Summary 3D Vision (SS2016)

by 1026803

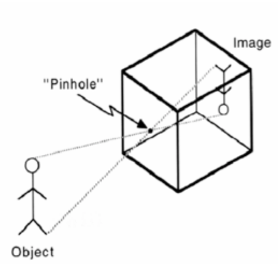
1. Image Acquisition

1.1. Human Eye

Retina, Fovea (center of view), Optic Disc, Lens, Cornea, Iris, Pupil, Muscles;
Rods (brightness), Cones (Colors)

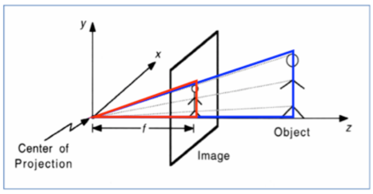
1.2. Pinhole Camera

- Aperture ≈ 0
- upside down and flipped image
- "Camera Obscura"
- Perspective Projection



1.3. Perspective Projection

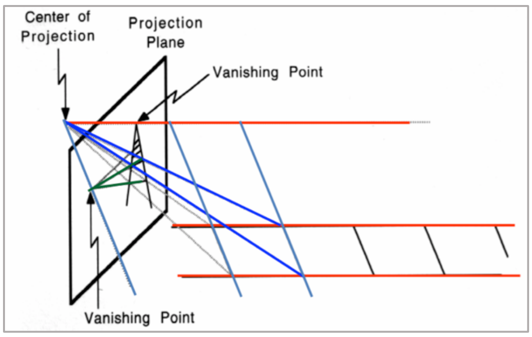
- Angles are not preserved
- Non-Linear Projection!
- Points → Points,
  Lines → Lines,
  Planes → whole/half image
- Degenerate cases:
  Plane through focal point → line, etc.
- Use case: Structure from motion



$$\frac{x}{X} = \frac{f}{Z}$$

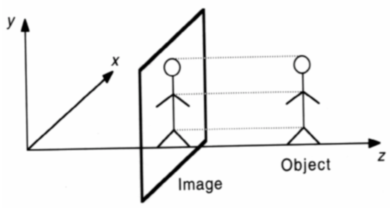
$$\frac{y}{Y} = \frac{f}{Z}$$

Vanishing Point:
Each set of parallel lines meet in a different point (Vanishing point)

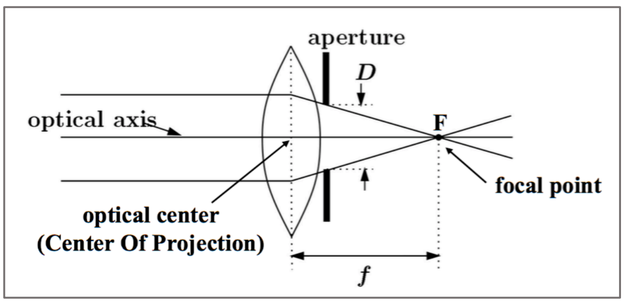


1.4. Normal Projection

$$x = X\,,$$
$$y = Y$$



1.5. Lens



Assumptions for Thin Lens Equation

- Spherical Lens surface
- Small angles of light rays to optical axis
- Small lens
- Same refractive index on both sides of lens

Then: $\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$

1.6. Focus and Depth of Field

- Only objects in a certain distance are imaged sharply at the image plane
- The bigger the aperture, the bigger the blur circles
- The smaller the aperture, the sharper is the image (too small → diffraction, dark)

1.7. Radiometry

- **Radiance** = Amount of light that is reflected by a surface point
- **Irradiance** = Amount of light that is projected from this point onto the image
- Unit = W/m^2

1.8. Camera Sensors

- CCD – integrative method
- CMOS – non-linear method, directly addressable

1.9. Shannon Theorem

Exact reconstruction of a continuous-time baseband signal from its samples is possible if the signal is band-limited and the sampling frequency is greater than twice the signal bandwidth.

$$f_{Sampling} > 2 \cdot f_{max}$$

1.10. Camera Problems

- Optical Distortion
  - Blur
  - Lens glare (tray inter-reflections of light when very bright sources are present)
  - Vignetting
  - Aberration (Geometrical, Chromatic)
  - Lens distortions
- CCD artifacts
  - motion blur
  - blooming (e.g. high dynamic light in CCD) / smearing (lines of image)
- Gamma distortion

Aberrations

2 Types:

- Geometrical (Spherical, Astigmatism, radial/tangential Distortion, Coma)
- Chromatic (refractive index is function of wavelength)

Further Parameters: White-Balance, Color, SNR, Resolution, Thermal-/ Photon Shot Noise

2. Camera Calibration

2.1. Calibration

Correct 3D information from 2D images

Extrinsic Parameters (location and orientation)

- 3 Euler angles
- 3 Translational vector components

Intrinsic Parameters (pixel coordinates w.r.t. camera reference frame)

- Focal length f (distance image plane – projection center)
- Lens distortion coefficient κ (κ_1, κ_2)
- Scaling Factor s (sampling factor in x-direction)
- Principal Point C_x, C_y (Intersection of optical axis with image plane)

