Machine Learning in Imaging

BME 590L
Roarke Horstmeyer

Lecture 24: Looking ahead and hyper-parameter optimization

Schedule for lectures and assignments

- Tuesday 4/9: Guest lecture – Dr. Nikhil Naik, MIT Media Lab, Harvard, Salesforce
- Thursday 4/11: ”Deep Imaging” Summary
- Tuesday 4/16: Guest lecture – Kevin Zhou, Deep Image Priors & Current GANs
- Saturday 4/20: Homework #5 due
- Monday 4/29, 9am – noon: Final project presentations

Components of final project

Total grade % (after re-distribution of quizzes) – 41%

- 5 minute presentation (share slides afterwards) – 8%
- 3-6 page write up with at least 3 figures and 5 references – 20%
    - Introduction, related work, methods, results, discussion
- share code for submission – 9%
- brief website template & permission to share results – 5%
- shared annotated datasets & permissions – no grade, but would be much appreciated

Final webpage – I'll try to find a suitable template for you guys to fill out

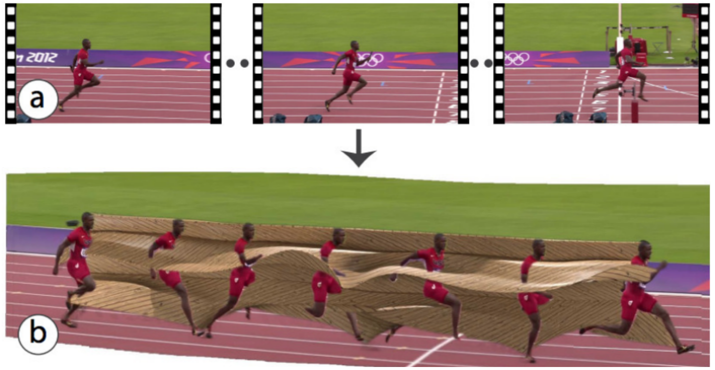
MoSculp: Interactive Visualization of Shape and Time

Xiuming Zhang [1] Tali Dekel [2] Tianfan Xue [2] Andrew Owens [3] Qiurui He [1] Jiajun Wu [1] Stefanie Mueller [1] William T. Freeman [1, 2]

[1] MIT CSAIL [2] Google Research [3] UC Berkeley

Abstract

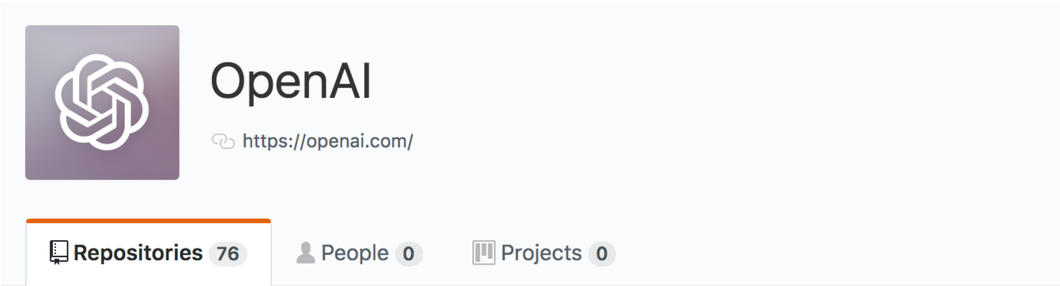
We present a system that visualizes complex human motion via 3D motion sculptures---a representation that conveys the 3D structure swept by a human body as it moves through space. Our system computes a motion sculpture from an input video, and then embeds it back into the scene in a 3D-aware fashion. The user may also explore the sculpture directly in 3D or physically print it. Our interactive interface allows users to customize the sculpture design, for example, by selecting materials and lighting conditions.

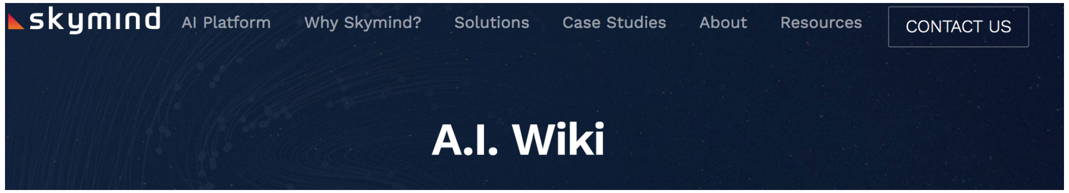


Where are things going with Machine Learning and Imaging in 10 years?

