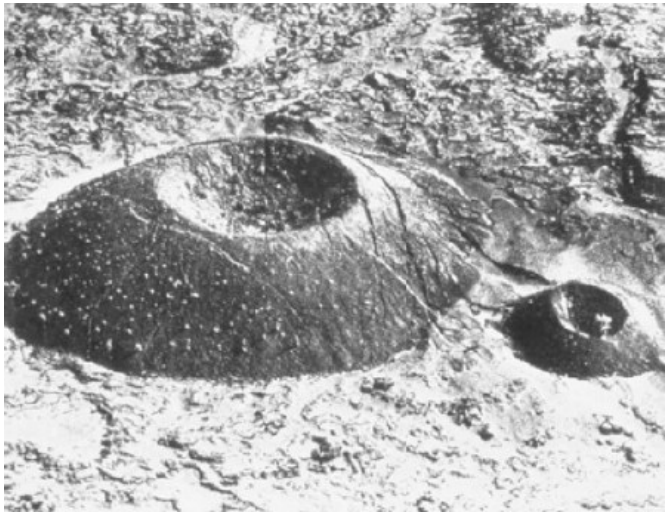
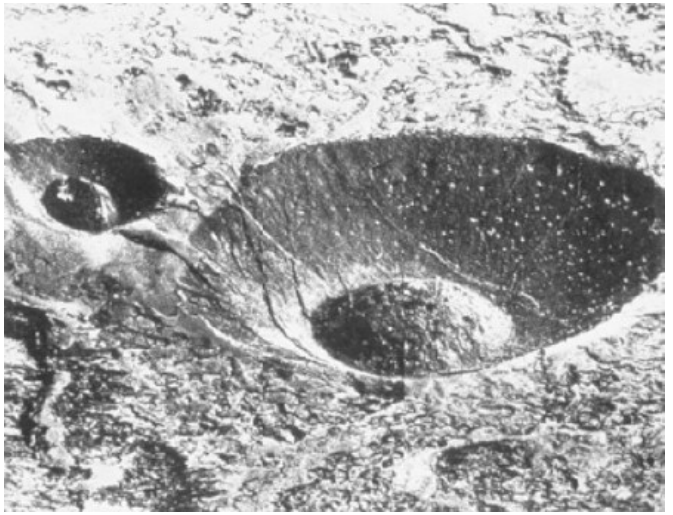
Artificial vision and pattern recognition

3D Vision

Introduction

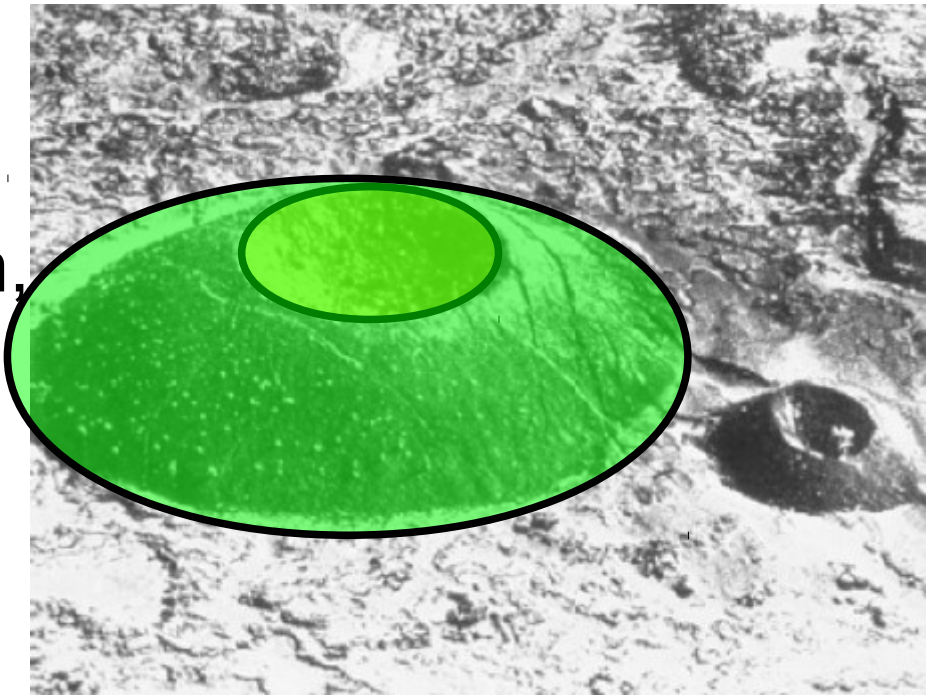
- Main problem: estimate 3D properties from 2D images (the human eye-brain system itself may not be reliable)





Introduction

- Estimate:
  - Geometry (shape, position, orientation, ...)
  - Dynamics (speed, path)
  - ... and sometimes try to recognize objects



- Can rely on detecting 2D shapes (another example was given with face detection)
- But the human eye-brain system gets informations from other sources...

Recognition clues

- Binocular vision (link: Lenticulations)

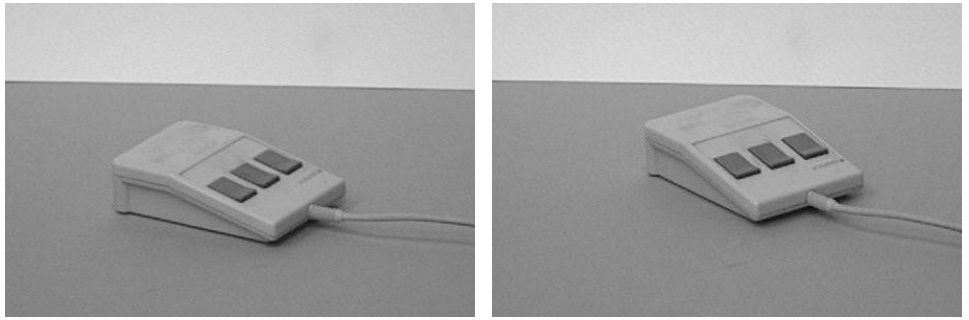
- HUGE knowledge base (context, components, ...)