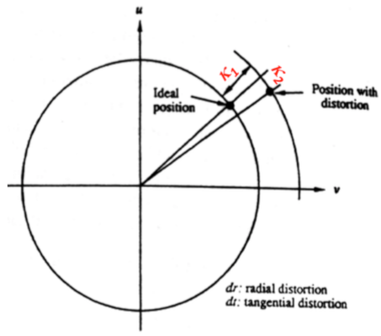
2.2. Geometrical Aberrations

- Spherical aberration
- Astigmatism
- Distortion
- Chromatic aberration

→ Can be reduced by combining lenses

2.3. Lens Distortion

- Radial distortion κ_1 (barrel/pincushion)
- Tangential distortion κ_2



2.4. Principal Point

Determination in **3 ways**

- Mathematical model fit (radial distortion)
- Direct optical method (Laser beam is projected through the lens onto the sensor)
- Variation of focal length (a: variation of image plane, b: Use 2 lenses)

2.5. Calibration Procedure

Linear with Target

- Use calibration plate with known structure
- Positioning in view of camera and find edges
- Calculate parameters

Image Processing

- Canny edge detection
- Straight line fitting to detect long edges
- Intersection of lines to find image corners
- matching of image corners and 3D checkerboard corners

Nonlinear Methods

Linear methods have problems: too many parameters, does not model lens distortion
→ Tsai Calibration. Covers all intrinsic and extrinsic parameters

4 Steps:

- World COS → Camera COS (R,T)
- Camera COS → undistorted image COS (f)
- Consider lens distortion (κ)
- Metric image COS → pixel image COS (C_x, C_y, s)

2D Motion

- Structure from motion
  - Track points over sequence of images
  - calibrate internal parameters beforehand
- Self-Calibration
  - Ultimate Solution (-;

Radiometric Calibration

Radiometric Errors:

- Different Sensitivity
- Same Brightness → Scaling coefficient for every pixel
- Different Brightness → Gamma correction

3. Range Scanner

3.1. Time of Flight Range Finder

Determine distance by runtime measurement (optical or ultrasound)

- Transmitter → Deflector → Receiver → Phase comparator
- Other method: Sending light pulses

3.2. Laser Radar ToF

$d = \frac{c \cdot t}{2} \Longrightarrow$ for 1 meter: t = 6.7 ns!

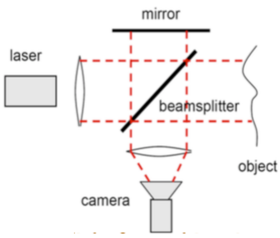
Phase detection leads to ambiguities by $n \cdot \frac{\lambda}{2}$ (solution: sweeping to finer wavelengths)

→ Continuous wave needed

Time of Flight	Pulse-based (AMCW)
Large distances: up to 1000 m	Distances up to 100 m
Lower measuring speed: 1000-10000 points / s	High measuring speed: up to 650000 points / s
Short pulse: eye save	Reflectance of material also determined
Absolute depth measured	Ambiguity of distance

	Measurement uncertainty
Triangulation Range camera	0.1 mm
TOF based laser scanner (AM, PM)	2-20 mm
Laser Radar (FM CW)	0.1 mm

3.3. Interferometry



- Coherent laser light is split into two paths (reference mirror, object)
- Beams are added again (Interference)

High accuracy (nm), works only with smooth/mirror-like surfaces

3.4. Ultrasound/Infrared Range Finder

Advantages

- Illumination independent
- Low speed of measurement beam

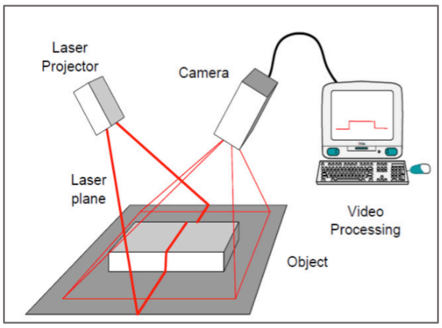
Disadvantage

- Poor resolution
- Low accuracy
- Deflector necessary

Applications: Car Parking radar, filling measurement, underwater meas.

3.5. Triangulation Range Finder

- Projection of light plane into measuring area
- Distortion defines distance from camera via triangulation



Coded light approach

- Moiré techniques
- Pattern projection
- Gray-Code
- Color coding

Problems

- Occlusions
- Contrast and sharpness of laser
- Speckle noise from laser

3.6. Types of Range Finder

- Spot Projection (point by point measurement and triangulation)
- Light point stereo Analysis (use 2 cameras, laser point)
- Light strip range finder (projecting light planes)
- Shadow scanning (like light strip but using shadows)
- Pattern projection (use pattern instead of point/line)
  - Random patterns
  - Encoded patterns (e.g. gray code, phase/freq./amp. modulation, color coding)

