Stereopsis

Computer Vision I
CSE252A
Lecture 13

Announcements

- Mailing list: You should be on
- HW3: Coming soon!! Stereo
- Final Exam: Friday, March 19, 11:30-2:30
- Last lecture: Color + Stereo
- Today: More stereo

Perceptual categories - a computational perspective
Pietro Perona
Professor, Caltech
Wednesday, February 18th
4:00pm, CSB 003

Abstract: When we are born we do not know about sailing boats, frogs, cell-phones and wheelbarrows. By the time we reach school age we can easily recognize these categories of objects and many more using our visual system; by some estimates, we learn around 10 new categories per day with minimal supervision during the first few years of our lives. How does this happen? I will outline a computational approach to the problem of representing the visual properties of object categories, and of learning such models without supervision from cluttered scenes using a minimum number of training examples. Both static images of objects and dynamic displays, such as the ones generated by human activity, are handled by the theory. Experiments on a large collection of object categories in real images will be used to illustrate the different aspects of the problem as well as the key properties of the theory.

RGB: primaries are monochromatic, energies are 645.2nm, 526.3nm, 444.4nm. Color matching functions have negative parts -> some colors can be matched only subtractively.

RGB Color Cube
- Block of colours for (r, g, b) in the range (0-1).
- Convenient to have an upper bound on coefficient of each primary.
- In practice:
  - primaries given by monitor phosphors
  - (phosphors are the materials on the face of the monitor screen that glow when struck by electrons)

CIE xyY (Chromaticity Space)
Variations in color matches on a CIE x, y space. At the center of the ellipse is the color of a test light; the size of the ellipse represents the scatter of lights that the human observers tested would match to the test color; the boundary shows where the just noticeable difference is. The ellipses on the left have been magnified 10x for clarity; on the right they are plotted to scale. The ellipses are known as MacAdam ellipses after their inventor. The ellipses at the top are larger than those at the bottom of the figure, and that they rotate as they move up. This means that the magnitude of the difference in x, y coordinates is a poor guide to the difference in color.

Binocular Stereopsis: Mars

Given two images of a scene where relative locations of cameras are known, estimate depth of all common scene points.

Two images of Mars

Need for correspondence

Reconstruction: General 3-D case

- Linear Method: find $P$ such that
  \[ \begin{align*}
  p \times M^TP &= 0 \\
  p' \times (M')^TP &= 0 \\
  \iff \quad (\frac{p}{|p'|, M})^TP &= 0
  \end{align*} \]
- Non-Linear Method: find $Q$ minimizing $d(p, q) + d(p', q')$

Two Approaches

- A) From each image, process “monocular” image to obtain cues.
- B) Establish correspondence between cues.
  - Directly compare image regions between the two images.
Human Stereopsis: Binocular Fusion

How are the correspondences established?
Julesz (1971): Is the mechanism for binocular fusion a monocular process or a binocular one?
• There is anecdotal evidence for the latter (camouflage).

Random dot stereograms provide an objective answer.

Random Dot Stereograms

A Cooperative Model (Marr and Poggio, 1976)

Epipolar Constraint
• Potential matches for \( p \) have to lie on the corresponding epipolar line \( l' \).
• Potential matches for \( p' \) have to lie on the corresponding epipolar line \( l \).

Epipolar Geometry
• Epipolar Plane
• Epipoles
• Epipolar Lines
• Baseline
Family of epipolar Planes
(standard approach)

Properties of the Essential Matrix

- \( p' E p = 0 \) with \( E = [t;] R \)
- \( E p' \) is the epipolar line associated with \( p' \).
- \( E' p \) is the epipolar line associated with \( p \).
- \( E e' = 0 \) and \( E' e = 0 \).
- \( E \) is singular.
- \( E \) has two equal non-zero singular values (Huang and Faugeras, 1989).

Calibration
Determine intrinsic parameters and extrinsic relation of two cameras

The Eight-Point Algorithm (Longuet-Higgins, 1981)
Much more on multi-view in CSE252B

Minimize:
\[ \sum_{i=1}^{n} (p_i^T \mathcal{F} p_i)^2 \]
under the constraint \( \| \mathcal{F} \|^2 = 1 \).
Example: converging cameras

Example: motion parallel with image plane

Example: forward motion

Rectification

Given a pair of images, transform both images so that epipolar lines are scan lines.

Rectification

All epipolar lines are parallel in the rectified image plane.

Image pair rectification

simplify stereo matching by warping the images

Apply projective transformation so that epipolar lines correspond to horizontal scanlines

map epipole e to (1,0,0)

try to minimize image distortion

Note that rectified images usually not rectangular
Rectification
Given a pair of images, transform both images so that epipolar lines are scan lines.

Input Images

Rectified Images
See Section 7.3.7 for specific method