Introduction

Masahiro Toyoura

Masahiro Toyoura

http://www.design.yamanashi.ac.jp/en/toyoura/

~2008 Ph.D. (Informatics) @ Kyoto University

3D Shape Reconstruction from Multiple Silhouettes for Objects in Rigid Motion

2008 Visiting Researcher @ UC Santa Barbara

Virtual Object Manipulation with Human Hand in Random Pattern Glove

2009 Assistant Professor @ University of Yamanashi

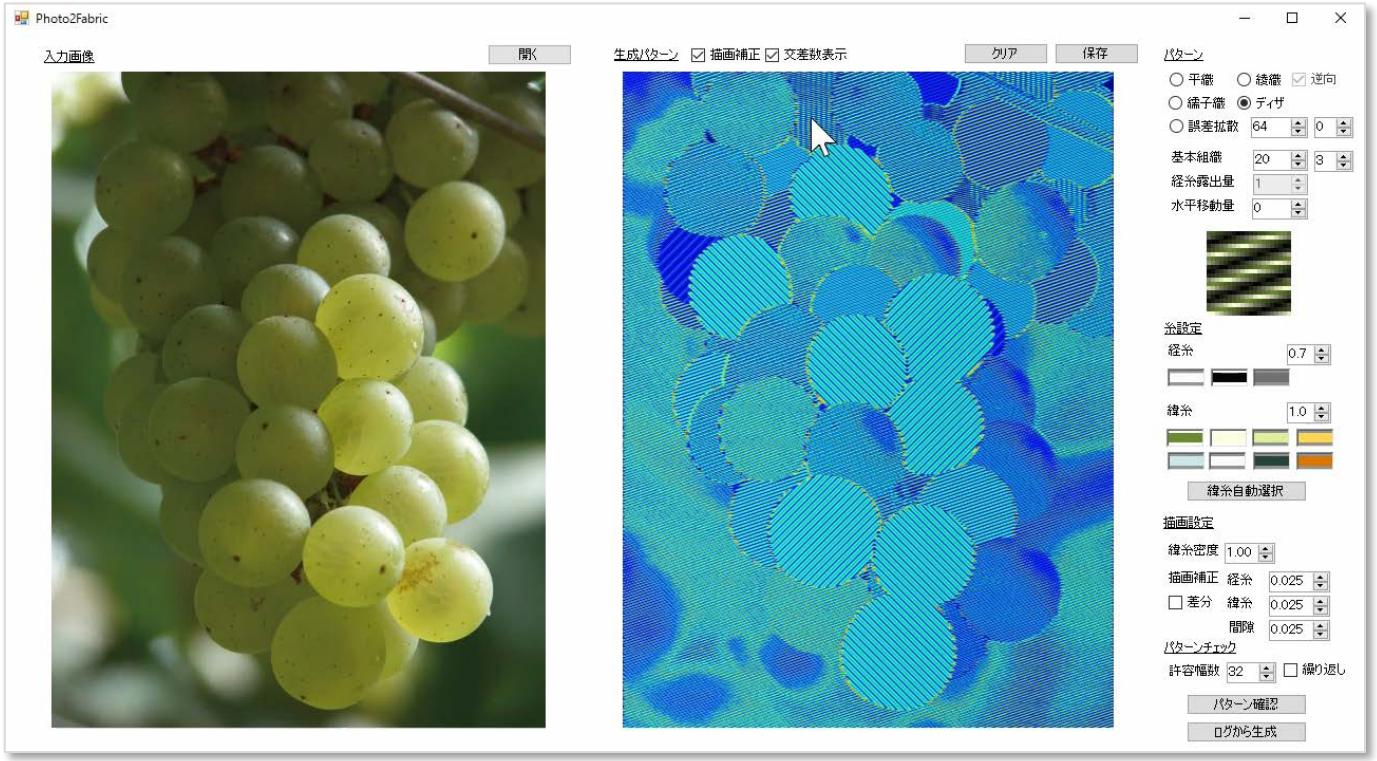
AR, CV & CG, digital fabrication, educational technology

