fNIRS 2022
October 9-12, Boston, MA

STORM-Net: Open-source toolbox for automatic subject-specific co-registration of probe placements for developmental and clinical populations

Based On

STORM-Net: Simple and Timely Optode Registration Method for Functional Near-Infrared Spectroscopy (fNIRS)

Yotam Erel[1] Sagi Jaffe-Dax[1,2], Yaara Yeshurun[1], Amit H. Bermano[1]

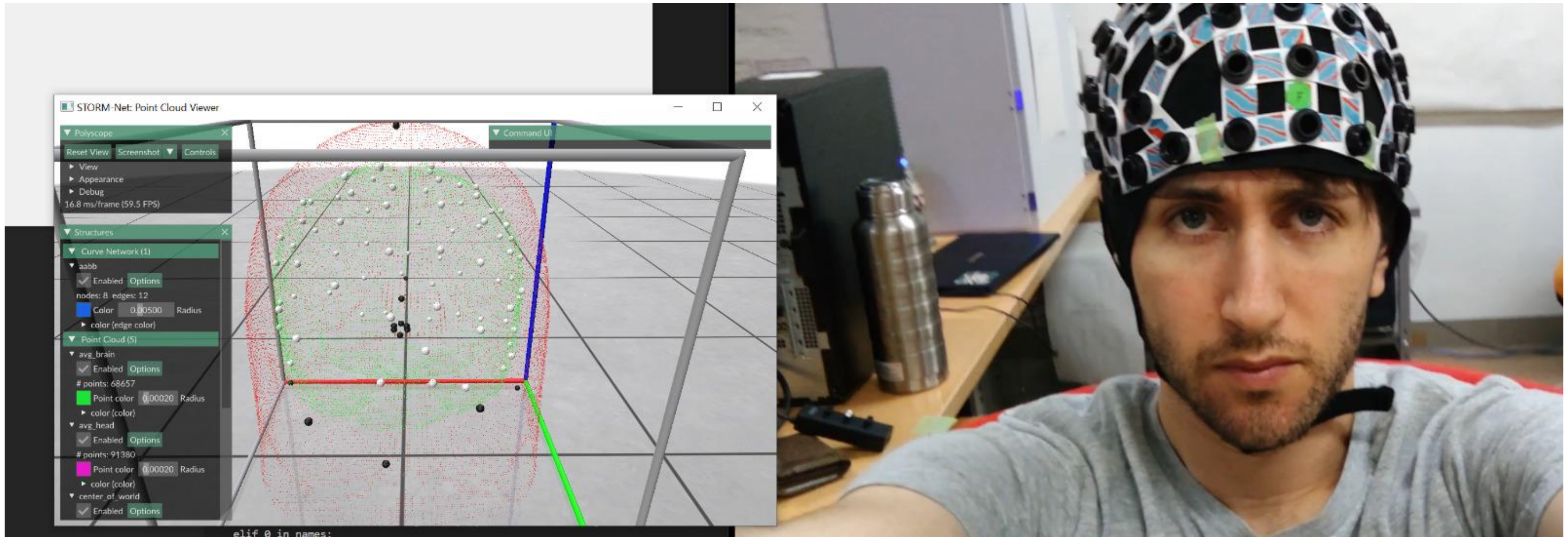
Tel-Aviv University[1], Princeton University[2]

First thing's first !

- Let's install Storm-Net (Windows / Linux)
  - We need git and Miniconda (free)
- https://github.com/yoterel/STORM-Net
- **Don't be afraid of terminal!**

1. Open Anaconda Prompt
2. Get source code (Python):
   - "git clone https://github.com/yoterel/STORM-Net.git storm"
3. Navigate to directory:
   - "cd storm/CapCalibrator"
4. Download dependencies:
   - "conda env create -n storm-net -f environment.yml"
5. Download model files (manually)
6. Activate virtual environment:
   - "conda activate storm-net"
7. Run:
   - "python main.py --mode gui"

Topics



- Introduction
- Method
- Results
- Live Demo
  - Need average sized head volunteers !

Introduction

- PhD candidate in Tel-Aviv University (currently visitor researcher in Osaka, Japan)

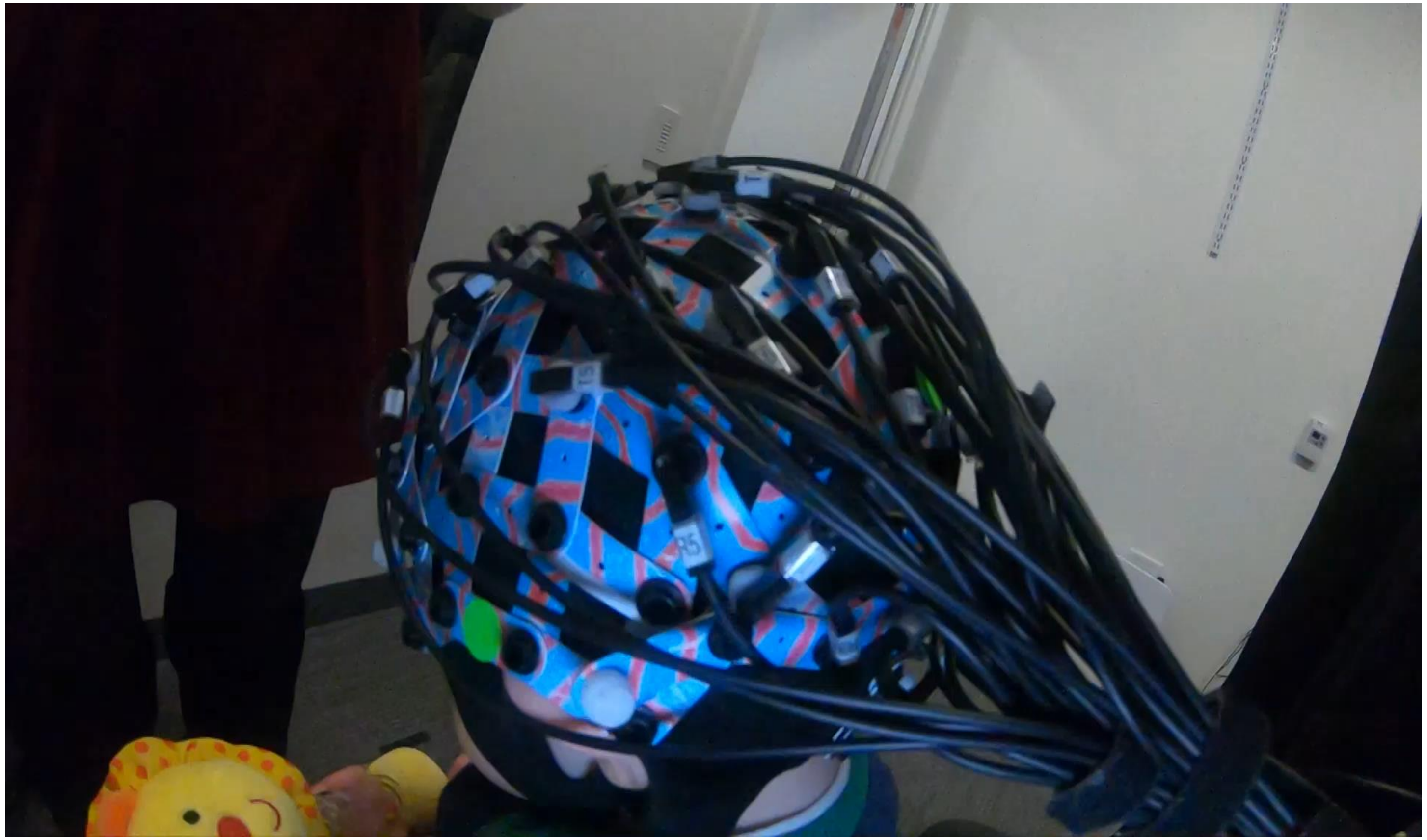
- Computer Science
  - Computer Vision (2D → xD, image understanding)
  - Graphics (3D → 2D, rendering)
  - Computational Geometry (3D → 3D, scene processing)

