A LOW-COST MOTION CAPTURE SYSTEM FOR BODY TRACKING USING SYNCHRONIZED AZURE KINECT SYSTEMS

PRESENTERS: JONGOH JEONG, YUE WANG
ADVISOR: MILI SHAH

THE COOPER UNION

SIGGRAPH 2020 19-23 JULY WASHINGTON DC THINK BEYOND

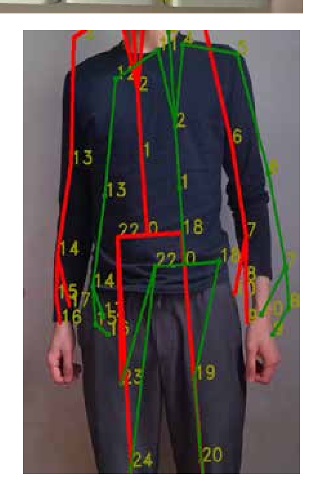
PROBLEM

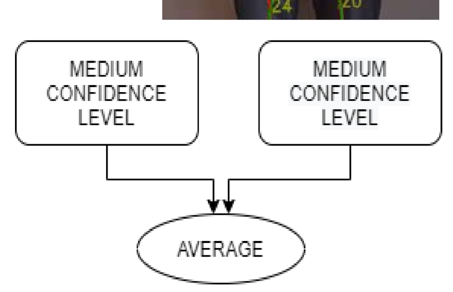
- Setting up a motion capture environment for human pose analysis is reliable, yet costly and typically restricted to a fixed lab space. On the other hand, a single camera vision-based system is less expensive and maybe prone to erros due to occlusion and illumination changes, but is highly portable.
- We extend the Azure Kinect DK's capability to track body joints by synchronizing multiple devices to mitigate the effects of occlusion and illumination at a much lower cost and with higher portability.

METHOD

We leverage the body data from each Kinect to output a more precise weighted average.

1. Configure devices in daisy-chain method (RS-232)
2. Track joints from each device
3. Compute average of the joint positions from all devices
4. Transform positions from subordinate mode devices onto master device RGB camera space
5. Project 3-D homogeneous points onto 2-D display window





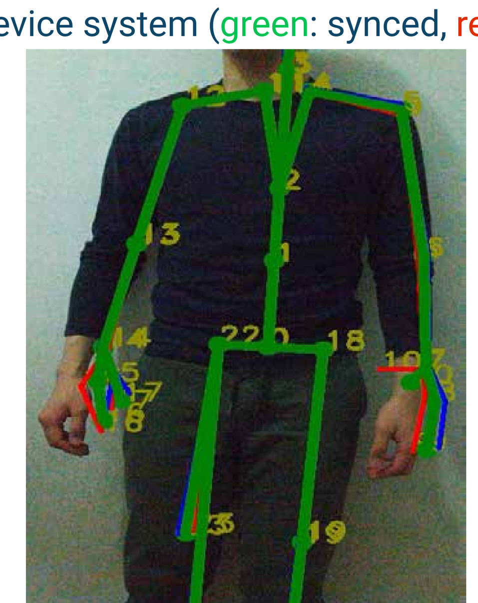
RESULTS

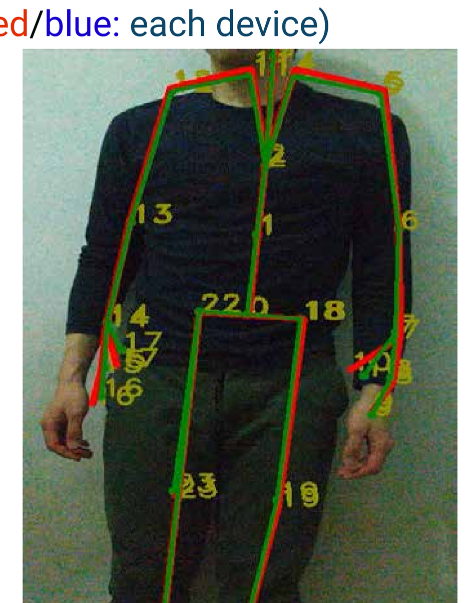
Synced joint estimates fall in between the estimated positions from each device.