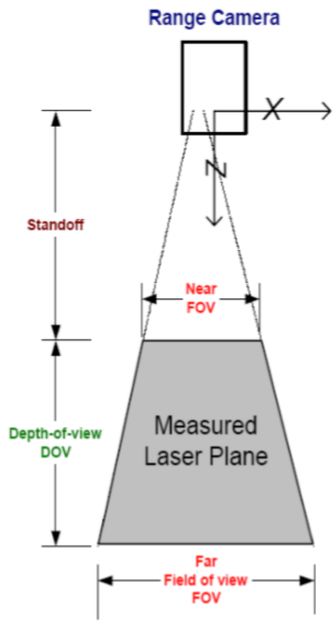
3.7. Errors in optical triangulation

- Laser width limits accuracy (e.g. at sharp edges)
- Different surface colors/reflectivity
- Occlusions (solution: use more lasers or cameras)

3.8. Specifications for 3D Scanners

- Standoff
- Depth of view (DOV)
- Near field of view (near FOV)
- Far field of view (far FOV)

- Accuracy
- Reproducibility
- Uncertainty of Precision
- Systematic Errors
- Random Errors
- Resolution



4. Shape from Monocular Images

Shape from	How many images	Method type
Stereo	2 or more	passive
Motion	a sequence	active/passive
Focus/defocus	2 or more	active
Zoom	2 or more	active
Contours	single	passive
Texture	single	passive
Shading	single	passive

4.1. Shape from Shading

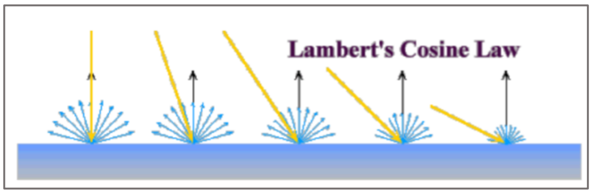
- Shading on the surface gives the depth information
- Surface reflection of untextured objects includes depth information
- Surface boundaries play crucial role in interpretation by humans

Mechanisms for Reflection

- Body Reflection: Diffuse, matte, non-homogeneous - (e.g. paper, clay)
- Surface Reflection: Glossy, specular (=mirror-like) - (e.g. metals)

  ➔ Many materials have both types

For simplification normal projection is always used (object far away and close to optical axis)



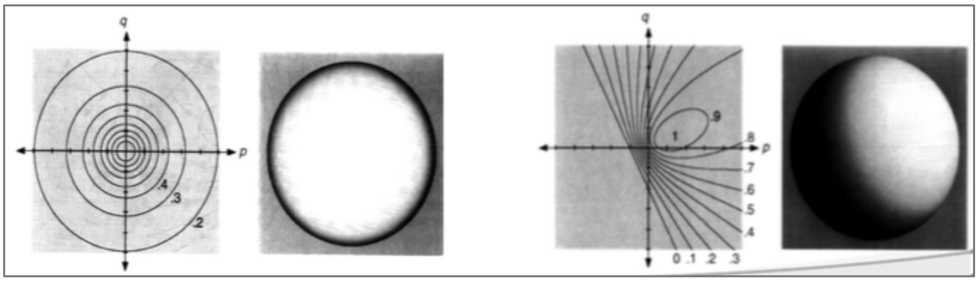
4.2. Reflectance Map

Lambertian Surface

Brightness depends only on the direction of illumination, not observation

Reflectance Map

2D plot of gradient space (p,q) of normalized image brightness of a surface as function of surface orientation.



Straight line is called Terminator and separates illuminated from shaded regions.

Problem: in reality combination of matte and specular reflection

- Weighted average of diffuse and specular component
- Reflectance map must be determined experimentally
- not possible for general shape from shading (with known standard forms like ellipse, parable, hyperbola and line/terminator)

Another Problem: Rounded corners lead to overshooting

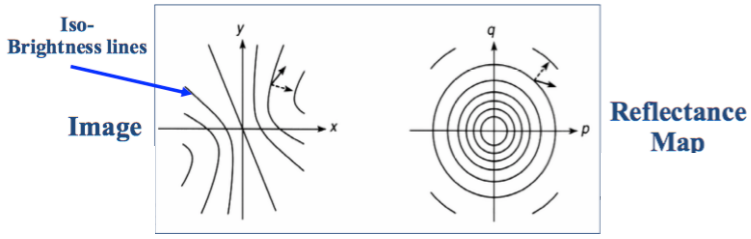
4.3. Shape determination in Shape from Shading

Methods: Strip Method, Photometric Stereo, Polarized Light

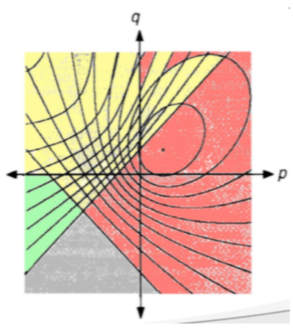
Strip Method

- For each brightness value of a pixel → Reflectance Map restricts surface orientation
- Strips of equal brightness in the picture = height lines
- Starting point with known surface normal
- Small movement in the direction of greatest change in brightness
→ small movement in the direction of greatest slope
- Requires one or more starting points