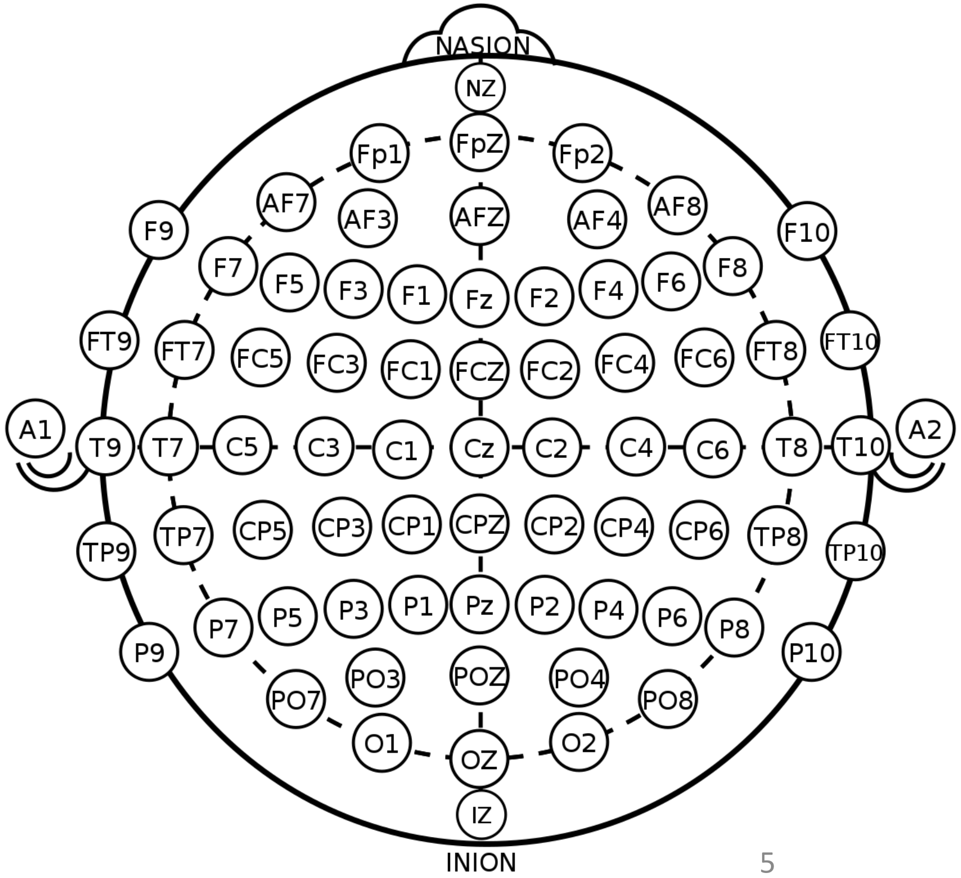
- I promise we will only have 1.5 math slides ☺

- STORM-Net developed to aid Sagi with registration with infants!

The Main Challenge

- **<u>Registering</u>** a cap to the scalp





Importance

Precise understanding of underlying cortical activity

- Reliable data
- Inter-subject variance reduction

Current Methods

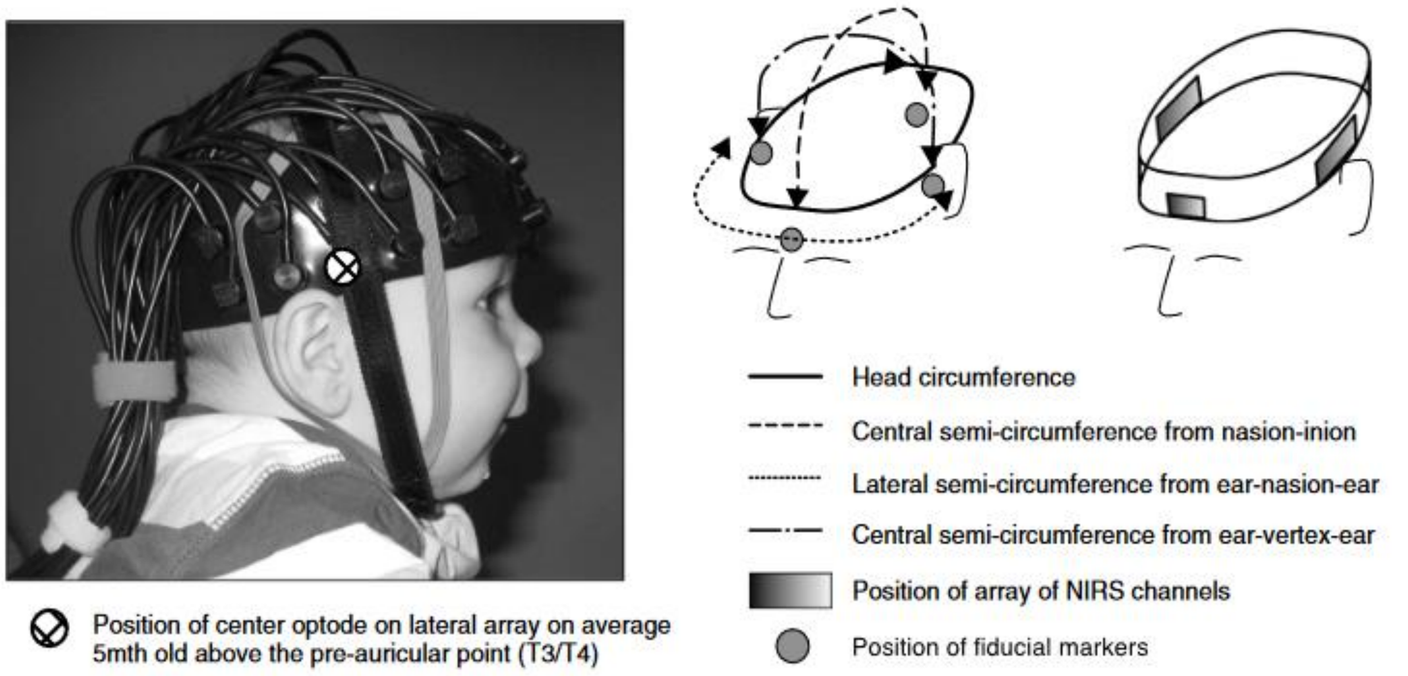
Do nothing !

- Surprisingly ok (sometimes)
- $Var = f(n1, n2, n3, \dots) =_{n1 \gg ni} \; n1$

Current Methods

Manual measurements

- Accuracy low
- Acquisition (~minutes)
- Computation (~minutes)
- Laborious



Lloyd-Fox et al. 2014

Current Methods

Digitizer

- Accurate* (~5mm)
- Acquisition (~minutes)
- Computation (~seconds)
- Expensive
- Physical Constraints
  - Metals
  - Movement
  - Location

Polhemus FASTRAK

*Mainly affected by human precision

Current Methods

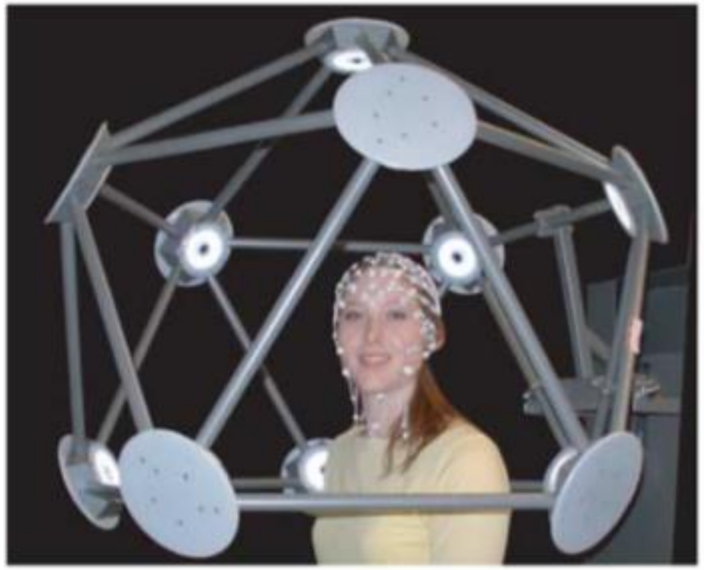
3D scanners (hardware-based)

