Image-Based Reconstruction

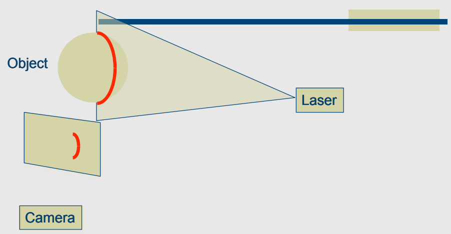
Asla Sá
Luiz Velho

IMPA – Instituto de Matemática Pura e Aplicada

Outline

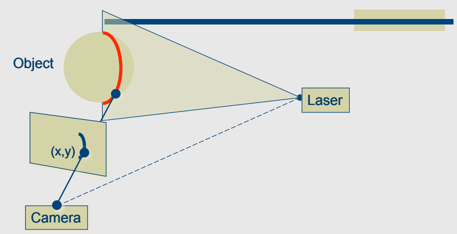
- Triangulation
- Pair-wise Scan Registration
- Global Registration
- Structured Light Principles
- Code Design
- BRDF Acquisition

Triangulation



- Project laser stripe onto object

Triangulation



- Depth from ray-plane intersection

General Ray-Plane Intersection

- Ray Equation
  $r(t) = o + td, \quad \text{with } o = (o_x, o_x, o_x) \text{ and } d = (d_x, d_x, d_x)$
- Plane Equation
  $ax + by + cz + w = 0$
- Intersection
  $a(o_x + td_x) + b(o_y + td_y) + c(o_z + td_z) + w = 0$
- Solving for *t*

$$t = -\frac{ao_x + bo_y + co_z + w}{ad_x + bd_y + cd_z}$$

Intersection in Camera Space

- Ray from the origin, with normalized direction

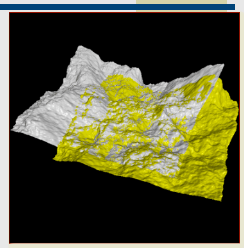
$$r(t) = o + td, \quad \text{with } o = (0,0,0) \text{ and } d = (d_x, d_x, d_x)$$

- Scene Depth

$$t = -\frac{w}{ad_x + bd_y + cd_z}$$

Pair-wise Scan Registration

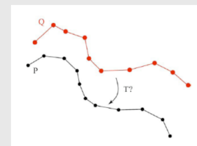
- Problem:
  - Given initial guess for relative transform
  - Align two scans partially-overlapping



Least Square Error

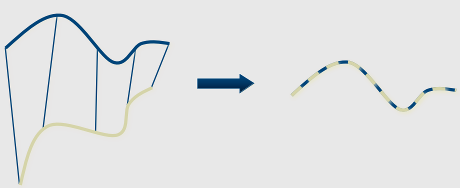
- Given Two Overlapping Scans, *P* and *Q*
  - Find Rigid Motion, *T =(R,t)*
  - Minimizing Distance $E_{p,q}$ Between *P* and *Q*

$$E_{p,q} = \sum_{i}^{N} \left\| p_i - (Rq_i + t) \right\|^2$$



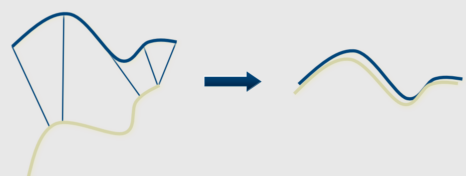
Aligning 3D Data

- If correct correspondences are known, can find correct relative rotation/translation



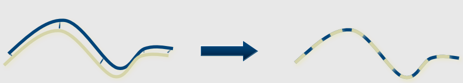
Aligning 3D Data

- How to find correspondences:
  - User input? / Feature detection?
- Alternative: assume ***closest*** points correspond



Aligning 3D Data

- … and iterate to find alignment
  - Iterative Closest Points (ICP) [Besl & McKay 92]
- Converges if starting position “close enough“



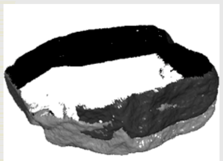
ICP

until *E is small*
Identify nearest points
Compute the optimal T
end until

* *Guaranteed to converge to a minimum (not necessarily global)*

Multiview Global Registration

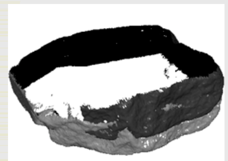
- The Problem:
  - Given: n scans around an object
  - Goal: align them all
  - First attempt: ICP each scan to one other

