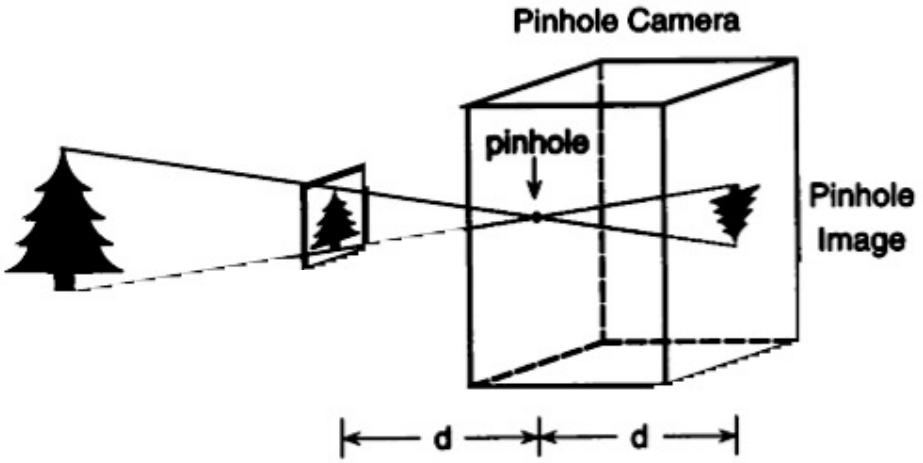
M30242 Graphics and Computer Vision

Lecture 12 Shape-from-X

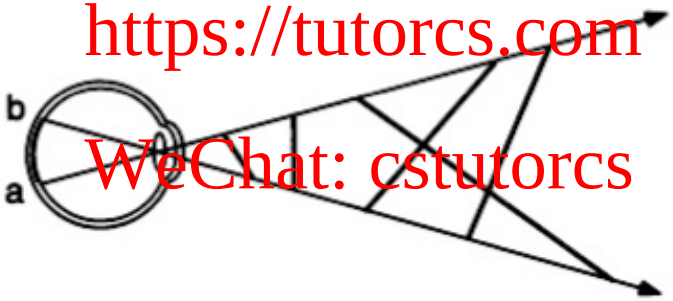
Imaging: From 3D to 2D

- The optics of formation of images suggests that only two of the three spatial dimensions of objects are explicitly present in the 2-D images.
- The third dimension – the distance of an object from the observer, i.e., depth, is lost.



Imaging: Many-to-One Mapping

- Once this information is lost, it is very difficult to regain with absolute certainty.



A single line segment on the retina (arc ***a-b***) can be the projection of an infinite variety of lines in the environment

Depth Perception

- However, people are very good at perceiving 3-D structures/depths **from 2d images**.
- This means that 3D surfaces can indeed be recovered from 2-D images.
- How is this possible?

- Vision research has revealed that different perceptual modalities are at working for depth perception in human vision system:
  - Motion, shading, texture, stereo, edge, etc

Motion Parallax

- Depth information arises from motion parallax – the differential motion of pairs of points due to their different depths relative to the fixation point.

What is the direction of observer's motion? Left or right?

Motion Parallax

- Depth information arises from *motion parallax*: the differential motion of pairs of points due to their different depths relative to the fixation point.

