PitcherNet: Powering the Moneyball Evolution in Baseball Video Analytics

Jerrin Bright, Bavesh Balaji, Yuhao Chen, David Clausi and John Zelek

Vision and Image Processing Lab

Department of Systems Design Engineering

University of Waterloo, Waterloo, ON, Canada

Date: Tuesday, June 18th, 2024

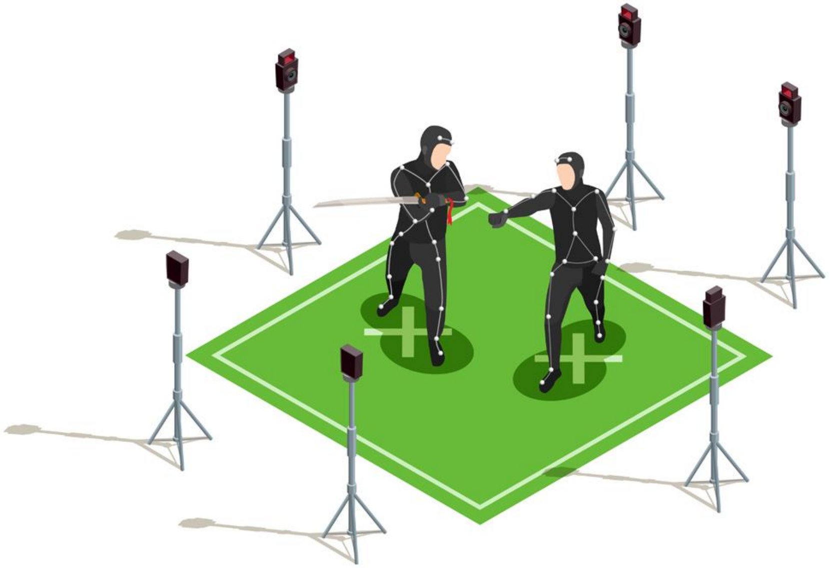
Motivation

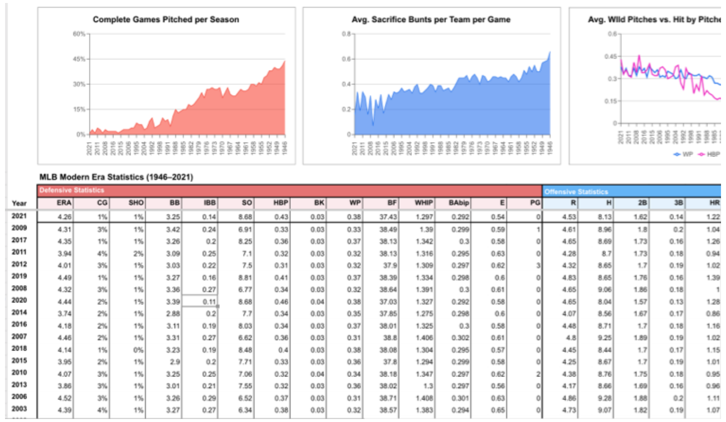
- Analysis from kinematic information.
- Performance optimization, injury prevention, quantitative analysis of the player mechanics.

Prior Research on Baseball Analysis

- Pre-recorded baseball databases (Pitch f/x).

- Controlled environments (MoCap Systems).





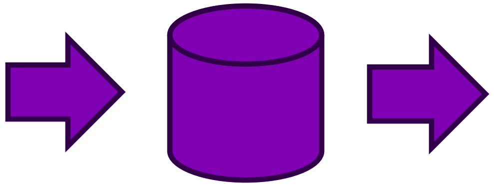
Complete Games Pitched per Season

Avg. Sacrifice Bunts per Team per Game

Avg. Wild Pitches vs. Hit by Pitche

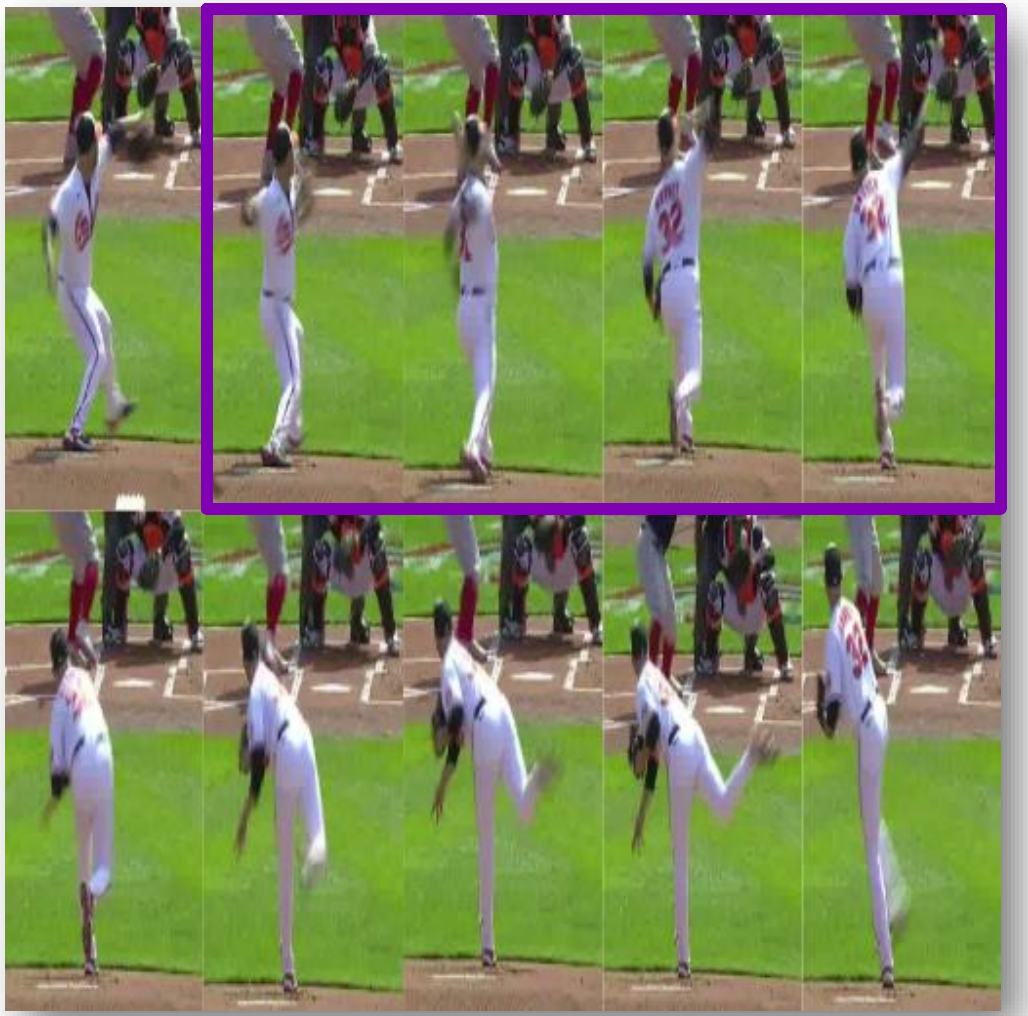
MLB Modern Era Statistics (1946–2021)