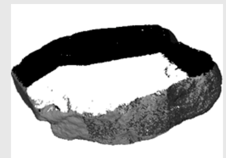


Solution

** Distribute Accumulated Error Among All Scans*

- Approaches:
  - Direct
  - Graph-Based





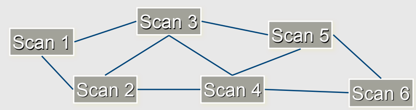
Direct Approach: *The Brute-Force Solution*

While not converged:
  For each scan:
    For each point:
      For every other scan
      - Find closest point
  Minimize error w.r.t. transforms of all scans

- Disadvantage:
  - Solve large $(np)\times(np)$ matrix equation, where n is number of scans and p is number of points per scan

Graph Methods

- Many global registration algorithms create a graph of pair-wise alignments between scans



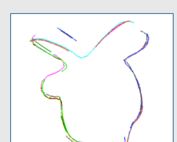
Kari Pulli's Algorithm

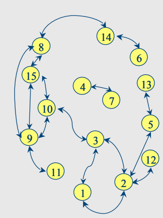
1. Perform pairwise ICPs, record sample of corresponding points (e.g. 200)
2. For each scan, starting with most connected
   - Align scan to existing set
   - While (change in error) > threshold
     - Align each scan to others

** All alignments during global registration phase use pre-computed corresponding points*

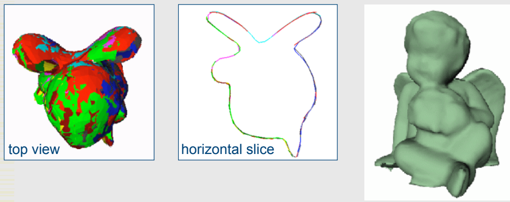
Example: *Aligning the Angel*


top view


horizontal slice


connectivity after local verification

The Final Angel Model



Structured Light Principles

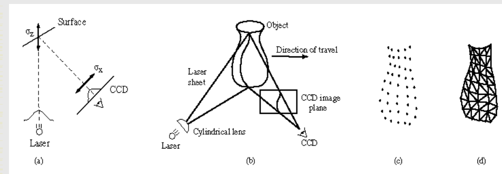
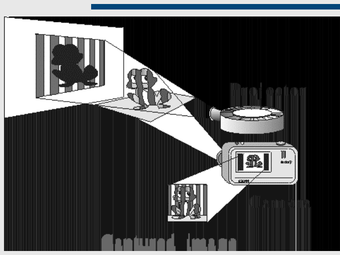


Figure from M. Levoy, Stanford Computer Graphics Lab

Coding structured light



Why coding structured light

- Point lighting - O(n²)
- “Line” lighting - O(n)
- Coded light - O(log n)*

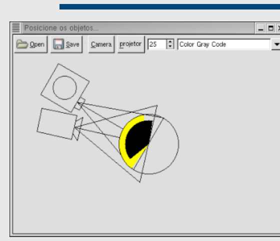
* Log base depends on code base

By reducing the number of captured images we reduce the requirements on processing power and storage.

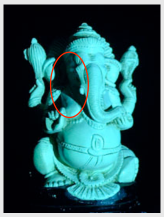
Pipeline

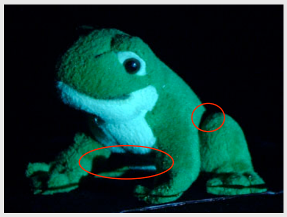
- Calibrate a pair camera/projector.
- Capture images of the object with projected patterns.
- Process images in order to correlate camera and projector pixels :
  - Pattern detection.
  - Decoding projector position.
- Triangulate to recover depth.

Working volume



Shadow areas





Our Goal

- How to correlate camera and projector pixels?
  - Different codes gives us several possibilities to solve correlation problem.
  - We are going to show you an overview of many different approaches.