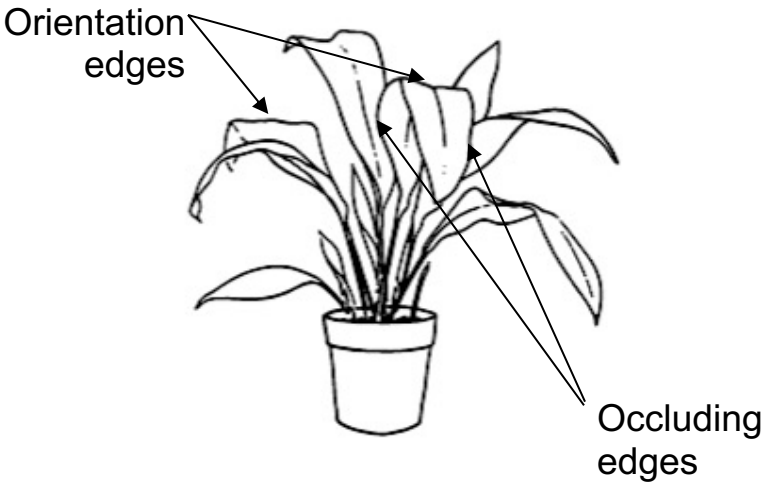
Kinetic Depth Effect

- Depth information about a specific object becomes available not only when the observer moves with respect to it, but also when it moves with respect to the observer.

video

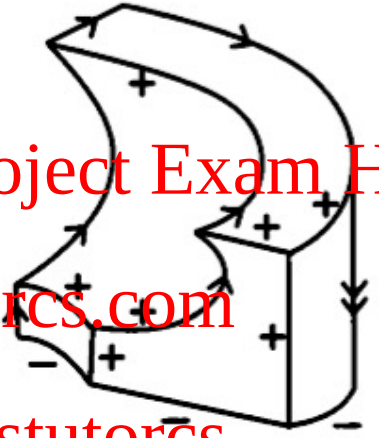
Edge Interpretation

- Different type of edges, labelling and reasoning
  - Orientation edges: edges where surface orientation changes.
  - Depth edges: place where one surface occludes another.
  - Illumination edges: place where there is a difference in the amount of light falling on the surface, such as the edge of a shadow.
  - Reflectance edges: edge where there is a change in the light-reflecting properties (e.g., colour/texture/material) of surfaces.
  - etc



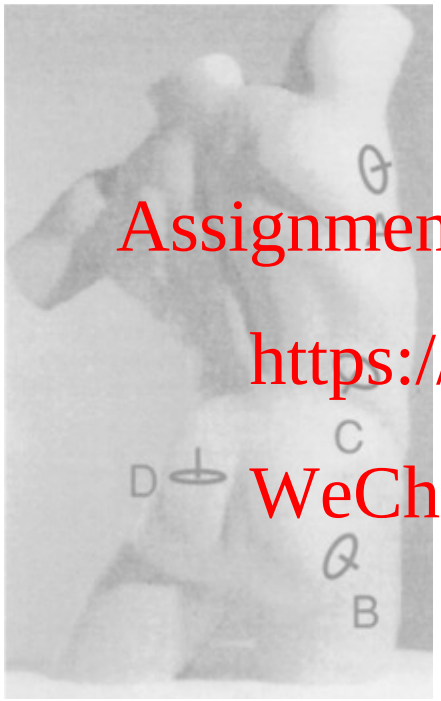
Edge Interpretation

Computer can interpret shape like this.