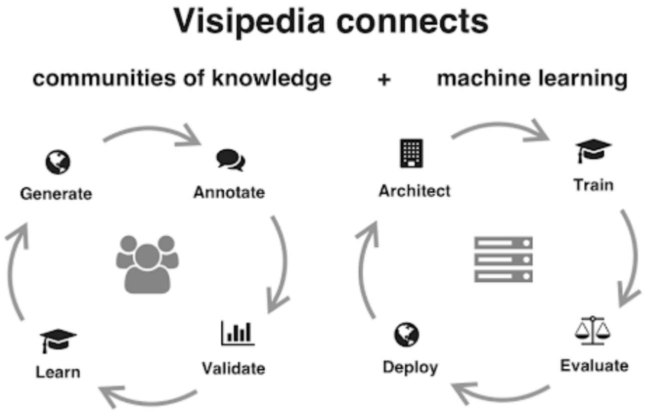
Where are things going with Machine Learning and Imaging in 10 years?

1. Proliferation of trained models, similar datasets and shared goals





Visipedia

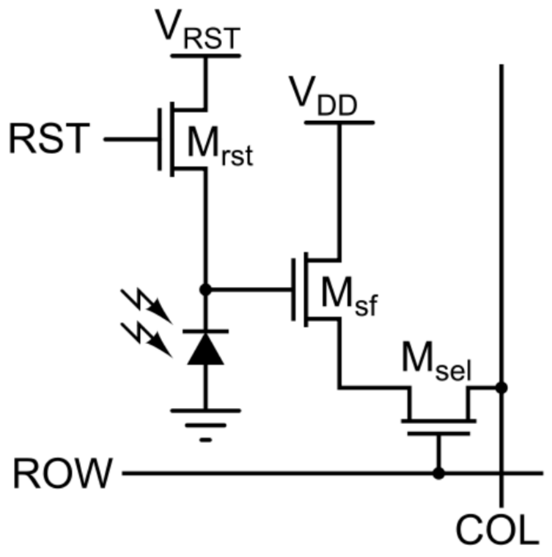


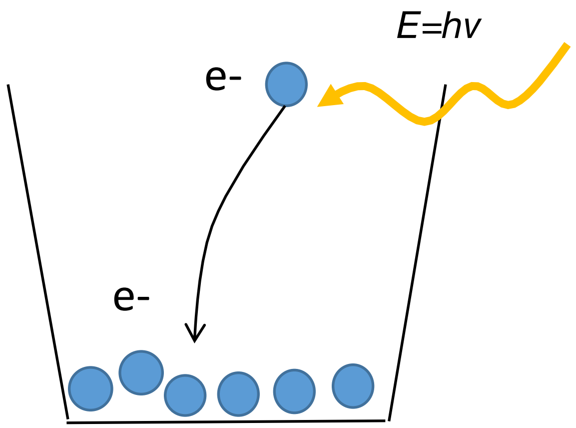
Caltech Visipedia

Where are things going with Machine Learning and Imaging in 10 years?

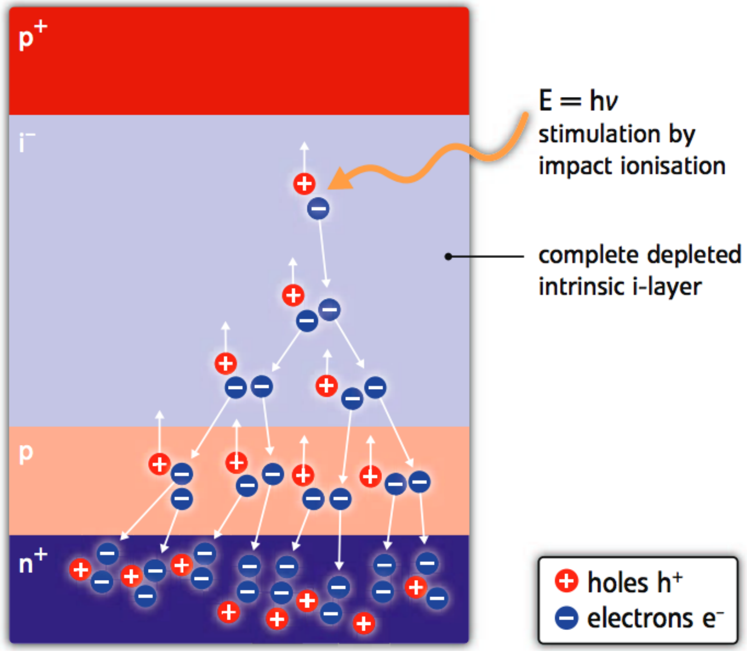
2. "Cameras" on many devices & new types of sensors