Year	ERA	CG	SHO	BB	IBB	SO	HBP	BK	WP	BF	WHIP	BAbip	E	PG	R	H	2B	3B	HR
	Defensive Statistics														Offensive Statistics				
2021	4.26	1%	1%	3.25	0.14	8.68	0.43	0.03	0.38	37.43	1.297	0.292	0.54	0	4.53	8.13	1.62	0.14	1.22
2009	4.31	3%	1%	3.42	0.24	6.91	0.33	0.03	0.33	38.49	1.39	0.299	0.59	1	4.61	8.96	1.8	0.2	1.04
2017	4.35	1%	1%	3.26	0.2	8.25	0.36	0.03	0.37	38.13	1.342	0.3	0.58	0	4.65	8.69	1.73	0.16	1.26
2011	3.94	4%	2%	3.09	0.25	7.1	0.32	0.03	0.32	38.13	1.316	0.295	0.63	0	4.28	8.7	1.73	0.18	0.94
2012	4.01	3%	1%	3.03	0.22	7.5	0.31	0.03	0.32	37.9	1.309	0.297	0.62	3	4.32	8.65	1.7	0.19	1.02
2019	4.49	1%	1%	3.27	0.16	8.81	0.41	0.03	0.37	38.39	1.334	0.298	0.6	0	4.83	8.65	1.76	0.16	1.39
2008	4.32	3%	1%	3.36	0.27	6.77	0.34	0.03	0.32	38.64	1.391	0.3	0.61	0	4.65	9.06	1.86	0.18	1
2020	4.44	2%	1%	3.39	0.11	8.68	0.46	0.04	0.38	37.03	1.327	0.292	0.58	0	4.65	8.04	1.57	0.13	1.28
2014	3.74	2%	1%	2.88	0.2	7.7	0.34	0.03	0.35	37.85	1.275	0.298	0.6	0	4.07	8.56	1.67	0.17	0.86
2016	4.18	2%	1%	3.11	0.19	8.03	0.34	0.03	0.37	38.01	1.325	0.3	0.58	0	4.48	8.71	1.7	0.18	1.16
2007	4.46	2%	1%	3.31	0.27	6.62	0.36	0.03	0.31	38.8	1.406	0.302	0.61	0	4.8	9.25	1.89	0.19	1.02
2018	4.14	1%	0%	3.23	0.19	8.48	0.4	0.03	0.38	38.08	1.304	0.295	0.57	0	4.45	8.44	1.7	0.17	1.15
2015	3.95	2%	1%	2.9	0.2	7.71	0.33	0.03	0.36	37.8	1.294	0.299	0.58	0	4.25	8.67	1.7	0.19	1.01
2010	4.07	3%	1%	3.25	0.25	7.06	0.32	0.04	0.34	38.18	1.347	0.297	0.62	2	4.38	8.76	1.75	0.18	0.95
2013	3.86	3%	1%	3.01	0.21	7.55	0.32	0.03	0.36	38.02	1.3	0.297	0.56	0	4.17	8.66	1.69	0.16	0.96
2006	4.52	3%	1%	3.26	0.29	6.52	0.37	0.03	0.31	38.71	1.408	0.301	0.63	0	4.86	9.28	1.88	0.2	1.11
2003	4.39	4%	1%	3.27	0.27	6.34	0.38	0.03	0.32	38.57	1.383	0.294	0.65	0	4.73	9.07	1.82	0.19	1.07