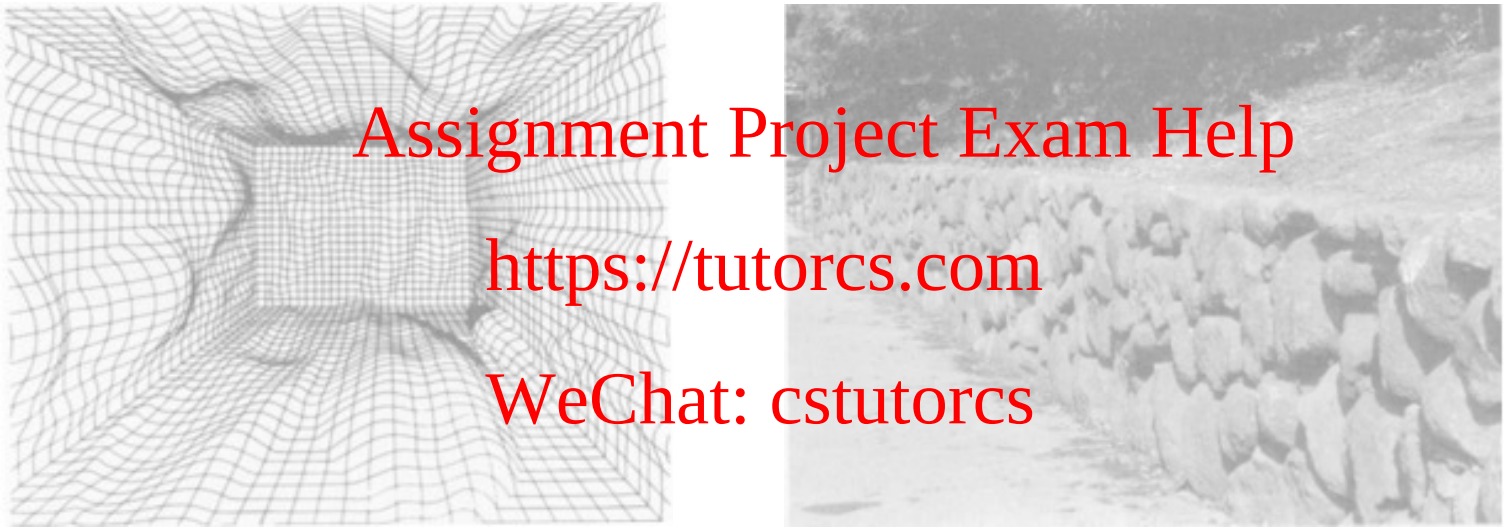
But not this

Shading



- One can judge the direction of surface normals according to shading info.
- Normal directions plus other constraints allow shape of the torso to be recovered.

Texture



- If one assume that the elements (texels) have the same shape and size, then depth can be correctly perceived.

Relative Size

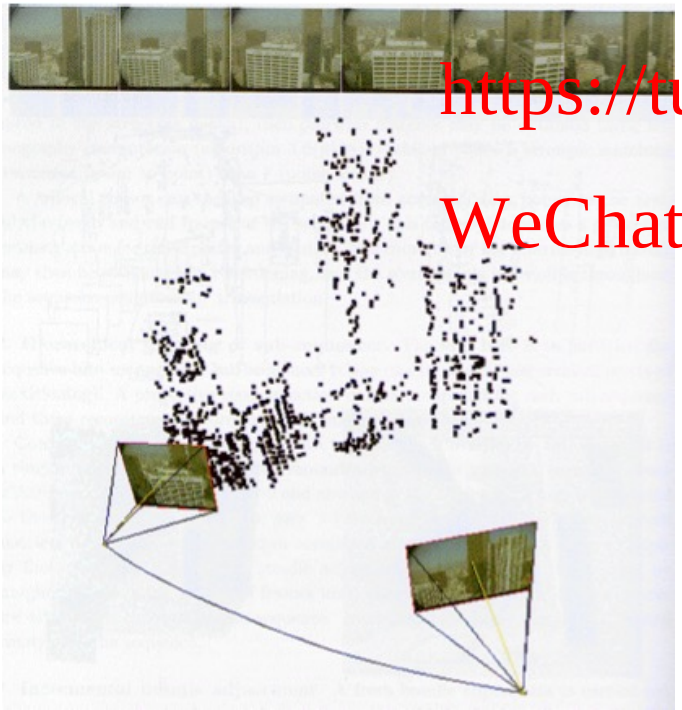
- Depth can be inferred from the known size of familiar objects in the scene

