Adv. Computer Graphics

NYU - Spring 2024

AI Graphics

Machine Learning for Media Applications

Luiz Velho
IMPA

Outline

- 3D Graphics and AI
- Basic Concepts for Modeling and Rendering
- 3D Deep Learning
- Course Affairs

3D Graphics & AI

Scope

Machine Learning Models for Media

3D

What we will study

- 3D Graphics
- Synthetic Images

Areas of Application

- Scientific Visualization
  - Medicine, Biology, Mathematics
- CAD / CAM
  - Engineering, Architecture, Design
- Entertainment
  - Television, Film, Games

Other Areas

- Vision
- 3D Animation
- Virtual Reality
- Multimedia

2D

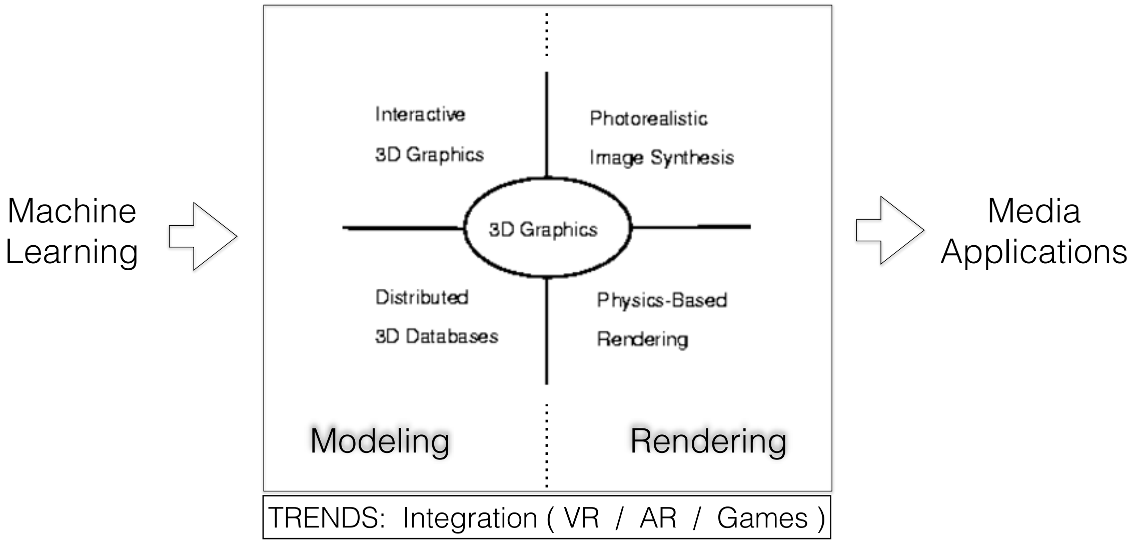
What we will not cover

- 2D Graphics
- Human-Computer Interaction

Areas of Applications

- Electronic Publishing
  - Illustration
  - Paint Systems
  - Page Layout
- Image Processing
  - Video Effects
  - Pattern Recognition
- Interface Design
  - Window Systems