Standard CMOS pixel = bucket that collects electrons



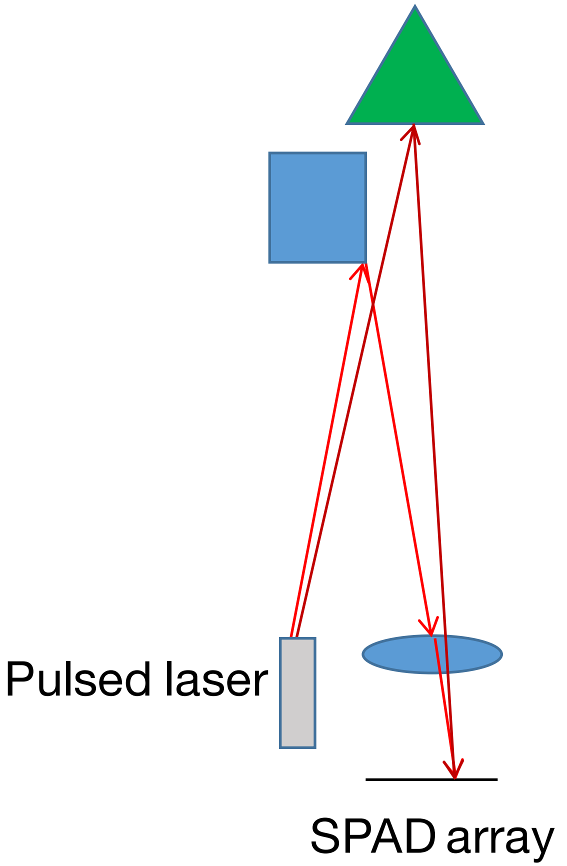


SPAD pixel: was there a photon or not?

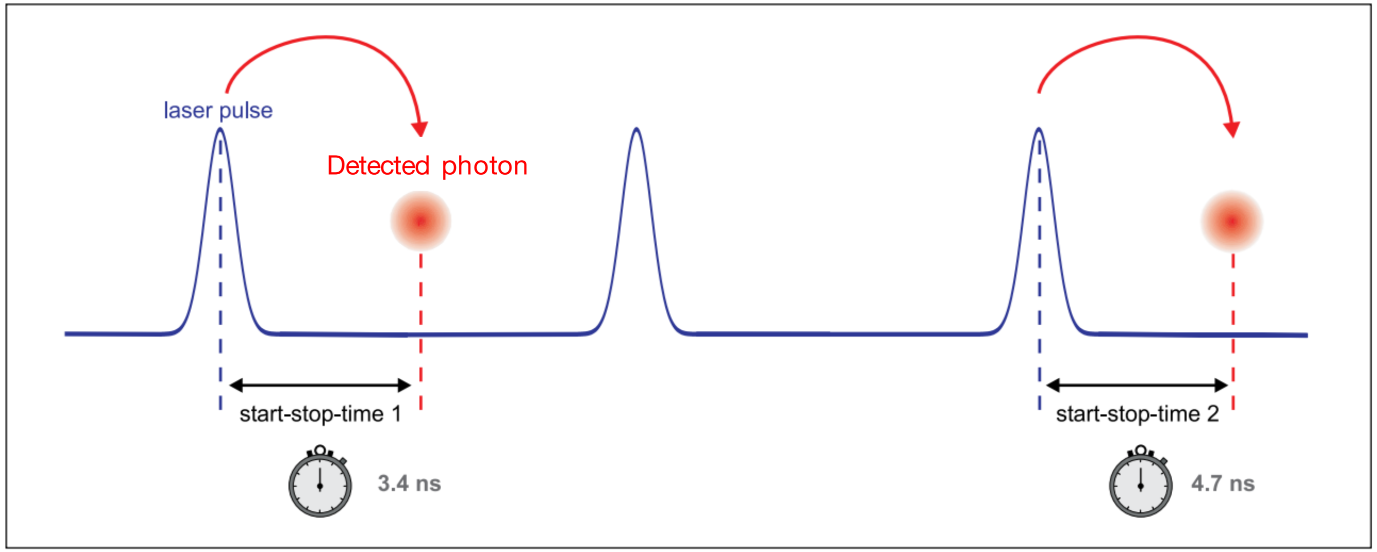


Where are things going with Machine Learning and Imaging in 10 years?