Challenges with Video Inputs

Motion Blur



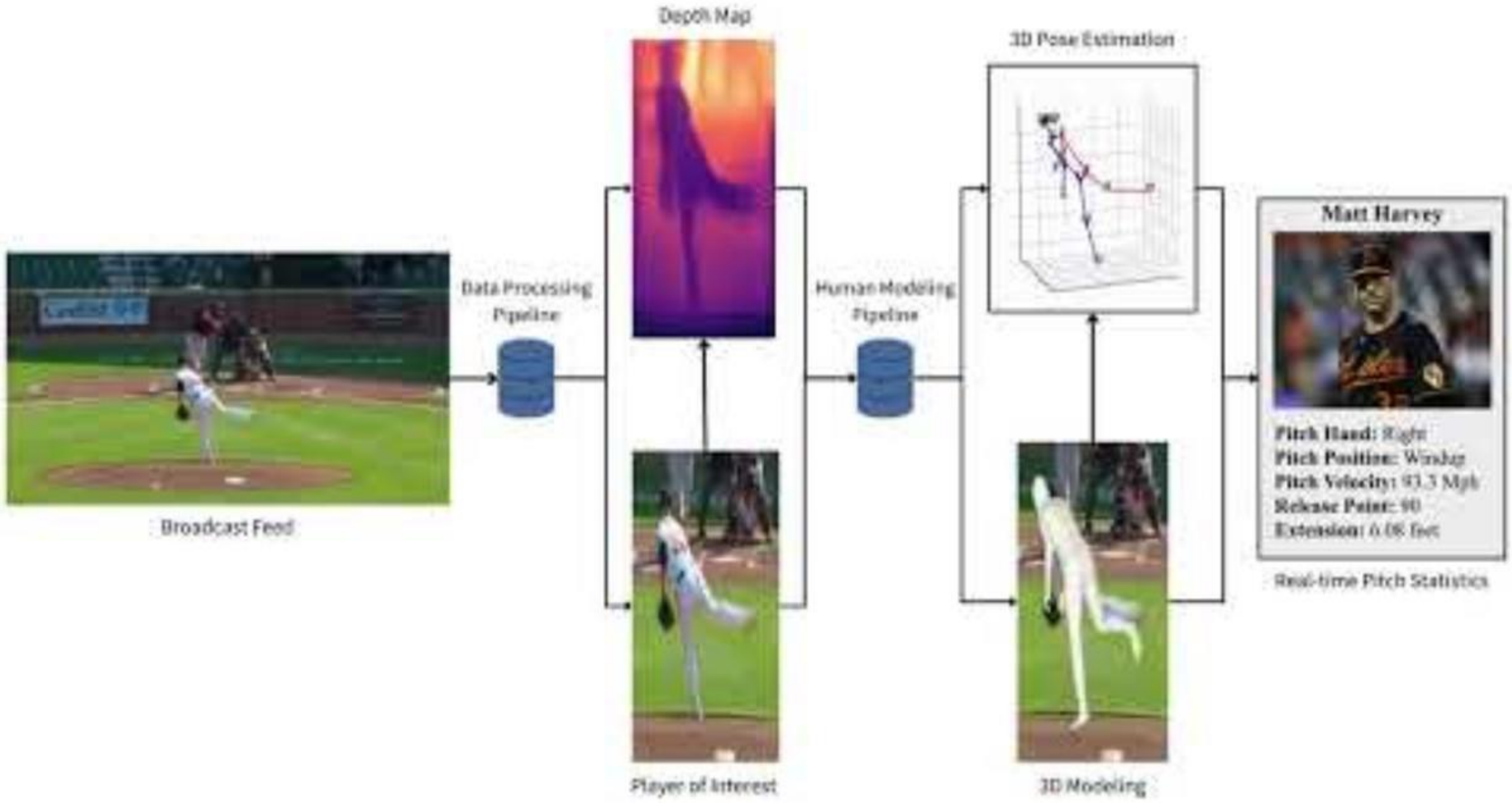
Self-Occlusion

Out-of-distribution

Objective

"Enable detailed analysis of pitcher dynamics from human models in 3D extracted solely from monocular broadcast feeds"