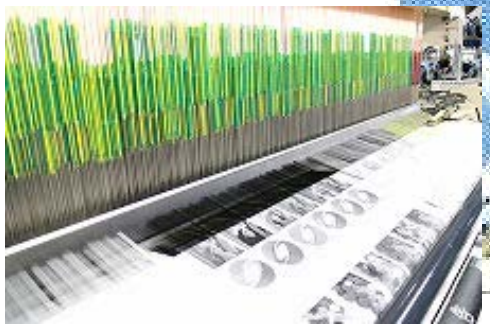
2017 Associate Professor @ University of Yamanashi

Design x IoT x AI (Design Lab. 2020~)

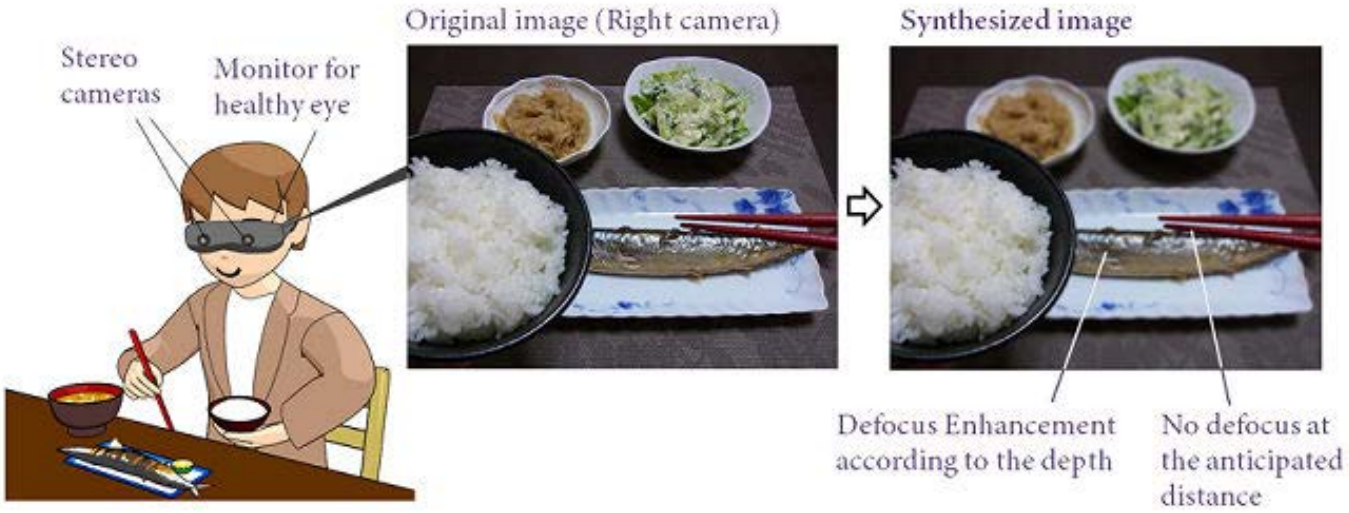
http://www.design.yamanashi.ac.jp/en/

Textile pattern generation





Mono-glass for Providing Distance Information

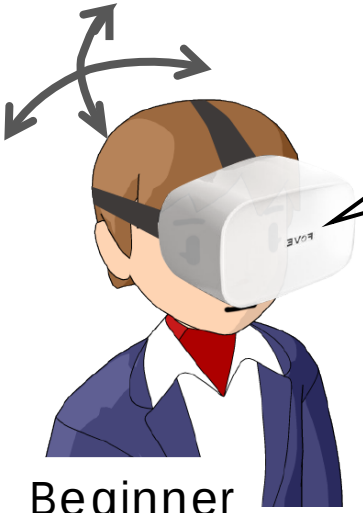


We propose **mono-glass** for providing distance information for people losing sight in one eye. We implemented a pilot system of mono-glass under careful consideration of precision, real-time processing and intuitive presentation. The loss of sight in one eye disables binocular disparity processing and makes short-range activities difficult. The proposed mono-glass is a wearable device with two cameras and one display. The two cameras capture the images on behalf of user's eyes. Depth information is then reconstructed from the captured images and visualized with defocusing for the healthy eye.