Recognition clues

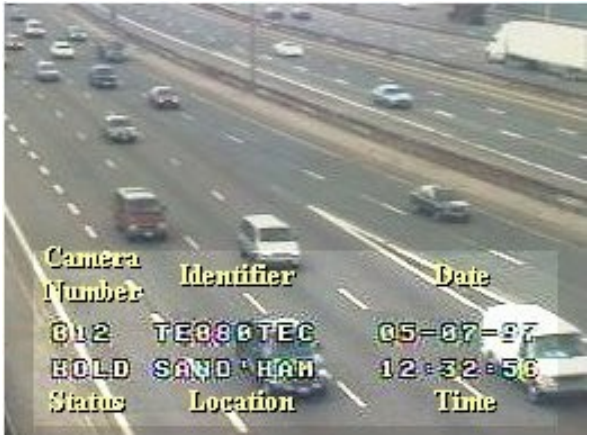
Color

Texture

Shadows



Multiple views



Motion

Major problems

- Why is it difficult to recognize an object ?
  - Shape changes due to different viewpoints
  - Color changes due to different illumination
  - Occlusions / Noise
  - Scale / Non-linear deformations
  - Background complexity
  - **Intra-class variation**
- Goal: find an *invariant* description of objects (using constraints to limit complexity)

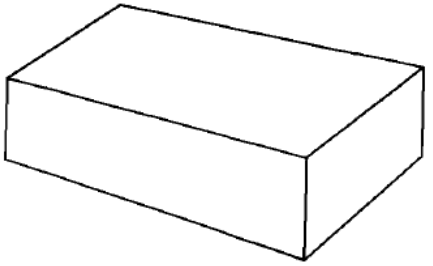
Practical considerations

- Sometimes life is made easier by *calibration*:
  - Rely on informations obtained from the camera which took the picture (EXIF)
  - A reliable calibration object (*eg* a cube, a gnomon...) is placed somewhere in the scene

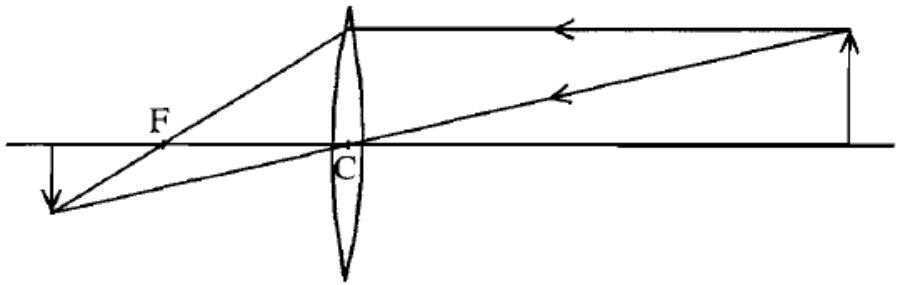
- Calibration helps to recover parameters of the camera (relative distance, orientation, Field Of View, ...)

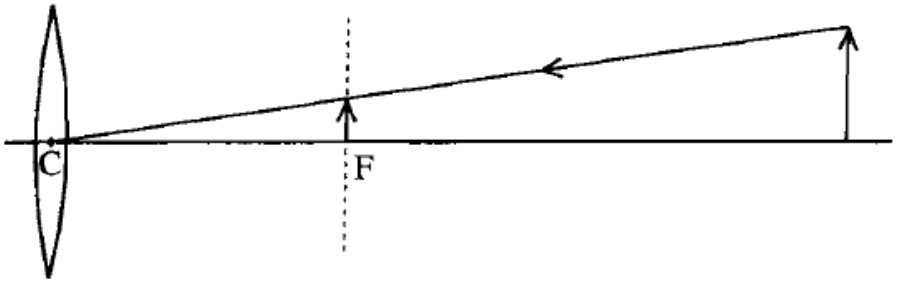
A 2D image from a 3D scene

- Useful informations (which ones ?) are lost because of perspective projection



- A simple projection model :
  - *C=(0,0,0)* is the center of projection and *F* is the focal length (*ie* distance from *C* to the image plane)
  - A point *(X,Y,Z)* appears at position *(x1,y1)* on the screen, with: *x1=fX/Z, y1=fY/Z*



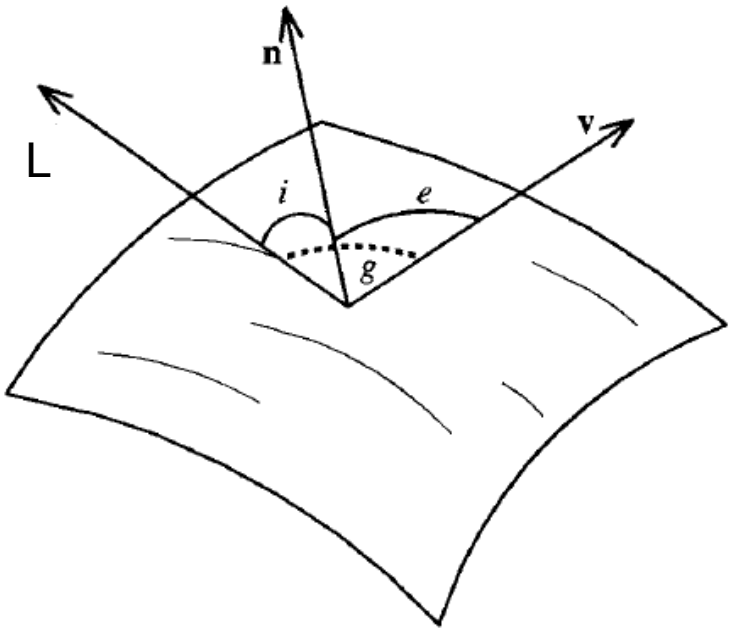


Shape from X

- From a single image, try to automatically infer the shape of an object from different aspects
- Actually the *shape* can be described from different parameters:
  - Depth
  - Surface normal
  - A set of contours
  - ...