- Accuracy high, medium*
- Acquisition (~seconds* - ~minutes)
- Computation (~minutes - ~hours*)
- Physical Constraints (sometimes)

*Requires manual stitching

Structure.io

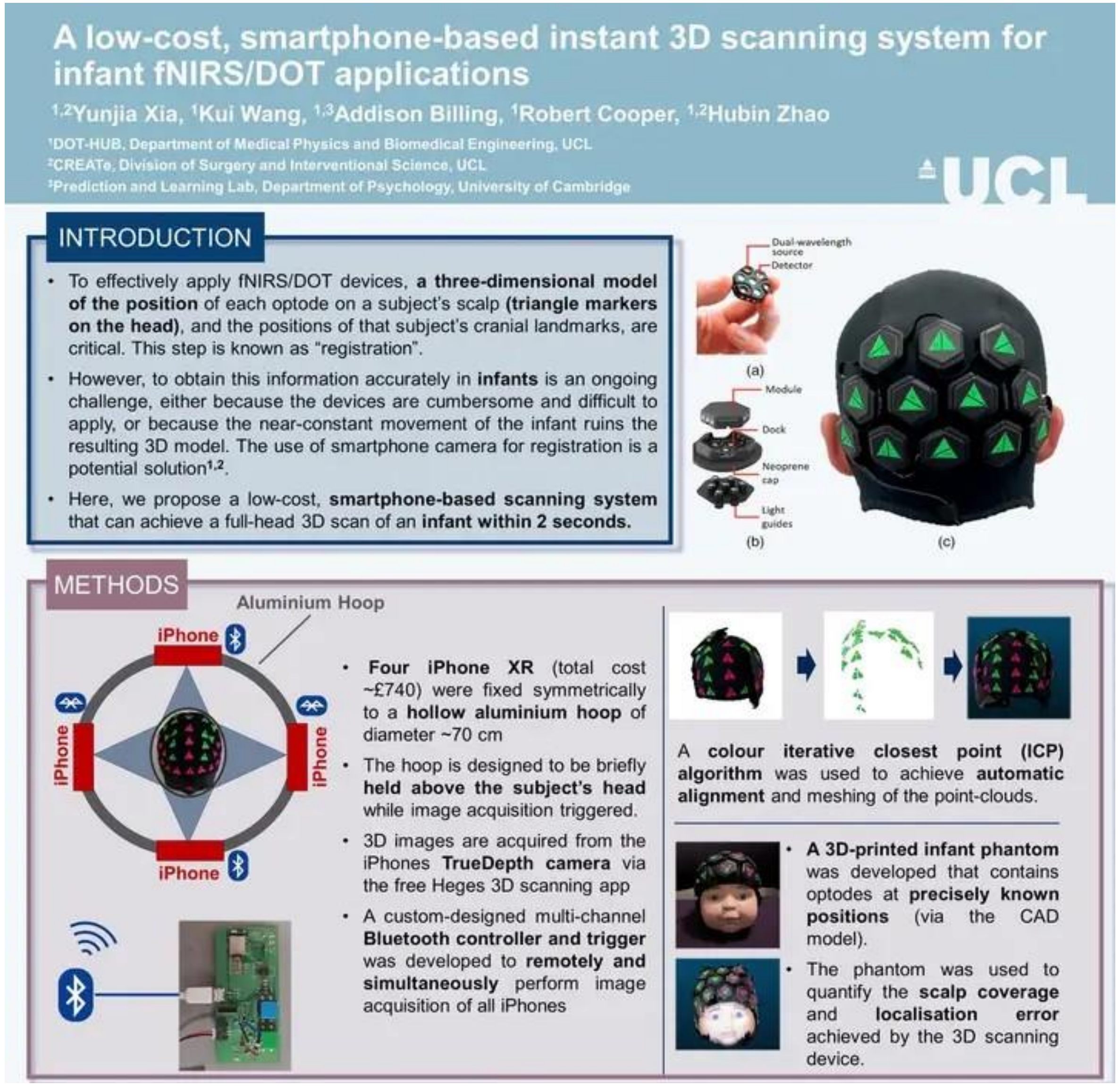
Artec



Geodesic Photogrammetry, Russell et al., 2005

fnirs 2022

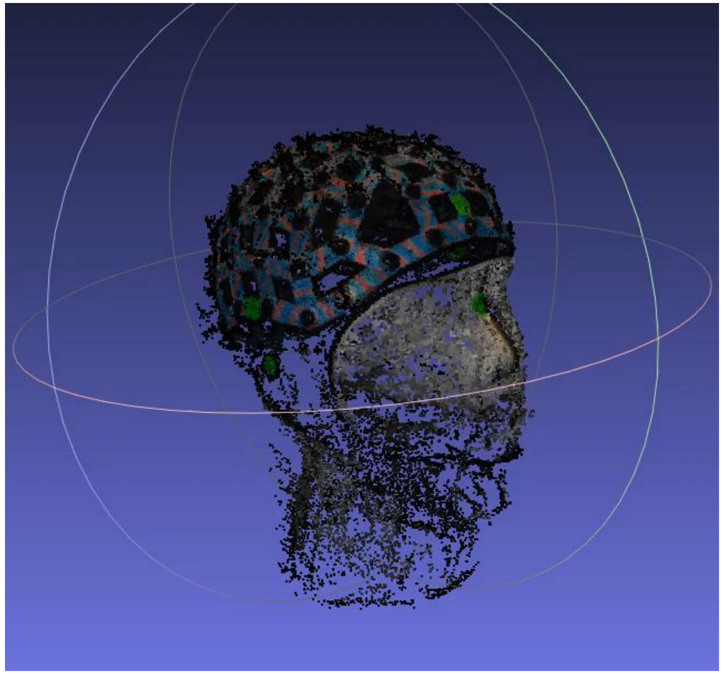
- Check out this poster !

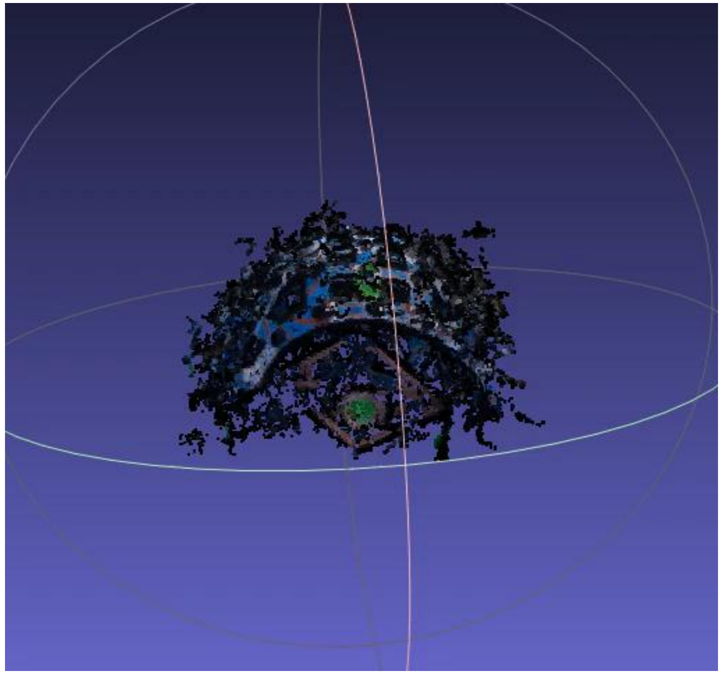


Current Methods

3D SfM from video

- Accurate*
- Acquisition (~minutes)
- Computation (~minutes)
- Cheap
- Physical Constraints (sometimes)





Jaffe-Dax et al. 2020

- Scanners perform SfM ! (fast)

Key Insights

- Goal
  - Reduce noise in measurements attributed to mislocated cap
- Constraints
  - Acquisition speed
    - Crucial for developmental / clinical populations
  - Computational speed
    - Avoid manual labor
    - Allow re-registration if possible
  - Robustness
    - Work in many scenarios
  - Price
  - Technical Expertise
    - Avoid the need for an expert's time

Questions ?