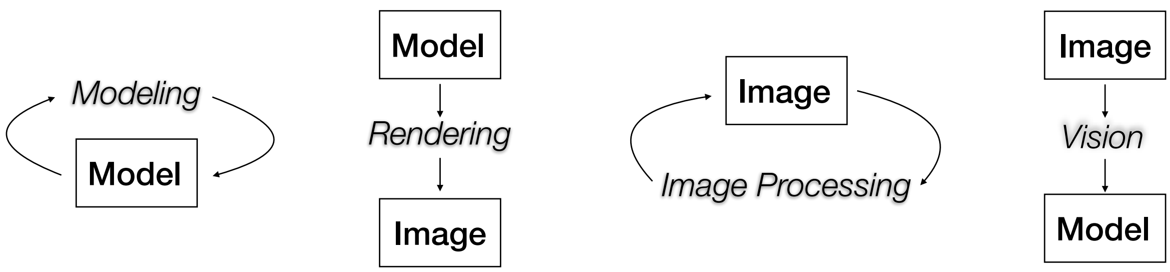
3D Graphics & Media Standards



Evolution of the Area

Traditional *Graphics*

- Four Separated Areas



Modern *Graphics*

- Integrated Areas

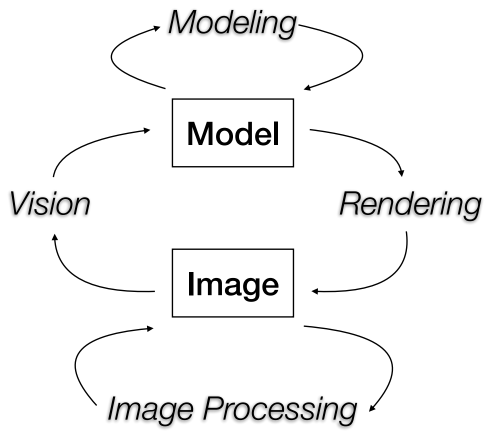
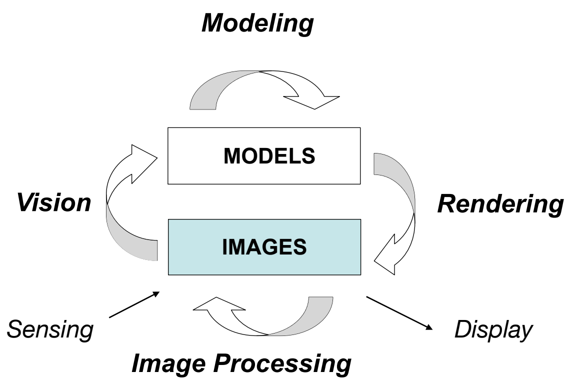