Shape from shading

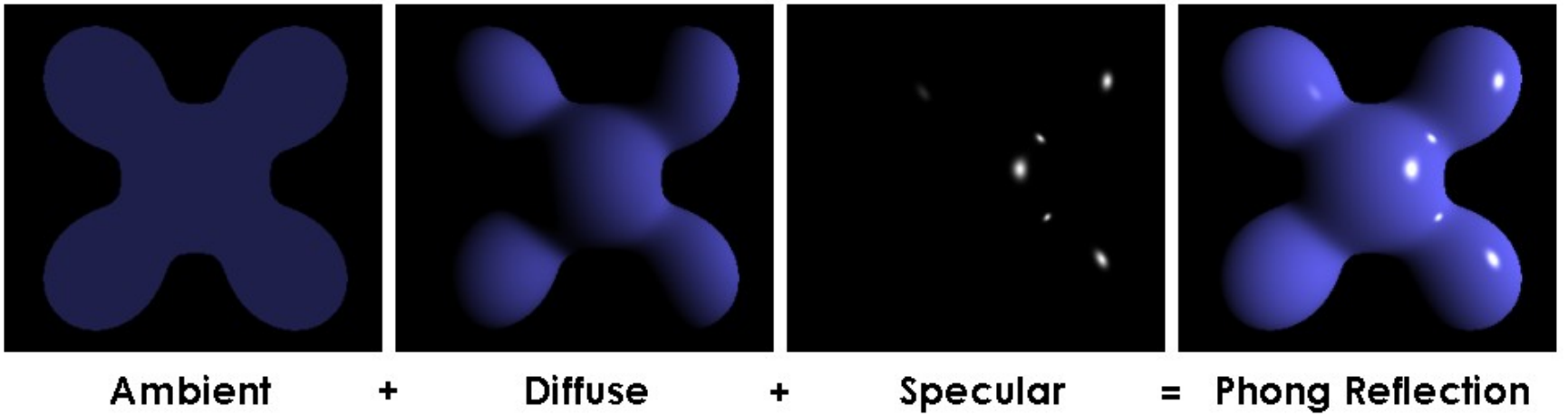
- The color of a pixel depends on the *reflectance* of the corresponding point on a surface



- *Phong shading model (R:* reflection vector):

$$I_p = k_a i_a + \sum_{\text{lights}} (k_d (L \cdot N) i_d + k_s (R \cdot V)^{\alpha} i_s).$$

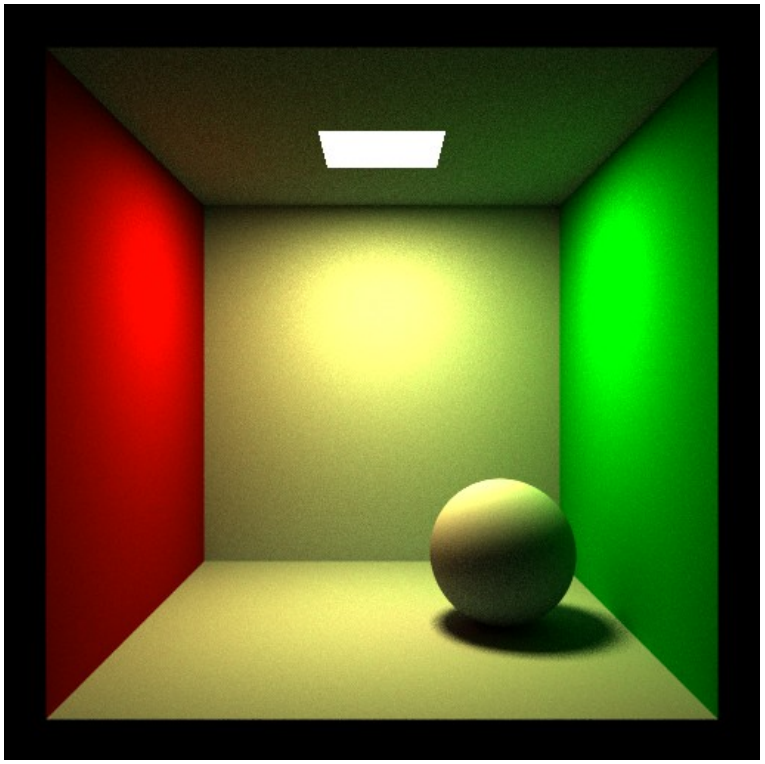
Shape from shading



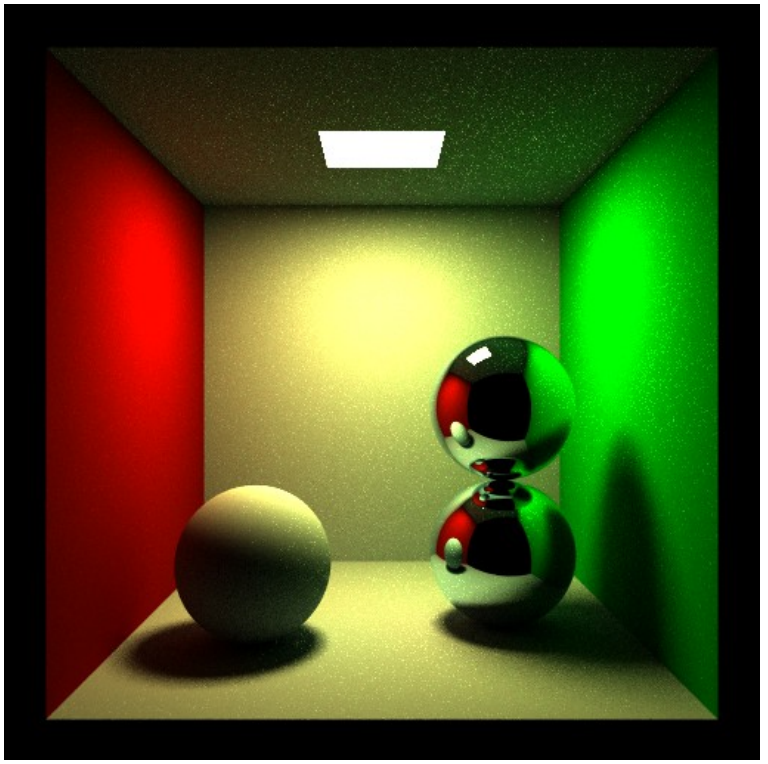
- Useful assumptions:
  - One light only
  - Constant parameters can be set to default values
- Goal: estimate surface orientation (*ie* vector *n)*