Caveats

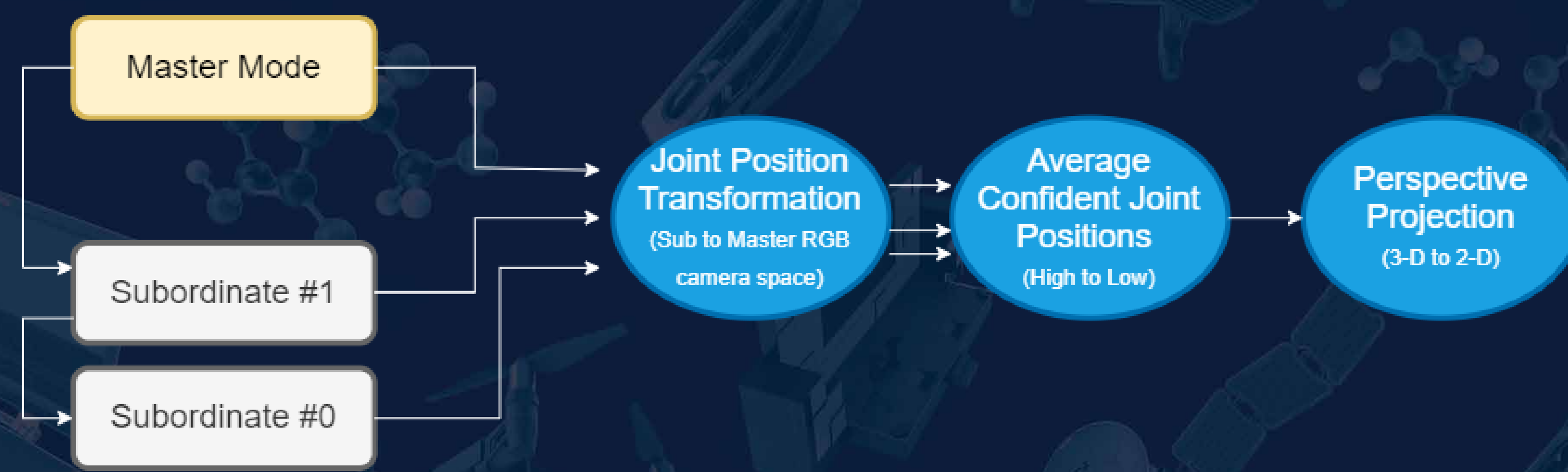
The estimates varies largely by environment and posture. There is a slight offset from actual present due to parallax.

2-device system (green: synced, red/blue: each device)