- Main challenge (informally)
- Current methods

Concentration time

Formal definition of problem

- Find F such that:

$$F(x_{expected}) = x_{actual}$$

- Assumption 1: we can describe this transformation with a ***linear transformation:***

$$F_{4x4}X_{4xn} = X'_{4xn}$$

- Assumption 2:
  - The cap is approximately rigid, rotating around some fixed point → F is a rotation matrix

- Assumption 3:
  - We don't need to worry about reflections → 3 rotational parameters (θ, ϕ, ξ)

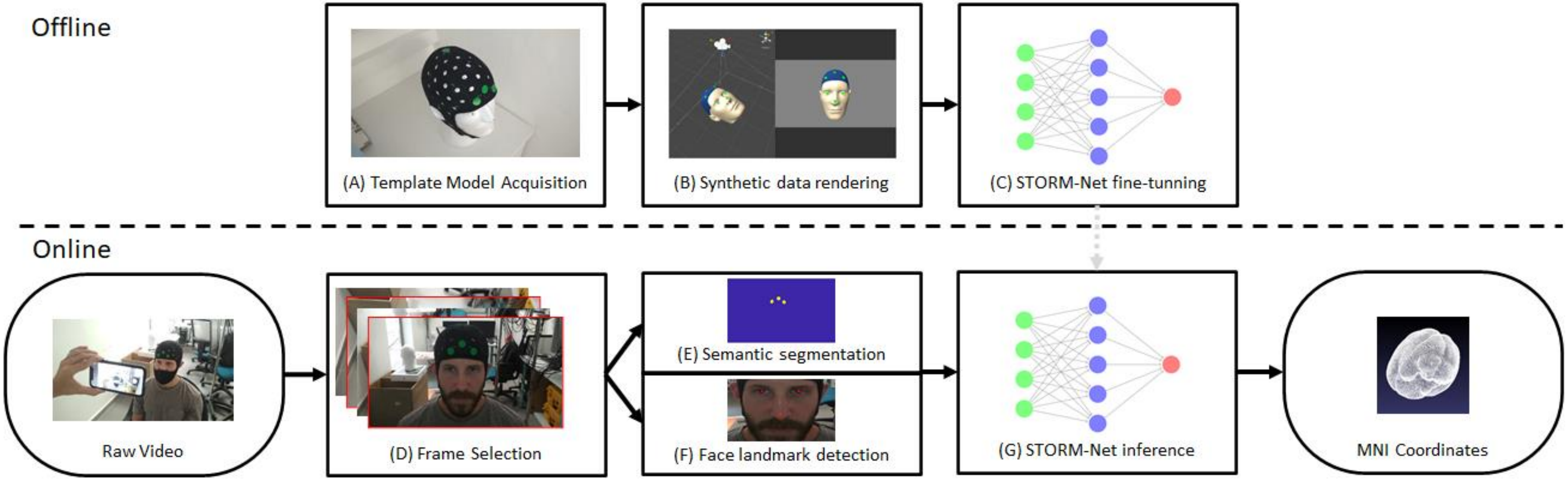
A sub problem

- We need to express x_{actual} in some known canonical coordinate system.

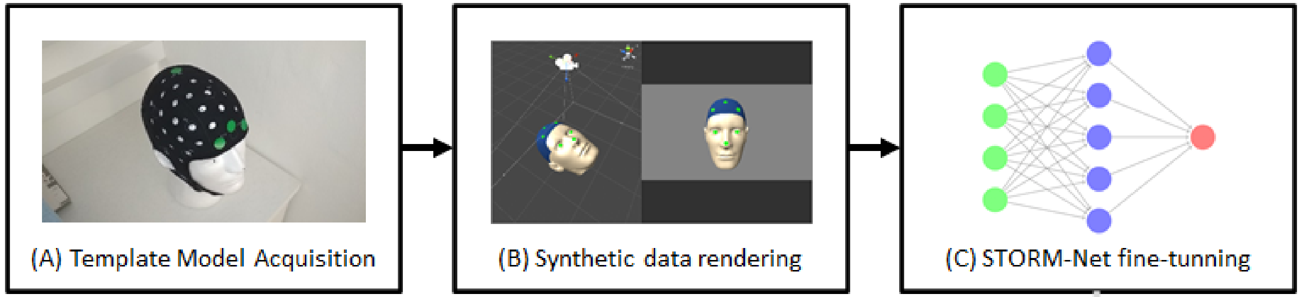
- If we didn't, they are in some arbitrary world coordinates, and we can't register them to some fMRI scan (why do this?)

- We use an anatomically correct puppet head to approximate subject anatomical landmarks
  - A possible source of error – lets discuss later !

STORM-Net


Offline
(A) Template Model Acquisition
(B) Synthetic data rendering
(C) STORM-Net fine-tunning
Online
Raw Video
(D) Frame Selection
(E) Semantic segmentation
(F) Face landmark detection
(G) STORM-Net inference
MNI Coordinates

Offline Step

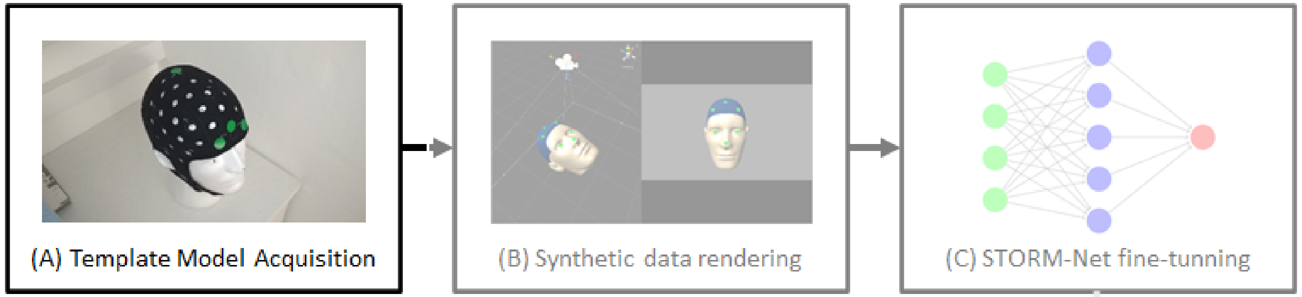


- Purpose:

  train a neural network

- Why use a neural network ?
  - Universal function approximators: $G(video) = (\theta, \phi, \xi)$
  - Does there exist such G?
    - Yes !
    - In theory we just need 2 frames capturing at least 4 known points
  - Can we find it?
    - Mostly yes