Disadvantage: Errors cumulate, no stable solution (depending on starting point)



Photometric Stereo



Basis: 2 images with same geometry but different illumination directions

- 1 Reflectance Map limits surface orientation only by one iso-brightness contour (many solutions possible)
- 2 RM restrict possible directions of a surface normal to 2 candidates (intersection of lines in gradient space)
- Clear solution by using third light source
- Practical application: using colored light or use of a chrome sphere

Procedure

1. Estimate light source directions
2. Compute surface normals
3. Compute albedo values
4. Estimate depth from surface normals
5. Relight the object (with original texture and uniform albedo)

Prerequisites for SfS

- Surfaces with constant albedo → rotationally invariant
- Orthographic projection
- Distant and calibrated sources of illumination
- No drop shadows
- No reflection illumination - inter reflection

4.4. Shape from Shading Variants

- Shape from Specularity (for highly reflective Surfaces)
- Extension for non-lambertian surfaces by using polarized light
  (exact surface normal but only approximate position)
- Shape from Shadow (Reconstruct surface topography from self-occlusion)

4.5. Shape from Texture

Texture

- Repetition of a basic pattern
- Pattern repetition neither regular nor deterministic (e.g. human made texture), only statistically regularly (e.g. grass, ocean, etc.)

Statistical Texture Analysis

- Suitable for all natural textures
- Used for classification rather than for shape determination

Structural Texture Analysis

For deterministic textures (mostly made by humans), made out of elements called texels.

Shape from Texels

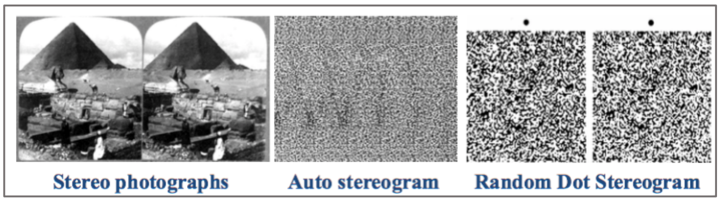
- Is based on the distortion of the single texel
- Texel must be clearly identifiable and must not overlap
- All texels have the same spatial extent
- Texel are "small", i.e. are planar and have unique surface normal

5. Shape from Stereo / Stereo Vision

Correct 3d information using 2d images:

- 2 or more images taken from different positions plus geometric calibration of camera
- Tries to imitate human visual system
- Is also used in the entertainment industry

Examples



5.1. Entertainment Industry

- Dual Displays (Oculus, HTC Vive, etc.)
- 3D Glasses
  - Anaglyph (cyan/red)
  - Polarized Displays
  - Active shutter/Field sequential
- Lenticular Display / Barrier strip Displays

5.2. Stereo Geometry

Objective

- Given two images of a scene acquired by known cameras compute the 3D position of the scene (structure recovery)
- Basic principle: triangulate from corresponding image points

Disparity

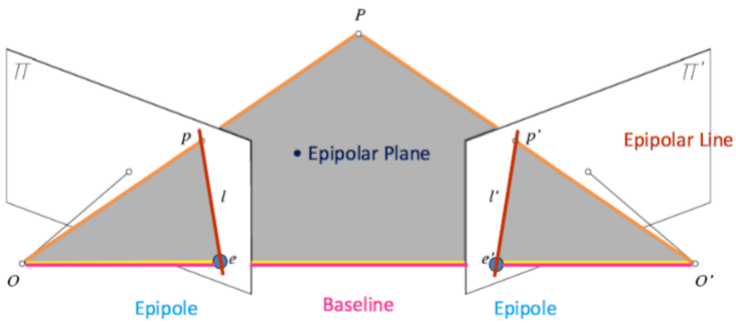
Disparity: $D = x_1 - x_2$

Distance to center of projection: $-Z = \frac{f \cdot B}{D}$

B…Baseline, f…focal length, Z…distance of object point

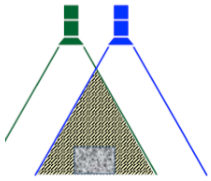
Epipolar Constraint

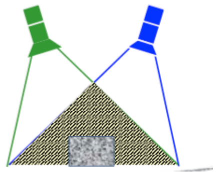
Each point of the left image can lie only on a specific line in the right image: the Epipolar Line



➔ Reduced to 1D search problem!

