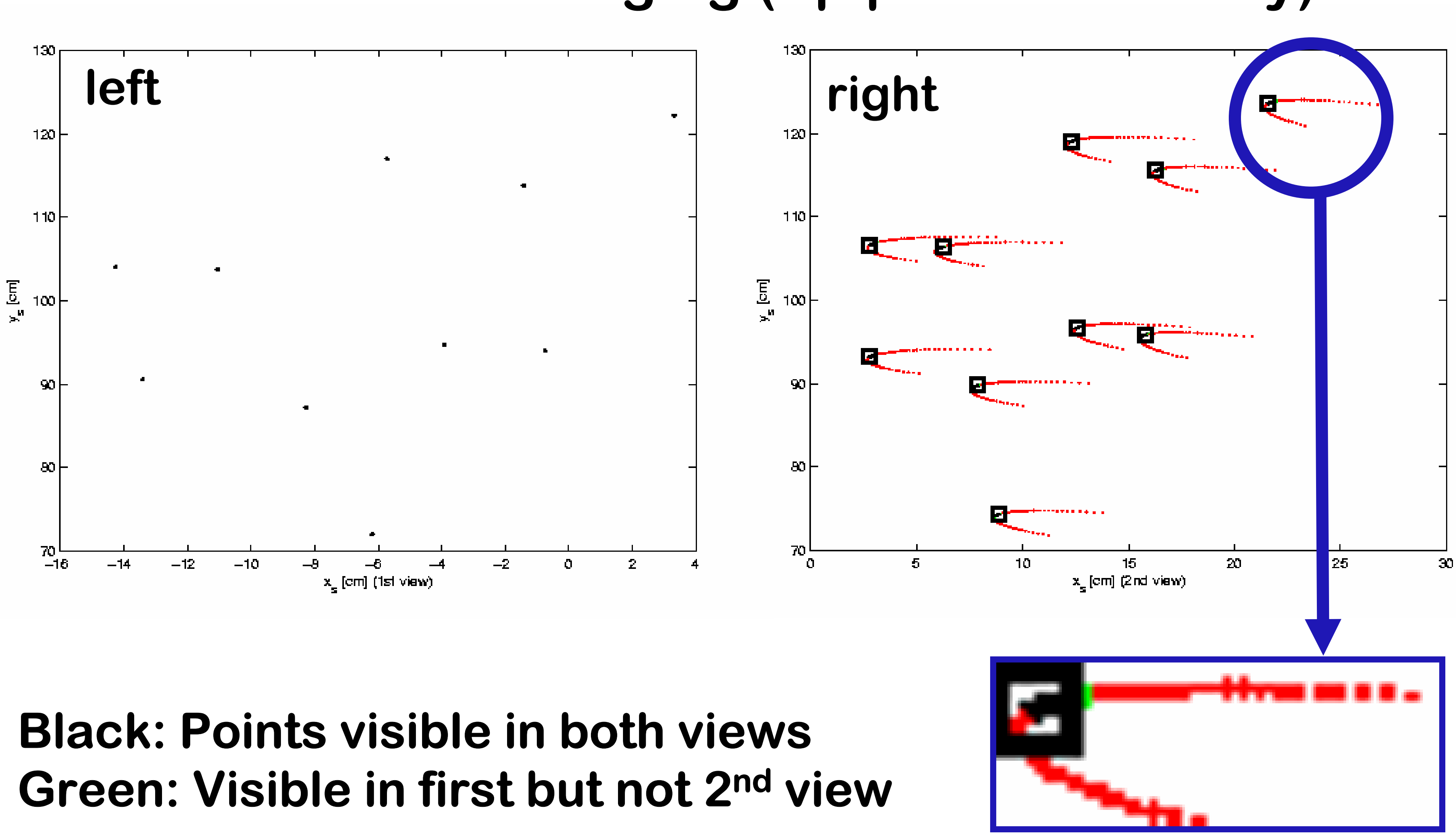
Sonar Stereo Imaging



$$\begin{bmatrix} -1 & \tan\theta & 0 \\ r_{21}\tan\theta' - r_{11} & r_{22}\tan\theta' - r_{12} & r_{23}\tan\theta' - r_{13} \\ \mathbf{t}\,.\,\mathbf{r}_1 & \mathbf{t}\,.\,\mathbf{r}_2 & \mathbf{t}\,.\,\mathbf{r}_3 \end{bmatrix} P = \begin{bmatrix} 0 \\ t_x - t_y \tan\theta' \\ \Re'^2 - \Re^2 - \|\mathbf{t}\|^2 \end{bmatrix} \quad \text{subject to } \|P\| = \Re$$

