Lecture 1: Introduction

Jun-Yan Zhu
16-726, Spring 2022
Teaching Staff

**Instructors**

Jun-Yan Zhu

**Teaching Assistants**

Sheng-yu Wang
Zhiqiu Lin
Jun-Yan Zhu

- Computer Vision, Computer Graphics, Machine Learning, Computational Photography
- Love pets (cat & dog)
- Gaming (mostly FIFA these days)
Cat Paper Collection

As reported by Cisco, 90% of net traffic will be visual, and indeed, most of the visual data are cat photos and videos. Thus, understanding, modeling, and synthesizing our feline friends becomes a more and more critical research problem these days, especially for our cat lovers.

Cat Paper Collection is an academic paper collection that includes computer graphics, computer vision, and machine learning papers that produce experimental results related to cats. If you would like to add/remove an article, please send an email to Jun-Yan Zhu (junyanz at cs dot cmu dot edu). We thank all the authors for their contribution and support.

See also GitHub | CSV file

Flow-edge Guided Video Completion
Chen Gao, Ayush Saraf, Jia-Bin Huang, Johannes Kopf
In ECCV 2020
[Paper] [Project]

Matching Guided Distillation
Kaiyu Yue, Jiangfan Deng, Feng Zhou
In ECCV 2020
[Paper] [Project]

Strong 3D Printing by TPMS Injection
Xin Yan, Cong Rao, Lin Lu, Andrei Sharf, Haisen Zhao, Baoquan Chen
In IEEE TVCG 2019
[Paper]
Zhiqiu Lin

- From Beijing, China
- Undergrad in CS&Maths at Cornell University
- Advised by Prof. Deva Ramanan
- Interested in visual recognition and continual learning
Sheng-Yu Wang

• Interested in generative modeling and model interpretation
• From Taiwan
• Undergrad in CS at Berkeley
• Play guitar for free time
Lecture 1: Introduction

Jun-Yan Zhu
16-726, Spring 2022
Visual Content Creation

Cave art

Time

32,000 BC
Visual Content Creation

- **Cave art**: 32,000 BC
- **Sculpture**: 1498
- **Painting**: 1872
Visual Content Creation

Cave art 32,000 BC

Sculpture 1498

Painting 1872

Computer Graphics 2012

Time
Who is creating visual content?

- **Cave art**
  - 32,000 BC
  - Michelangelo

- **Sculpture**
  - 1498
  - Claude Monet

- **Painting**
  - 1872
  - Paul Cezanne

- **Computer Graphics**
  - 2012
  - George Lucas
  - Ang Lee
Who is creating visual content?

100+ visual artists
12 months
$60+ million

Skeleton ➔ Geometry ➔ Texture ➔ Details ➔ Image

Ang Lee

Idea

Visual Content
Who is creating visual content?
Who is creating visual content?

Kid’s drawing

Photoshop result by his father

© Tom Curtis
Creating Visual Realism Manually

CG office

CG office (more details)

My advisor’s office
Data-Driven Graphics (2000s)

Graphics → Image Retrieval

Picture from James Hays
Data-Driven Graphics (2000s)

Compositing multiple parts

User Input

Database images

Output

Sketch2Photo [Tao et al. SIGGRAPH Asia 2009]
Data-Driven Graphics (2000s)

- Hard to combine pieces
- No understanding of visual realism
Help everyone easily create visual content

Teach machines how to create realistic content
Image-to-Image Translation with pix2pix

pix2pix [Isola, Zhu, Zhou, Efros. CVPR 2017]
#edges2cats with pix2pix

By Christopher Hesse

Ivy Tasi @ivymyt

@matthematician

Vitaly Vidmirov @vvid

https://affinelayer.com/pixsrv/
GauGAN input

GauGAN result

By Darek Zabrocki, Concept Designer and Illustrator
Collection Style Transfer

Photograph ©Alexei Efros

Monet

Van Gogh

Cezanne

Ukiyo-e

CycleGAN [Zhu, Park, Isola, Efros. 2017]
Monet’s paintings → photographic style

CycleGAN [Zhu, Park, Isola, Efros. 2017]
Horse → Zebra

-cycleGAN [Zhu, Park, Isola, Efros. 2017]
Failure case
Failure case
Swapping Autoencoder For Deep Image Manipulation
Taesung Park¹, Jun-Yan Zhu², Oliver Wang², Jingwan Lu², Eli Shechtman², Alexei Efros¹, Richard Zhang²
¹UC Berkeley, ²Adobe Research

Swapping Autoencoder [Park et al. NeurIPS 2020]
Photoshop 2021 Neural Filters
Research Highlights
from other universities & industry labs
Synthesizing High-res Portraits