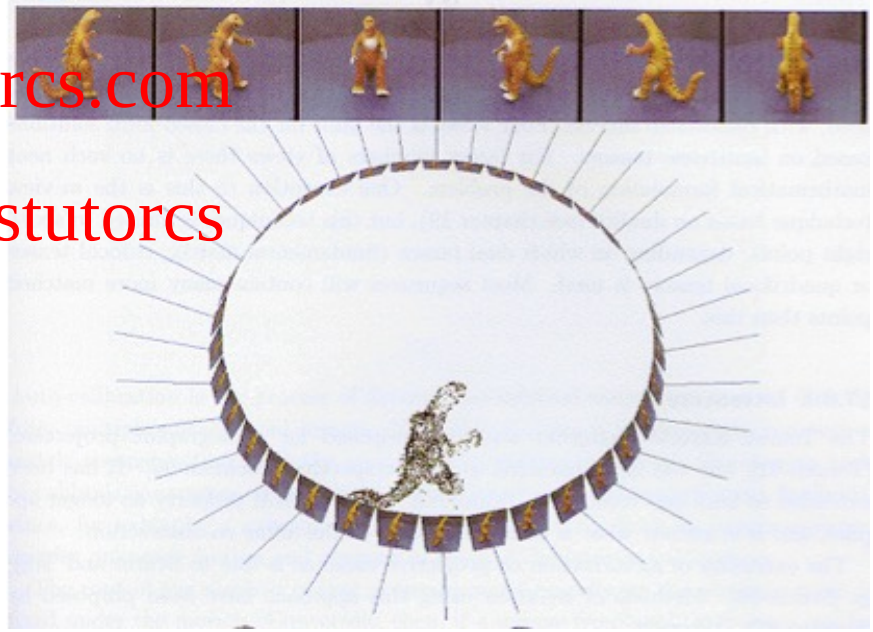
Shape-from-X

- To find a solution for computerised depth perception, scientists have tried different methods to replicate the perceptual process of human vision system.
- A category of methods that utilise different visual cues have been investigated.
- The methods in this category have a common name: Shape-from-X, where X stands for various visual cues.

Shape-from-Motion

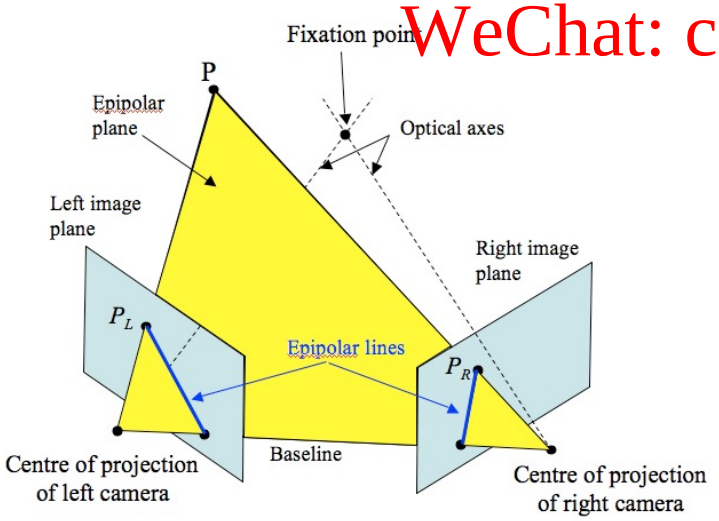
- Compute 3D shape from the relative motion between the cameras and the scene
  - Normally camera is moving

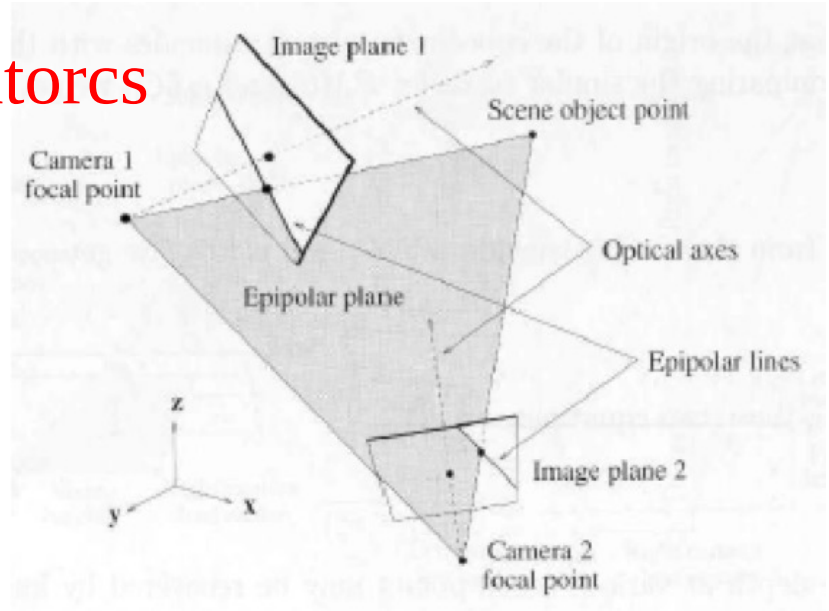




Shape-from-Motion

- Conceptually, shape-from-motion is an extension to the verged binocular stereo.
- The main difference between the two approaches is that in binocular stereo the camera configuration is fixed whereas in shape-from-motion the configuration keeps changing.