CSL research

- Early approaches (the 80's)
- Structuring the problem (the 90's)
- New Taxonomy
- Recent trends
- Going back to colors

Main ideas

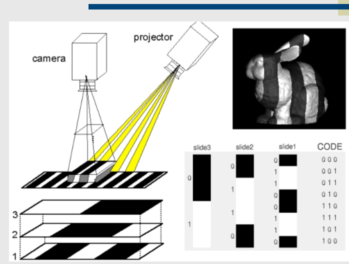
Structured Light Codes can be classified observing the restrictions imposed on objects to be scanned:

- Temporal Coherence.
- Spatial Coherence.
- Reflectance restrictions.

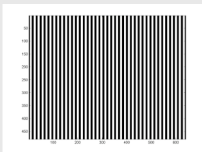
Temporal Coherence

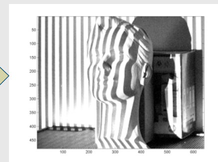
- Coding in more than one frame.
- Does not allow movement.
- Results in simple codes that are as less restrictive as possible regarding reflectivity.

Gray Code



…in practice:

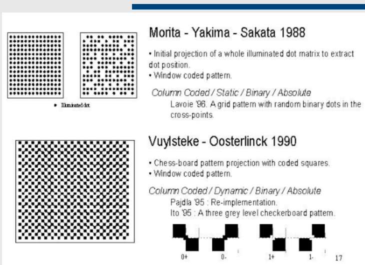




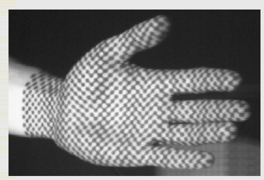
Spatial Coherence

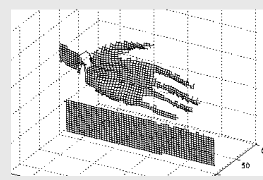
- Coding in a single frame.
- Spatial Coherence can be local or global.
- The minimum number of pixels used to identify the projected code defines the accuracy of details to be recovered in the scene.

Some examples



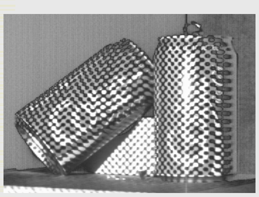
Binary spatial coding

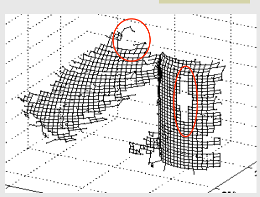




http://cmp.felk.cvut.cz/cmp/demos/RangeAcquisition.html

Problems in recovering pattern

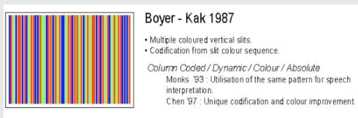


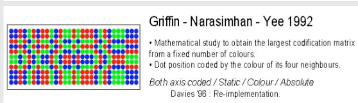


Introducing color in coding

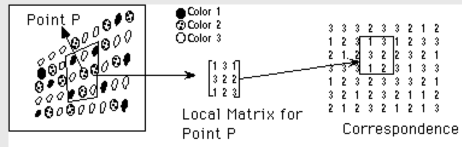
- Allowing colors in coding is the same as augmenting code basis. This gives us more words with the same length.
- If the scene changes the color of projected light, then information can be lost.
- Reflectivity restrictions (neutral scene colors) have to be imposed to guarantee the correct decoding.

Examples





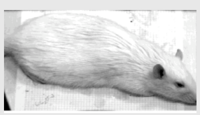
Local spatial Coherence

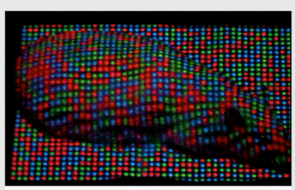


http://www.mri.jhu.edu/~cozturk/sl.html

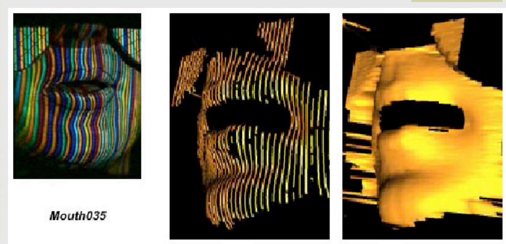
•Medical Imaging Laboratory
Departments of Biomedical Engineering and Radiology
Johns Hopkins University School of Medicine
Baltimore, MD 21205

Dot Matrix



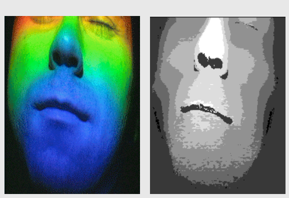


6 different colors used to produce sequences of 3 without repetition



Images from: Tim Monks, University of Southampton

Rainbow Pattern



Assumes that the scene does not change the color of projected light

http://cmp.felk.cvut.cz/cmp/demos/RangeAcquisition.html

Noisy transmission channel

- There is a natural analogy between coded structured light and a digital communication system.
- The camera is recieving the signal transmitted through object by the projector.