2. “Cameras” on many devices & new types of sensors



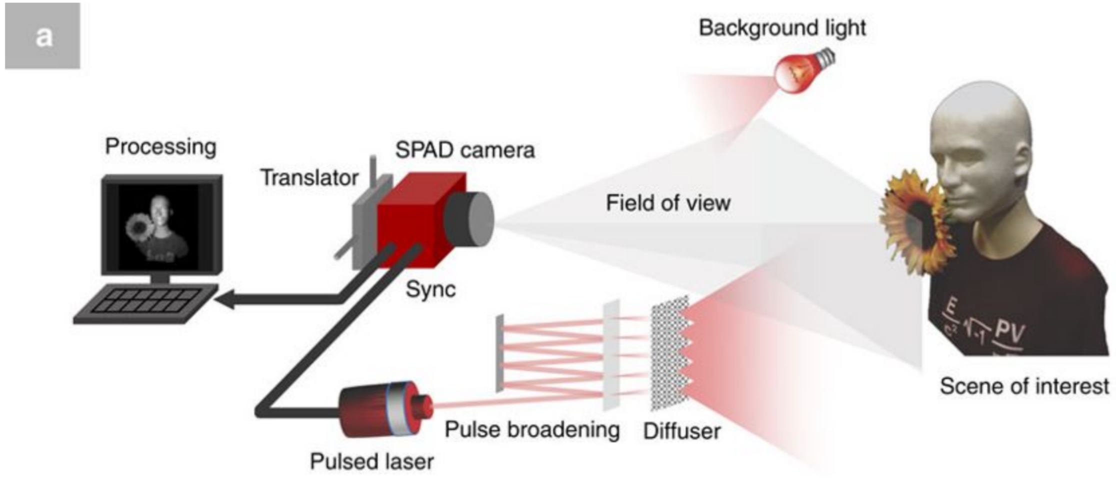
- Light travels 1 ft in 1 ns.
- SPADs can precisely photon arrival time to measure travel distance (TOF)

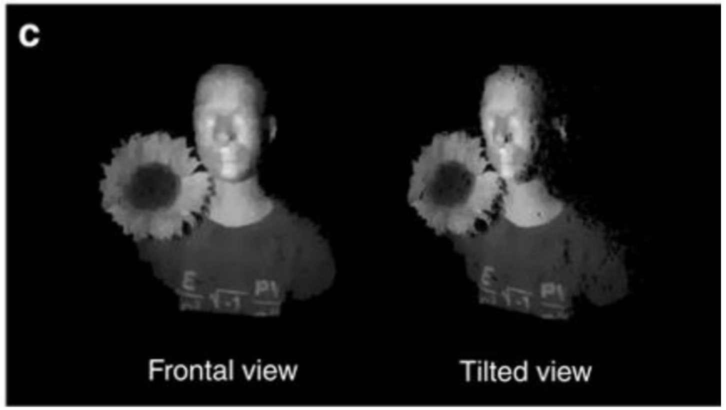


https://www.picoquant.com/images/uploads/page/files/7253/technote_tcspc.pdf

Where are things going with Machine Learning and Imaging in 10 years?

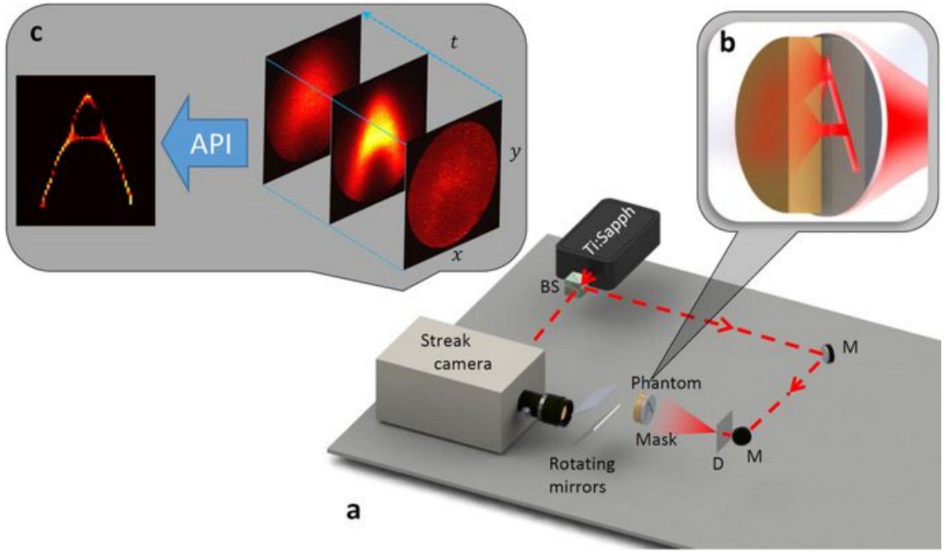
2. “Cameras” on many devices & new types of sensors





D. Shen et al, https://www.nature.com/articles/ncomms12046

Figure 1: Imaging Through Thick Scattering.