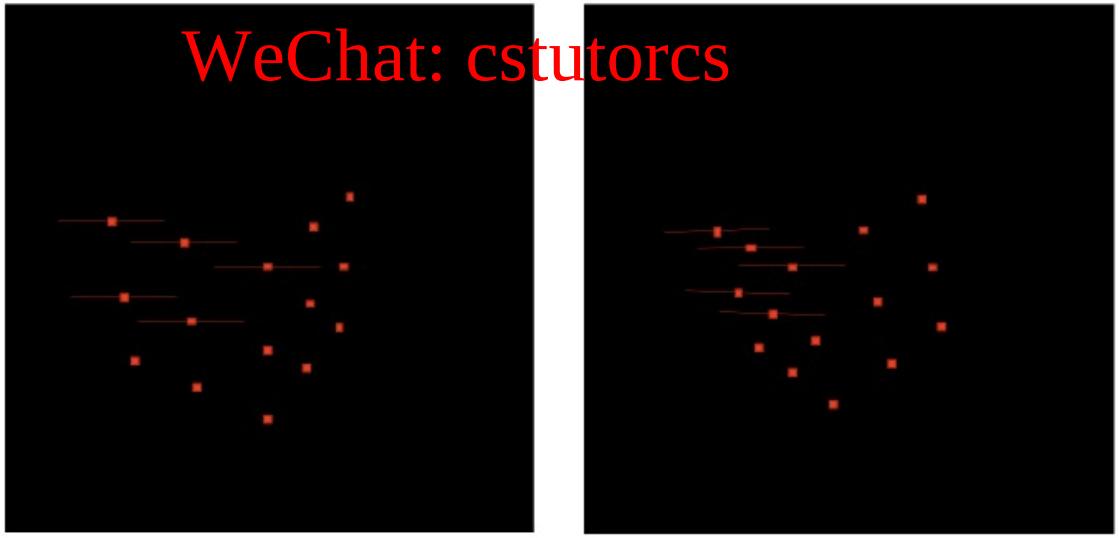


Essential Matrix

- To apply the triangulation algorithm of stereo vision, one needs to work out the configurations of two camera positions at any moment.
- The configuration is determined by two transformations (relative motions) between the two cameras:
  - A translation & a rotation. (In the tutorial on stereo vision, the nonverged system has a displacement d the baseline, but has no rotation involved).
- These transformations can be encoded in a single matrix called *essential matrix.*

Essential Matrix

- Essential matrix is a 3x3 matrix with 9 unknowns
- It can be calculated from 8 or more corresponding points or features from two images.
- So, identifying the correspondence is the very first thing to do in the calculation of the essential matrix.



Procedure

- Find interest points in each image.
- Establish the correspondence between the interest points in image pairs.
  - Use any methods mentioned last week
- Compute essential matrices from the correspondence pairs.
  - algorithms for this are readily available.
- Compute depth by triangulation
  - matching (again) features along epipolar lines.

Summary on Shape-from-Motion

- Which cues are utilised?

- What assumptions are made?