Opti-Acoustic Stereo Imaging;
New Paradigm for Underwater
3-D Target Mapping

Sonar Stereo Imaging (Epipolar Geometry)



Black: Points visible in both views
Green: Visible in first but not 2nd view

Concluding Remarks

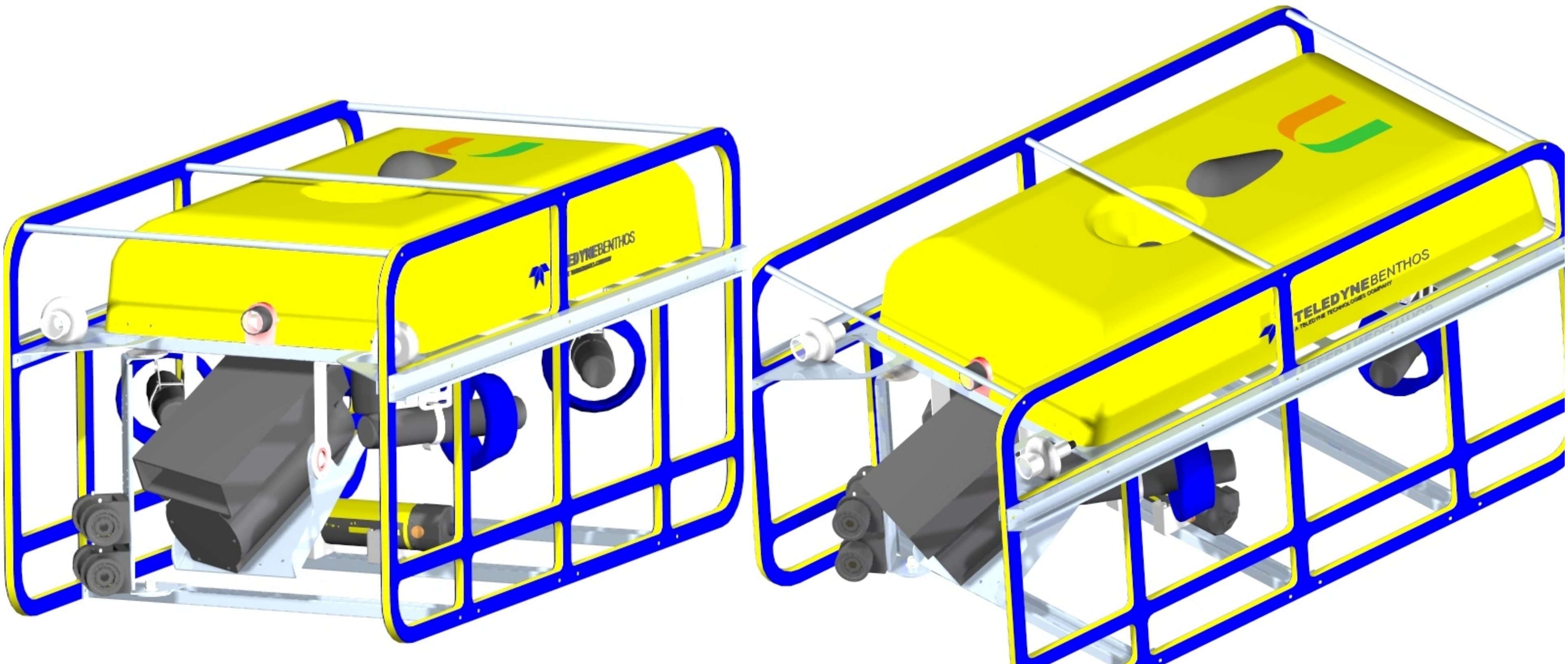
- Underwater vision research is challenging, but also fun and rewarding (lot of diving in great locations);
- Theoretical developments of vision in scattering media will play a critical role in underwater vision;
- Underwater vision can truly benefit from contributions from a larger body of CV community;
- Perhaps, we can establish assessment criteria of 2-D and 3-D reconstruction methods, not merely for performance on existing data bases but also on underwater imagery;
- To promote this, UM has valuable resources and facilities for underwater research to be made available to researchers who are interested in collaborative activity.

Resources

Phantom XTL with mega-pix digital stereo cameras, analog CCD camera, compass and pitch/roll sensors

Opti-Acoustic Stereo Imaging; New Paradigm for Underwater 3-D Target Mapping

Resources



Benthos Custom-designed Stingray ROV with stereo Camera sonar Camera and 6 d.o.f. acoustic positioning

Opti-Acoustic Stereo Imaging; New Paradigm for Underwater 3-D Target Mapping

Resources

Of course, large experimental tank in our backyard

The End