Shape from shading

- Non-specular surfaces

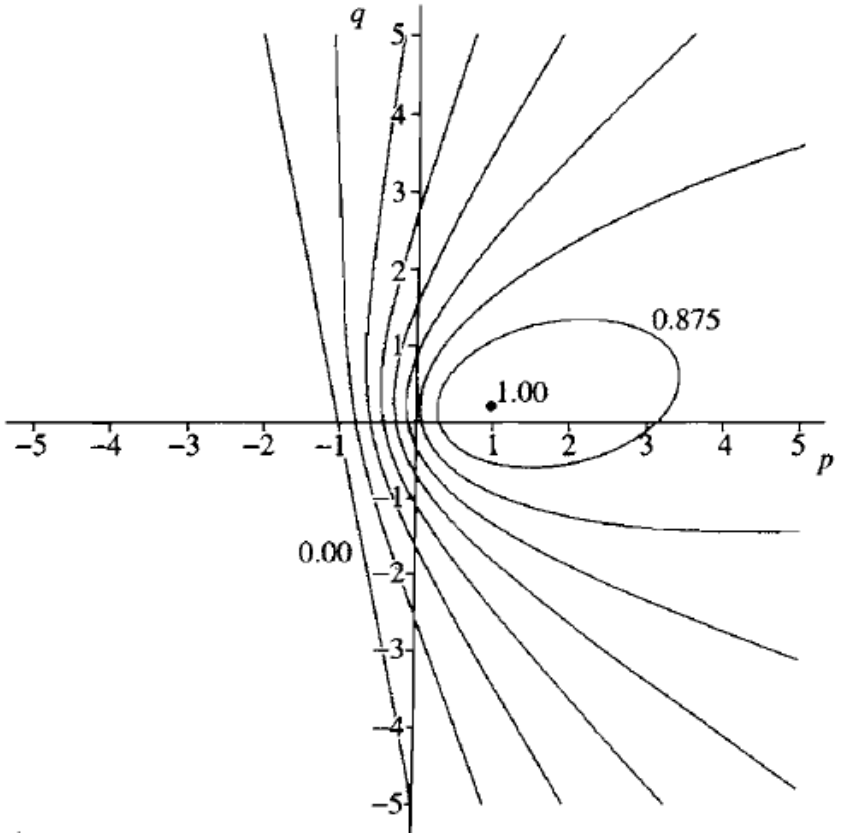


- Specular surfaces

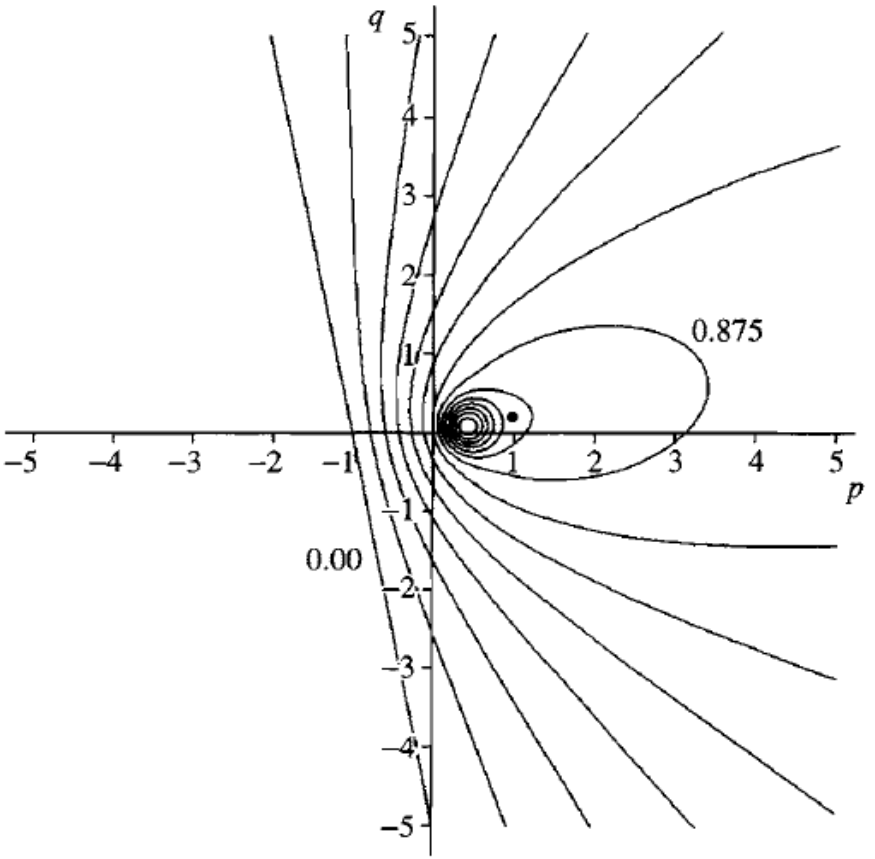


Shape from shading

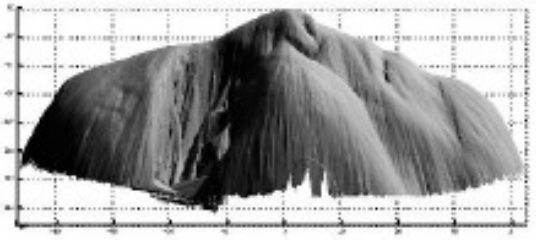
- Non-specular surfaces

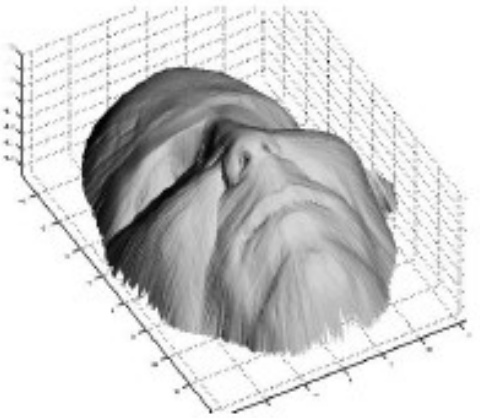