Rectification





Normal Case

- Disadvantage: small distance between the centers of projection
- Advantage: low computational complexity

General Case

- the larger the distance between projection centers, the more accurate
- but larger distance leads to large occlusion areas

We can always get to the normal case by image re-projection

- Re-project image planes onto common plane parallel to line between optical centers
- Notice, only focal point of camera really matters

5.3. Correspondence Analysis

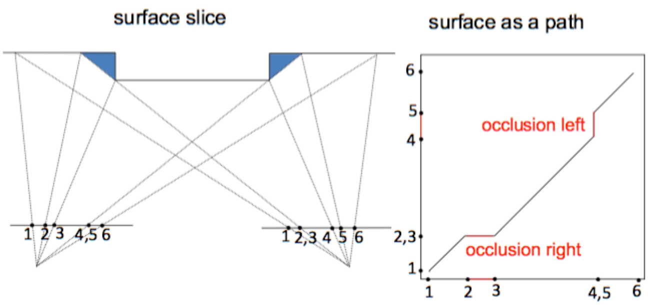
Area Based

- Compare intensity levels of left and right image
- Correspondence due to similarity of intensity levels
- Correspondence for each pixel

Feature Based

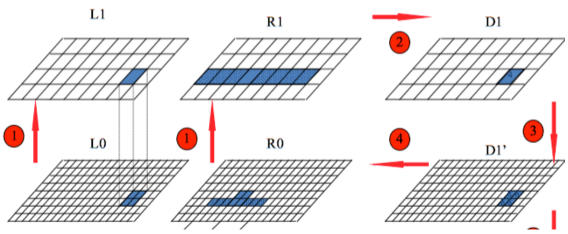
- Compare features of left and right image
- Correspondence on basis of selected characteristics (edge, gradient, etc.)
- Correspondence only for selected Pixels
- more accurate (sub-pixel positioning)

Problem: Point does not exist or is not distinct



5.4. Hierarchical Stereo Matching

- Faster Computation
- Deals with large disparity ranges



5.5. Energy Minimization

- Matching pixels should have similar intensities
- Most nearby pixels should have similar disparities

→ Labeling problem

5.6. Feature-based Correspondence Analysis

- Look for a feature in an image that matches a feature in the other
- Set of geometric features is used (e.g. edges, line segments, corners, etc.)
- Need for interpolation if only sparse set of points available

5.7. Active Stereo

- Feature-based methods cannot be used when objects have smooth surfaces or surfaces of uniform intensity
- Patterns of light can be projected onto the surface of objects, creating "interesting" points even in regions which would be otherwise smooth

Problem: Ambiguity

→ Using multiple cameras reduces likelihood of false matches

5.8. Components of Stereo Vision Systems

- **Camera calibration**: Find inner and outer parameters of cameras
- **Image rectification**: simplifies the search for correspondences
- **Correspondence**: which item in the left image corresponds to item in the right image
- **Reconstruction**: recovers 3-D information from the 2-D correspondences

6. Shape from Multiple Images

6.1. Depth from Focus

- Range from focus using $\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$
- Take pictures along axis
- find image having highest frequency (best focus)
- more than 10 images needed (monocular)
- use Gaussian interpolation to form a set of approximations

Disadvantage: many images must be captured to find best focus → Depth from defocus

6.2. Depth from Defocus

- Assume a blurring function (blur model)
- Diffusion parameter is related to blur radius
- Depth can be computed from the two measurements (2 unknowns: depth and image frequencies)
- Needs textured surfaces
- Active depth from defocus: project pattern onto surface (frequency of scene is then known)

Problem: Ambiguity (object too near or too far)
Solution: Use two sensor planes

If no texture available → project structured lighting onto surface (Active Depth from Defocus)

6.3. Shape from Motion

- Motion of an observer relative to the environment
- Problem: moving direction and amount of camera movement
- Prerequisites:
  - Known moving direction of camera
  - Known speed of camera3rd Dimension = time dimension
- Correct Assumptions → Depth calculation possible

Motion Field
…is characterized by vectors that represent the movement of the corresponding scene points.
If camera does not rotate:

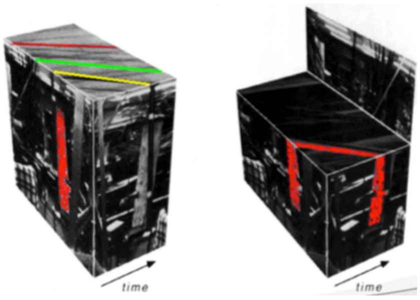
- Vectors point radially to or from a focus
- Point where motion vector of camera intersects image plane:
  - FOE: Focus of Expansion (forward movement)
  - FOC: Focus of Contraction (backward movement)
- Length of the vector is:
  - inversely proportional to distance of point
  - proportional to sine of the angle between moving direction and image point

→ movement zero = FOE or FOC (except for points at infinity!)

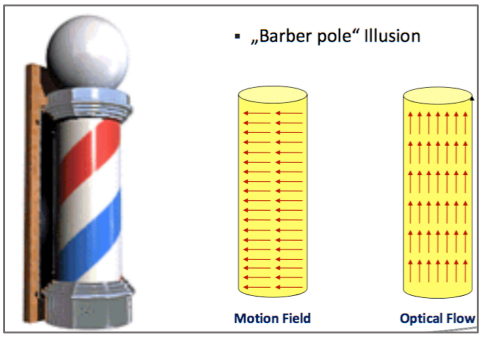
Motion Field Determination
Task: Determination of corresponding points in two images
- sparse vector field
- same problem as stereo - only moving direction of camera not known
- Epipolar line not known at the beginning