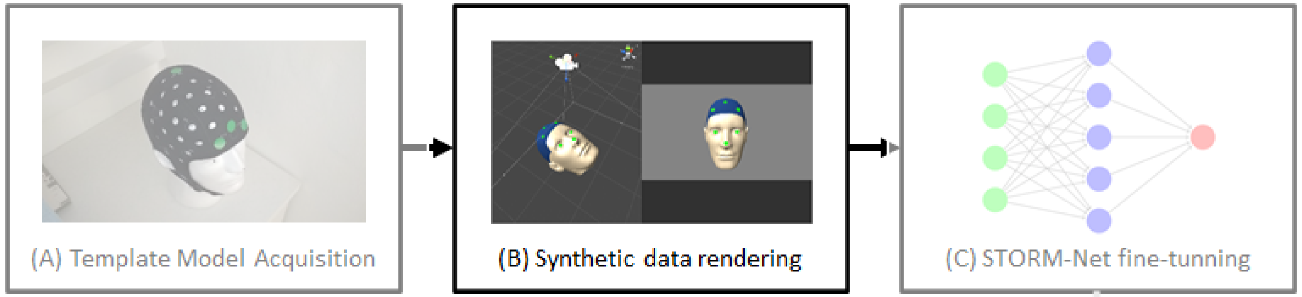
Offline Step



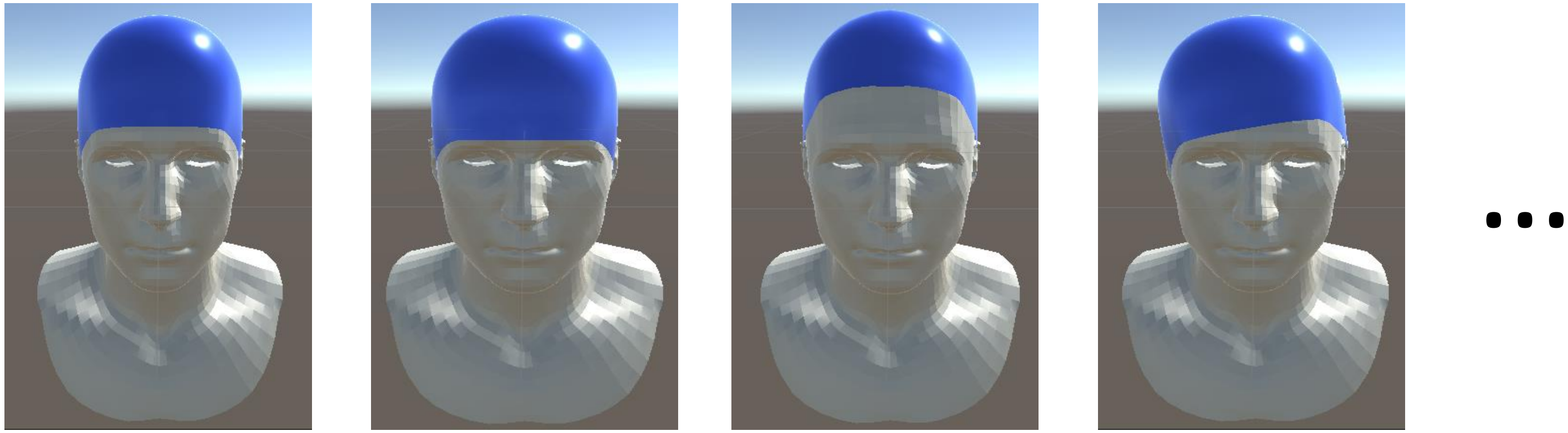
Measure points of interest on a model (30-60 minutes, once per cap)

- SfM (software)
- Digitizer
- 3D scanner

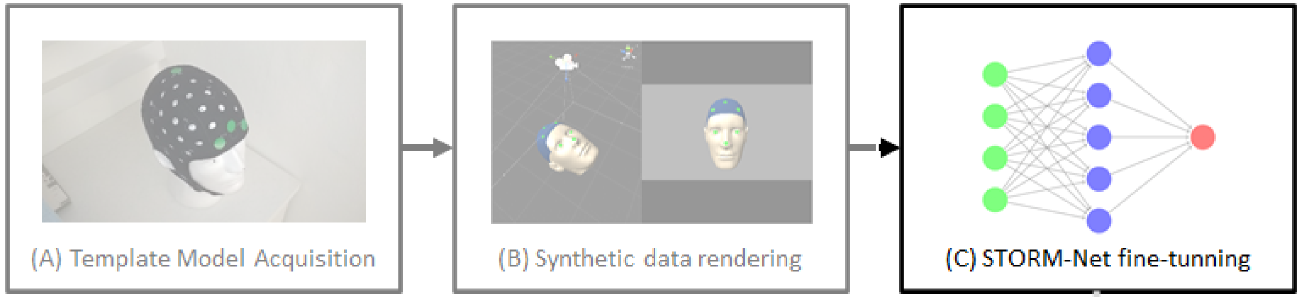
Offline Step



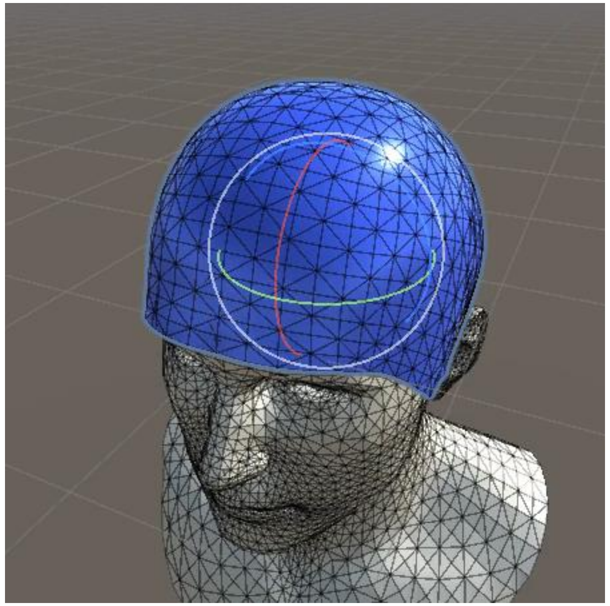
- Automatically render 100,000 synthetic “videos” (1 – 4 hours)
  - Simulate videos
  - Output consist of the “dataset” for training STORM-Net



Offline Step

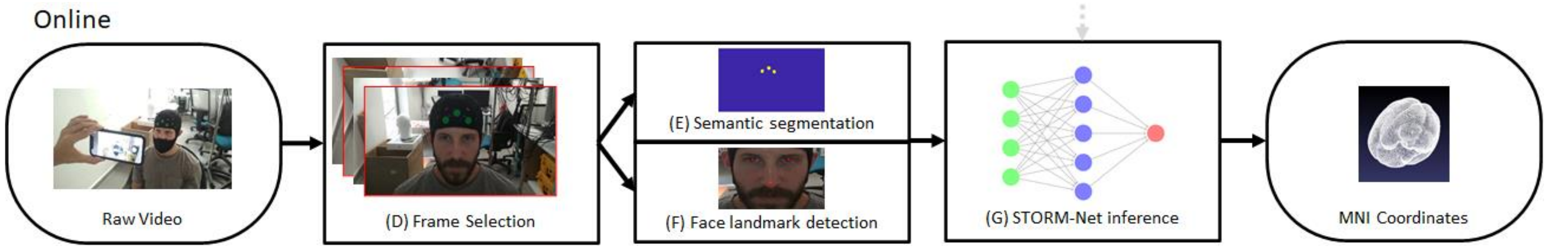


- Automatically train STORM-Net using the synthetic data
  - 30 minutes with GPU
  - 6 hours without
- STORM-Net is trained to predict rotation of the cap
  - Euler Angles ϕ, θ, ξ



Concentration time

Online Step

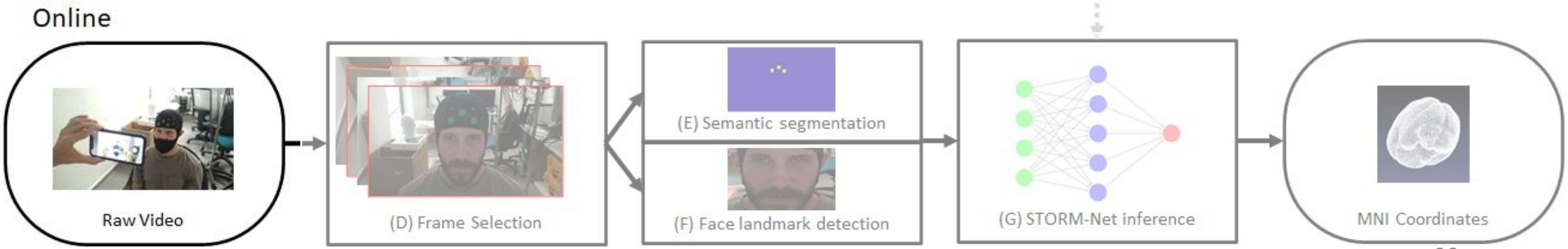


- Purpose:
  - Find parameters from video
  - Video dimensions: $(n, H, W, 3)$

- Video is reduced to its essentials:
  - $G(video) = (\theta, \phi, \xi) \rightarrow \mathrm{G} \stackrel{\text{def}}{=} G_2 \circ G_1$
  - $G(video) = G_2\big(G_1(video)\big)$
  - $G_1(video) = \text{landmarks}$
  - $G_2(landmarks) = (\theta, \phi, \xi)$

