IEEE 15th Annual International Conference on Wearable and Implantable Body Sensor Networks BSN'18

Markerless Gait Analysis Based on a Single RGB Camera

Xiao Gu

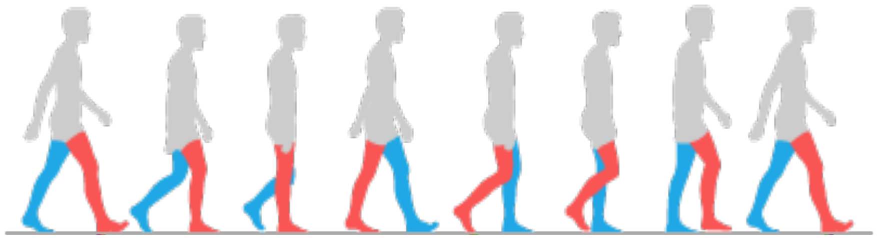
gux14@fudan.edu.cn

Xiao Gu, Fani Deligianni, Benny Lo, Wei Chen and Guang-Zhong Yang

The Hamlyn Centre
The Institute of Global Health Innovation

CIME
智慧医疗电子中心
CENTER FOR INTELLIGENT MEDICAL ELECTRONICS

Gait Analysis

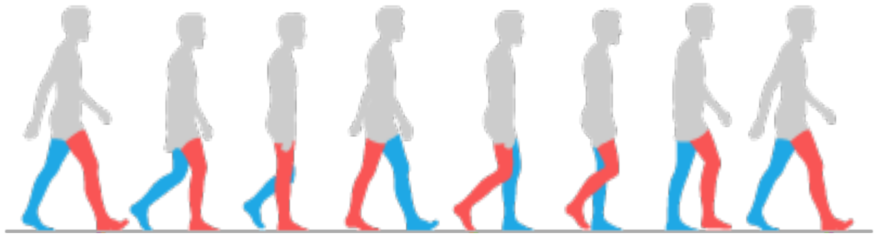


A Gait Cycle

- **Abnormal walking conditions:**
  - Limping,
  - Supination,
  - Pronation

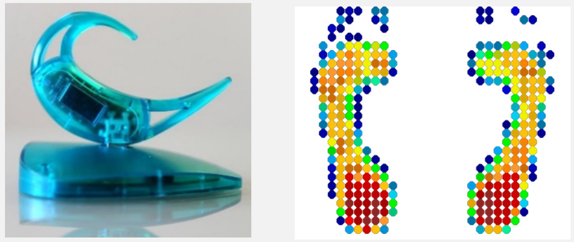
- **Important Gait Parameters:**
  - Foot Progression Angle,
  - Ankle angle,
  - Inversion/Eversion,
  - Dorsiflexion/Plantarflexion

Gait Analysis



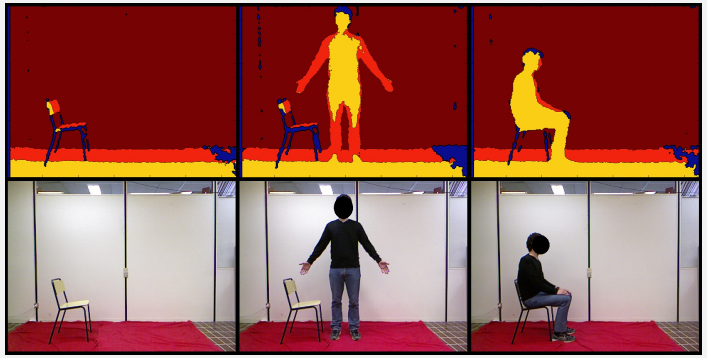
A Gait Cycle

- **Abnormal walking conditions:**
  - Limping,
  - Supination,
  - Pronation
- **Important Gait Parameters:**
  - Foot Progression Angle,
  - Ankle angle,
  - Inversion/Eversion,
  - Dorsiflexion/Plantarflexion



e-AR sensor [1, 2] Foot pressure insole [2]

Wearable Sensor



Depth image [3]