(Articles on Communications of the ACM, MIT Technology Review Online, GIZMODO.com, EXTREME TECH, phys.org, and ksa)

Clip image from 360 degrees video

+ Gazing points



Beginner

Video with gazing points

Master

IoT Sensor Analysis for Lessons

Assignment 3: Literature Review – 20pts

1. Select one paper from CVPR 2021

   https://openaccess.thecvf.com/CVPR2021?day=all

   CV (Geometry, Modeling) > PR (Pattern Recognition)

2. Make the list for selected papers (by Feb 28)

   Put the title which your selected as an open comment

3. Introduce the paper in 7 mins with slides (9th lesson~)

   (Abstract, the last part of Introduction, Result, Core idea & algorithms …)

4. Questions and discussions in 3 mins

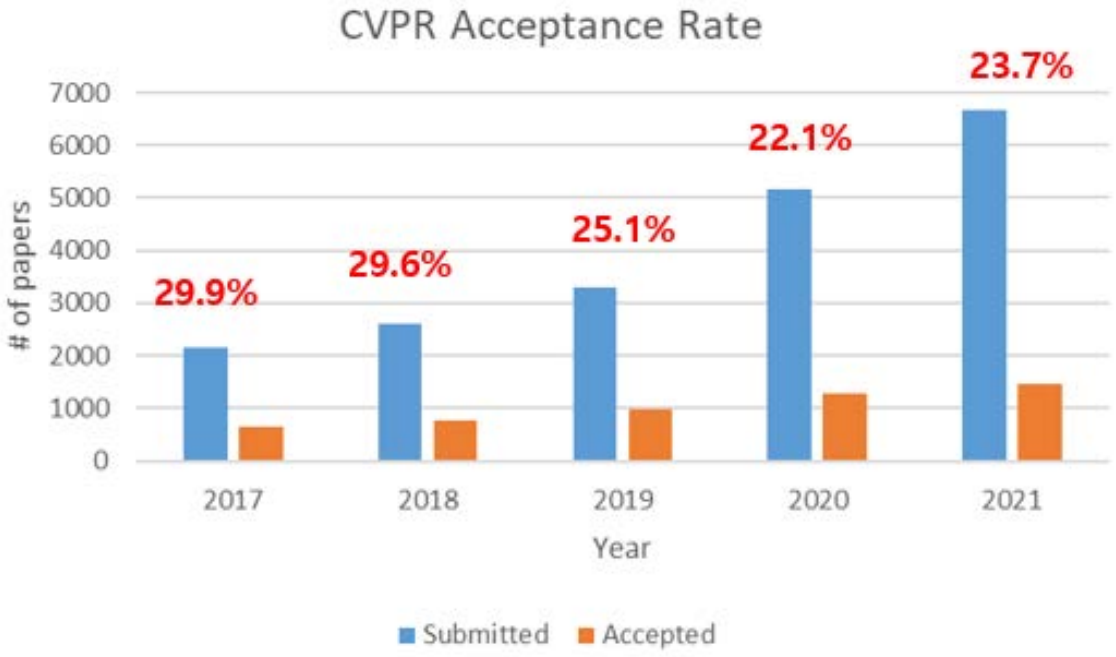
5. Add additional slides which describe on Q&A