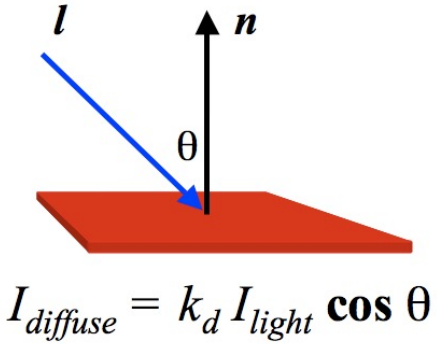
- Compare with binocular stereo

Summary on Shape-from-Motion

- Which cues are utilised?
  - Features (lines, corners, etc) and their displacements.
- What assumptions are made?
  - Objects/scenes are rigid (i.e., the features do not change their relative positions when images are taken).
  - Otherwise we cannot find the camera parameters (essential matrix).
- Compare with binocular stereo
  - Similar:
    - Use multiple images, information of cameras (position & orientation), principle of triangulation, correspondence problem.
  - Different:
    - Camera positions are arbitrary.
    - The relative camera position must be found through feature matching.
    - Need to solve essential matrix.

Shape-From-Shading

- Shape-from-shading uses the pattern of shading in ***a single image*** to infer the shape of the surface.
- The method utilises the relationship between image intensity and surface orientation - Lambert's cosine law for matte surfaces.
- Lambert's Law
  - For a Lambertian (matte) surface, the brightness of the surface depends only on the angle between the surface normal and the light source.



Shape-From-Shading

Read the handouts and answer these questions:

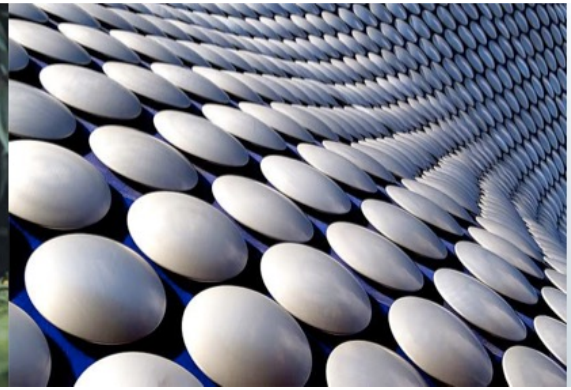
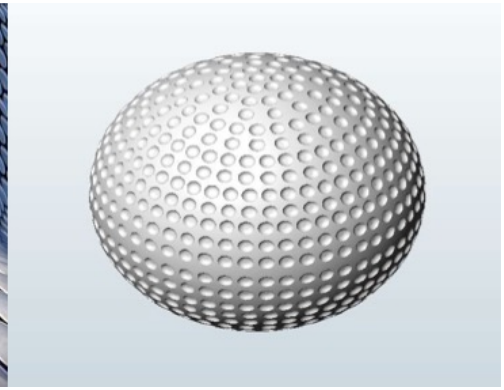
- What cues are utilised?
- How the cues are related to the shapes of surfaces?
- What assumptions are made?

Shape-From-Shading

- What cues are utilised?
  - Image intensity
- How the cues are related to the shapes of surfaces?
  - Changes in intensity (gradients) are related to surface normal (orientation)
- What assumptions are made?
  - Lambertian surface: (diffusely) reflects light equally in every direction.
  - Light source is known.
  - Both assumptions are too strong.