High-level Workflow of PitcherNet

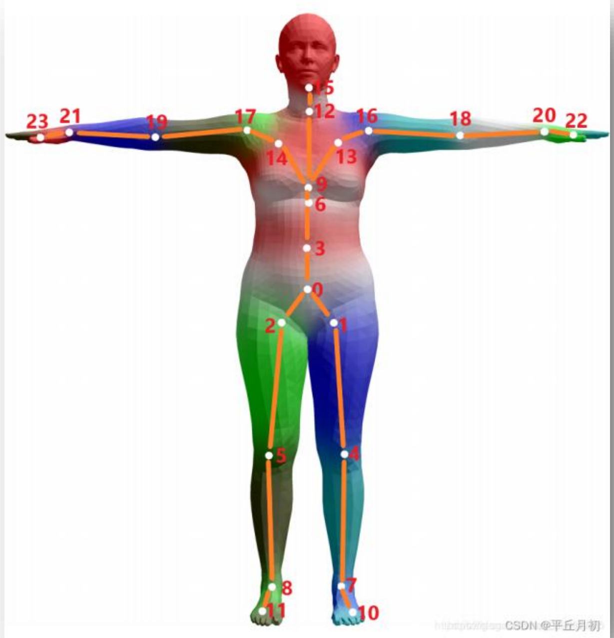


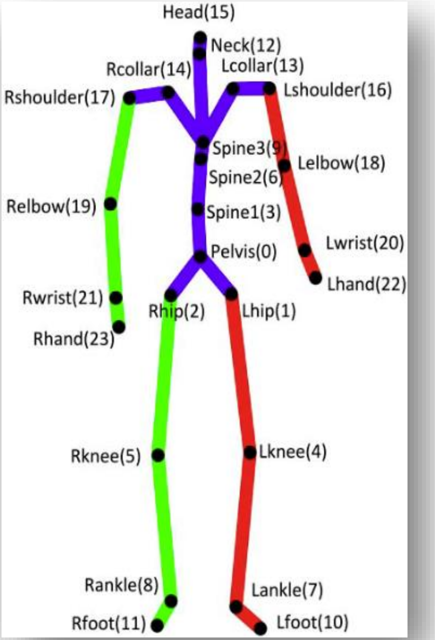
https://youtu.be/wKBOtDPPxws

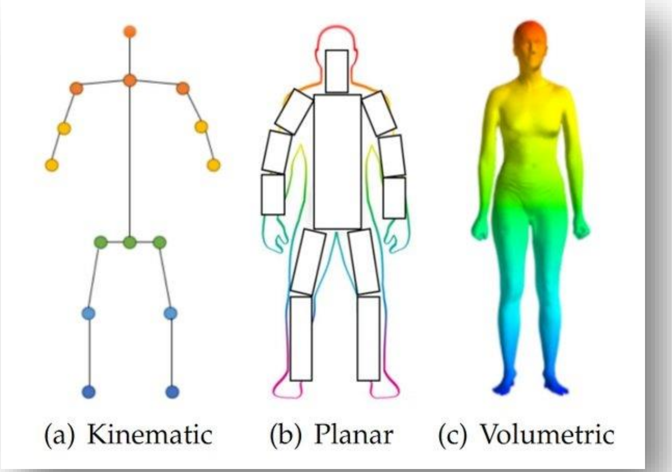
Background

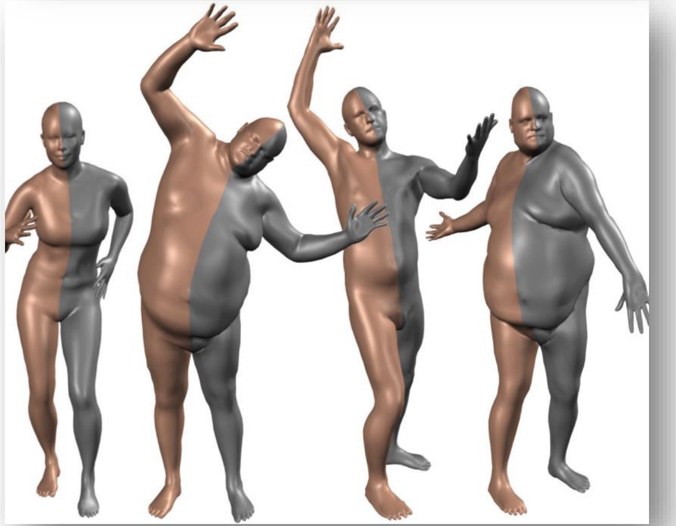
3D Human Modeling - SMPL

- Skinned **M**ulti-**P**erson **L**inear model[1].
- 72 joint and 10 shape parameters -> 6890 vertices.



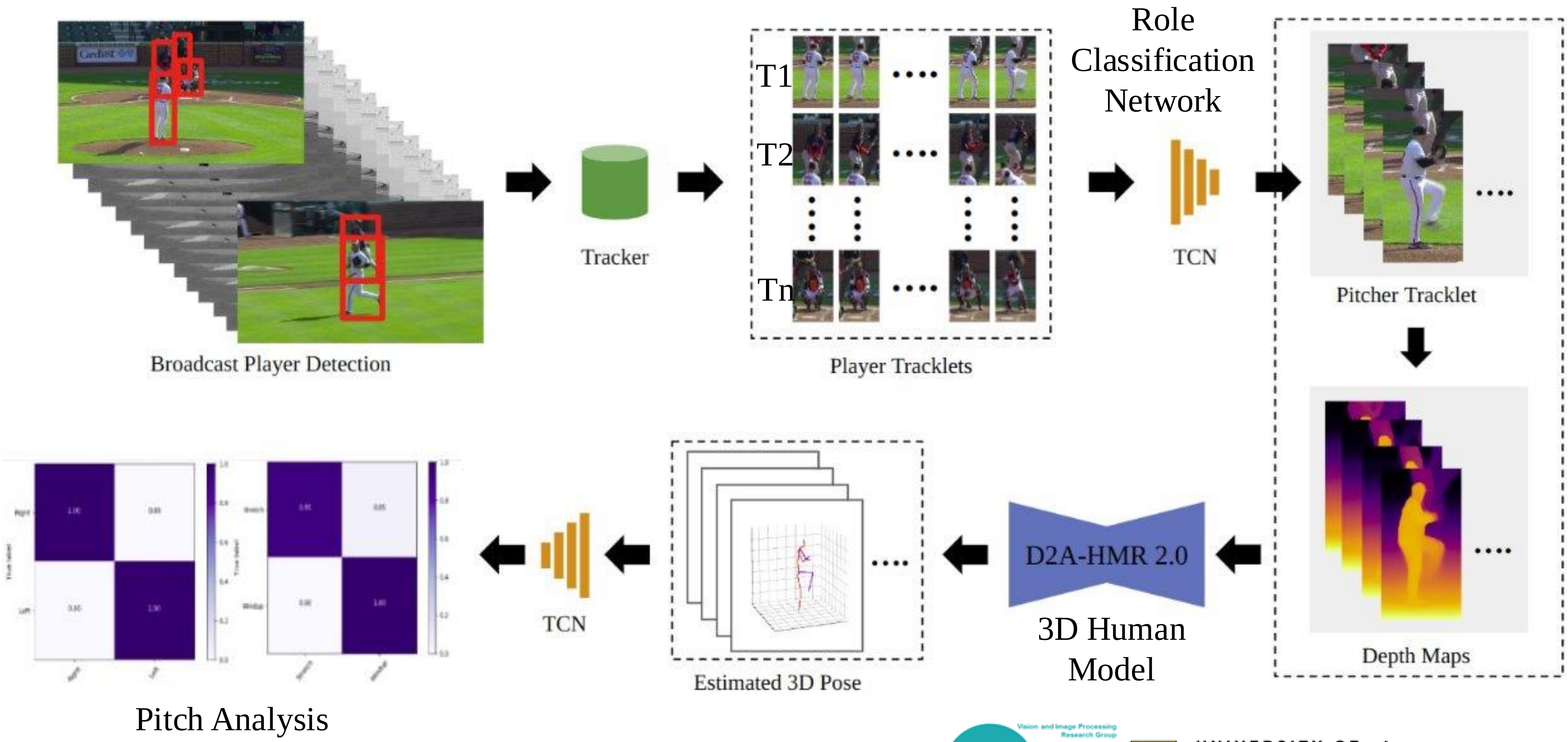






Credits:
[1] Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael J. Black. SMPL: a skinned multi-person linear model. ACM Transactions on Graphics, 2015.