New Taxonomy

Method	Number of slides	Number of characters of the alphabet	Neighbor-hood	Resolution (number of words)
Gray Code	n	Binary(2)/ monochromatic	Single pixel	2^n per line
Rainbow pattern	1 for coding plus 1 used in decoding	2^8 per channel - RGB	Single pixel	$(2^8)^3$ per line
Dot matrix	1	3	4 neighbors (N,S,E,W)	3^5

Designing codes

Goal: design a light pattern to acquire depth information with minimum number of frames without restricting the object to be scanned (impose only minimal constraints on reflectivity, temporal coherence and spatial coherence.)

What is minimal?

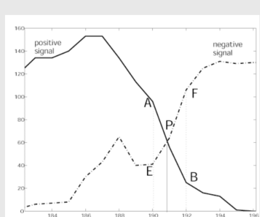
- Temporal Coherence: 2 frames.
- Spatial Coherence: 2 pixels.
- Reflectivity restrictions: allow non neutral objects to be scanned without loosing information.

Processing images

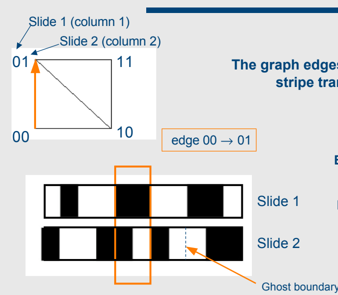
- To recover coded information a pattern detection procedure has to be carried out on captured images.
- The precision of pattern detection is crucial to the accuracy of resulting range data.
- Shadow areas also have to be detected.

Edge Detection

- Stripes transitions produce edges on camera images.
- Transitions can be detected with sub-pixel precision.
- Projecting positive and negative slides is a robust way to recover edges.



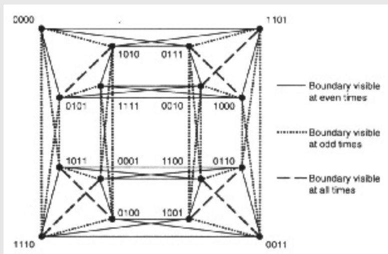
Boundary Coding



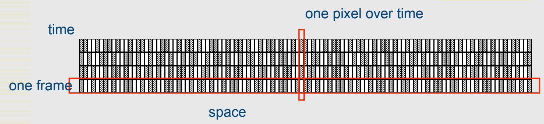
The graph edges correspond to the stripe transition code.

The maximal code results from an Eulerian path on graph.
Obs.: 2 frames of Gray code gives us 4 stripes. In this case we have 10.

Rusinkiewicz 4 frames graph

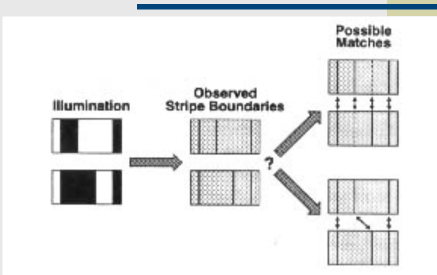


4 frames stripe boundary code

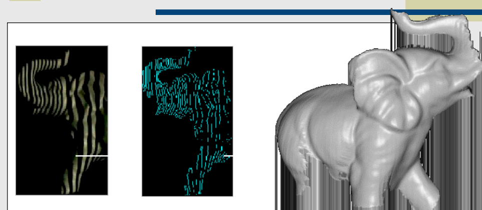


- When using a binary base, ghost boundaries have to be allowed in order to obtain a connected graph.
- The decoding step isn't straightforward, due to the presence of ghosts. A matching step have to be carried out.