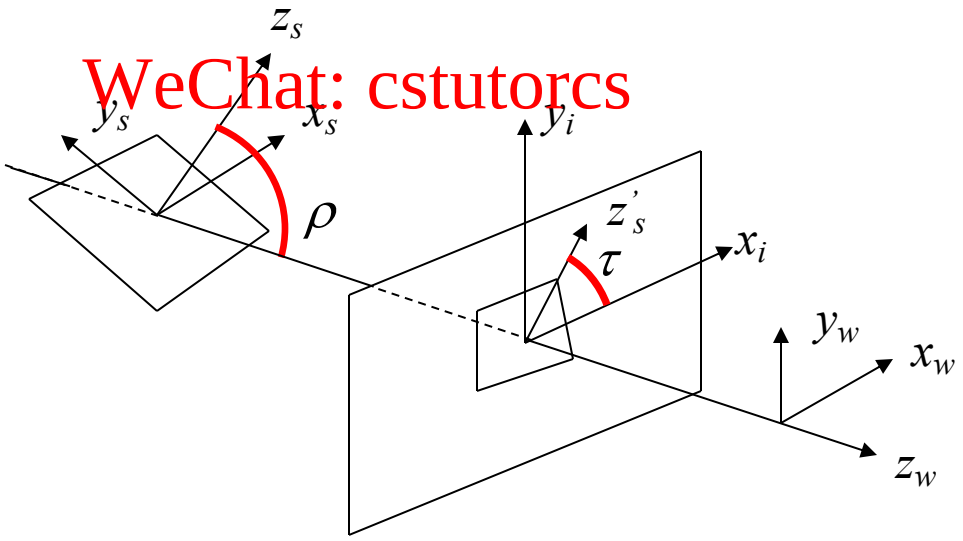
Shape-From-Texture

- Shape-from-texture is also a 3D reconstruction method that works with ***a single image***.
- It analyses the elements of repetitive patterns, called texel, to recover 3D information.
- The shape of a texel (or its deformation) gives its orientation. In this sense, texels work in the same way as shades.

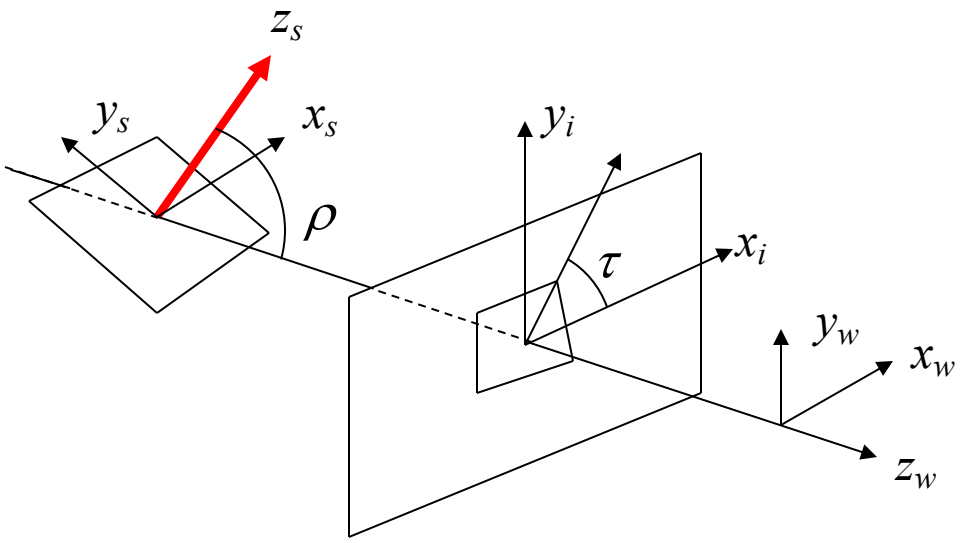
Shape-From-Texture

- In order to measure the orientation of the texels, we need to find the *slant* and *tilt* angles, ρ and τ.



Shape-From-Texture

- With the slant and tilt of a texel known, the direction of the normal z_s can be found.
- From the normals of all texels, the surface can be reconstructed in a way similar to shape-from-shading method.



Shape-From-Texture

Read the handouts and answer these questions:

- What cues are utilised?
- How the cues are related to the shapes of surfaces?
- What assumptions are made?

Shape-From-Texture

- What cues are utilised?
  - Texture elements – texels.
  - Distortion of texels: size, foreshortening.
  - Rate of change of texel distortions - texture gradient.
- How the cues are related to the shapes of surfaces?
  - Distortions -> surface orientation/normals.
  - Texture gradient -> surface curvature.
- What assumptions are made?
  - Texels have the same shape and size – stationary.
  - Evenly distributed (smooth texture) – homogeneity.
  - The distribution of edges of texels are equal in different orientations – isotropy.

Summary on Shape-from-X

- Each of shape-from-X methods uses a simple source of information:
  - Shading,
  - Texture,
  - Motion,
  - Boundary and etc.
- Each is based on a mathematical model, which specifies the relationships between a type of visual cues and a small patch of surface (or its orientation).
- From the visual cues, the math model and other constraints, 3D surfaces can be reconstructed.
- Many publications on the subject in journals and online.