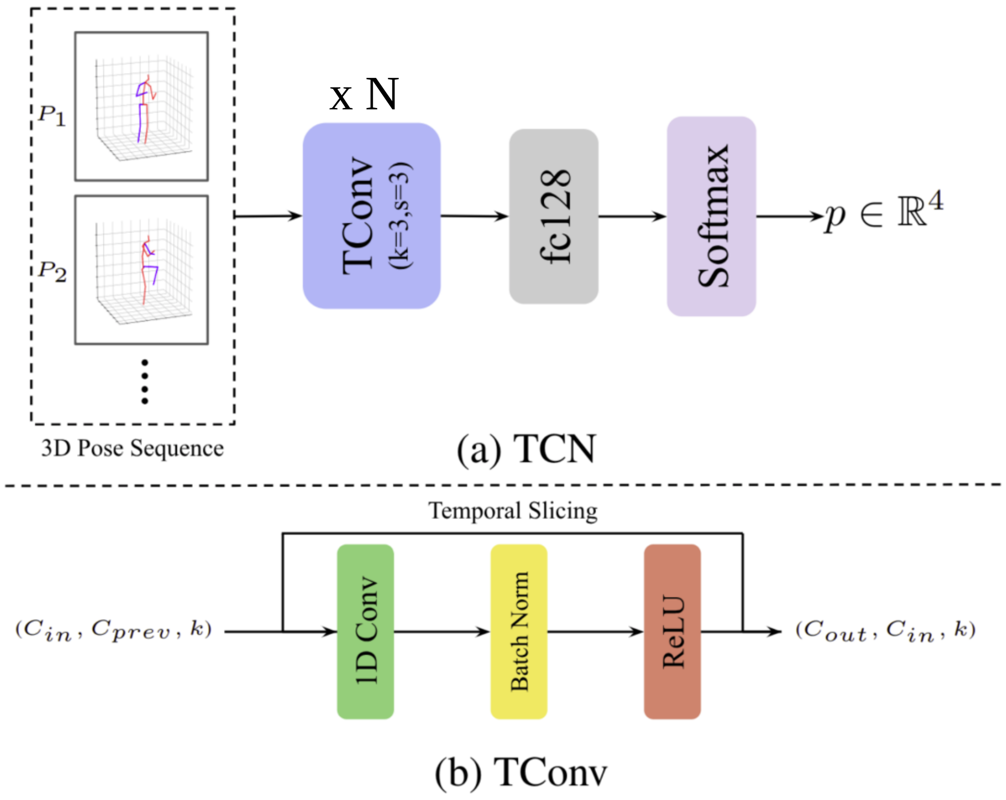
PitcherNet System



Role Classification

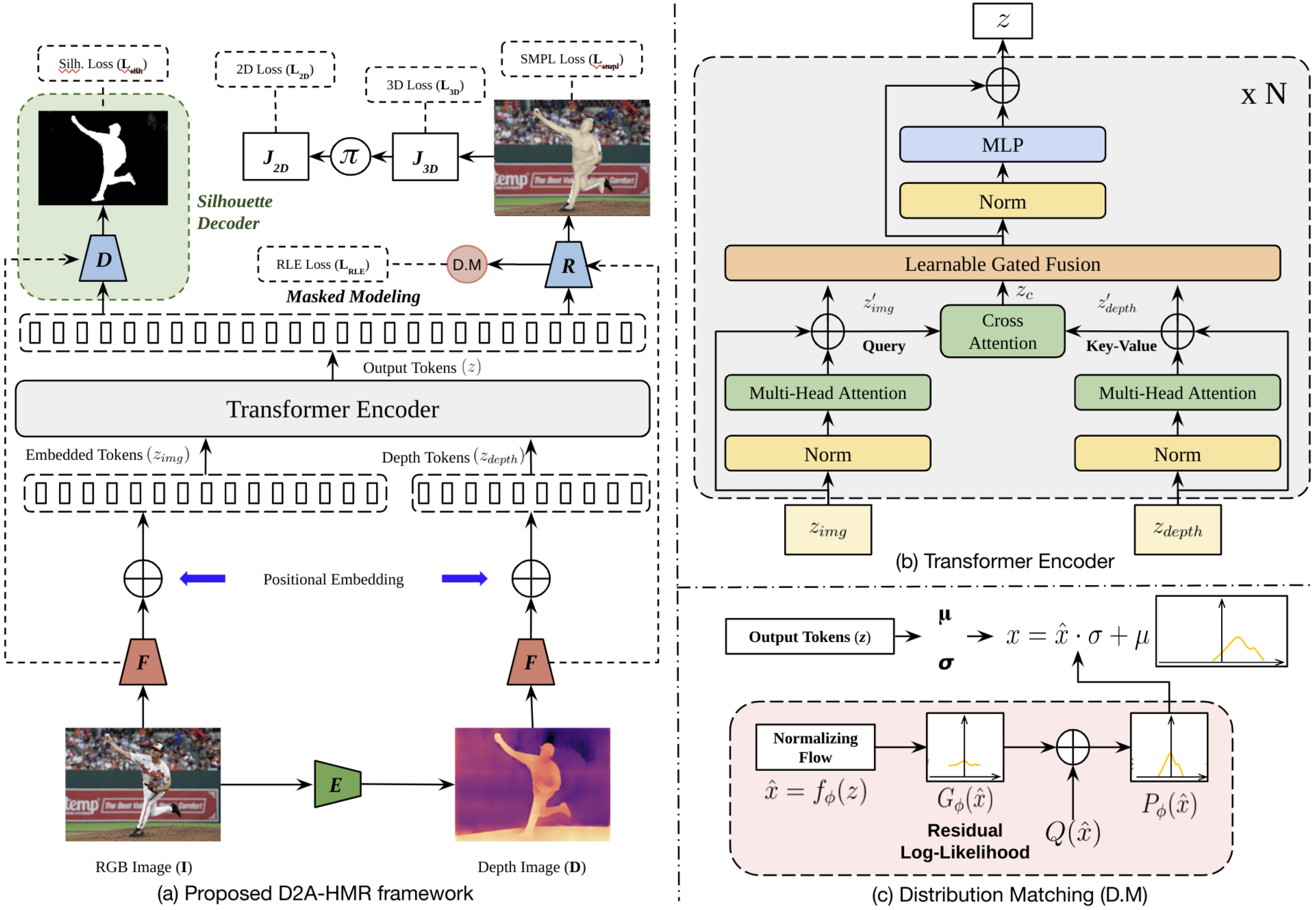
- Decouples action from player kinematics.
  - Input: Pseudo-pose from estimated tracklets.
  - Output: Player role.

- Invariant to viewpoint/ facial features/ player jersey numbers.



3D Human Model

- Distribution and depth aware 3D modeling [2].
- Motion blur and in-the-wild data augmentation.
- Generalizable, reliable 3D human models.