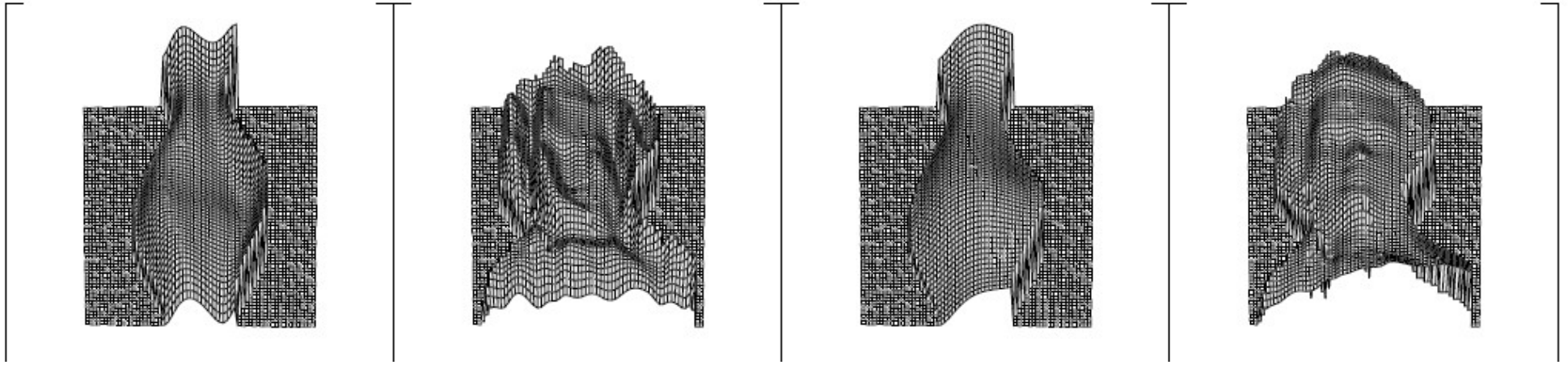
- Specular surfaces



Shape from shading

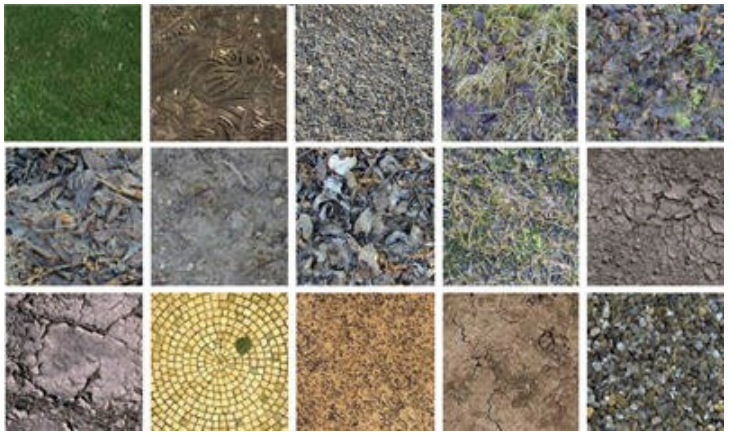




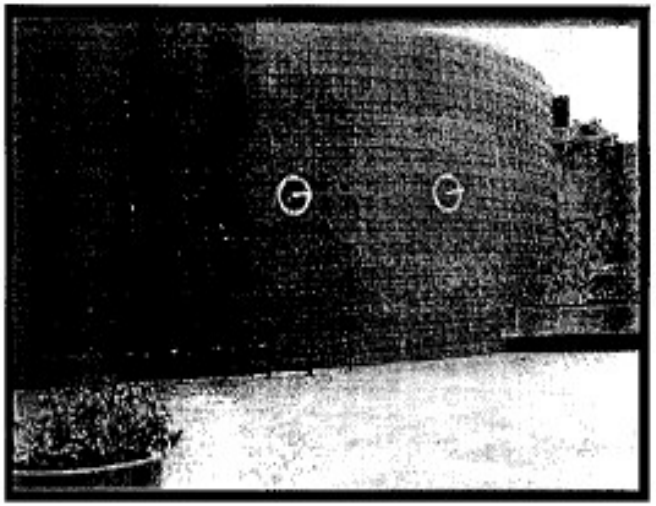


Shape from texture

- Various sizes, orientations, degrees of uniformity, …
- Different approaches (FFT, autocorrelation, ...)