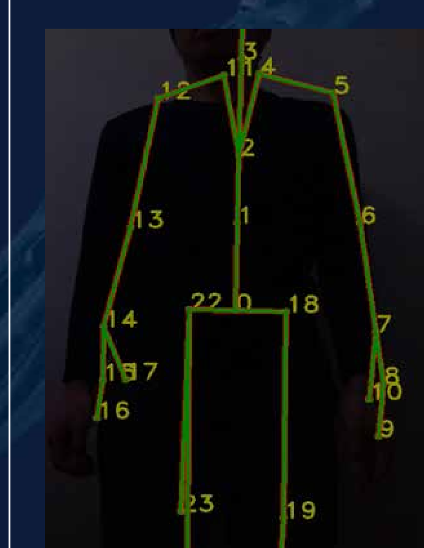


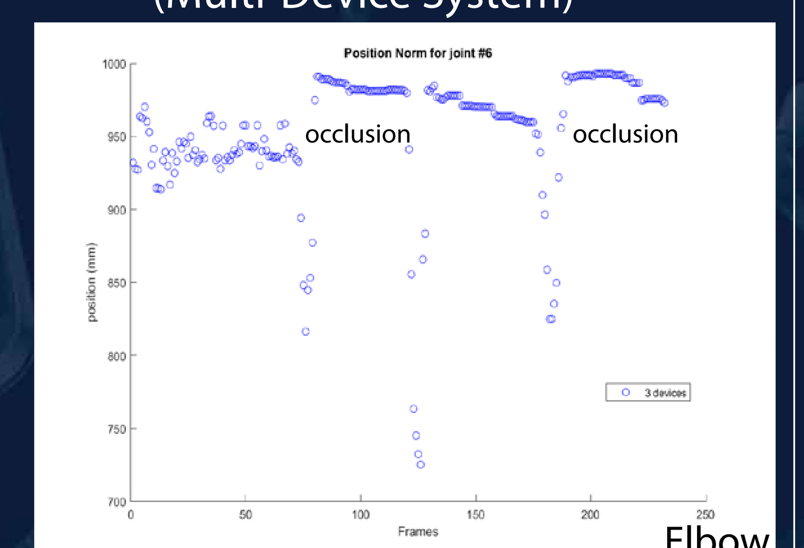


3-device system (green: synced, red: single)



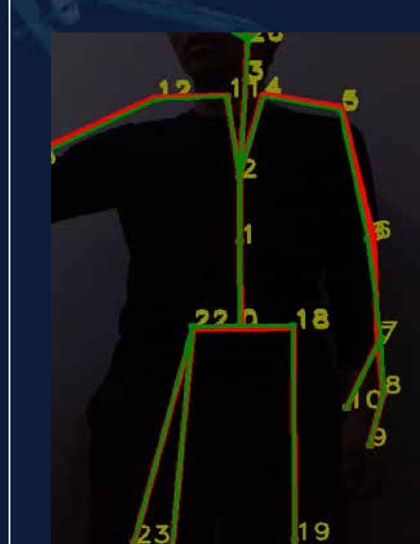
Consistent Positions in Occlusion (Multi-Device System)

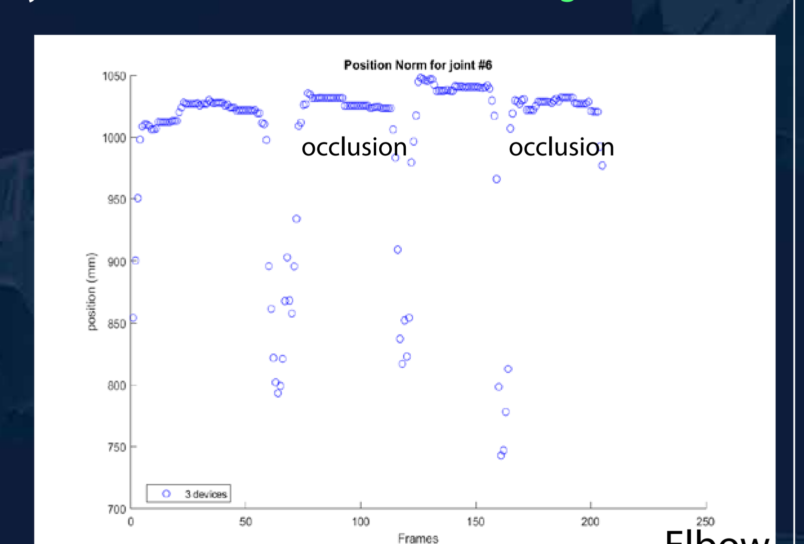




Sub #0 in Occlusion

We observed consistent position norm estimates in both open and occluded (at sub device #0) cases with a small error (a few cm). In this test, it relied mostly on the master device estimates (green).