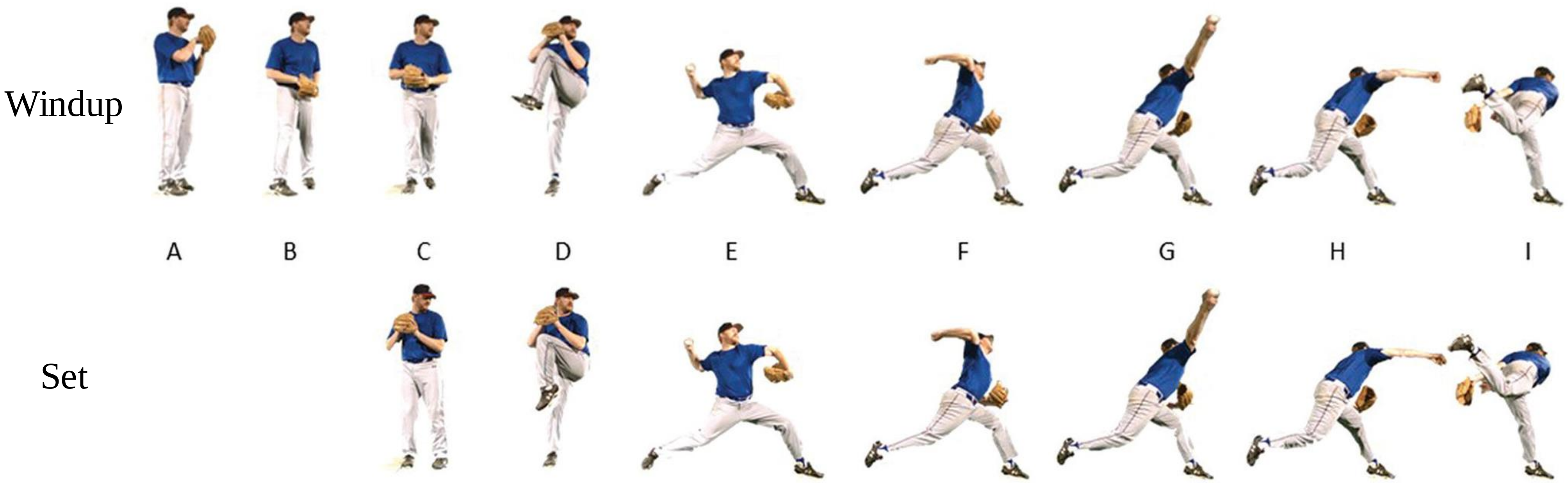


[2] **Jerrin Bright**, Bavesh Balaji, Harish Prakash, Yuhao Chen, David A Clausi, and John Zelek. 2024. Distribution and Depth-Aware Transformers for 3D Human Mesh Recovery. In 21st Conference on Robots and Vision - ORAL