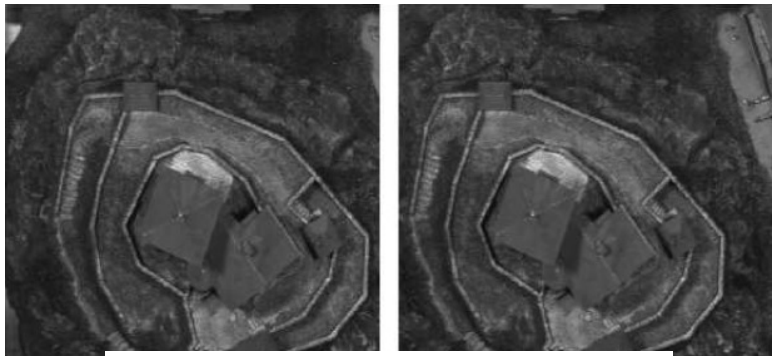


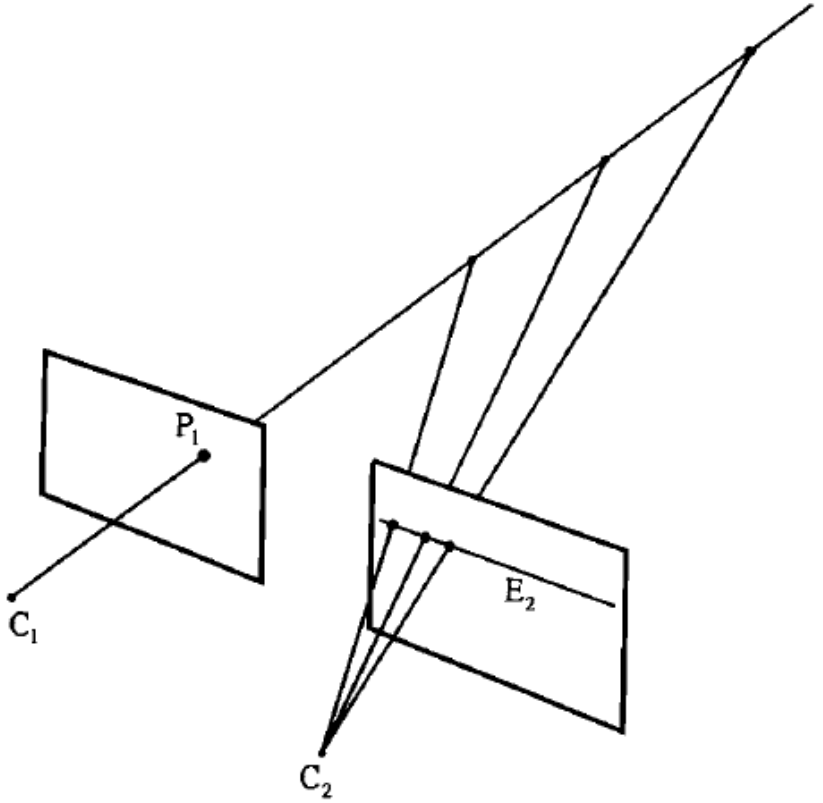
- Can also help to find surface orientation



Shape from stereo

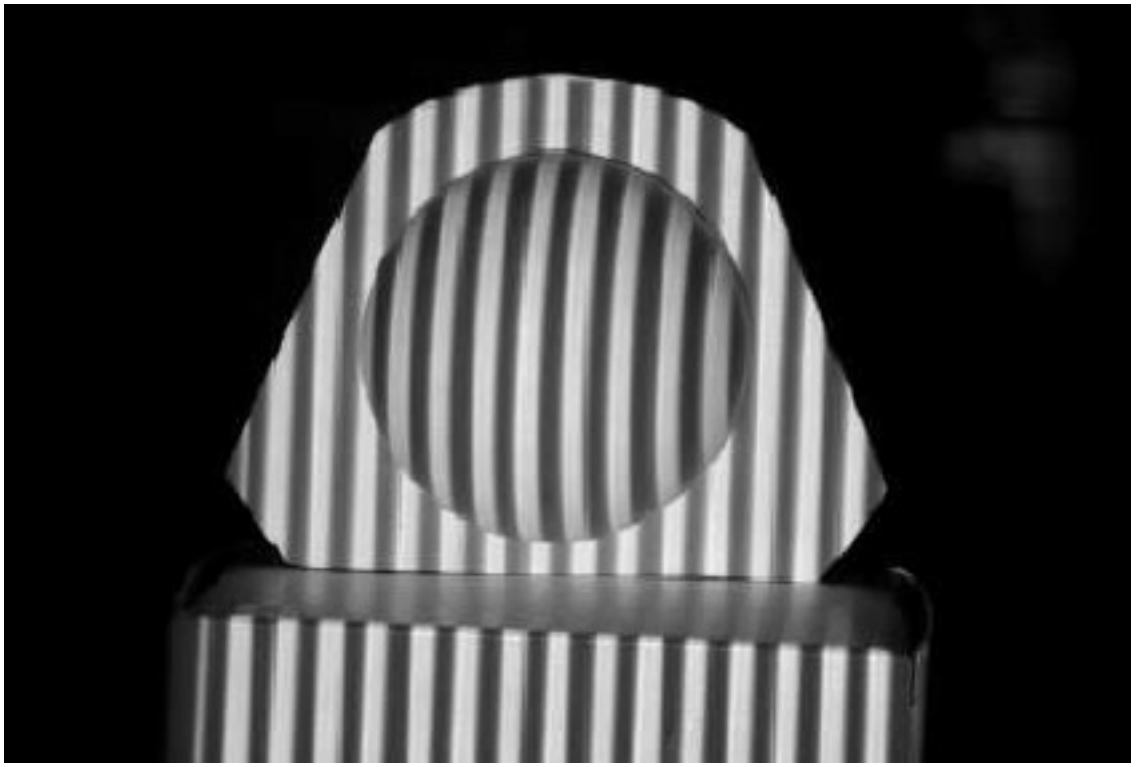
- Try to determine the relative depth corresponding to each pixel by matching two neighbouring images (no moving objects)

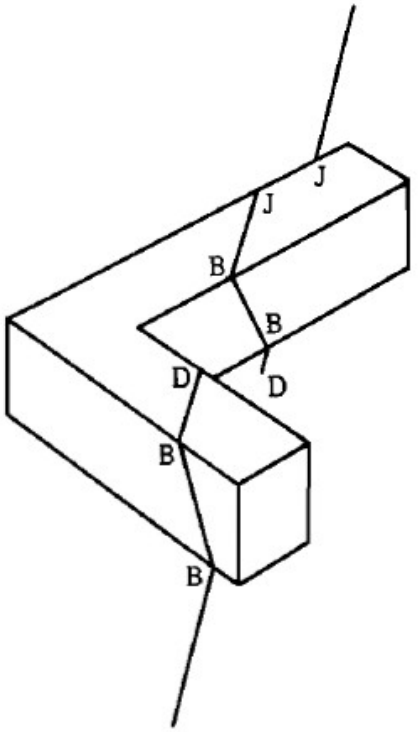




Shape from structured lighting

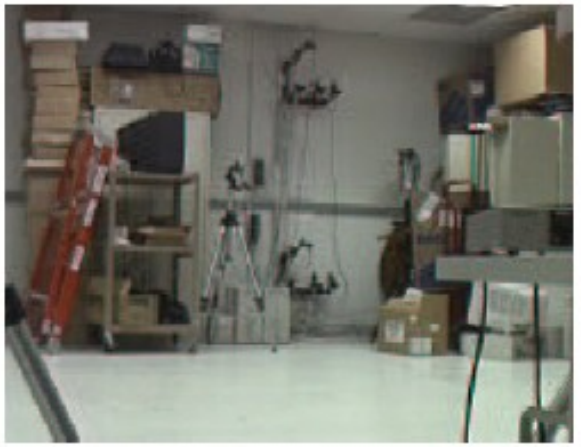
- Use a lighting *pattern* to determine the shape by estimating deformations of this pattern