[G. Satat et al, https://www.nature.com/articles/srep33946](https://www.nature.com/articles/srep33946)

Where are things going with Machine Learning and Imaging in 10 years?

3. Beyond convolutions - new constructs for deep networks

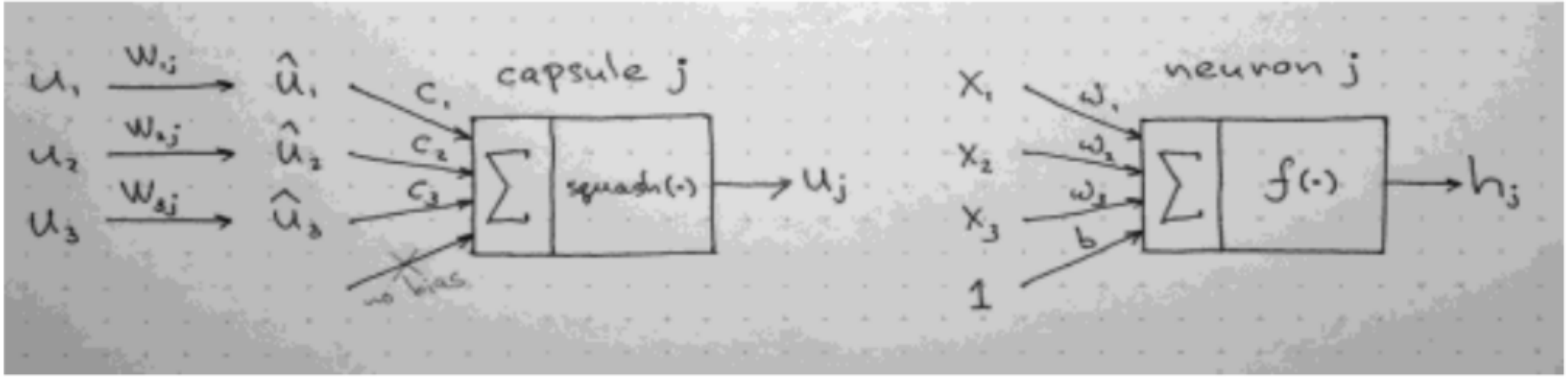
Where are things going with Machine Learning and Imaging in 10 years?