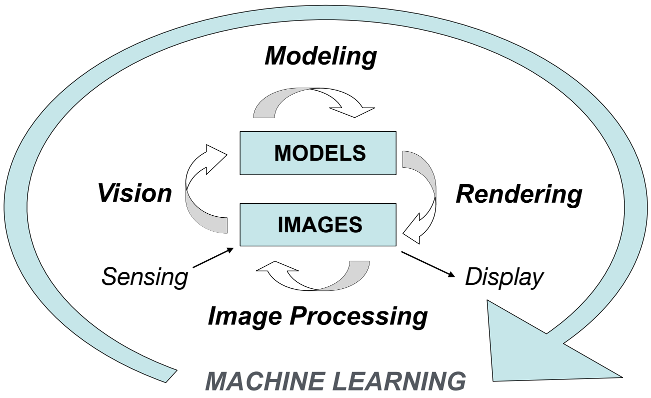
Image-Based *Graphics*

- From Appearance to Algorithms



A.I. *Graphics*

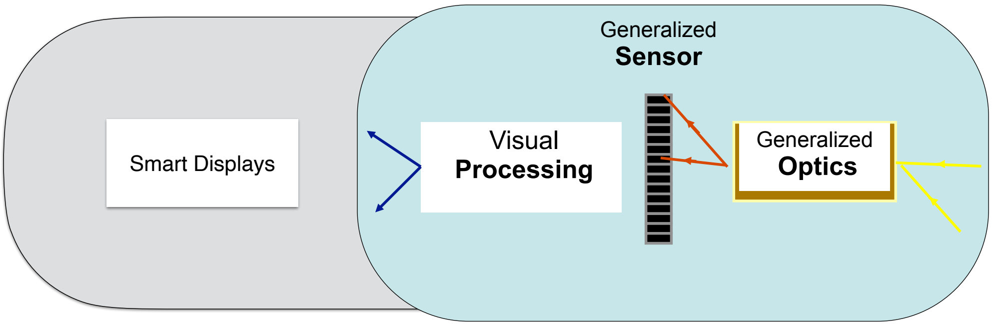
- From Data to Networks

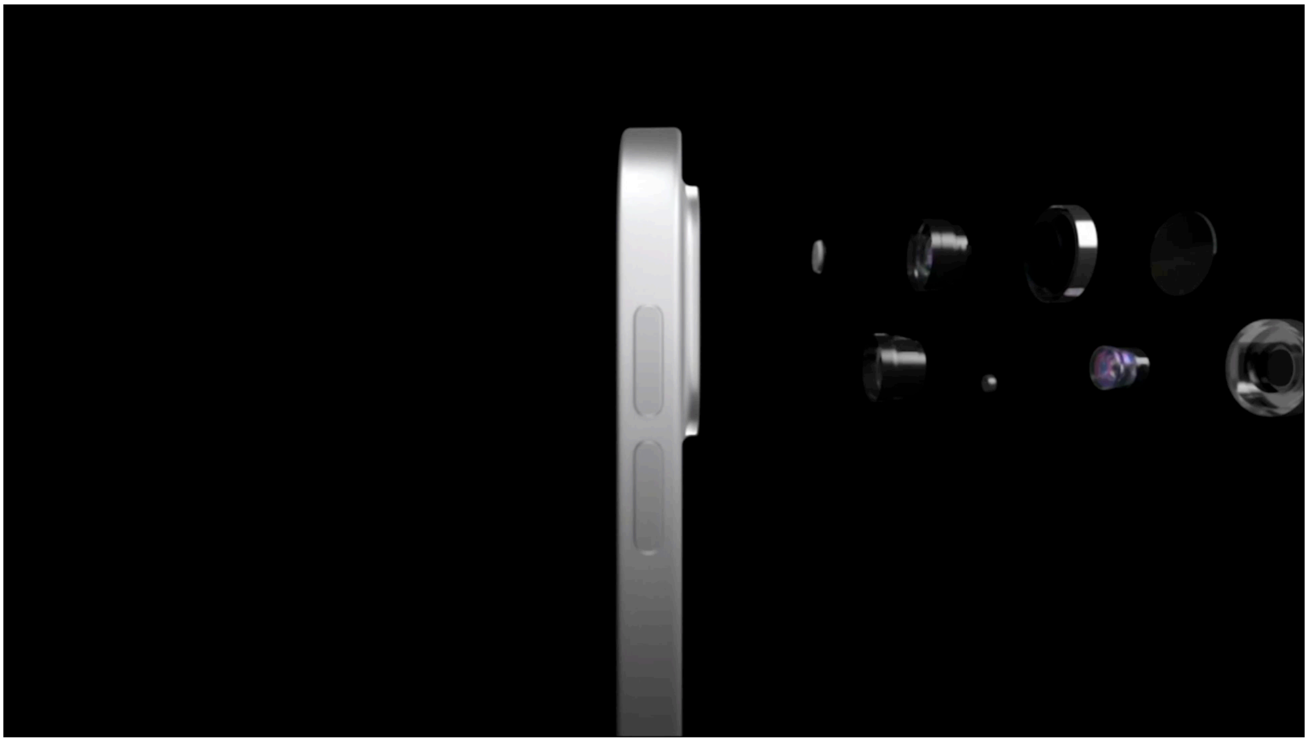


Media Technology

Computational Imaging

- Intelligent Acquisition / Display Devices





Case Study

GANCraft: Rendering Realistic Worlds

Pause for a Break

"At least this supports my theory that we live in a computer simulation."

Basic Concepts for 3D Graphics Systems

Concepts

- Paradigm of the 4 Universes
- Media Objects

4 Universes Paradigm

- Levels of Abstraction

Physical Universe

↕

Mathematical Universe

↕

Representation Universe

↕

Implementation Universe

(Gomes and Velho, 1995)

Example: Numbers

Scalar Quantities	length, area
Real Numbers	7.598
Floating Point Representation	$7598 * 10^{-3}$
IEEE Standard	mantissa exponent

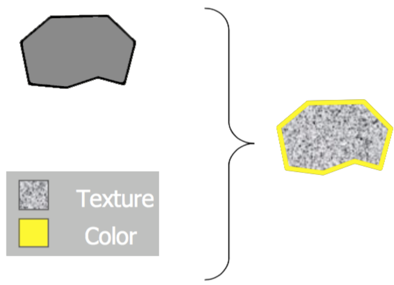
Issues

- Study each Level
    - Objects in level
- Relation between Levels
    - How to Map (e.g. Correspondences)
    - Equivalences and Losses

