FOVEA

Foveated Vergent Active Stereo System

for

Dynamic Three-Dimensional Scene Recovery

William Klarquist
Al Bovik

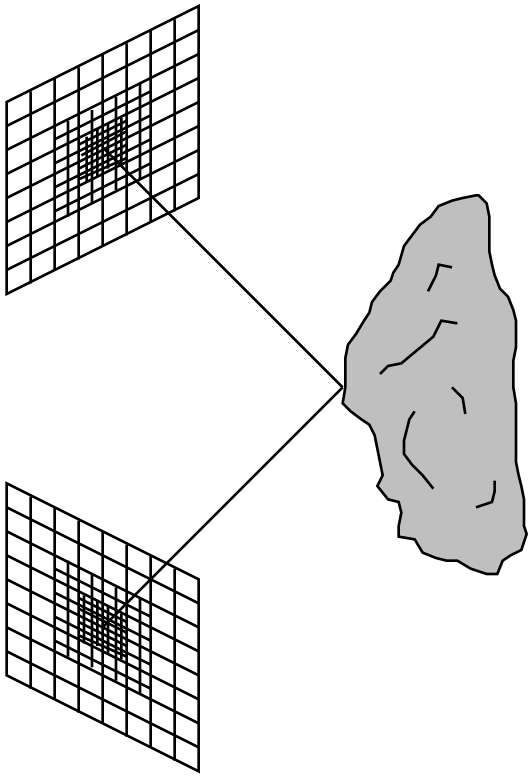
Center for Vision and Image Sciences
The University of Texas at Austin

WHAT FOVEA DOES

GOAL: **Direct** the **attention** of **an active camera platform** to object surfaces and **recover 3-D** scene structure using **multiple fixations**

RESULT: **FO**veated **VE**rgent **A**ctive stereo system

FOVEA



METHOD

What composes **FOVEA** ?

- **Active** camera control
- Initial **Depth Recovery** ("bootstrapping")
- **Foveated** Vergent Stereo
- **Fixation** (where to look next)
- **Multi-Fixation** Dynamic **Data Accumulation**

MOTIVATION

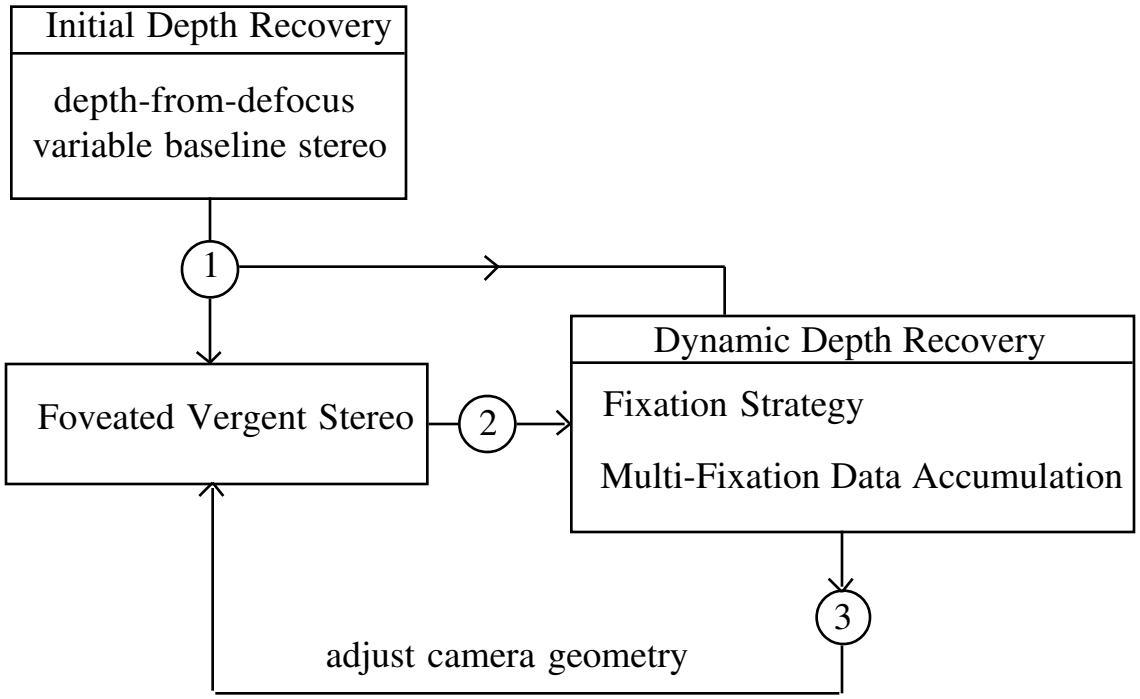
- Why is **FOVEA** useful?

  1. Tasks: 3-D mapping
     surface analysis
     navigation / hazard avoidance
     object recognition
  2. Better **constrain depth recovery** process
  3. **Aligned view** of 3-D region of interest
  4. **Integrate** information over **multiple fixations**
  5. Better **depth resolution**
  6. **Efficient** use of data via **foveal coding**

- What makes this problem challenging?