[Kerras et al., CVPR 2020] @ NVIDIA
Everybody Dances Now

Caroline Chan, Shiry Ginosar, Tinghui Zhou, Alexei A. Efros, ICCV 2019. @ UC Berkeley
Neural Talking-Head Synthesis

face-vid2vid: One-Shot Free-View Neural Talking-Head Synthesis for Video Conferencing
Ting-Chun Wang, Arun Mallya, Ming-Yu Liu. CVPR 2021 @ NVIDIA
NeRF in the Wild

[Martin-Brualla et al., CVPR 2021] @ Google Research
Text-based Image Editing

StyleCLIP: Text-Driven Manipulation of StyleGAN Imagery
Or Patashnik*, Zongze Wu*, Eli Shechtman, Daniel Cohen-Or, Dani Lischinski
Hebrew University of Jerusalem, Tel-Aviv University, and Adobe Research
Course preview

- A modern machine learning perspective
- Widely-used learning algorithms
- Interactive content creation tools
Logistics
Course objectives

1. You will get a foundation in image editing and synthesis.
   o Texture synthesis and style transfer.
   o Face modeling and synthesis.
   o Image colorization and inpainting.
   o Video generation and editing.
   o Image-to-image translation.
   o Image and video editing. (warping, morphing, compositing)
   o Image and video forensics.
Course objectives

2. You will get a foundation of machine learning concepts
   o (fast) Nearest neighbor search.
   o Principal component analysis, Gaussian Mixture model. Markov Random Field (MRF)
   o Convolutional neural networks.
   o Conditional generative models.
   o Neural Radiance Fields (NeRF)
Course objectives

3. You will have some cool results with your own photos
Textbook

https://szeliski.org/Book/ (2021 edition”)

Grading

• Emphasis on programming projects (65%).
  – Classic: 1. image alignment. 2. image blending
  – Deep learning: 3. neural style transfer. 4. GANs and conditional GANs.
  – 5. reconstructing and editing an image with GANs.

• Late Policy for programming assignments.
  – Five (5) emergency late days for semester, to be spent wisely
  – 10% of penalty per day afterwards

• One paper presentation (10%):
  – 10-20 min, 1-2 people in a group.
  – Need to answer questions about this paper from now on.

• Final Project (25%)
  – A webpage-based report + a presentation.
  – No late day.
  – 2-3 people per group.
Assignments

Assignment #0 - How to submit assignments?

Assignment #1 - Colorizing the Prokudin-Gorskii Photo Collection
Winner: [Korwoo Kim]
Honorable Mentions: [Juyong Kim] [Zihang Lai] [Manuel Rodriguez]

Assignment #2 - Gradient Domain Fusion
Winner: [Manuel Rodriguez]
Honorable Mentions: [George Cazanavette]

Assignment #3 - When Cats meet GANs
Winner: [Jun Loo]
Honorable Mentions: [George Cazanavette]

Assignment #4 - Neural Style Transfer
Winner: [Zihang Lai]
Honorable Mentions: [Zijie Li] [Tarang Shah]

Assignment #5 - GAN Photo Editing
Winner: [George Cazanavette]
Honorable Mentions: [Manuel Guevara] [Zijie Li] [Zhe Huang]

For each assignment

• Derive the math, implement stuff from scratch (+ starter code), and apply it to your own photos
• Every person does their own project (except final)
• Reporting via web page (+ submit code to Canvas)
• Afterwards, vote for class favorite(s)! Gift!
• Programming Language:
  – Python and PyTorch
  – you can use other languages, but you are on your own
Academic Integrity

• Can discuss projects, but don’t share code

• Don’t look up code or copy from a friend

• If you’re not sure if it’s allowed, ask

• Acknowledge any inspirations

• If you get stuck, come talk to us
Getting help outside of class

• Course Web Page
  – https://16726-image-synthesis.github.io/sp22/

• Discussion board:
  – Piazza.com

• Office hours (EST)
  – Zhiqiu: 2-3 pm Friday
  – Sheng-Yu: 2-3 pm Tuesday
  – Jun-Yan: 11 am-12 pm Monday
Zoom links: see the Piazza post
Why you should NOT take this class

• Project-based class
  – No canned problem sets.
  – Not theory-heavy.
  – will read many research papers.
  – Open-ended by design.

• Need time to think, not just hack
  – Creativity is a class requirement.

• Not worth it if you don’t enjoy it.
Now... reasons TO take this class

• Not too many similar courses at other places.
• You get to create pictures and unleash your creative potential.
• Interested in grad school and research? 😊
• Interested in industry jobs? 😊
Thank You!

16-726, Spring 2022
https://learning-image-synthesis.github.io/sp22/