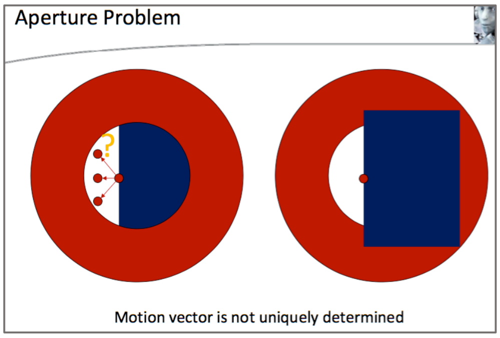
To find correspondence between images:
- high temporary sampling = slight differences
- either unchanged intensities in both images, or unchanged edges in both images

2 Strategies for determination:
- **Spatio-temporal derivation**:
  Intensities do not change in timeline,
  gray values are continuously differentiable
- **Spatio-temporal coherence:**
  Intensity and edges are preserved



Motion Field vs. Optical Flow
Motion Field: Projection of movement onto the image plane
Optical Flow: Observed flow in the image plane (constant brightness constraint)
Assumption: motion field = optical flow





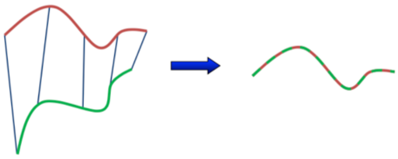
We can only determine the motion parallel to the gradient but not normal to the gradient!

7. Registration
Range Images are 2.5 D – Full 3D is made of multiple range images.
Process of putting single images together → **Registration**

The result of a single scan is called Range Map (depth value for each pixel → point grid) and is an incomplete 3D Model. Multiple shots are needed!

The Scanning Pipeline
1. Scanning (data acquisition)
2. Alignment of data
3. Merging to get single surface
4. Manipulation (simplification, coloring, mesh cleaning)
5. Visualization

7.1. Alignment



- Each part has its own COS
- Objective: bring all parts in common reference system
- First step: roughly positioning to have an overlap region

2 Approaches

- Target based registration
- Surface based registration

7.2. Target based Registration

- Align scans by using reference markers
- Automatic matching possible

Advantage

- Less computational effort
- Geo-Referencing to a higher reference system

Disadvantage

- Longer fieldwork time
- Accuracy

7.3. Surface based Registration

- Only point cloud data used for registration
- uses more scans

Advantage

- Better accuracy
- Optimizing project cost and duration

Disadvantage

- Computationally expensive
- Not well-suited for geo-referencing

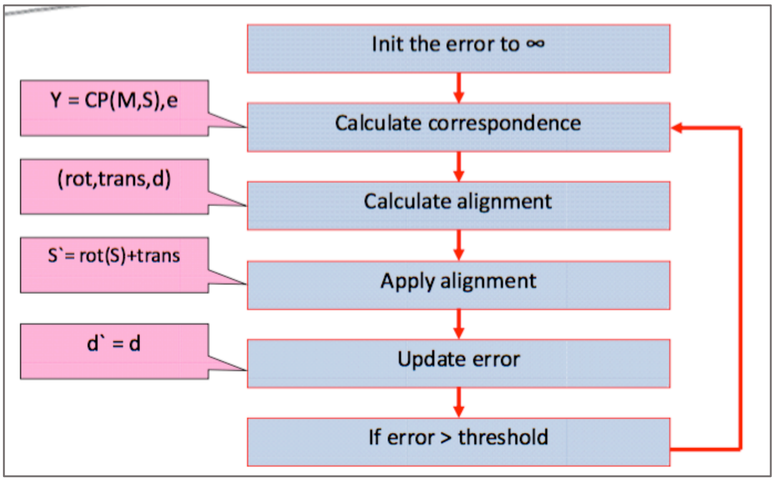
Methods

- Zippering
  - Scans are simply joined to one surface
  - Simple and fast
  - does not use redundancy to eliminate sampling error
- Volumetric Methods
  - Range Maps are mapped in a volumetric grid
- Marching Cube
  - Mostly used of merging software

7.4. Iterative Closest Points

General: Closest Point approach converges if starting position is “close enough”

Algorithm



The ICP algorithm always converges to a local minimum

Variations

- Selection of Points: all available, random samples, uniform subsampling
- Matching: Closest point, normal shooting, reverse calibration, include color/intensity
- Weighting of Pairs: w.r.t. distance, compatibility of normal vectors, scanner uncertainty
- Rejecting Pairs: w.r.t. distance, worst n%, points at and of lines, not consistent neighboring pairs
- Error Metric: sum of squared distance, SVD, orthonormal matrices

8. Space Carving

Algorithm

- Initialize a volume V containing the true scene
- Choose a voxel on the surface
- Project to visible input images
- Carve if not photo-consistent
- Repeat until convergence