Pitch Analysis

- Pitch Position

$$PP(windup, set) = \sigma(TCN(X))$$

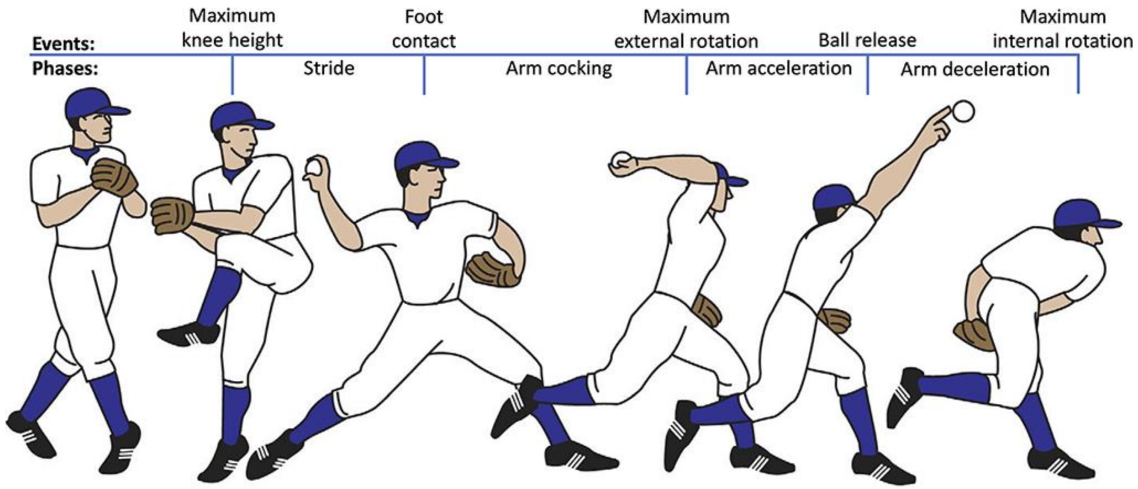


Pitch Analysis

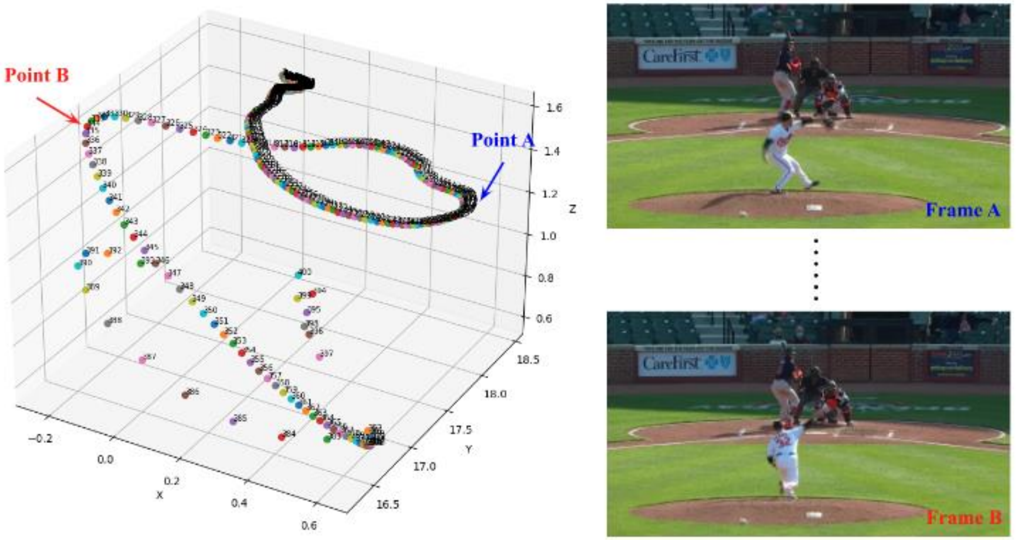
- Release Point

$$P_{rel} = argmax(v(i)|i \in [P_b - n/2, P_b + n/2])$$

Point A- Arm Cocking
Point B- Arm Deceleration



6 phases of pitching action



Trajectory of the right wrist joint in 3D space

Pitch Analysis

- Pitch Velocity

$$v_p = \omega \times l = \{(atan(w_y^r, w_x^r) - atan(w_y^{r-1}, w_x^{r-1})) \times T\} \times l$$

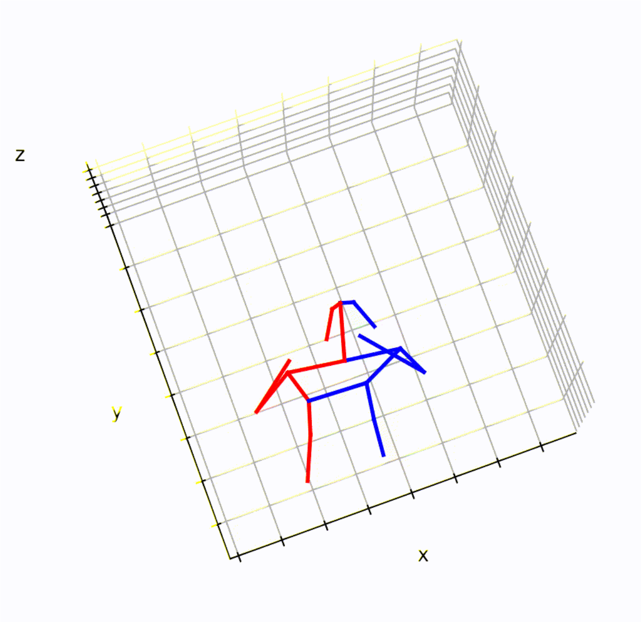
- Release Extension

$$E_{rel} = \sqrt{(w_x - a_x)^2 + (w_y - a_y)^2 + (w_z - a_z)^2}$$

MLBPitchDB Dataset

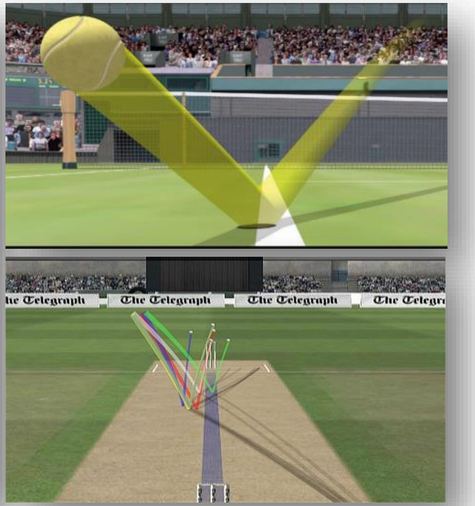
What we have?

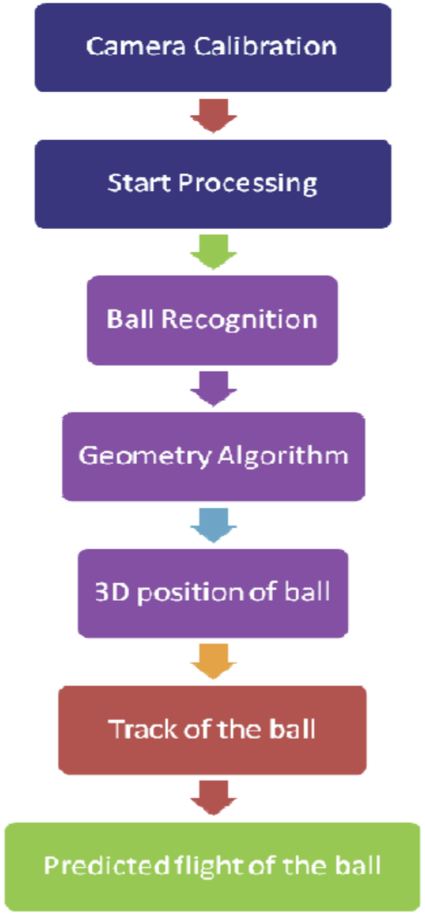
- 1000+ games
- 3D Hawk-Eye pose data
- Various pitch metrics



What is Hawk-Eye Camera System?

- Triangulation with many cameras around the playing area
- Applications include pose estimation, tracking, etc.





Quantitative Results of Role Classification

Table I. Comparison with baselines

	Test Accuracy ↑
LSTM	85.55
Transformer	91.11
Ours	**96.66**

Table II. Comparison with jersey identification techniques

	Test Accuracy ↑
Gerke *et al.* [21]	64.47
Li *et al.* [30]	88.29
Vats *et al.* [48]	89.46
Balaji *et al.* [2]	93.68
Balaji *et al.* [3]	94.70
Ours	**96.82**

Quantitative Results of 3D Human Modeling

Table III. Comparison of D2AHMR 3D model

Method	Human3.6M		3DPW	
	mPJPE	PA-mPJPE	mPJPE	PA-mPJPE
HMMR'19	-	58.1	116.5	72.6
TCMR'21	62.3	41.1	95.0	55.8
VIBE'20	65.6	41.4	93.5	56.5
SPIN'21	62.5	41.1	96.9	59.2
PyMAF'21	57.7	40.5	92.8	58.9
ROMP'21	-	-	105.6	53.5
HMREFT'20	63.2	43.8	85.1	52.2
PARE'21	76.8	50.6	82.0	50.9
ProHMR'21	-	41.2	95.1	59.5
P2M'20	64.9	47.0	89.2	58.9
METRO'21	54.0	36.7	77.1	47.9
Ours	53.2	35.9	78.7	46.9