Media Objects

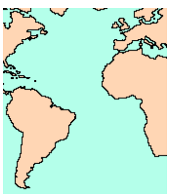
$$o = (U, f) \qquad\qquad f : U \subset \mathbb{R}^m \to \mathbb{R}^n$$

- Support $\boldsymbol{U}$
  - Shape / Geometry
- Properties $\boldsymbol{f}$
  - Attributes / Appearance

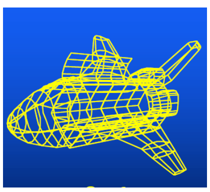


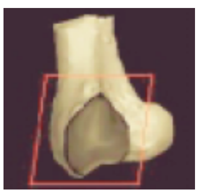
Unifying Concept

Examples


drawings


images


surfaces

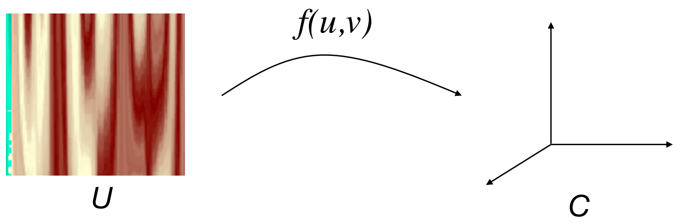

volumes

time

- General Encompassing Definition (Gomes and Velho, 1996)

Defining Attributes

- Ex: Image
  - Simple Shape: $U = [0,1] \times [0,1]$
  - Complex Attributes: (color, density, etc) $f(u, v) = [r, g, b, a, ..]$



Defining Shape

- Ex: Curves

  (unit circle)

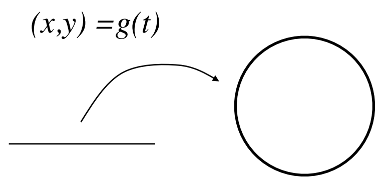
  - Parametric Description

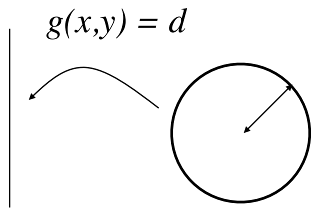
    $g(\theta) = (\cos\theta, \sin\theta)$

    $\theta \in [0,2\pi]$

  - Implicit Description

    $g^{-1}(0) = \{(x, y) \mid x^2 + y^2 - 1 = 0\}$





Putting it All Together

- Ex: 3d Object
  (Diana statue)
  - Shape:
    - Scanned 3D Surface
    - Solid Object
  - Properties:
    - Material Marble

(Wann Jensen et al.)

Questions

- How to Define:
  - Function (model)
  - Support (domain)
  - Attribute (range)

Mathematical Models

- Deterministic
  - Function (one object)
- Probabilistic
  - Stochastic Process (class of objects)
- Procedural
  - Generators + Operators (algebra / expression)

Computation and Representation

- How to Compute
- Conversion between Models / Representations

✳ Hybrid Models

Media Objects Operators

- Spaces of Media Objects

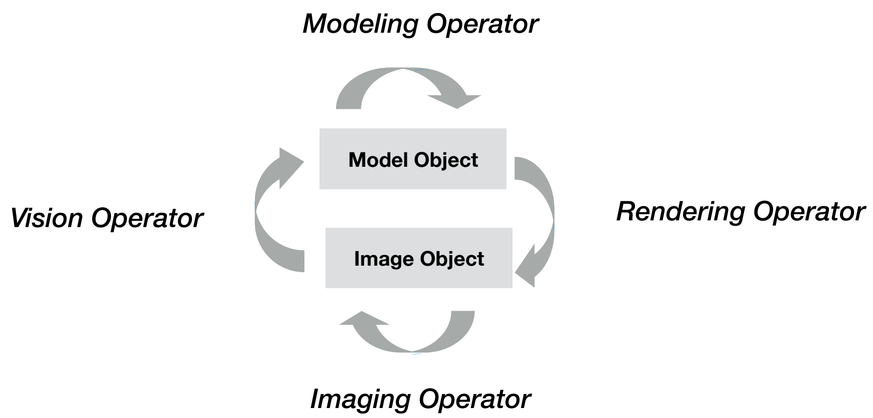
$$o \in \mathcal{O}$$

- Media Objects Operators

$$T : \mathcal{O}_1 \mapsto \mathcal{O}_2$$

- Intra $(i = j)$ / Inter Levels $(i \neq j)$
- Representation

Landscape of Operations



Problems

Direct Problems

- Given T and x, find y

$$y = Tx$$

- Ex: Visualization
  - x is the scene (geometry, lighting, camera)
  - T is the rendering operator
  - y is the rendered image