Ambiguity in decoding



Real-time range scanning



video frame range data merged model (159 frames)

Comments

- It scans moving objects.
- It is designed to acquire geometry in real-time.
- Some textures can produce false transitions leading to decoding errors.
- It does not acquire texture.

Revisiting Colors

- Taking advantage of successively projecting positive and negative slides, reflectivity restrictions can be eliminated.
- To solve the problem of allowing ghost boundaries we have to augment the basis of code, that is, allowing colors.

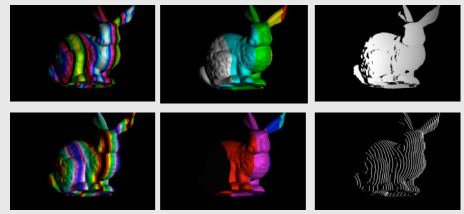
Recovering colored codes

$$I_R = u_R + r_R p_R$$
$$I_G = u_G + r_G p_G$$
$$I_B = u_B + r_B p_B$$

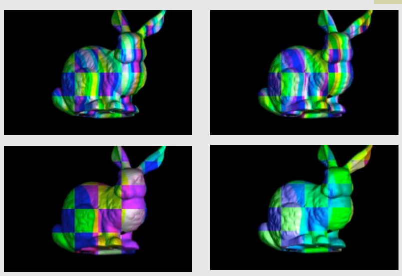
- **u** is the ambient light
- **r** is the local intensity transfer factor mainly determined by local surface properties
- **p** is the projected intensity for each channel

$$I_i = \begin{cases} u_i, & if \;\; p_i = 0 \quad \leftarrow \text{Negative slide} \\ u_i + r_i, & if \;\; p_i = 1 \quad \leftarrow \text{Positive slide} \end{cases}$$

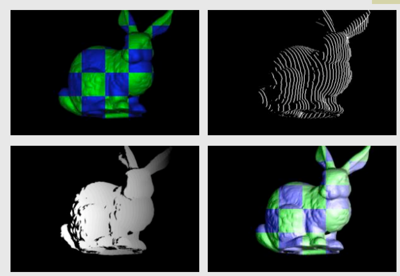
Colored Gray code



Confusing colors!



Recovering texture

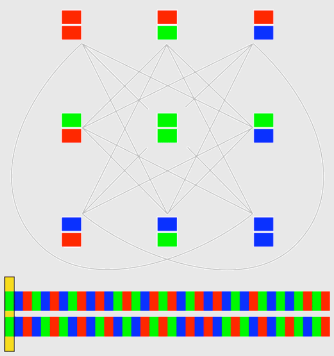


(b,s)-BCSL

- Augment basis of Rusinkiewicz code eliminating ghost boundaries.
- We proposed a coding scheme that generates a boundary stripe codes with a number *b* of colors in *s* slides, it is called (b,s)-BCSL.

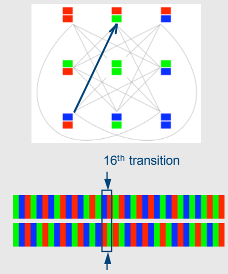
Generating code

- Choose the basis *b*>2 and the number of slides *s*>1 of the coding scheme.
- Construct the (b,s)-graph eliminating ‘ghost’ edges.
- Choose an initial vertex v_0 in V(G) and construct an eulerian circuit starting from v_0.
- Complete the decoding table and generate the final coded slides.

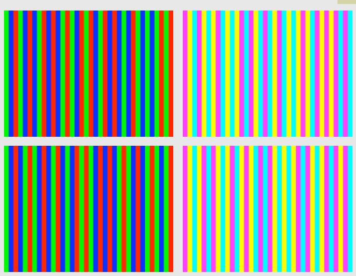


Decoding table

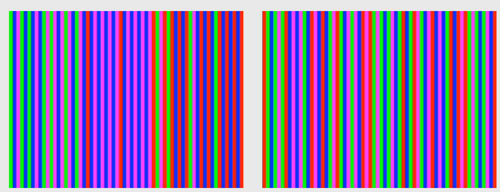
vertices	d(0)	d(1)	d(2)	d(3)
V(00)	0	3	6	9
V(01)	14	17	19	11
V(02)	28	34	22	24
V(10)	26	29	18	21
V(11)	1	31	33	35
V(12)	15	4	8	13
V(20)	16	23	32	12
V(21)	27	5	7	25
V(22)	2	10	20	30