Shape from silhouette

- Two steps:
  - Silhouette extraction
  - Posture interpretation

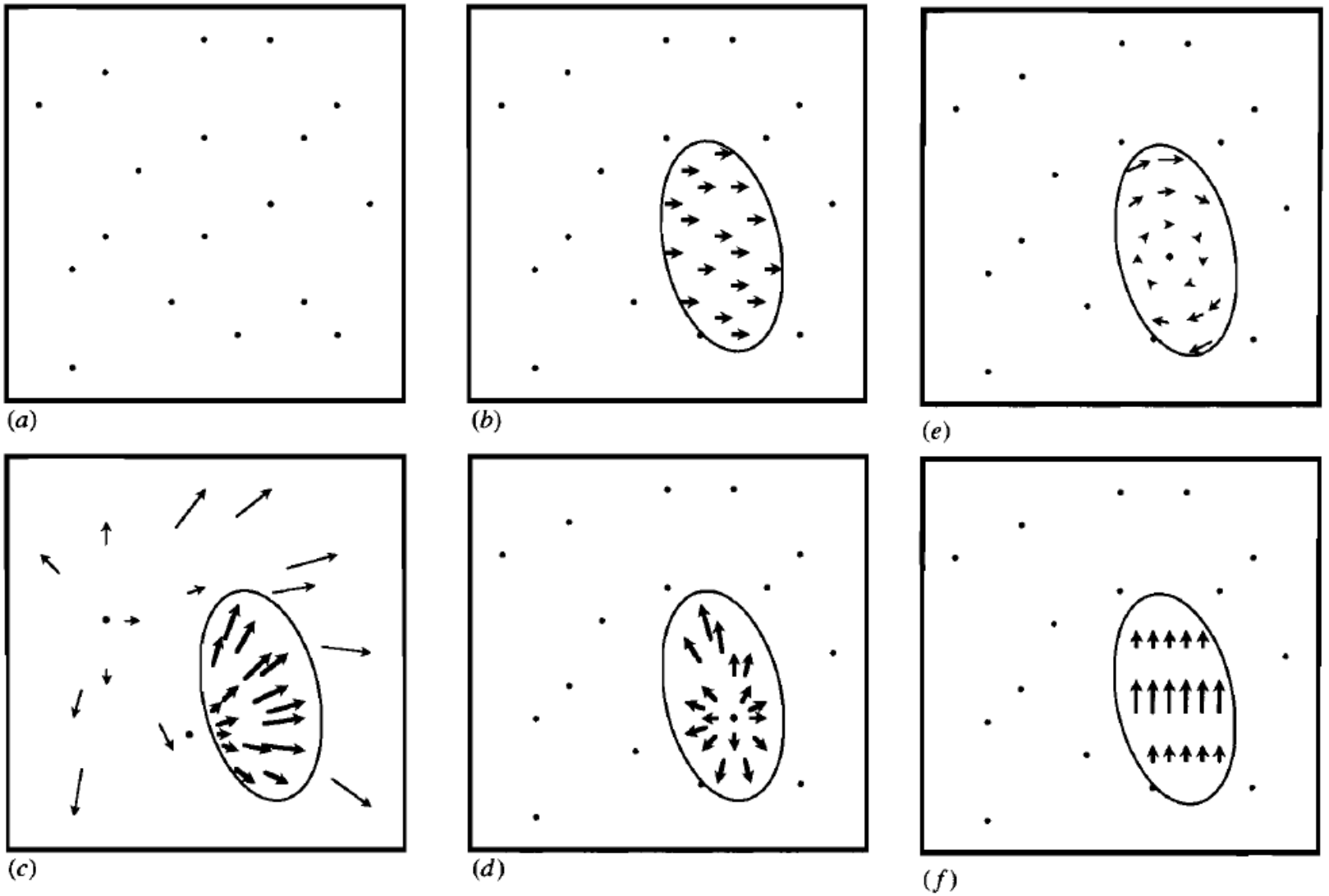


- Video (accv2007)

Optical flow

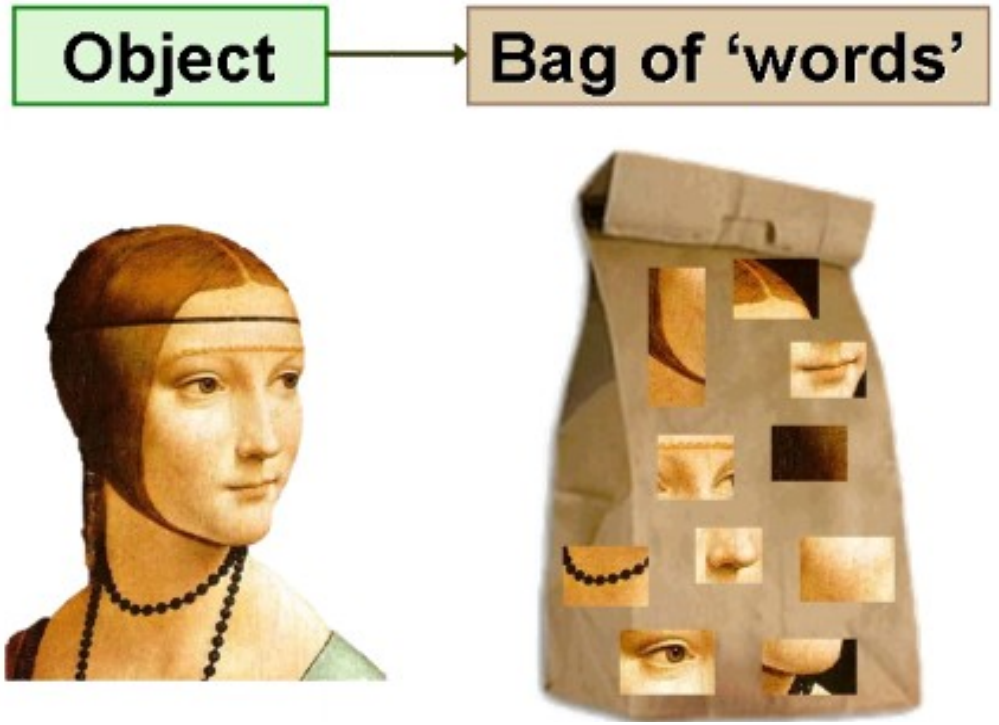
- Try to determine the motion of each pixel in an image *sequence* (fixed camera)
- Assume very small displacements and lighting variations between successive images
- Videos