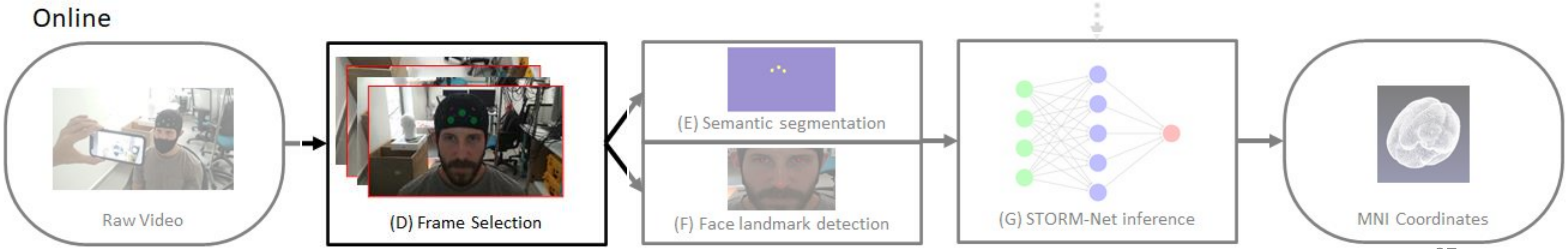
Online Step

- Take a video of the subject (5 seconds)



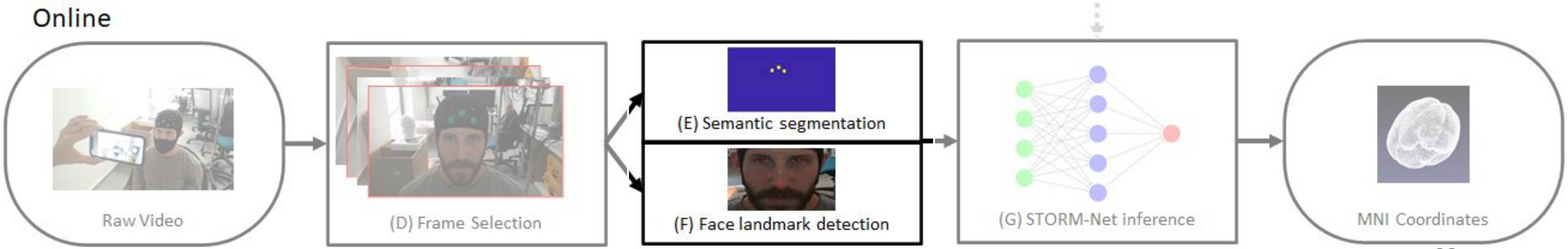
Online Step

- Frames are selected **manually** (1 minutes) / **automatically** (<1 second)
  - For automation, video is split into n bins
  - For each bin, least blurred frame is selected using variance of Laplacian (high pass filter).
    - Variance high – many edges and non edges as in a normal image
    - Variance low – tiny spread of responses, small amount of edges
  - We use n=10 for all experiments



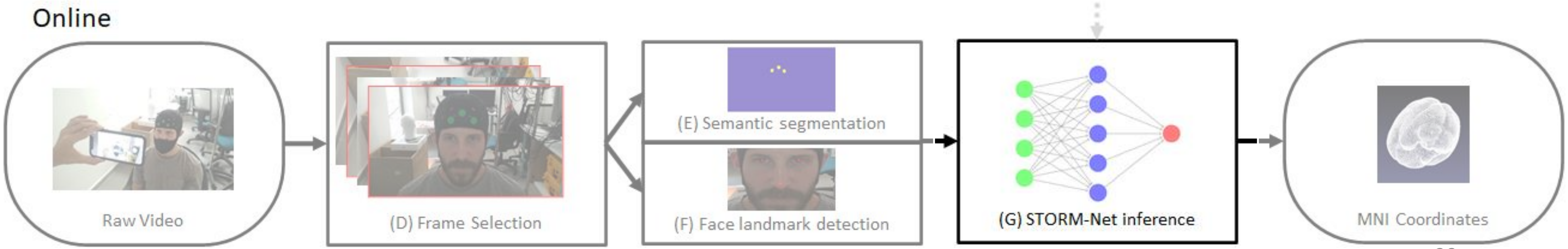
Online Step

- Facial & Cap landmarks are extracted **manually** (2 minutes) or **automatically** (< 1 second)
  - For automation, we use a pretrained segmentation neural network, and a landmark detection neural network.



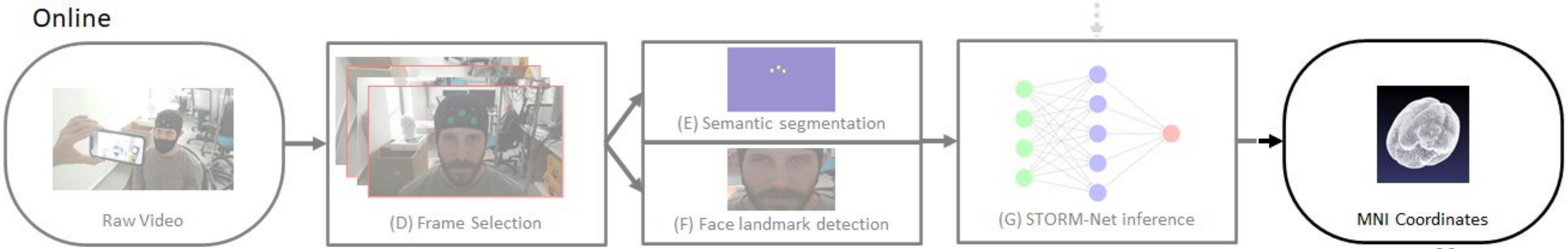
Online Step

- STORM-Net predicts registration parameters (< 1 second)
  - Reminder: network was trained to predict these in the offline step



Online Step

- We apply parameters to the template model (< 1 second)
  - It can further be transformed to a (statistical, or not) MNI coordinate system (head surface or projected to cortex).