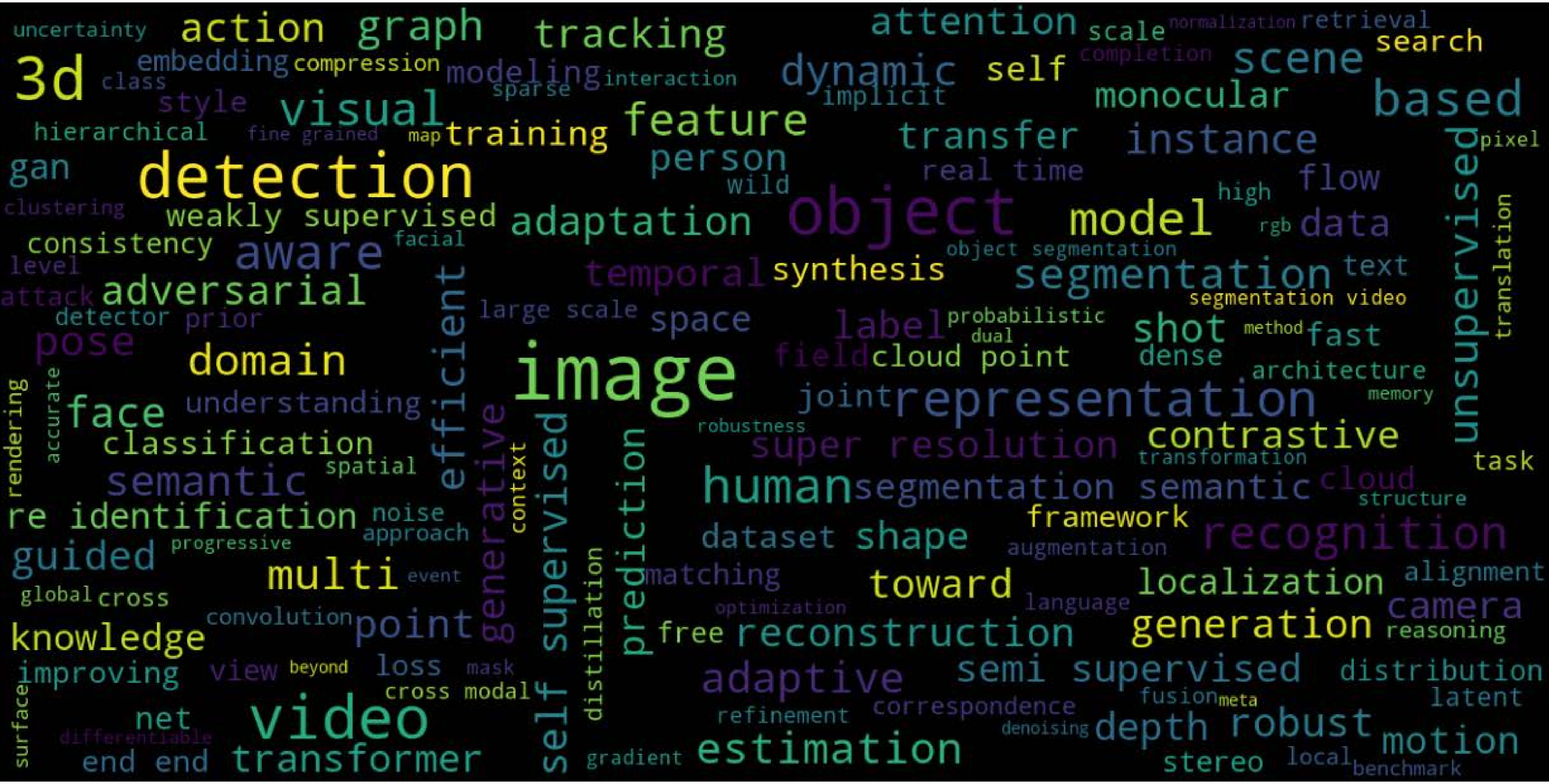
Presentation (5pts) & slides (10pts), Discussion (5pts)

Computer Vision Papers

https://yassouali.github.io/ml-blog/cvpr2021/

The 2021 CVPR conference, one of the main computer vision and machine learning conferences, concluded its second 100% virtual version last week with a record of papers presented at the main conference. Of about 7500 submissions, 5900 made it to the decision making process and 1660 papers (vs 1467 papers last year) were accepted with an acceptance rate of 23.7% (vs 22.1% last year). Such a huge (and growing) number of papers can be a bit overwhelming, so to get a feel of the general trends at the conference this year, I will present in this blog post a quick look of the conference by summarizing some papers (& listing some) that seemed interesting to me.




uncertainty
action
graph
tracking
attention
scale
normalization
retrieval
search
3d
embedding
compression
class
modeling
interaction
sparse
dynamic
self
completion
scene
implicit
monocular
based
style
visual
feature
hierarchical
fine grained
map
training
transfer
instance
pixel
gan
detection
person
real time
wild
flow
clustering
high
weakly supervised
adaptation
object
model
rgb
data
consistency
facial
aware
object segmentation
unsupervised
translation
level
temporal
synthesis
segmentation
text
attack
adversarial
segmentation video
detector
prior
large scale
space
label
probabilistic
efficient
pose
dual
shot
method
fast
domain
field
cloud point
dense
image
architecture
rendering
accurate
face
understanding
joint
representation
memory
robustness
generative
classification
super resolution
contrastive
transformation
context
spatial
task
semantic
human
segmentation
semantic
cloud
structure
prediction
re identification
noise
approach
framework
self supervised
dataset
shape
recognition
guided
progressive
augmentation
distillation
multi
event
matching
toward
localization
alignment
global
cross
optimization
language
camera
convolution
point
knowledge
free
reconstruction
generation
reasoning
improving
view
beyond
loss
mask
adaptive
semi supervised
distribution
cross modal
fusion
meta
latent
surface
net
video
refinement
correspondence
differentiable
denoising
depth
robust
motion
end end
transformer
gradient
estimation
stereo
local
benchmark