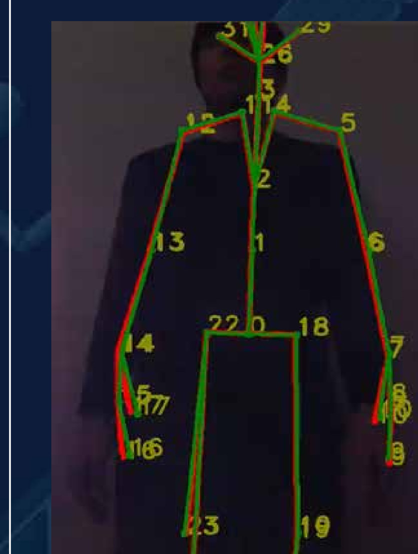


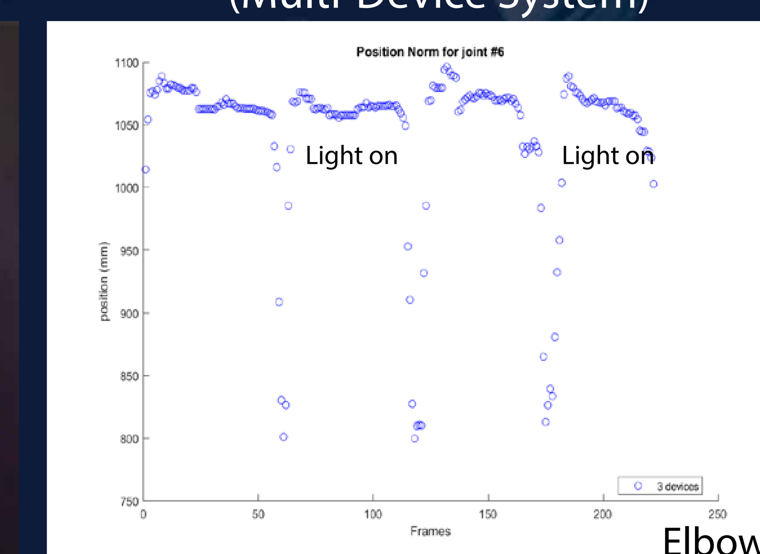


Sub #1 in Occlusion

We observed consistent position norm estimates in both open and occluded (also at sub device #1) cases with a small error (a few cm). In this case, it output weighted average (green) of master and sub device #0, deviating from a single device estimate slightly (red).

Consistent Positions in Illumination (Multi-Device System)

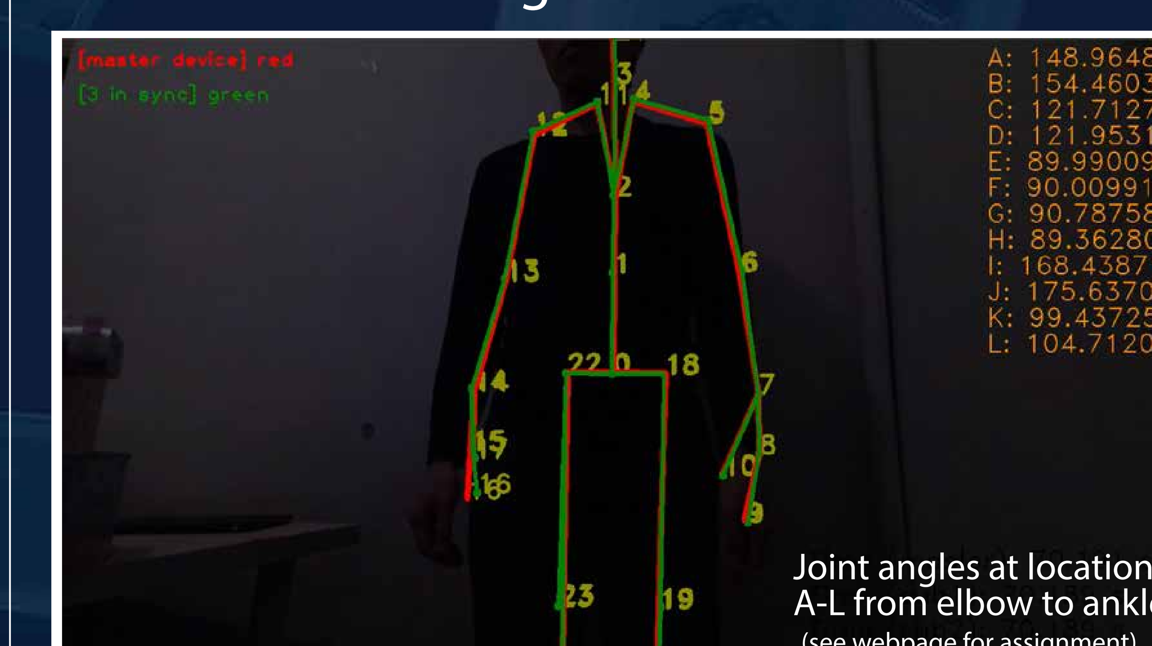




Additional Light Directed Toward Master

We observed consistent position norm estimates in both ambient light-only and additional light cases with a very small error (a couple cm). In this experiment, there was not much illumination distortion perhaps due to the strength of the light source.