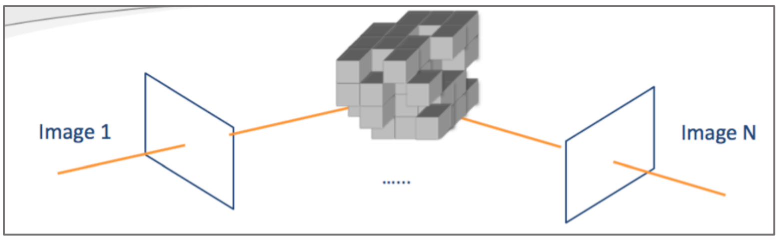
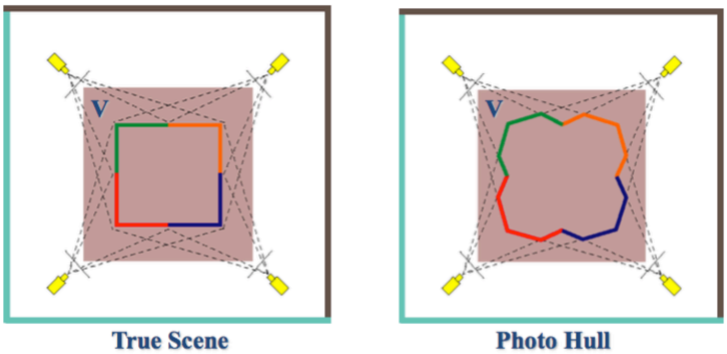


Photo Hull

…is the union of all photo-consistent scenes in V (tightest possible bound on the scene)

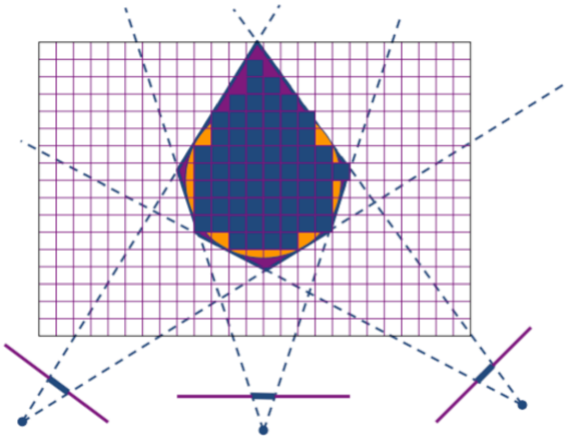


9. Shape from Silhouette

Basics

- Silhouette of object contains 3D information
- Only binary images used
- Voxel is photo-consistent if it lies inside silhouettes of all views

<u>Final model</u>: intersection of all conic volumes



Visual hull is a good starting point for optimizing photo-consistency

- Easy to compute
- Tight outer bound
- parts already lie on the surface and thus are photo-consistent

Algorithms

- Voxel based method (standard)
- Marching intersections
- Image-Based visual hulls
- Exact polyhedral methods

Strengths

- Reconstructs visual hull of object that is never larger than model
- can reconstruct handle (of a cup?)

Weaknesses

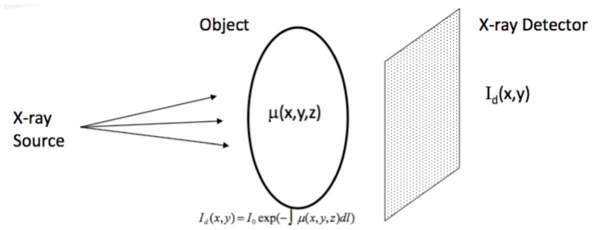
- Unable to reconstruct concavities
- Flat surfaces reconstructed poorly
- Sufficient for convex, oval objects

10. Medical 3D Scanning / Volume Scanners

Types

- X-Ray Projection (Radiography)
- X-Ray Computed Tomography (CT)
- Magnetic Resonance Imaging (MRI)
- Nuclear Medicine (PET)
- Ultrasound

10.1. Radiography

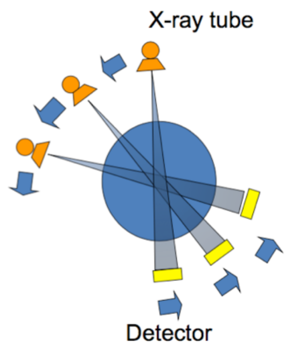


X-Ray density increases for different tissues

- Air (minimal absorption)
- Adipose tissue
- connective tissue of organs
- Bones (large nuclei, high density)

Application: Medicine, Safety

10.2. Computed Tomography

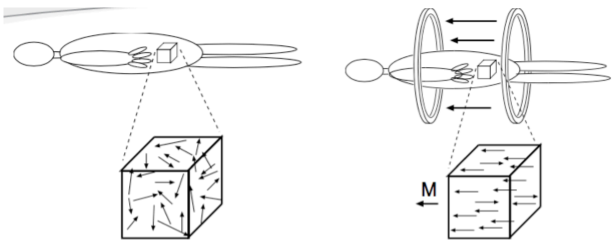


- X-rays that scan axial sections/layers of body
- Propagation delays produce scan data
- Computer calculates image