https://scholar.google.co.jp/citations?view_op=top_venues&hl=en

Google Scholar

Top publications

Categories English

	Publication	h5-index	h5-median
1.	Nature	414	607
2.	The New England Journal of Medicine	410	704
3.	Science	391	564
4.	IEEE/CVF Conference on Computer Vision and Pattern Recognition	356	583
5.	The Lancet	345	600
6.	Advanced Materials	294	406
7.	Cell	288	459
8.	Nature Communications	287	389
9.	Chemical Reviews	270	434
10.	International Conference on Learning Representations	253	470
11.	JAMA	253	446
12.	Neural Information Processing Systems	245	422
13.	Proceedings of the National Academy of Sciences	245	337

TED ~ Fei-Fei Li

https://www.ted.com/speakers/fei_fei_li

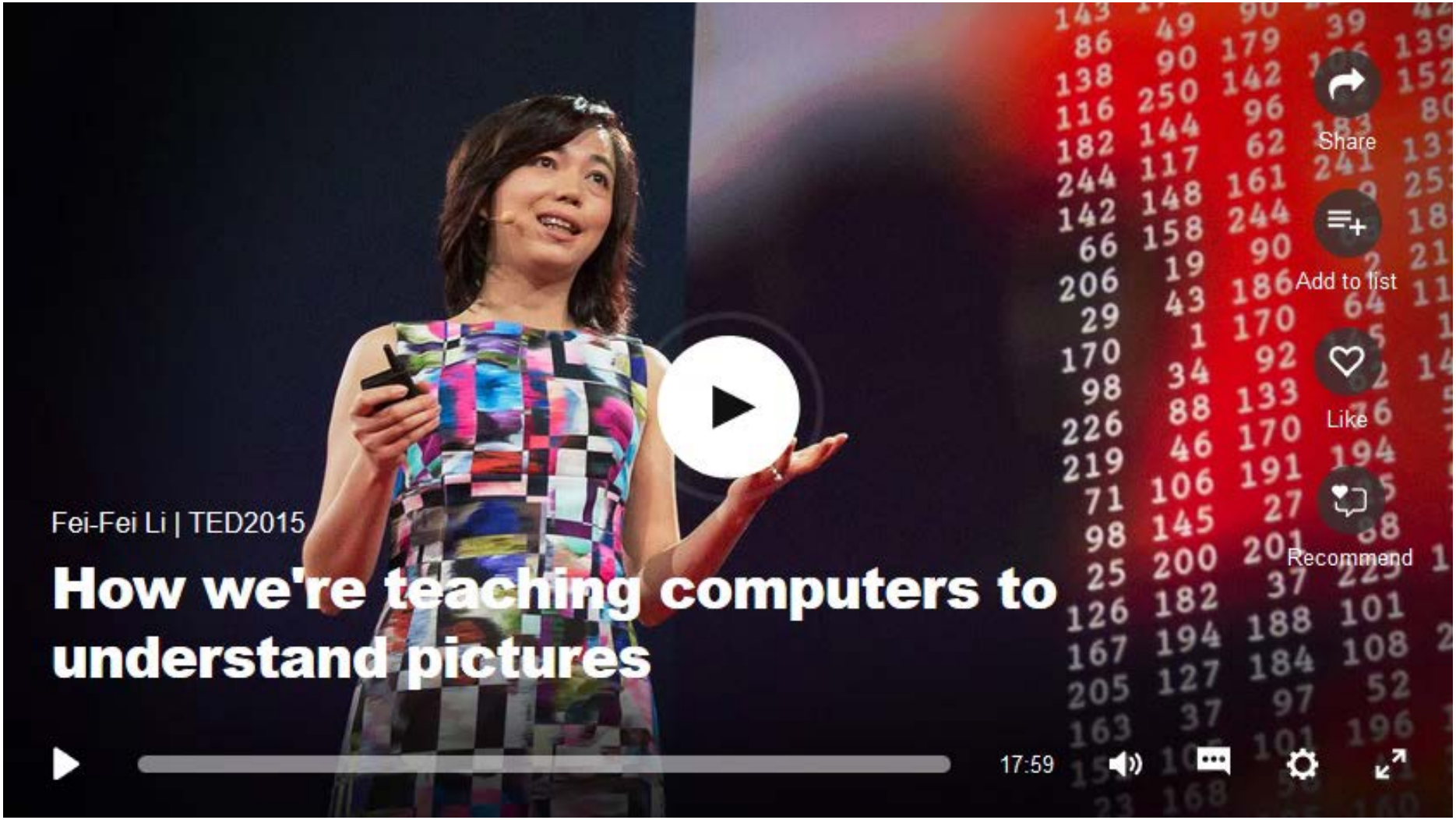


Image Processing vs. Computer Vision

Image processing

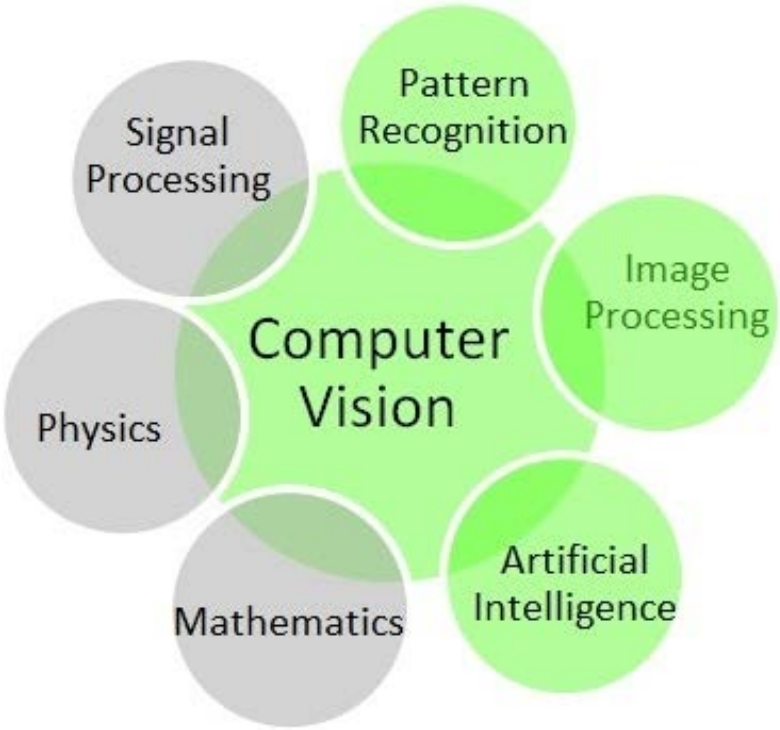
- Input: image
- Output: image
- e.g. de-noising, mpeg, edge, recoloring, …

Computer vision

- Input: image
- Output: symbol, shape, …
- e.g. face recognition, object tracking, shape reconstruction, …

“ill-posed” problem

- Mathematically unsolvable
- ex) 2D-3D



OpenCV

OpenCV (Open Source Computer Vision Library) is released under a BSD license and hence it's free for both academic and commercial use. It has C++, Python and Java interfaces and supports Windows, Linux, Mac OS, iOS and Android. OpenCV was designed for computational efficiency and with a strong focus on real-time applications. Written in optimized C/C++, the library can take advantage of multi-core processing. Enabled with OpenCL, it can take advantage of the hardware acceleration of the underlying heterogeneous compute platform.

DNN platforms always require Python & OpenCV !

Homework

1. Select one paper from CVPR2021 (1,660 papers!)
   Put the title in Google Classroom > Assignment 3,
   two or more students cannot read the same paper
2. Check your environment – sample code outputs image
3. Check the availability of video recording in ZOOM for Assignment 3
   (Slide show + your voice)