  1. Accurate camera fixation
  2. **Vergent stereo** correspondence
  3. Fixation strategy
  4. Data accumulation over **multiple fixations**

FOVEA Block Diagram



1 initial coarse depth map

2 foveated vergence depth map

3 accumulated depth map

OUTLINE

I. INITIAL DEPTH RECOVERY

II. FOVEATED VERGENT STEREO

III. DYNAMIC DEPTH RECOVERY

IV. CONCLUSION

INITIAL DEPTH RECOVERY IN FOVEA

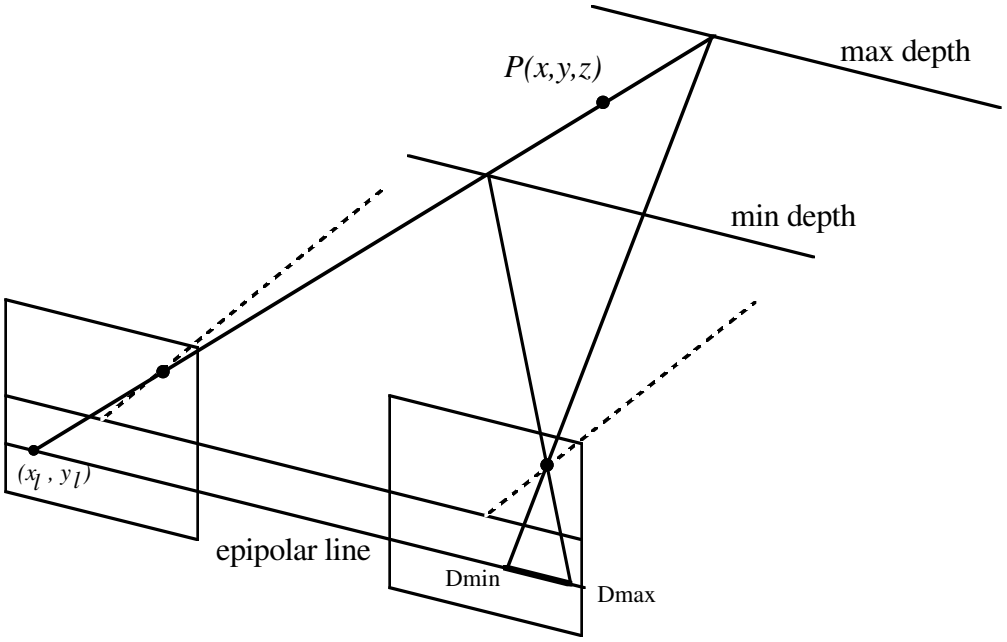
OBJECTIVES:

1. **Make vergent fixation possible**
2. **Adapt** to current scene structure

NEW TECHNIQUE -- Integrated IDR process:

1. **Adaptive Variable Baseline Stereo** (discussed here)
2. Maximum-Likelihood Depth-from-Defocus

PARALLEL BASELINE STEREO CORRESPONDENCE



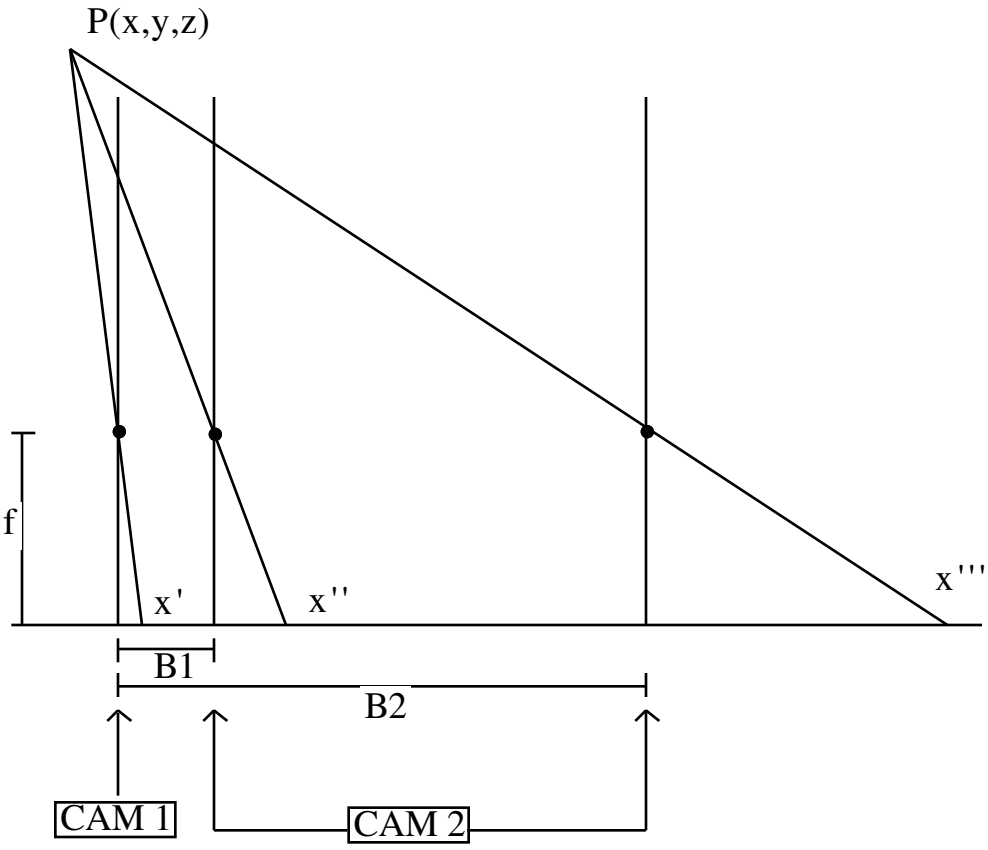
VARIABLE BASELINE STEREO

- Active control of baseline separation

- Fundamental tradeoffs:

$$\text{Baseline} \propto (\text{ease of matching})^{-1}$$

$$\text{Baseline} \propto \text{depth resolution}$$



BASELINE SELECTION IN FOVEA

OLD APPROACH:

- Conservative "worst case" strategy to guarantee unique matching

$$B = \frac{(D+f)\Delta x}{2f} \text{ with } \Delta x = 1 \text{ pixel}$$

where,

B = Stereo Baseline
D = Minimum distance to platform
Δx = disparity

- Expand by a fixed amount at each stage

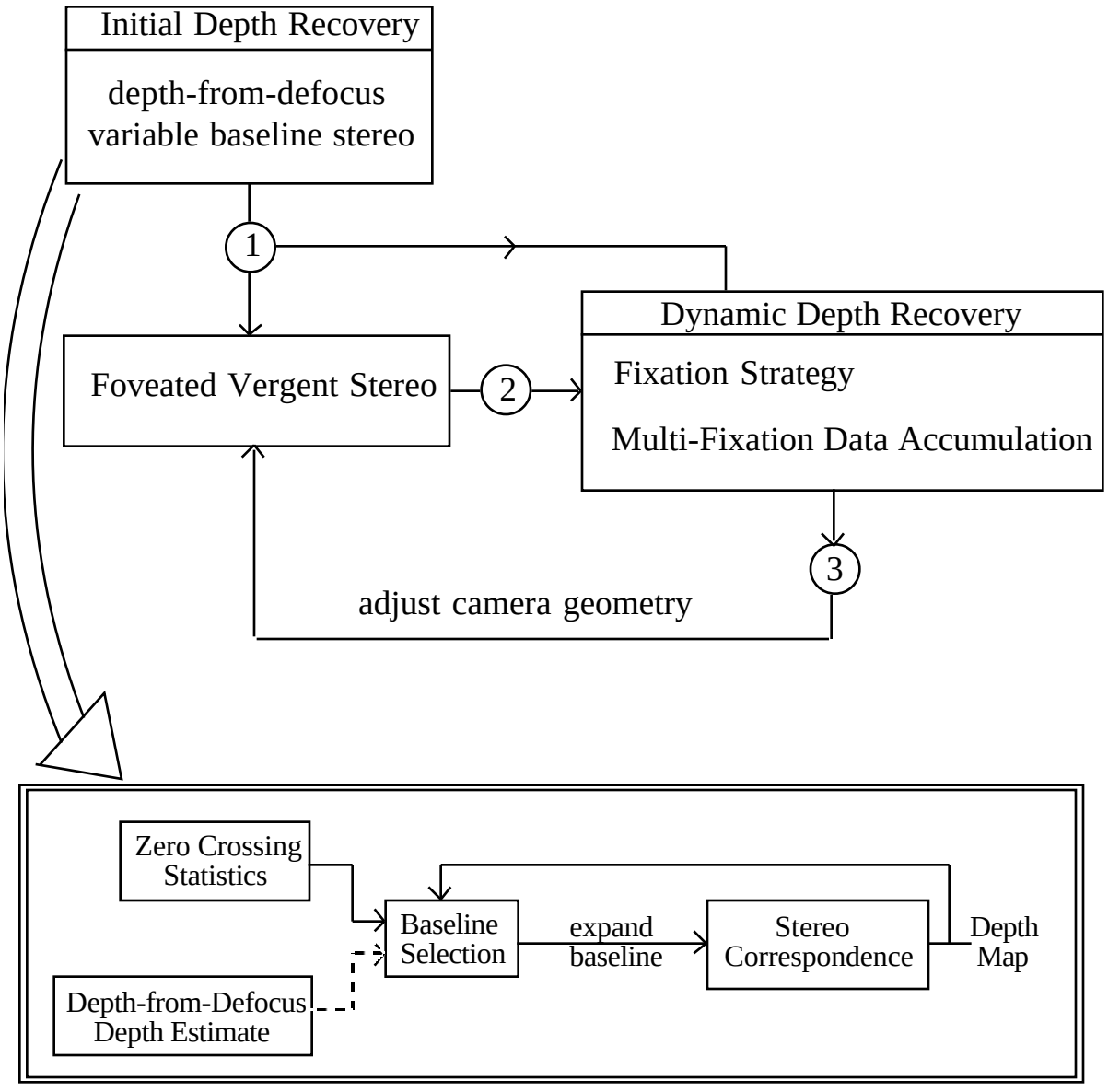
FOVEA'S APPROACH:

- *Adapt* to current scene information
- Use the image information to determine the minimum matching primitive separation (T_{min})
- Set baseline accordingly

$$B = \frac{(D+f)(T_{min}-1)}{2f}$$

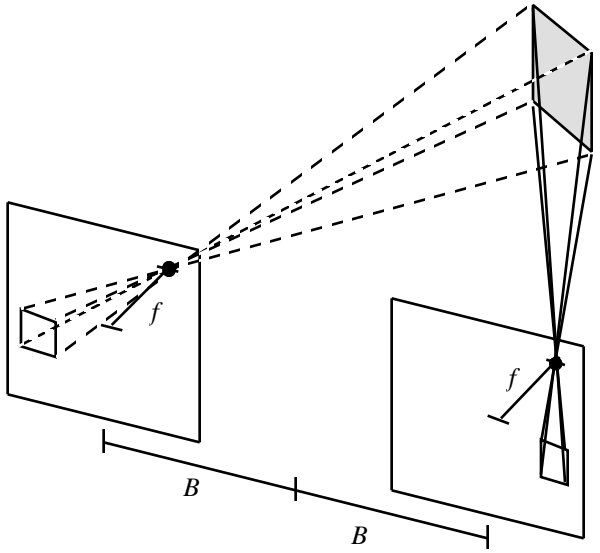
- Set expansion factor according using current disparity map to guarantee unique matching

INITIAL DEPTH RECOVERY



SURFACE PATCH DEPTH MAP

- Begin with disparity map from IDR process
- Disparity map can be converted to depth directly or converted to surface depth patches