(3,2)-BCSL

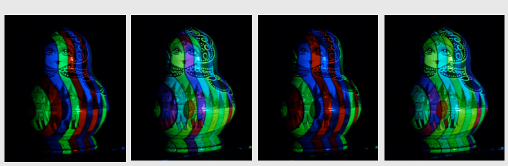


(4,2)-BCSL



144 stripes (here the code was truncated before using all stripes)

In practice …

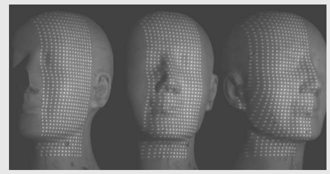


…after processing:

How to reconstruct the entire object?

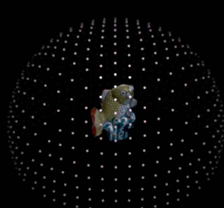
- Capturing images from many different points of view.
- The resultant clouds of points have to be aligned to be unified.
- The clouds of points can be processed to become a mesh.

Projected pattern changing object’s position



From: Medical Imaging Laboratory
Departments of Biomedical Engineering and Radiology
Johns Hopkins University School of Medicine
Baltimore, MD 21205

Take photographs



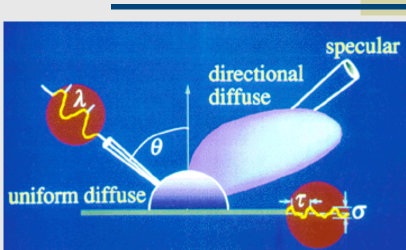
Camera positions

Photographs

Aquisição de dados fotométricos

Comparison between a photograph (left) and a rendered image of a 3D model under similar lighting conditions.

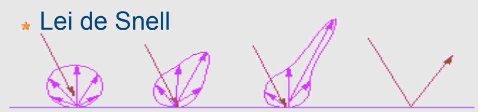
BRDF – *bidirectional reflectance distribution function*



$\rho(\theta_i, \phi_i, \theta_r, \phi_r, \lambda)$

Simplificações da BRDF

- Modelo difuso:
  - Reflectância igual em todas as direções.
- Modelo espectral:
  - Lei de Snell

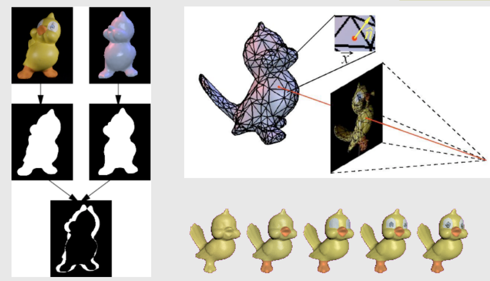


Trabalhos relacionados

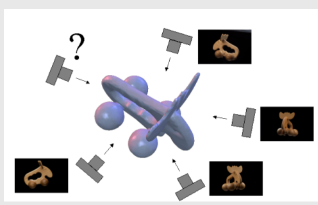
- Dados fotométricos densos:
  - Surface light field (Curless)
  - Light stage (Debevec)
- Dados fotométricos esparsos:
  - Image based reconstruction (Lensch, MPI)
- Aquisição de BRDF dos materiais:
  - Keldysh Institute of Applied Mathematics (Moscow, Russia)

Acquisition pipeline at MPI

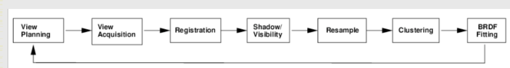
http://www.mpii.mpg.de/departments/d4/areas/scenedigitization/



Planejamento de aquisição

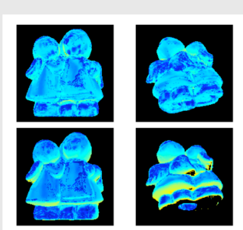


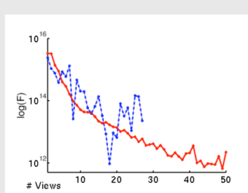
Pipeline de aquisição planejada



- A aquisição é incremental para que o erro entre o ponto de vista planejado e a imagem realmente adquirida seja considerado, e não acumulado para as próximas vistas.

Ganhos do planejamento





Modelos virtuais