Inverse Problems

- Given T and y, find x such that $T(x) = y$

  Ex: Object Recognition

  - y is a captured image
  - T is an acquisition system
  - x is a template

- Given x and y, find T

  Ex: Camera Calibration

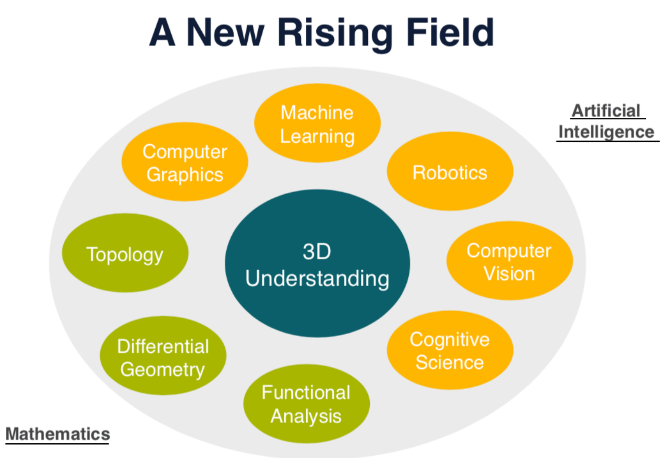
  - x is a calibration pattern
  - y is a transformed pattern
  - T is a projective transformation

Q&A

3D Deep Learning

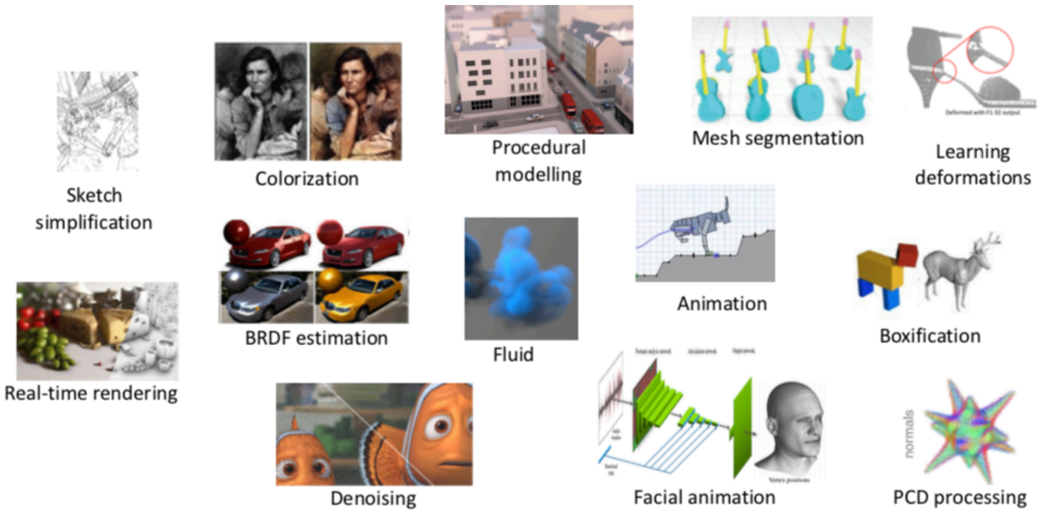
Slides from: (Mitra et. al, "CreativeAI: Deep Learning for Computer Graphics", 2019)

Deep Learning for 3D Data



(Mitra et. al, 2019)