Quantitative Results of Pitch Analysis

Table IV. Performance of our pitch statistics modules

(a) Handedness

	Accuracy ↑	F1 Score ↑	Precision ↑
LSTM	85.0	85.7	90.0
Ours (TCN)	**100.0**	**100.0**	**100.0**

(b) Pitch Position

	Accuracy ↑	F1 Score ↑	Precision ↑
LSTM	81.3	82.5	85.0
Ours (TCN)	**97.5**	**97.4**	**95.0**

(c) Release Point

	A_1 ↑	A_2 ↑	A_5 ↑
LSTM	31.3	46.4	63.5
TCN	43.4	51.5	77.6
Ours	**80.8**	**85.8**	**97.9**

(d) Pitch Velocity

	$A_{1\%}$ ↑	$A_{2\%}$ ↑	$A_{5\%}$ ↑
LSTM	5.1	13.1	22.2
TCN	10.1	18.1	48.4
Ours	**43.4**	**68.6**	**94.9**

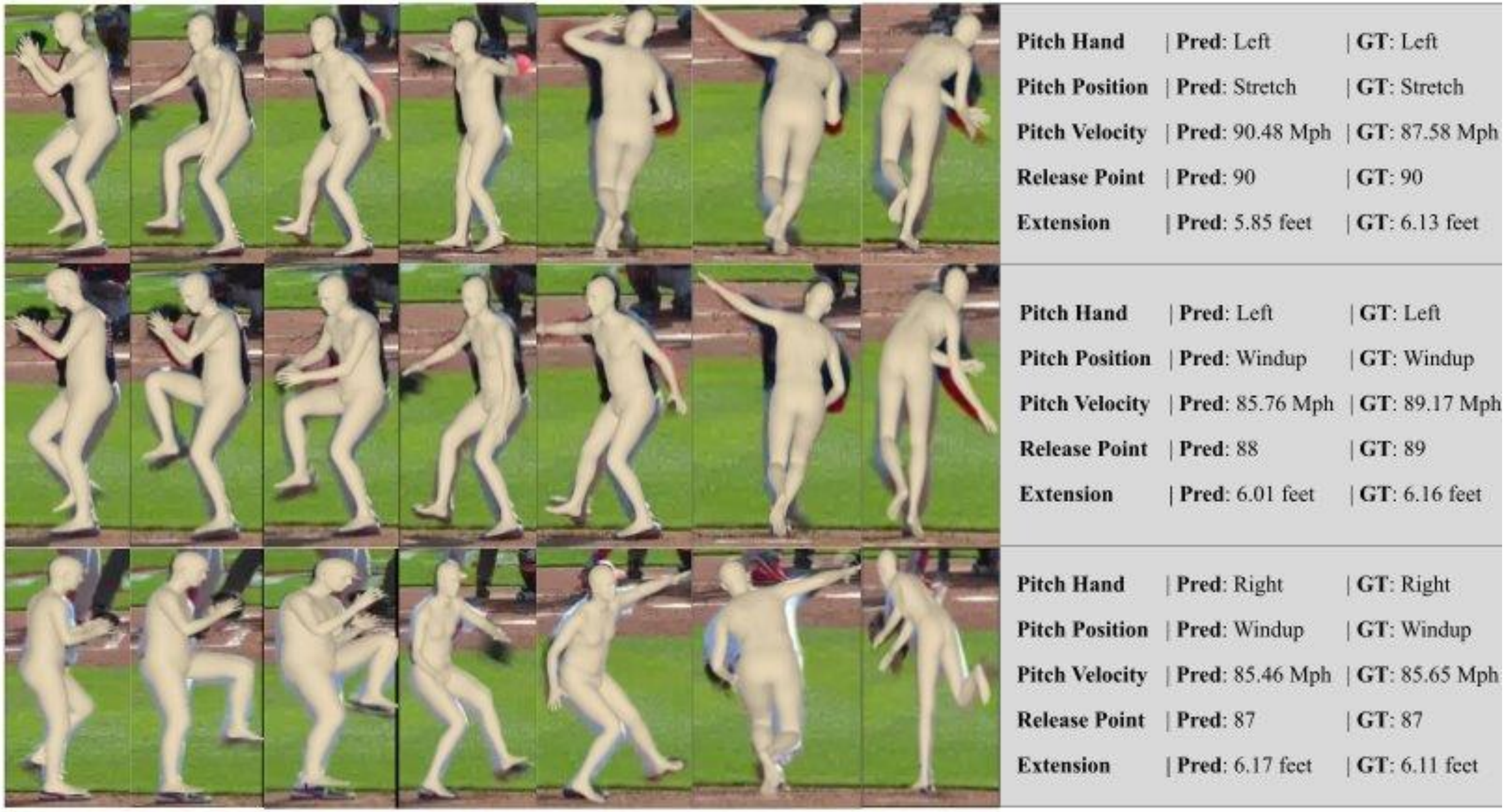
(e) Release Extension

	$A_{5\%}$ ↑	$A_{8\%}$ ↑	$A_{10\%}$ ↑
LSTM	4.0	7.1	11.1
TCN	14.1	19.1	25.2
Ours	**24.2**	**31.3**	**37.3**

Qualitative Results (3D Human Model)

https://www.youtube.com/watch?v=TsA6bOcaaiU

Qualitative Results (Pitch Analysis)



VIPlab
UNIVERSITY OF WATERLOO | FACULTY OF ENGINEERING

Summary

- Reliable pitch analysis driven by player kinematics and human model priors.

- Role classification aiming to classify players by decoupling actions.

- D2A-HMR v2 which improves 3D human modeling in degraded image quality.

Thank you!

Supported by: