Patterns of Binocular Disparity for a Fixating Observer

Miles Hansard & Radu Horaud

INRIA Rhône-Alpes, 655 Avenue de l'Europe, Montbonnot 38334, France.

miles.hansard@inrialpes.fr radu.horaud@inrialpes.fr

1. Introduction

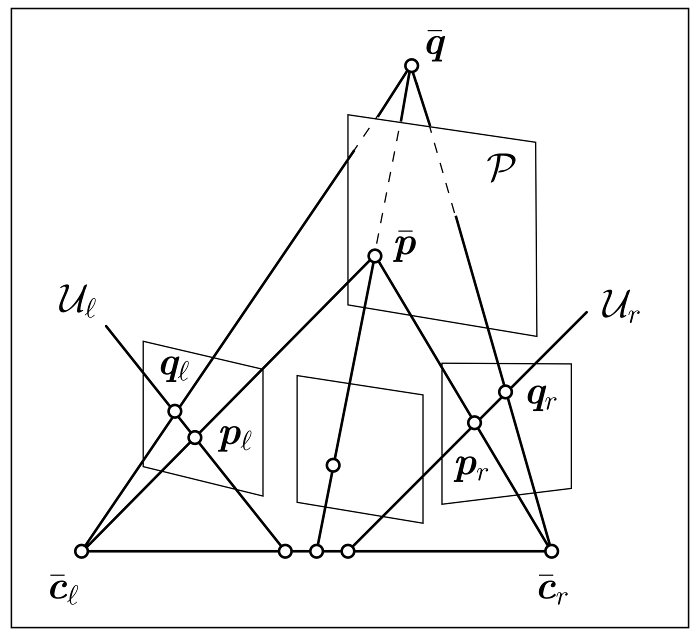
- The observed binocular disparity field is determined by the scene structure and the gaze configuration.
- Individual binocular mechanisms in area V1 are tuned to a small range of disparities.

2. Questions

- How does the 3-D geometric representation influence the observed disparity field?
- How can the gaze component of the disparity field be analyzed in a physiologically plausible way?

3. Representation

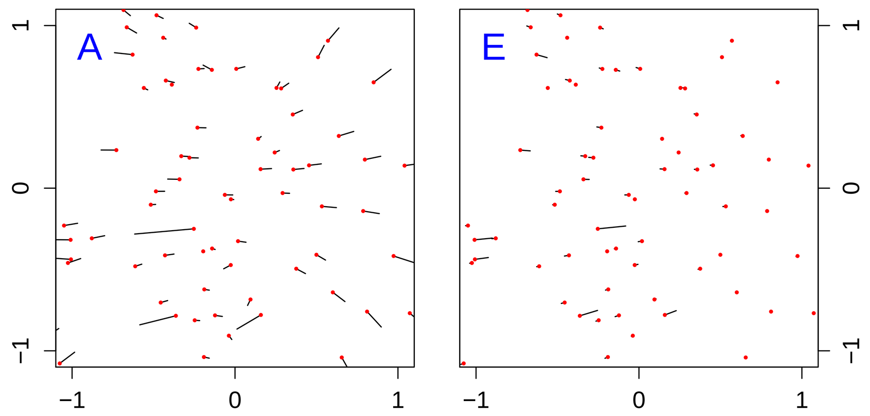
- The fixation point defines a reference plane $\mathcal{P}$, orthogonal to the Cyclopean gaze direction.
- The Cyclopean direction of scene point $\bar{q}$ determines a reference point $\bar{p} \in \mathcal{P}$.
- The scene point $\bar{q}$ projects to q_ℓ and q_r.
- The reference point $\bar{p}$ projects to p_ℓ and p_r.
- The geometry is shown below, with optical centres c_ℓ and c_r, where $|c_r - c_\ell| = 1$:



- The absolute disparity, at retinal location q_ℓ, is $q_r - q_\ell$.
- The epipolar disparity is $d = q_\ell - p_\ell$.
- The local direction of the epipolar disparity is u_ℓ, which determines an epipolar line $\mathcal{U}_\ell$.

4. Experiments

- Scene Points $i = 1 \cdots M$ were Normally distributed with high variance; $M = 1000$, mean $= (0, 0, 5)^\top$, cov $=$ diag$(100, 100, 100)$.
- Fixation points $j = 1 \cdots N$ were Normally distributed with low variance; $N = 100$, mean $= (0, 0, 5)^\top$, cov $=$ diag$(1, 0, 1)$.
- Retinal Samples $k = 1 \cdots K$ were arranged on a Cartesian lattice; $K = 17 \times 17 = 289$.
- Two example disparity patterns (M reduced for plotting) are shown below:



- Absolute (A) and epipolar (E) disparity fields, for the same scene, plotted in the left view.
- Use of a reference plane (E) reduces the average magnitude of the disparity.

5. Sampling

- The retinal distribution of disparity information varies according to the gaze.
- Problem: Direct smoothing introduces spurious minima in the flow, owing to the cancellation of oppositely directed vectors.
- Solution: The disparity vector d_{ij} at point q_{ij} is represented by a rank-one 2×2 matrix:

$$D^{q}_{ij} = d_{ij} d_{ij}^\top$$

- The matrices encode orientation (mod π), rather than direction (mod 2π).
- The matrices are smoothed and re-sampled onto the retinal lattice s_k:

$$D^{s}_{jk} = \sum_{i}^{M} \alpha^{k}_{ij} D^{q}_{ij}$$

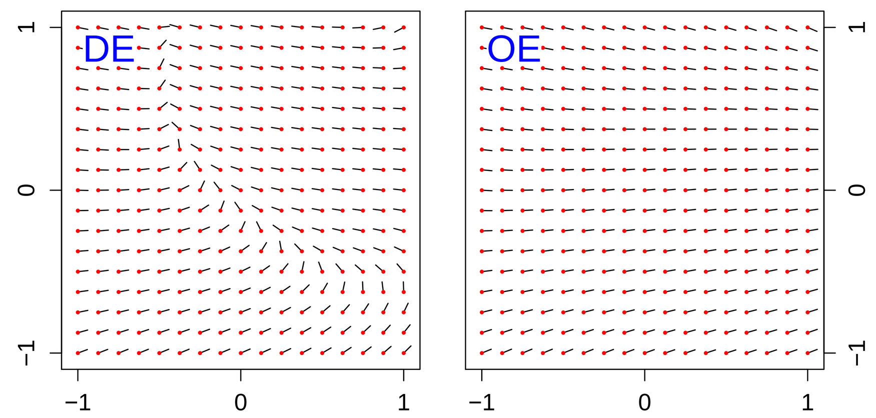
- The spatial weighting has an isotropic Gaussian profile:

$$\alpha^{k}_{ij} \propto \exp\left(\tfrac{-1}{2\delta^2} |q_{ij} - s_k|^2\right)$$

- Each average disparity matrix is rank-two, with eigenvectors u_{jk} and v_{jk}, and corresponding eigenvalues $\lambda_{jk} \geq \mu_{jk}$.
- For fixation j, the disparity matrix at retinal position s_k has the decomposition:

$$D^{s}_{jk} = \lambda_{jk} u_{jk} u_{jk}^\top + \mu_{jk} v_{jk} v_{jk}^\top$$

- The vectors u_{jk} approximate the epipolar geometry of fixation j, independently of the scene-structure.
- Examples of the two averaging procedures are shown here:



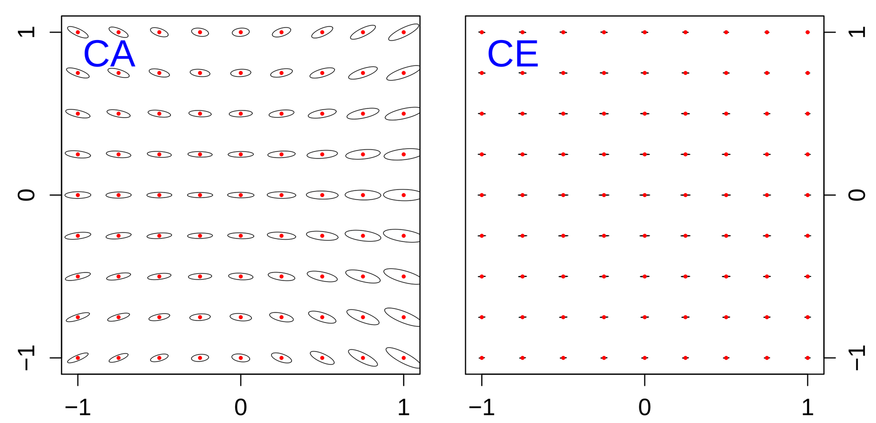
- Direction-average (DE), and orientation average (OE) of the epipolar disparity field shown in (E).
- Note the spurious contour in (DE), caused by averaging oppositely-directed vectors.

6. Local Structure

- Question: What is the typical distribution of disparity at a given retinal point, over many fixations?
- The gaze-averaged flow at retinal point s_k is:

$$D^{s}_{k} = \sum_{j}^{N} \beta_j D^{s}_{jk}$$

- The gaze weighting is defined here as $\beta_j = \frac{1}{N}$.
- D^{s}_{k} encodes the disparity fluctuation at retinal position s_k, as the gaze varies.
- The inverse of D^{s}_{k} can be used to construct a prior probability for the local binocular correspondence.
- Iso-probability contours (K reduced for plotting) are shown below for absolute disparity (CA), and epipolar disparity (CE):



- Note that the variability is greatly reduced by the use of a reference plane (CE).

7. Global Structure

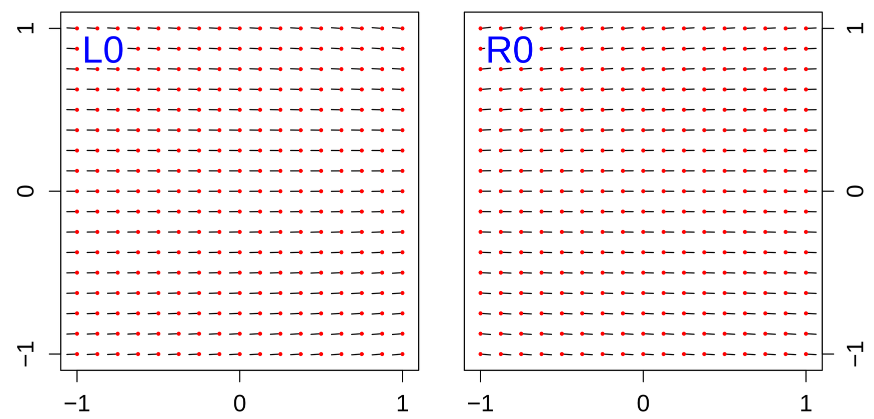
- The functional dependence between the left and right views is only approximate, after the resampling procedure.
- Hence each binocular flow is represented as a column-vector of length $2K + 2K$, containing u_{jk} from the left and right images:

$$f_j \in \mathbb{R}^{4K}$$

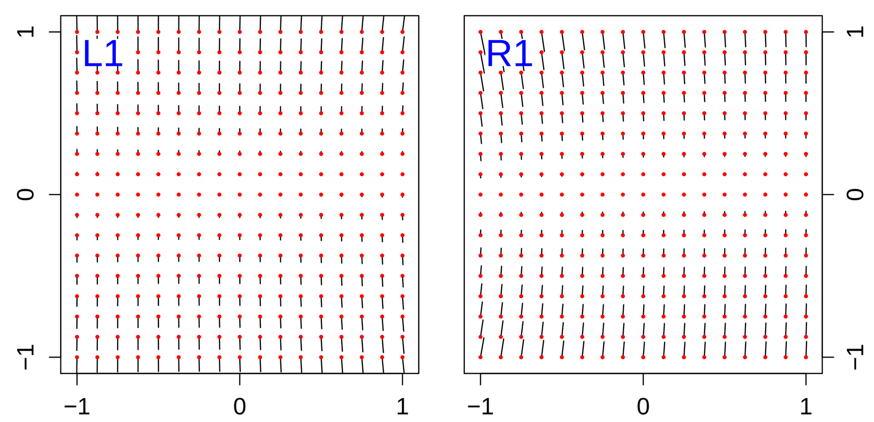
- The mean, f_0, is subtracted from each vector.
- The N vectors are collected into a $4K \times N$ data matrix F.
- The singular value decomposition of the data is

$$F = G \Sigma H^\top$$

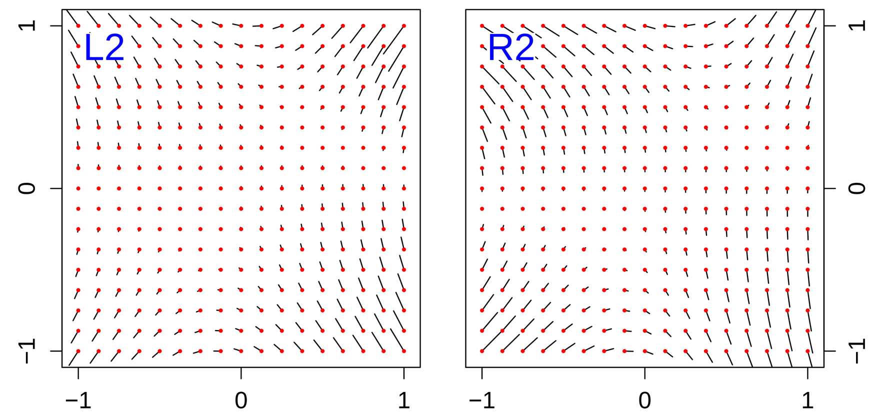
- G is a $4K \times N$ matrix, the columns of which are basis flows g_j.
- Σ is an $N \times N$ diagonal matrix of singular values, and $H^\top$ is an $N \times N$ orthogonal matrix.
- Left (L0) and right (R0) parts of the mean flow, f_0, are shown below:



- The pair (L0,R0) represents the average epipolar geometry, over $N = 100$ random fixations.
- Left (L1) and right (R1) parts of the first basis function, g_1, are shown below:



- The pair (L1,R1) represents a differential scaling of the disparity field, over $N = 100$ random fixations.
- Left (L2) and right (R2) parts of the second basis function, g_2, are shown below:



- The pair (L2,R2) represents a hyperbolic deformation of the disparity field, over $N = 100$ random fixations.

8. Dimensionality

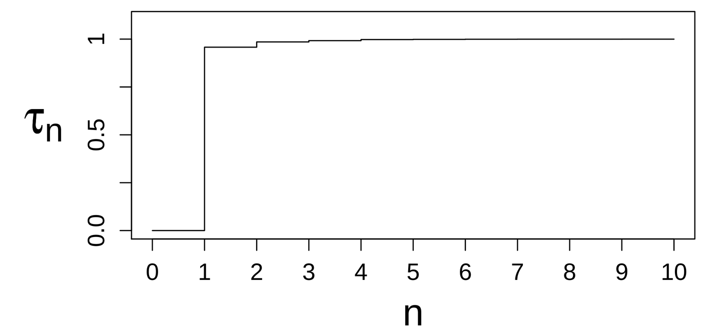
- A given flow f can be approximated by a linear combination of the first $n < N$ basis flows:

$$f \approx f_0 + \sum_{j}^{n} \gamma_j g_j$$

- The weights are projections $\gamma_j = (f - f_0)^\top g_j$.
- The proportion of variance explained by the first n components is:

$$\tau_n = \sum_{j}^{n} \sigma_j^2 \Big/ \sum_{j}^{N} \sigma_j^2$$

- The proportion of gaze-variance explained by the first $n = 10$ components, with $N = 100$, is shown below:



- The low dimensionality of the model reflects the fixation geometry (not the scene structure).

9. Conclusions

- The geometric representation of disparity is important.
- The gaze component of retinal disparity can be locally represented.
- The distribution of the gaze component, over many binocular fixations, is low-dimensional.

10. Acknowledgments

- This work is part of the Perception on Purpose project, supported by EU grant 027268.