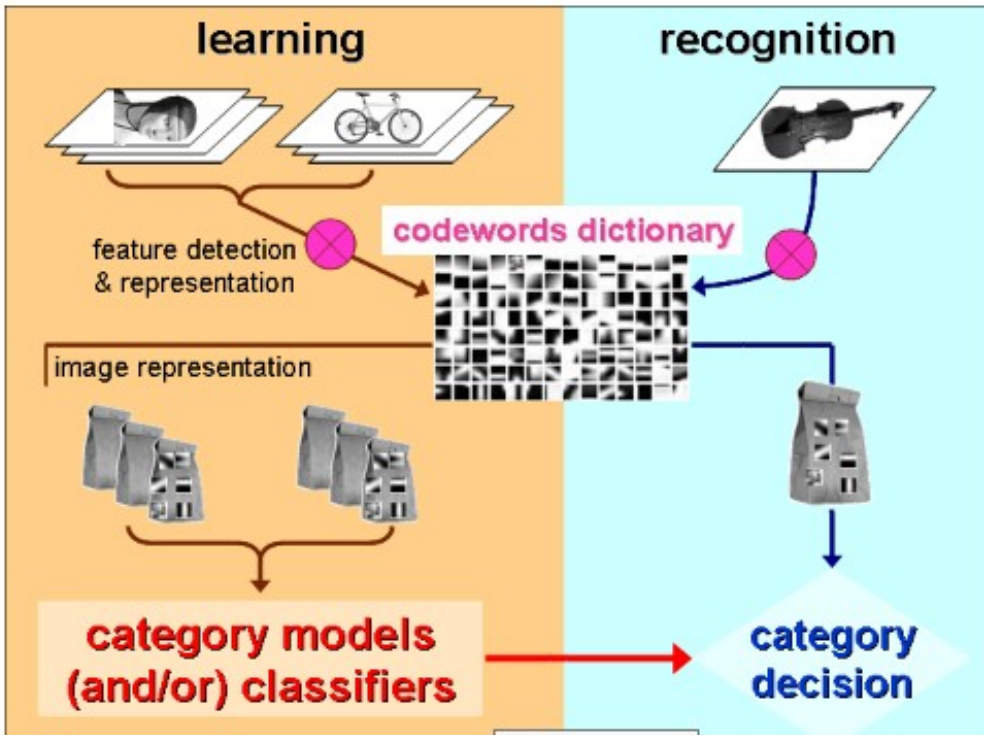
Optical flow interpretation



Bag-of-Words model

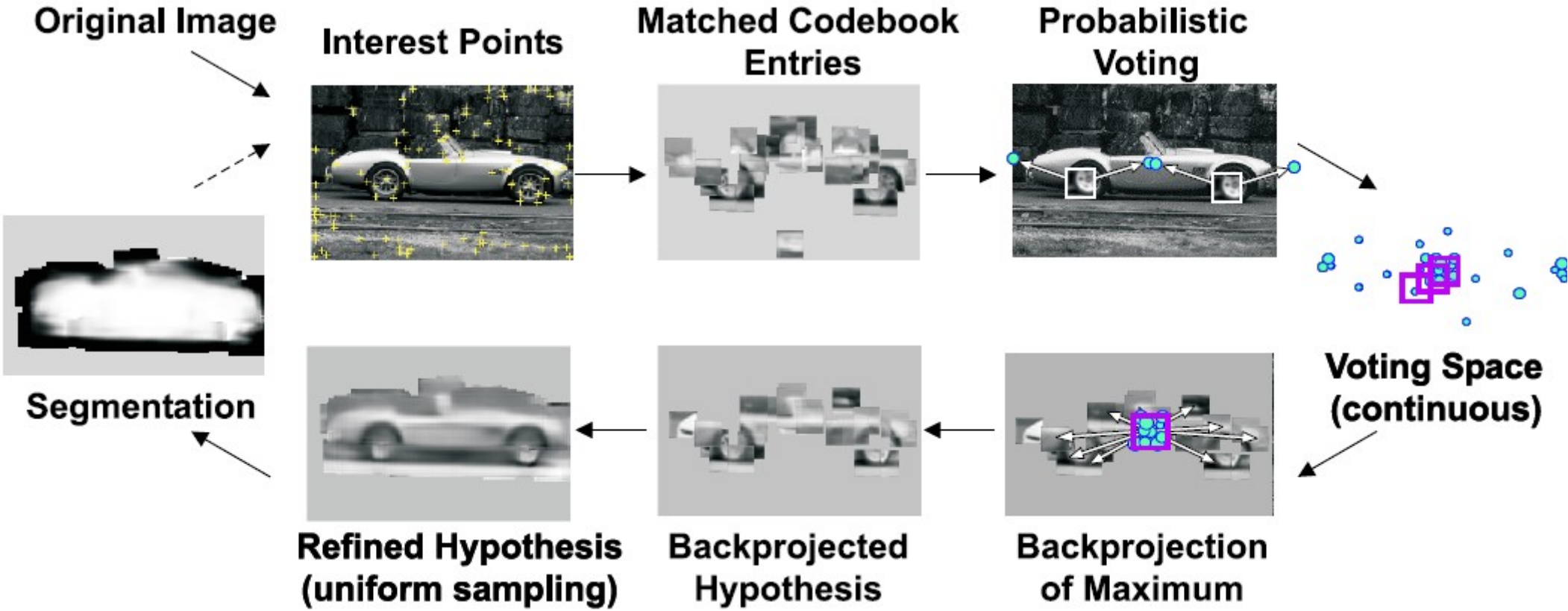
- Analogy with document statistics: a given text can be classified by considering words occurrences only





Combining words and structures

- Components may have structural (spatial) relationships, represented by graphs



Applications

- Position estimation w webcam
- Augmented Reality with ocean waves