- Disparity patches constructed using least square error fitting of MxM pixel patches
- Planar surface patches parameterized by (p, q, d_0):

$$pi + qj + d_0 = d$$

- 32x32 pixel patches used
- minimum data threshold used

OUTLINE

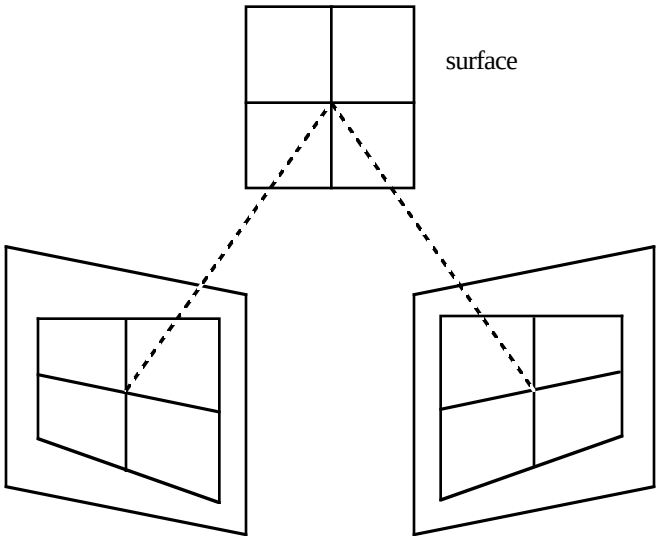
I. INITIAL DEPTH RECOVERY

II. FOVEATED VERGENT STEREO

III. DYNAMIC DEPTH RECOVERY

IV. CONCLUSION

VERGENT STEREO



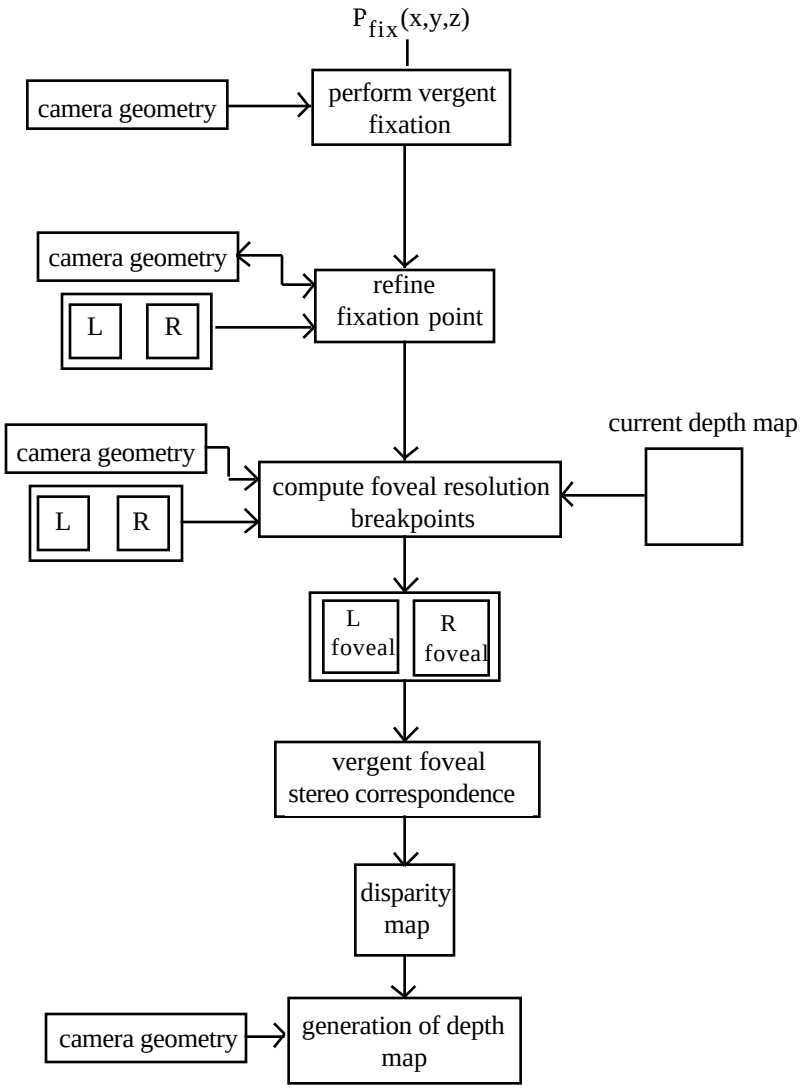
Advantages:

- **Aligned / Overlapping views** of surface
- **Natural correspondence** between central image regions
- **Improved** depth **resolution**
- Melds with **fixation strategies**
- Melds with **foveal processing** for reduced data load

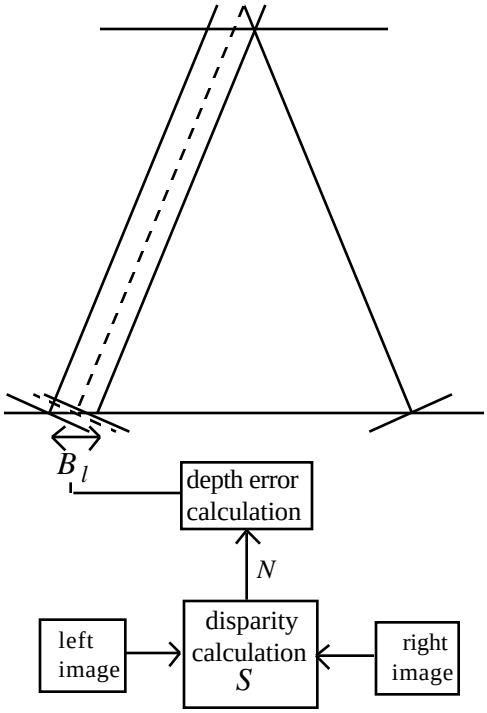
Complications:

- **Need support** to perform accurate vergence
- Peripheral **correspondence complicated** by **perspective projection**

FOVEATED VERGENT STEREO PROCESS



REFINE FIXATION POINT



- **Control baseline** position to **align cameras** on surface
- Expression for **current depth**

$$Z = \frac{\left(f\cos\theta - D_h N_x \sin\theta\right)}{\left(D_h N_x \cos\theta + f\sin\theta\right)}\left(f\sin\theta - \frac{\left(B_l + B_r\right)}{2}\right) - f\cos\theta$$

- Move to reduce D_h to 0

VERGENT STEREO IN FOVEA

To perform correspondence must either

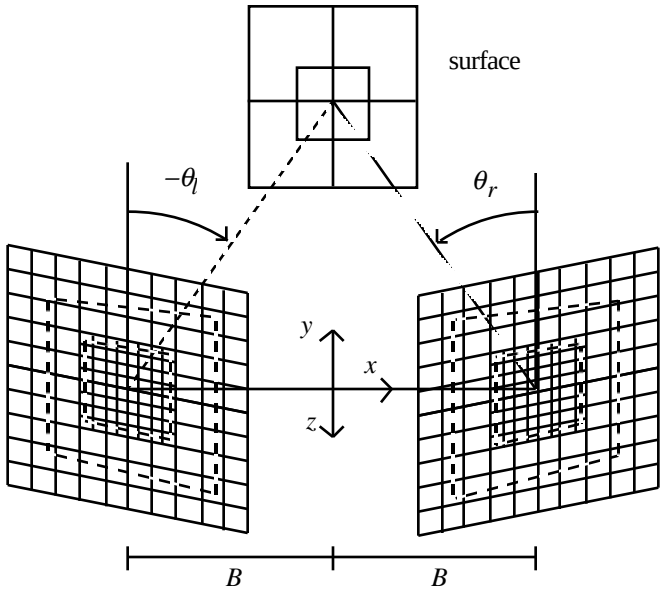
- **change correspondence process** to match new epipolar lines
- change images (*rectify*) to fit parallel baseline scanline correspondence

OR

- Use a multiresolution *foveal* processing structure to maintain scanline matching

FOVEATED VERGENT STEREO

- **Trade depth resolution** for **matching simplicity** in the periphery
- Maintain **high resolution** in **central foveal** region

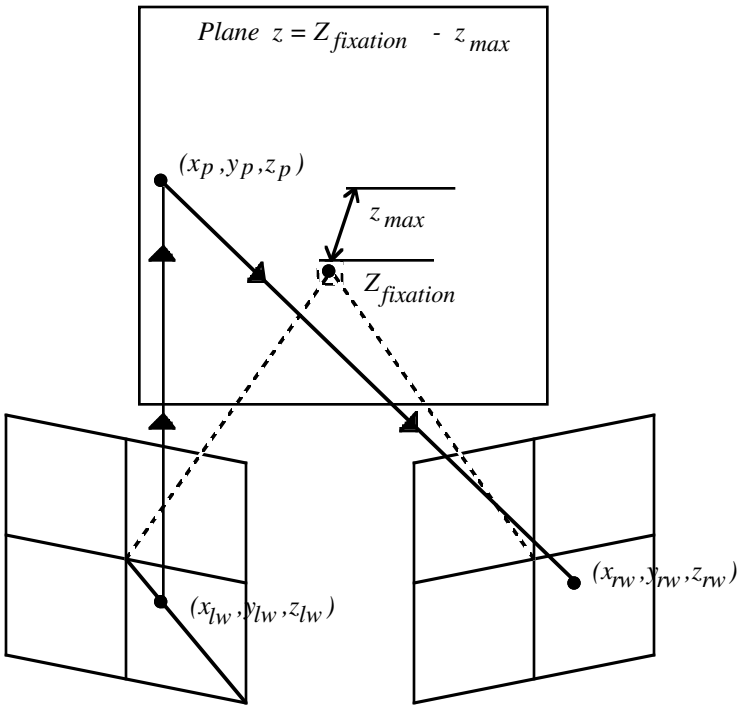


- The critical question:

  How to determine the foveal processing structure?

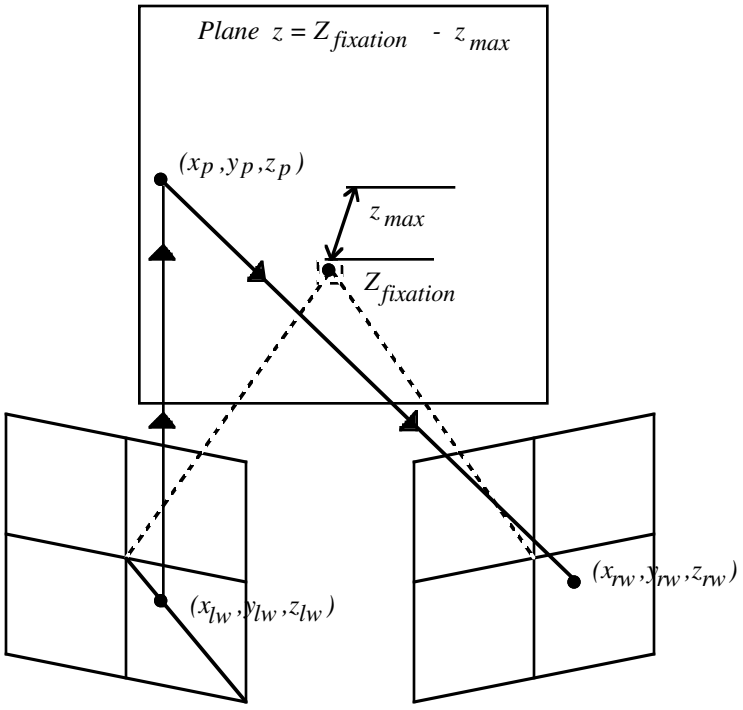
CONSTRUCTION OF FOVEAL ARRAY

- Depends on **camera geometry** and surfaces in view
- Must **constrain local scene** structure to guarantee correspondence
- Constraint placed on range of furthest surface from fixation point (z_{max})
- Given current camera geometry can compute the *breakpoints* that divide different resolution levels



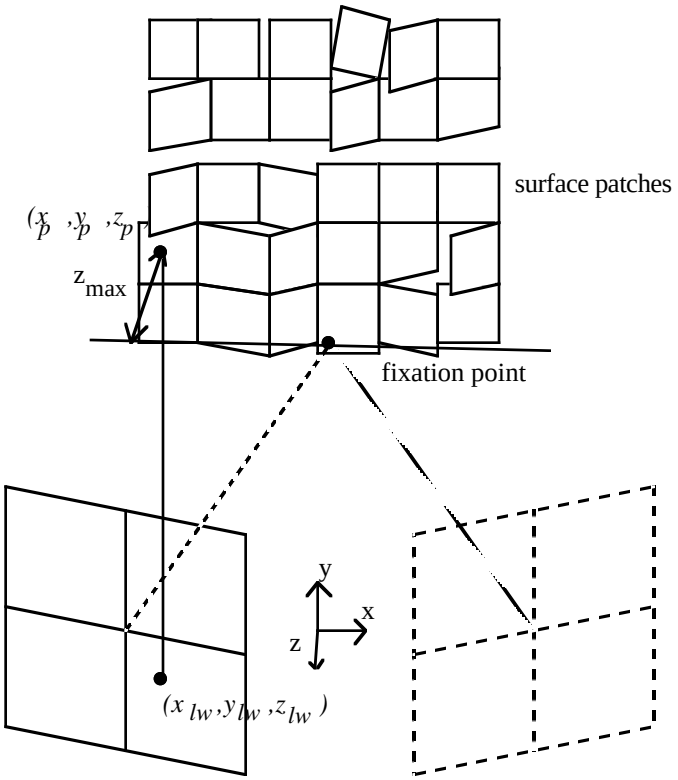
COMPUTATION OF RESOLUTION BREAKPOINTS

- For a chosen z_{max} :
  - **compute intersection** for backprojected image point with plane
  - **compute intersection** of forward projection of this intersection with right image plane
  - **Set breakpoints** when disparity increases past one-half the current resolution



REFINEMENT OF FOVEAL DESIGN

- Use **current / accumulated depth map** to determine z_{max}
- **Raytrace current surface** at low resolution
- Take maximum distance from fixation point as z_{max}



STEREO CORRESPONDENCE IN FOVEA

Two Algorithms tested with system:

- **Deterministic Search**

  For each feature point $L(i,j)$ compute l_2-norm $S(i,j,d)$:

$$S(i,j,d) = \sum_{u=-\frac{M}{2}}^{\frac{M}{2}} \sum_{v=-\frac{M}{2}}^{\frac{M}{2}} \left[\frac{L(i+u, j+v)}{\overline{L}} - \frac{R(i+u.\, j+v+d)}{\overline{R}} \right]^2$$

  S is computed for each disparity over range
  $(D_{min}<d<D_{max})$

  Minimum $S(i,j,d)$ chosen

- **Simulated Annealing**

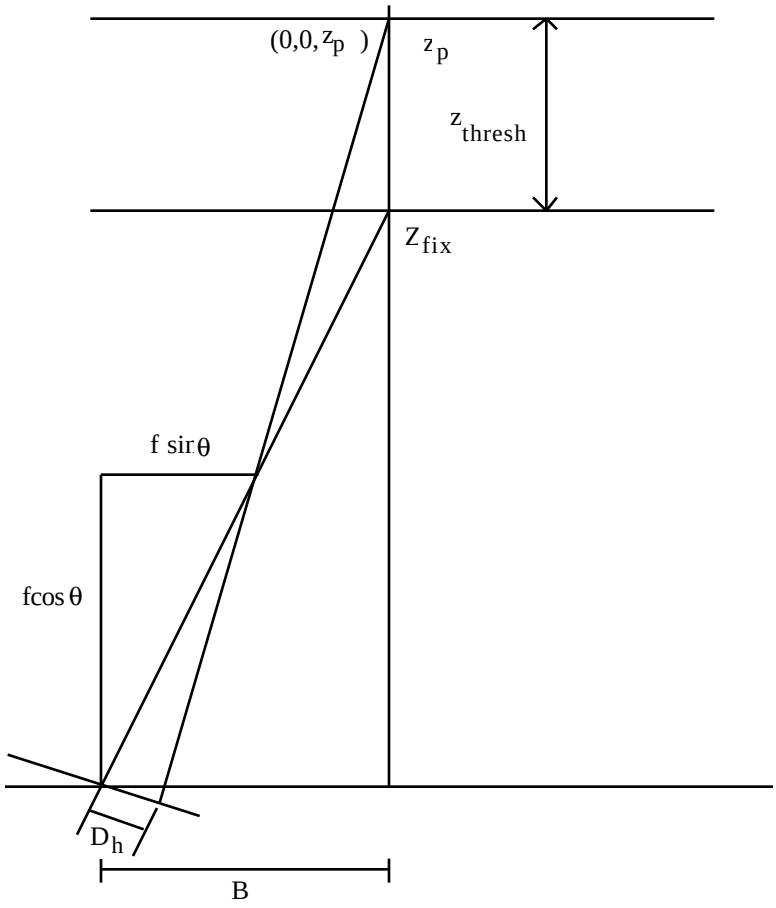
  Correspondence defined as energy minimization

  E = photometric term + (λ)(disparity smoothness)

$$E = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \left\{ \left| L(i,j) - R[i, j + D(i,j)] \right| + \lambda \sum_{u=-M}^{M} \sum_{v=-M}^{M} \left| D(i,j) - D(i+u, j+v) \right| \right\}$$

REFINEMENT OF FOVEAL DESIGN

- Need a method to appropriately set z_{max}
- Set a threshold (z_{thresh}) for the computation of depth



- Compute correspondence only within this range

OUTLINE

I. INITIAL DEPTH RECOVERY

II. FOVEATED VERGENT STEREO

III. DYNAMIC DEPTH RECOVERY

IV. CONCLUSION

DYNAMIC DEPTH RECOVERY PROCESS IN FOVEA

Composed of two components:

- **Probabilistic Fixation Strategy**

  --evaluation of competing criteria for fixation
  --**selection** of **new fixation** site

- **Multi-Fixation Data Integration and Access**

  --integrate new surface patches
  --provide access to data for outside processes

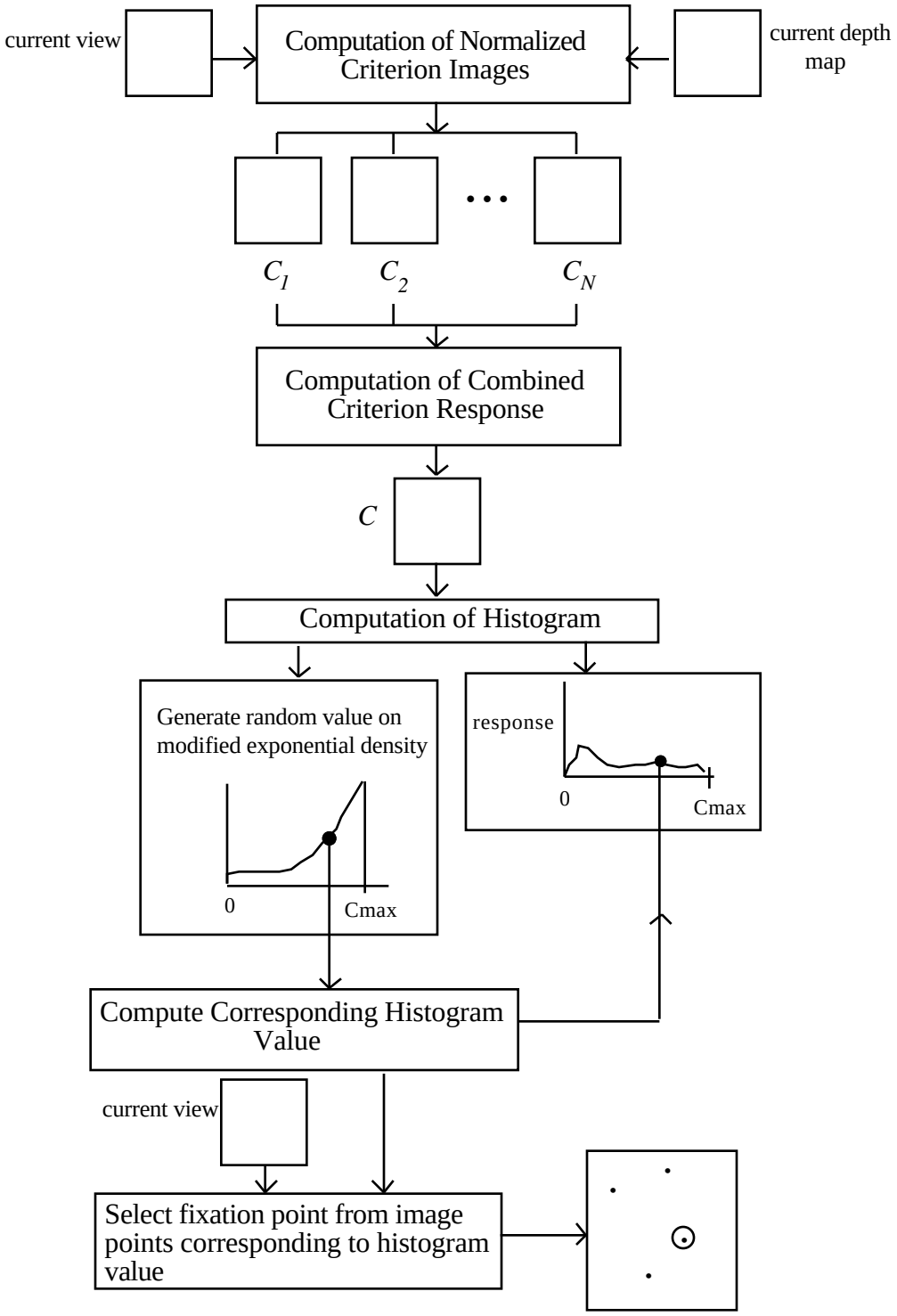
FOVEA'S PROBABILISTIC FIXATION STRATEGY

GOAL: Select the **next site** for fixation, based on the **current view** and the **accumulated depth map**

Advantages:

- Integration of 2-D and 3-D criteria in a consistent fashion
- Dynamic selection of next fixation point
- Modular Structure

PFS BLOCK DIAGRAM



FIXATION POINT SELECTION CRITERIA

- **Gradient attraction**

$$C(i,j)_{grad} = \left|\nabla I(i,j)\right|^2 = \left[I(i,j) - I(i,j-1)\right]^2 + \left[I(i,j) - I(i-1,j)\right]^2$$

- **Local Suppression**

$$(i,j)_{local\ sup} = \begin{cases} 1.00 & \text{if the point has not been mapped at foveal resolution} \\ 0.01 & \text{otherwise} \end{cases}$$