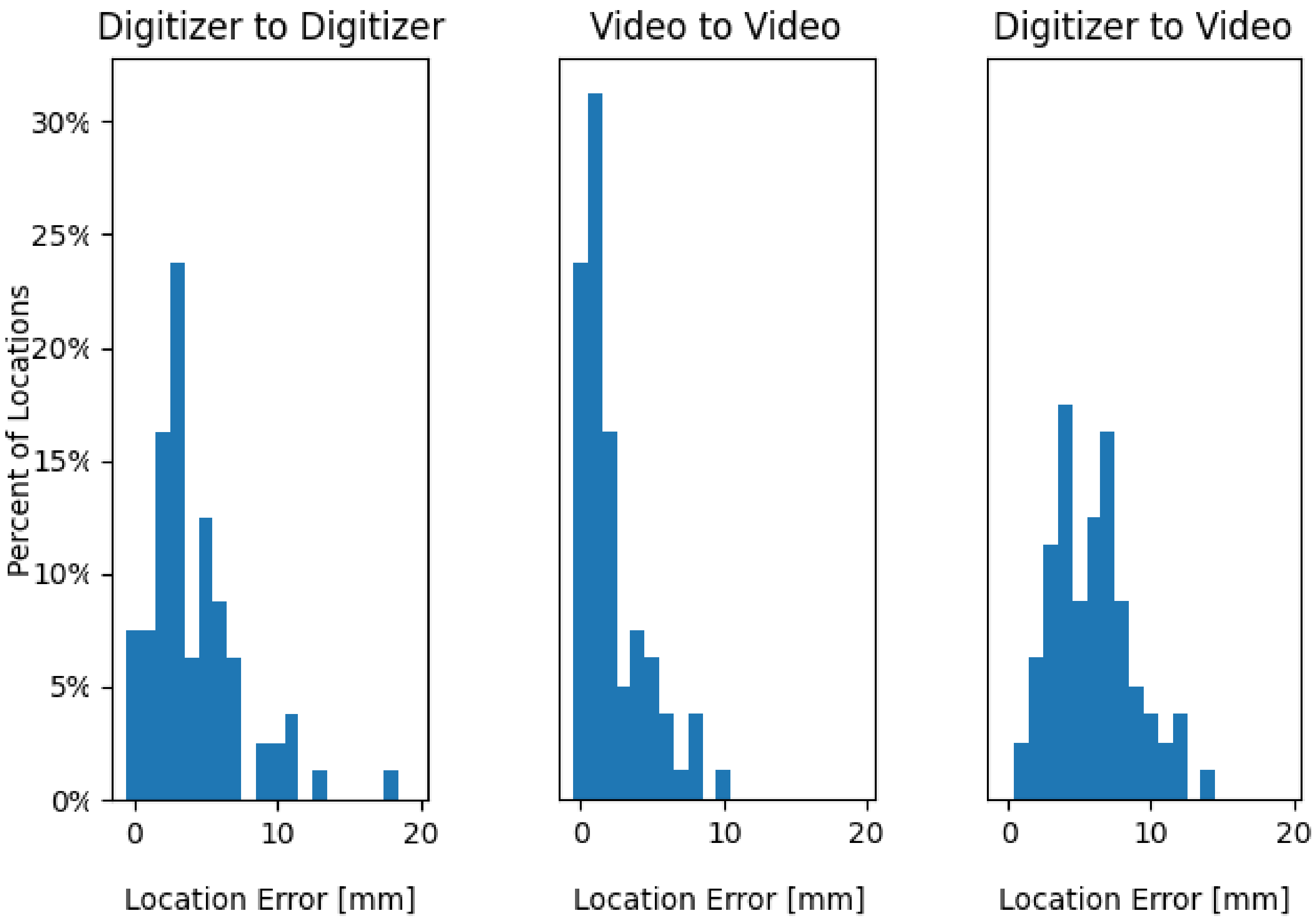
Qualitative Evaluation

$\theta = -4.3, \phi = -10.8, \xi = -3.4$

$\theta = -10.8, \phi = 9.3, \xi = 6.1$

$\theta = 2.4, \phi = 1.1, \xi = 5.4$

Quantitative Evaluation



Quantitative Evaluation

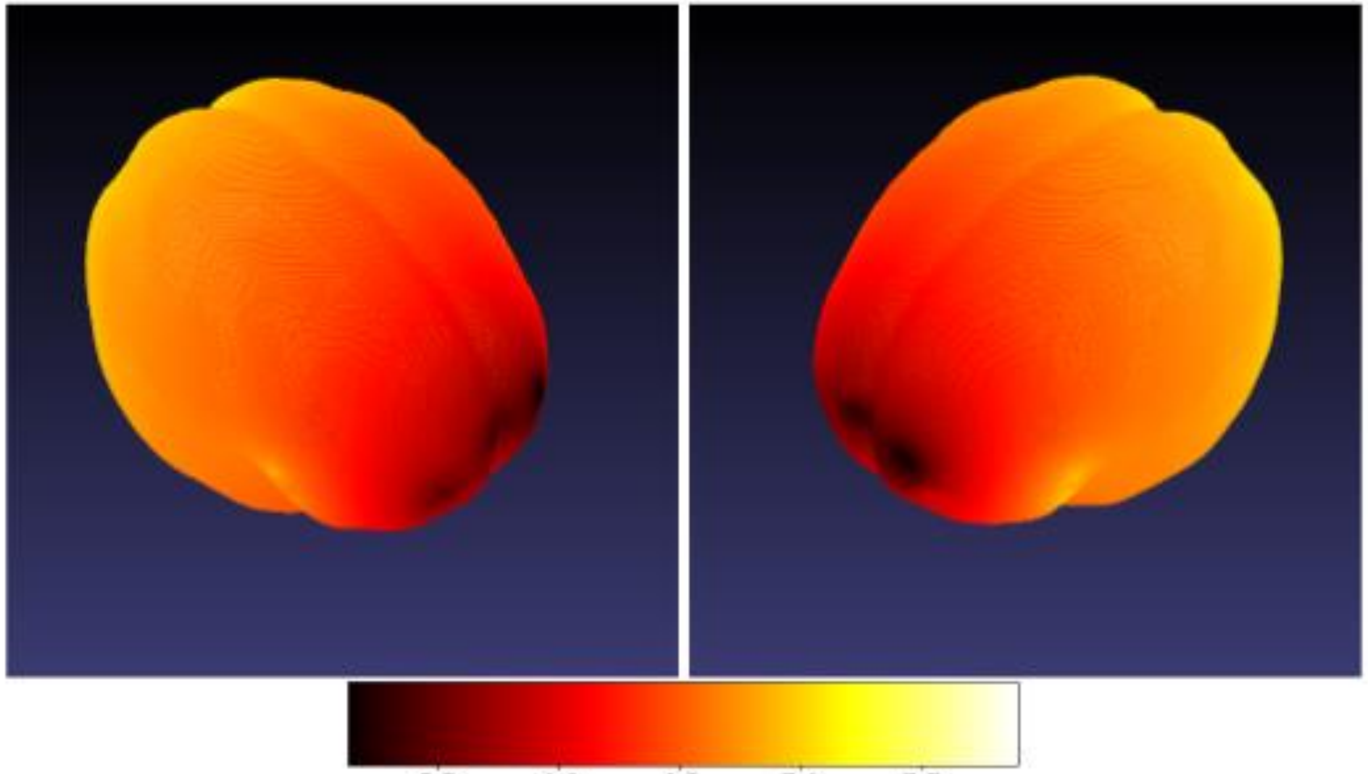
Table 1: Comparison of photogrammetry based approaches

	Hu et al. [9]	Jaffe-Dax et al. [11]	**Ours**
Intra-method error (adults, $AVG \pm STD$)	Not Reported	$2mm \pm 0.5mm$	$3mm \pm 1.6mm$
Inter-method error (adults, $AVG \pm STD$)	$6.66mm \pm 3.30mm$	$3.40mm \pm 0.90mm$	$6.47mm \pm 1.32mm$
Acquisition Time **Processing Time**	$\sim$ 1 minute $>$ 10 minutes	$\sim$ 1 minute $\sim$ 10 minutes	$\sim$ 5 seconds $<$ 1 seconds
Automation of method	semi-automatic	fully-automatic	fully-automatic
Camera specifications	8MP resolution per image	2.1MP resolution per image, 240 fps	2.1MP resolution per image 30 fps.
Environmental constraints	mono-colored walls or cloth/curtains covering the background.	None	None
Sensitivity to cap appearance	High	Medium	Low

Spatial Errors

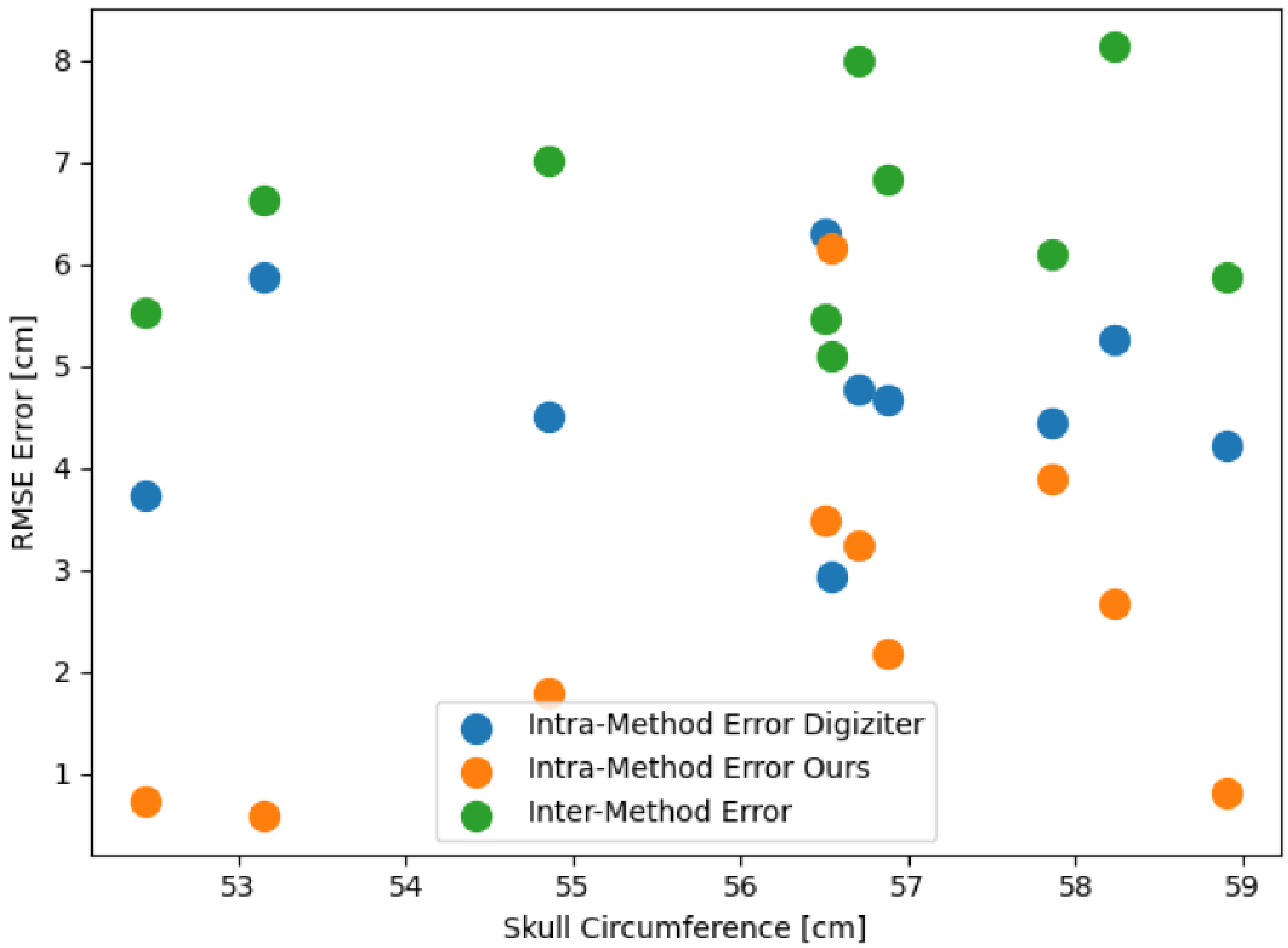
- Spatial accuracy deteriorates in the back side