3. Beyond convolutions - new constructs for deep networks

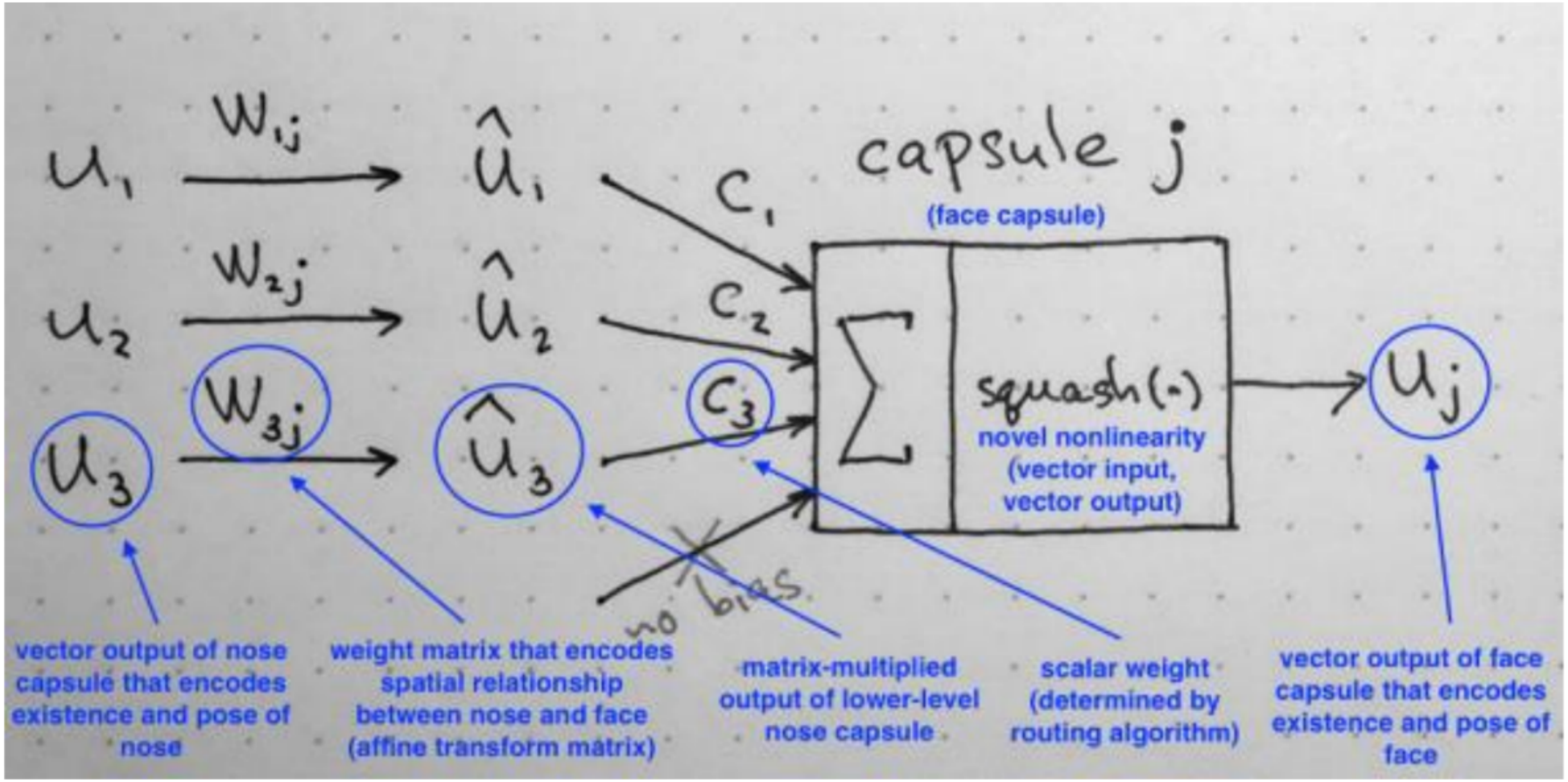
Dynamic Routing Between Capsules

Sara Sabour **Nicholas Frosst**

Geoffrey E. Hinton
Google Brain
Toronto
{sasabour, frosst, geoffhinton}@google.com



capsule j
u_1 W_{1j} $\hat{u}_1$ c_1
u_2 W_{2j} $\hat{u}_2$ c_2
u_3 W_{3j} $\hat{u}_3$ c_3
$\sum$ squash(·)
u_j
no bias
neuron j
x_1 w_1
x_2 w_2
x_3 w_3
1 b
$\sum$ f(·)
h_j



capsule j
(face capsule)
u_1 W_{1j} $\hat{u}_1$ c_1
u_2 W_{2j} $\hat{u}_2$ c_2
u_3 W_{3j} $\hat{u}_3$ c_3
$\sum$ squash(·)
novel nonlinearity (vector input, vector output)
u_j
no bias
vector output of nose capsule that encodes existence and pose of nose
weight matrix that encodes spatial relationship between nose and face (affine transform matrix)
matrix-multiplied output of lower-level nose capsule
scalar weight (determined by routing algorithm)
vector output of face capsule that encodes existence and pose of face

Where are things going with Machine Learning and Imaging in 10 years?

3. Beyond convolutions - new constructs for deep networks

Capsule vs. Traditional Neuron			
Input from low-level capsule/neuron		vector($\mathbf{u}_i$)	scalar(x_i)
Operation	Affine Transform	$\hat{\mathbf{u}}_{j\|i} = \mathbf{W}_{ij}\mathbf{u}_i$	–
	Weighting	$\mathbf{s}_j = \sum_i c_{ij}\hat{\mathbf{u}}_{j\|i}$	$a_j = \sum_i w_i x_i + b$
	Sum		
	Nonlinear Activation	$\mathbf{v}_j = \frac{\|\mathbf{s}_j\|^2}{1+\|\mathbf{s}_j\|^2}\frac{\mathbf{s}_j}{\|\mathbf{s}_j\|}$	$h_j = f(a_j)$
Output		vector($\mathbf{v}_j$)	scalar(h_j)

3. Therefore it will not assign one pixel to two digits if one of them does not have any other support.