Used especially for the analysis of bone structures, low contrast in tissue

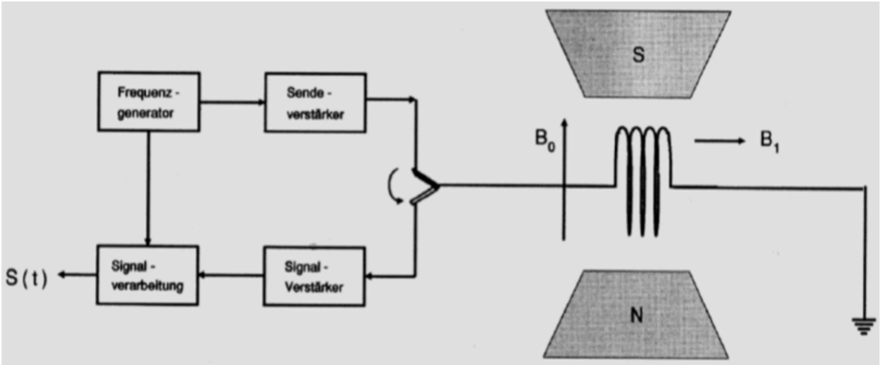
10.3. Magnetic Resonance Imaging

- Atomic nuclei and hydrogen nuclei, 1H, in particular, have a magnetic moment
- Moments tend to become aligned to applied field
- Creates magnetization, m(x,y,z) (a tissue property)
- MRI makes images of m(x,y,z)



Working Principle

- object is located in a homogeneous static magnetic field B_0
- B_1 is radiated perpendicular to B_0
- Record the magnetic resonance signal



10.4. Nuclear Medicine / Positron Emission Tomography (PET)

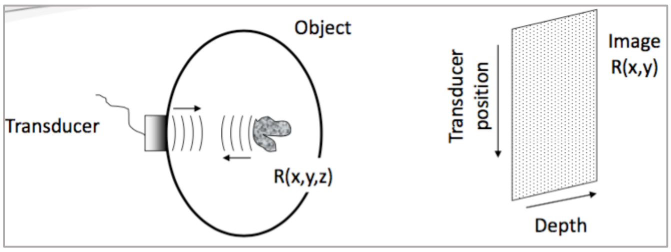
- Radioactive test solutions are given to patients
- Evaluating the radioactive radiation

<u>Example</u>: administration of radioactively enriched oxygen to verify the brain activity zones

Often combined with CT → PET/CT

10.5. Ultrasound Tomography

- Image reflectivity of acoustic wave
- Depth – function of time
- Lateral – focusing of wavefronts



10.6. Comparison

	X-Ray	CT	MR	US
Representation of				
• Bones	+ + +	+ + +	+	–
• Tissue	– / +	–	+ +	+
• Vessels	+ +	+ +	+ +	+
• Function	–	–	+ +	++
• Volumes	–	+ +	+ +	+
Real-time	*	+	+	+ +
Psychological stress	low	medium	high	low
Physical stress	high	high	low	low
Invasive	no	no	no	no
Cost (EUR)	~ 40	~ 100	~ 400	~ 10

11. 3D Printing

- Also known as: rapid prototyping, additive manufacturing
- Biggest market: Motor vehicles, consumer products
- Form of additive Manufacturing (joining material layer-by-layer)

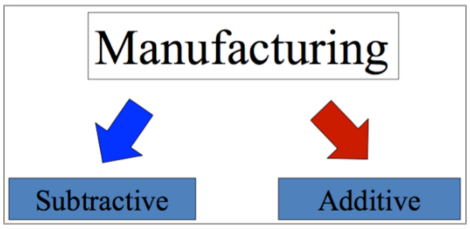
11.1. 3D Printing Techniques

Subtractive

- Milling
- Turning
- Drilling
- CNC

Additive

- Glue slices of object back together



Types of Additive Manufacturing

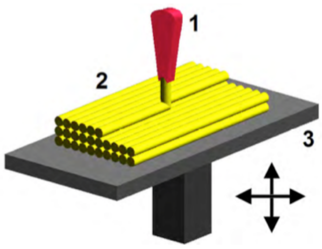
- SLS (Selective Laser Sintering)
- FDM (Fused Deposition Modeling)
- SLA (Stereo lithography)
- DLP (Digital Light Processing)
- EBM (Electron Beam Melting)

11.2. Selective Laser Sintering

- Platform with layer of powder
- Fuse powder with laser or by adding binder
- Lower platform, add powder and repeat

11.3. Fused Deposition Modeling

- Squirt semi-liquid material (plastic, wax, chocolate, etc)
- Add layer by layer
- Nozzle is heated to melt material and is moved horizontally and vertically



11.4. Stereo Lithography

- Tank of liquid polymer
- harden (polymerize) with laser beam
- Accurate, relatively fast

11.5. Digital Light Processing

Same as Stereo Lithography, but instead of a laser a DLP projector is used

11.6. Electron Beam Melting

- Power source: Electron beam
- Melting metal powder layer by layer in vacuum
- Parts are fully dense, void-free and extremely strong

11.7. Applications

- Medical procedures
- Advances in research
- Product prototyping
- Historic Preservation
- Architectural Engineering Construction
- Advanced Manufacturing
- Food Industries
- Automotive
- Accessories