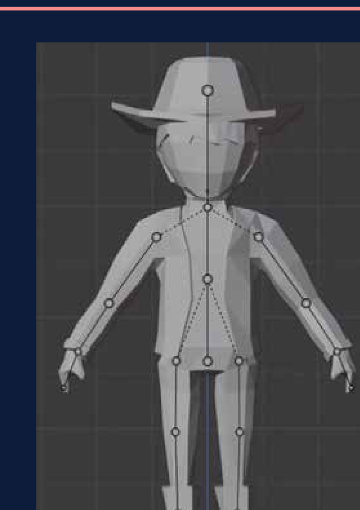
Joint Angle Estimation



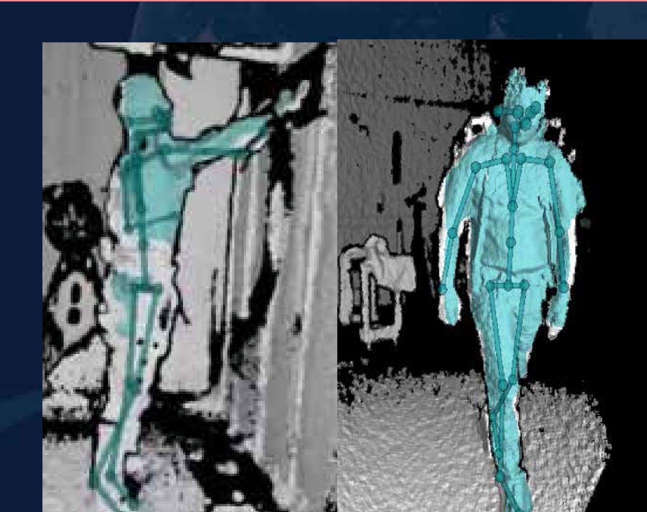
Joint angles at locations A-L from elbow to ankle (see webpage for assignment)

Future Applications

Human Body-Controlled Drone Tracking



Graphical Figure



Gait Analysis (Exoskeletons)

APPROACH

- Our approach solves the problem of optical occlusion and illumination effects in body joint tracking. We calibrate each pair of Kinects using the extrinsics computed by least-squares fitting of the two 3-D joint positions [1].
- The final estimated joint positions is selected by the confidence of each Kinect, from high to low. If there are multiple devices with the same level of confidence, the average is output. Then these 3-D points are projected onto display by perspective projection.

RELATED WORK

- Islam *et al.* uses joint angle estimates from a single Kinect (v2) to recognize a certain Yoga posture. They base their error on a reference model developed from joint data of gymnastics. We differ in that we utilize multiple devices to minimize angular distances of joints and increase accuracy in case of an occlusion.
- Napoli *et al.* compares joint angles calculated from Kinect (v2) positions and orientations with a professional motion capture system (Qualisys) at various postures and planes. Similarly, we extend this to a multi-device system for higher precision.

FUTURE DIRECTIONS

- Code optimization in time delay (currently ~1 sec).
- Validation with other MOCAP systems in market.
- This generalized low-cost motion capture system can be applied to the exoskeleton gait analysis study at NIST to provide more precise measurements of joints.
- Human body-controlled graphical media art and drone control applications using unobstructed, more precise joint data from our synchronized system.

REFERENCES

[1] Arun, K.S. & Huang, T.S. & Blostein, Steven. (1987). Least-squares fitting of two 3-D point sets. IEEE T Pattern Anal. Pattern Analysis and Machine Intelligence, IEEE Transactions on. PAMI-9. 698 - 700.

[2] M. U. Islam, H. Mahmud, F. B. Ashraf, I. Hossain and M. K. Hasan, "Yoga posture recognition by detecting human joint points in real time using microsoft kinect," 2017 IEEE Region 10 Humanitarian Technology. Conference

[3] Napoli, Alessandro, et al. "Performance Analysis of a Generalized Motion Capture System Using Microsoft Kinect 2.0." Biomedical Signal Processing and Control, vol. 38, 2017, pp. 265–280.

ADDITIONAL RESOURCES AVAILABLE AT:
https://github.com/andyj1/kinect