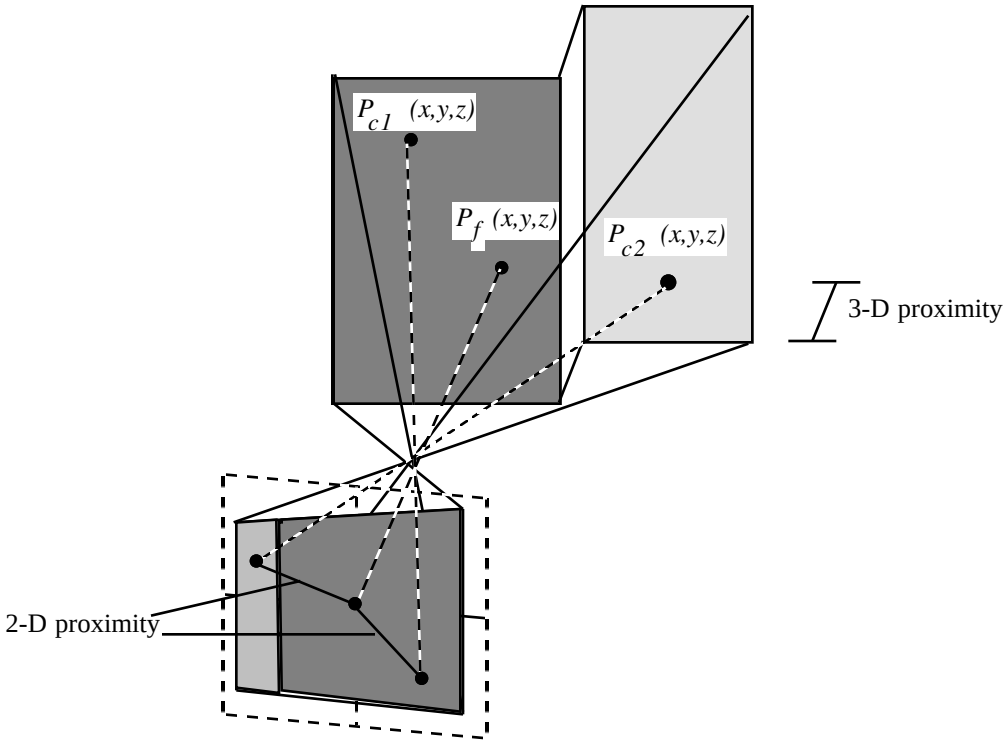
$$C(i,j)_{wander} = \begin{cases} 1.00 & \text{if the point is in the mapping range} \\ 0.01 & \text{otherwise} \end{cases}$$

- **Proximity**

$$C(i,j)_{2-D\ Proximity} = e^{-\alpha\sqrt{(i-i_c)^2+(j-j_c)^2}}$$

$$C(i,j)_{3\text{-}D\ proximity} = \begin{cases} 1.0 & \text{if within } z_{thresh} \text{ of fixation point} \\ 0.1 & \text{otherwise} \end{cases}$$

SELECTION CRITERIA-- PROXIMITY



- **2-D proximity criterion** keeps fixation **point near current position**
- **3-D proximity criterion** keeps fixation point on current surface

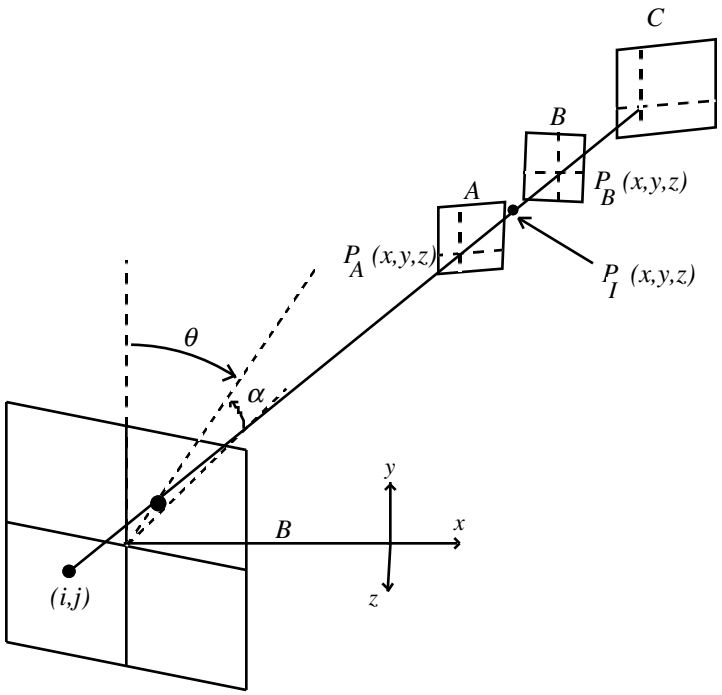
DATA INTEGRATION AND ACCESS

Two Approaches:

- Accumulated Surface Patch Structure

  --All patches are included in data structure

  --Depth is disambiguated by raytracing current view:



- Condensed Surface Patch Structure

  --Only most recent / highest resolution data integrated

  --Data integrated in a hierarchical fashion

CONCLUSION

CONTRIBUTIONS:

- The creation of an autonomous system for 3-D scene recovery with

  --integration of surface information over multiple fixations

  --depth recovery using a vergent geometry

  --integration of multiresolution surface information

- Two new methods for Initial Depth Recovery
  - -Adaptive Variable Baseline Stereo
  - -Maximum-Likelihood Depth-from-Defocus

- Novel technique for performing vergent stereo correspondence
  - --Foveated Vergent Stereo

- Adaptive fixation strategy for computing next view
  - --Probabilistic Fixation Strategy