  Add more stickers or spread them to balance spatial errors

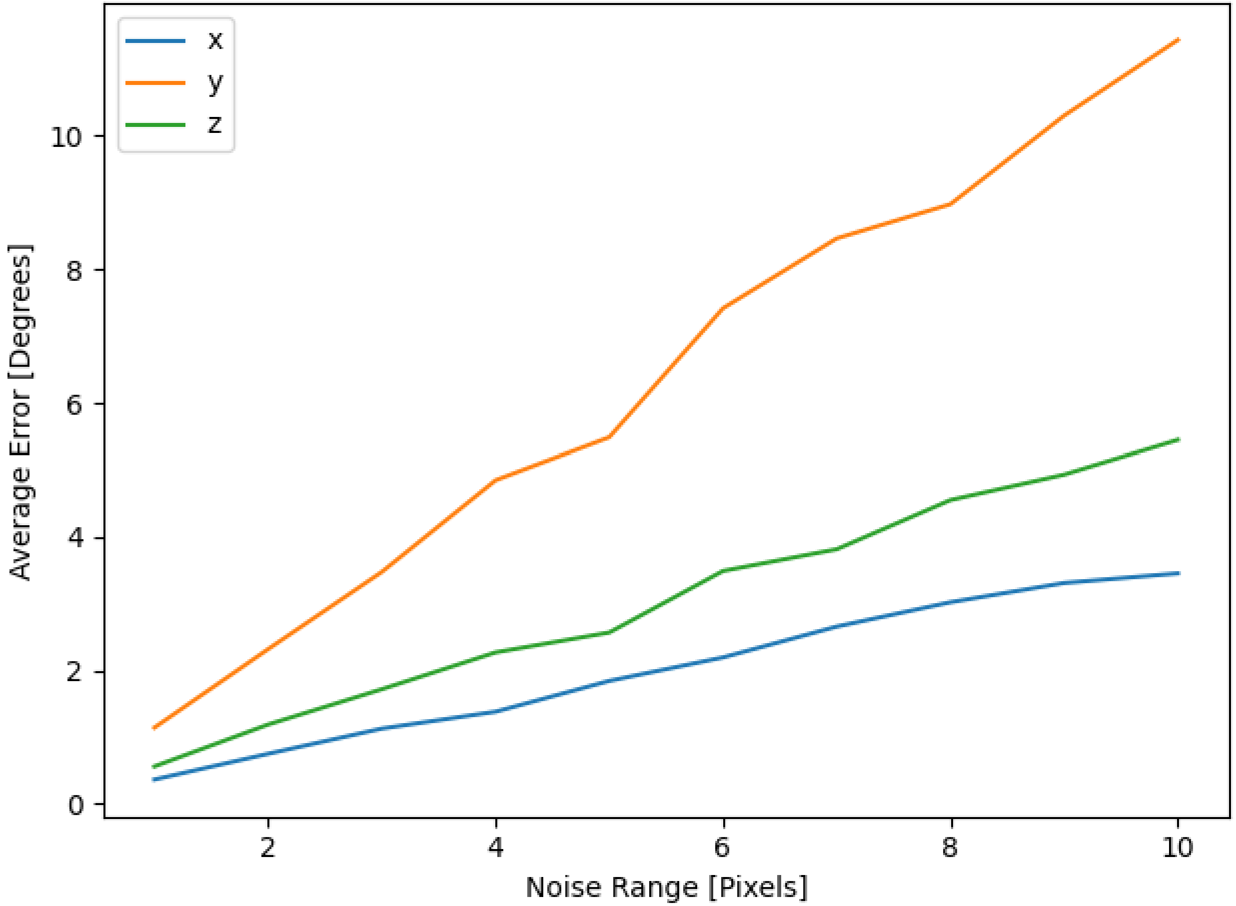


(a) Mean Inter-Method Validation Error [mm]

Skull size effect on RMSE



Robustness to Annotation Mistakes



Discussion

- Speed is crucial when dealing with developmental & clinical populations

- Subject specific anatomical positions are estimated using the puppet head
  - Found to be negligible relative to other sources of error, skull sizes did not affect quality of registration

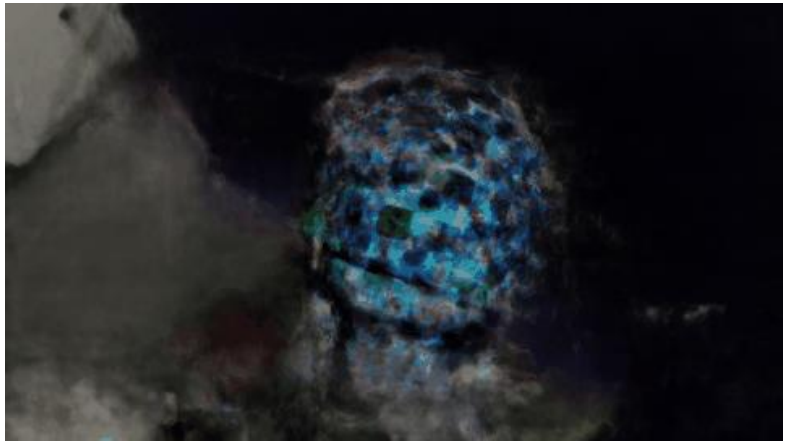
- Skipping reconstruction
  - A great way to obtain speed, on the expense of accuracy
  - **<u>Use this solution until variance is reduced from other sources</u>**

    Error is bound from above by distance between source and detector

Future Work

- Reducing variance by using real data (Let’s talk !)
  - Synthetic data has its limits

- Speed up reconstruction using more modern tools



Q & A
Thanks for Listening !

- Preprint
  - https://doi.org/10.1101/2020.12.29.424683
- For academic usage
  - https://github.com/yoterel/STORM-Net
- Contact
  - erelyotam@gmail.com

Live Demo

- Model Visualizer
  - Original coordinate system
  - Difference between anchors and sensors
  - Statistical MNI coordinate system
- Offline Step
  - Produces a network model file
  - Can take 10 hours (on top of acquisition time)
- Online Step
  - Capture video using camera path, keep head centered
  - Trim video manually if needed
  - Auto annotate, and complete with manual annotation
  - Co-register
    - Possibly transform to statistical MNI coordinate system