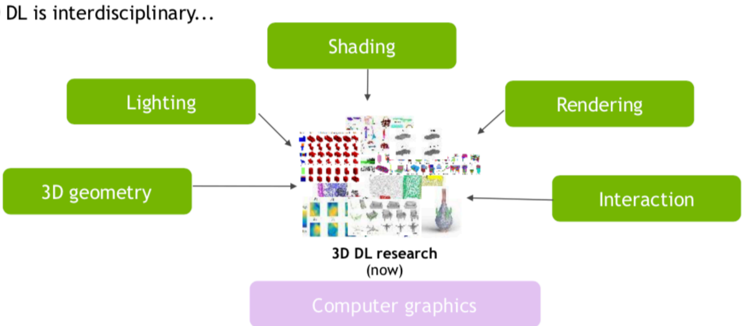
Applications



(Mitra et. al, 2019)

Computer Graphics

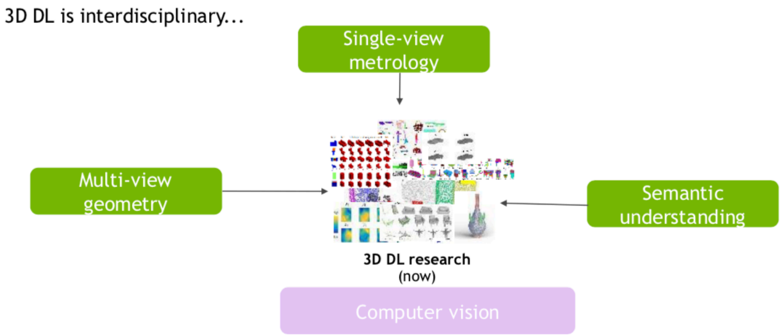
3D DL is interdisciplinary...



(Synthesis)

(Mitra et. al, 2019)

Computer Vision



(Analysis)

(Mitra et. al, 2019)

Synergistic Opportunities

What is Special about CG?

1. **Regular data structure** and easy to parallelize
   (e.g., image translation)

2. Many sources of input data — **model building**
   (e.g., images, scanners, motion capture)

3. Many sources of **synthetic data** — can serve as supervision data
   (e.g., rendering, animation)

4. Many problems in **generative models** and need for **user-control**

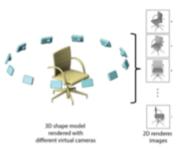
(Mitra et. al, 2019)

Representations in CG

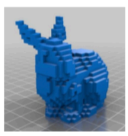
- Images (e.g., pixel grid)
- Volume (e.g., voxel grid)
- Meshes (e.g., vertices/edges/faces)
- Point clouds (e.g., collection of points)
- Animation (e.g., skeletal positions over time; cloth dynamics over time)

(Mitra et. al, 2019)

Representation Challenge



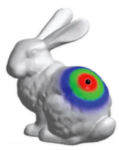
Multi-view



Volumetric

Part Assembly

Point Cloud



Mesh (Graph CNN)

$$F(x) = 0$$

Implicit Shape

(Mitra et. al, 2019)

3D Deep Learning

Goal: Learn a Parametric Function

$$f_\theta : \mathbb{X} \longrightarrow \mathbb{Y}$$

θ: function parameters, these are learned

$\mathbb{X}$: source domain

$\mathbb{Y}$: target domain

Examples:

Image Classification: $f_\theta : \mathbb{R}^{w \times h \times c} \longrightarrow \{0, 1, \ldots, k-1\}$

$w \times h \times c$: image dimensions k: class count

Image Synthesis: $f_\theta : \mathbb{R}^n \longrightarrow \mathbb{R}^{w \times h \times c}$

n : latent variable count $w \times h \times c$: image dimensions

(Mitra et. al, 2019)

Problems in GC

analysis:

- Feature detection (image features, point features) $\mathbb{R}^{m \times m} \to \mathbb{Z}$
- Denoising, Smoothing, etc. $\mathbb{R}^{m \times m} \to \mathbb{R}^{m \times m}$
- Embedding, Metric learning $\mathbb{R}^{m \times m, m \times m} \to \mathbb{R}^d$

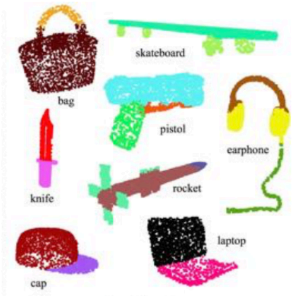
synthesis:

- Rendering $\mathbb{R}^{3n} \to \mathbb{R}^{m \times m}$
- Animation $\mathbb{R}^{3m \times t} \to \mathbb{R}^{3m}$
- Physical simulation $\mathbb{R}^{3m \times t} \to \mathbb{R}^{3m}$
- Generative models $\mathbb{R}^d \to \mathbb{R}^{m \times m}$

(Mitra et. al, 2019)

3D Analysis

Classification



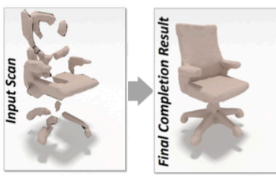
Parsing
(object/scene)

Correspondence

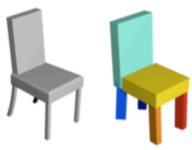
(Mitra et. al, 2019)

3D Synthesis

Monocular
3D reconstruction



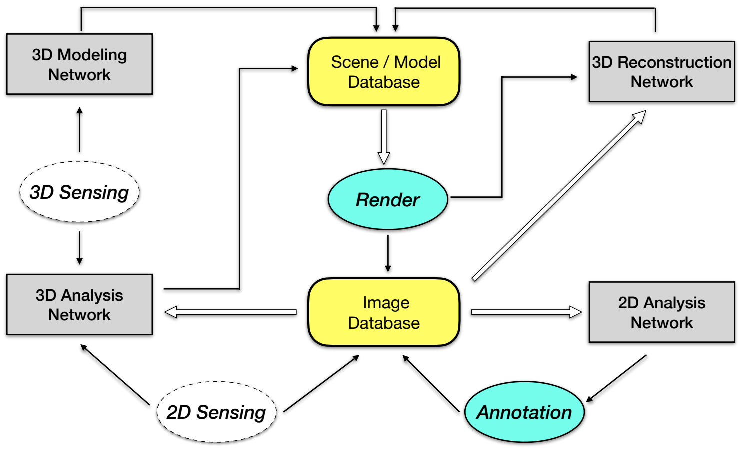
Shape completion



Shape modeling

(Mitra et. al, 2019)

A.I. Graphics

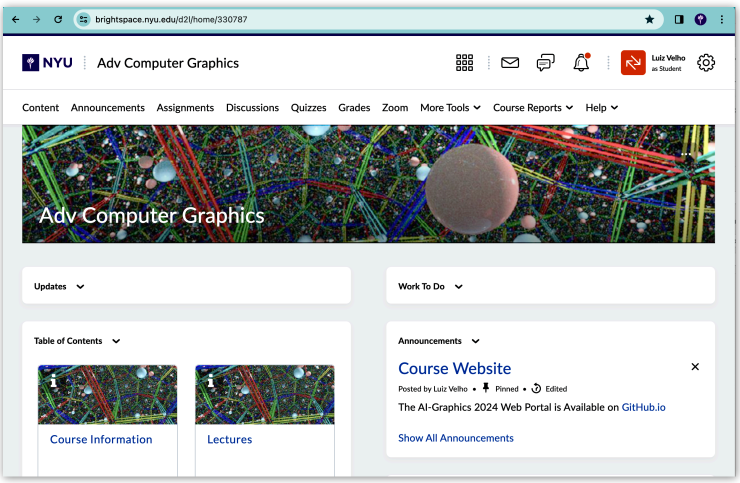


The Course

Course Affairs

- Lectures (Brightspace)
  - Summary / Slides / Recording
- Reading (Web Portal)
  - Bibliography / Material
- Grading
  - 2 Mandatory Assignments, 1 Optional (80%)
  - Final Exam (20%)

NYU Course Site



Course Portal

AI Graphics

Machine Learning for Media Applications

Spring 2024

About the Course

- Course Description
- Instructor: Luiz Velho
- TAs: _
- Classes: Thursdays, 13:30 to 15:30

Course Material

- Reading
- Software
- Datasets

Lectures

- Schedule

Coursework

- Instructions
- Assignments

Resources

- Elsewhere
- Resources
- Etc..

https://lvelho.github.io/ai-graphics-2024/

Assignment 1

Q&A