Depth Camera

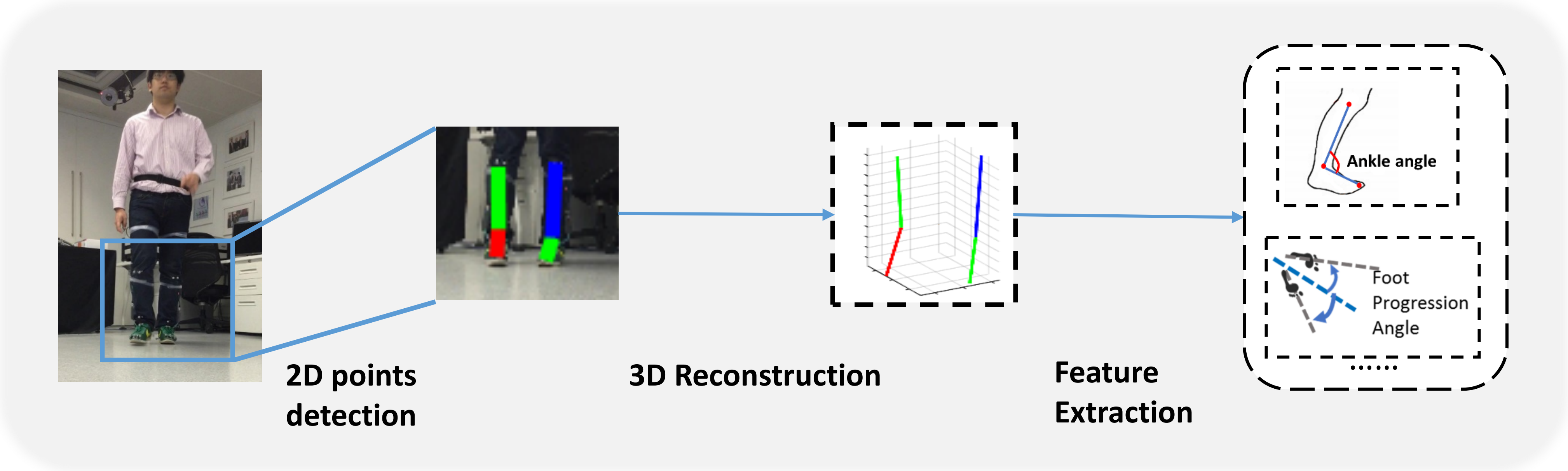
Laboratory setting [4]

Multi-Camera System

[1] **Jarchi et al., IEEE Transactions on Biomedical Engineering, 2014**
[2] **Deligianni et al., Information Fusion, 2018**
[3] **Cippitelli et al., Sensors (Basel), 2015**
[4] **Wong et al., IEEE Sensors Journal, 2015**

Framework - Markerless Gait Analysis Based on a Single RGB Camera

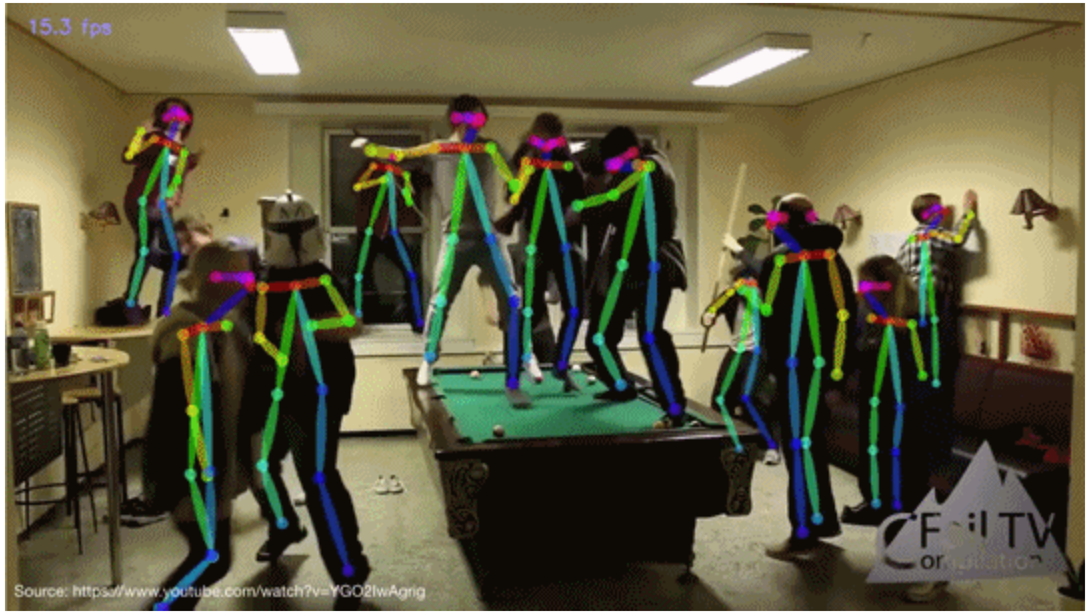
- **Based on a single RGB camera system in a cell phone**
- **No strict standards to camera position and background settings**
- **Focus on the lower limbs with six key points (Left and Right Knee, Ankle & Toe)**



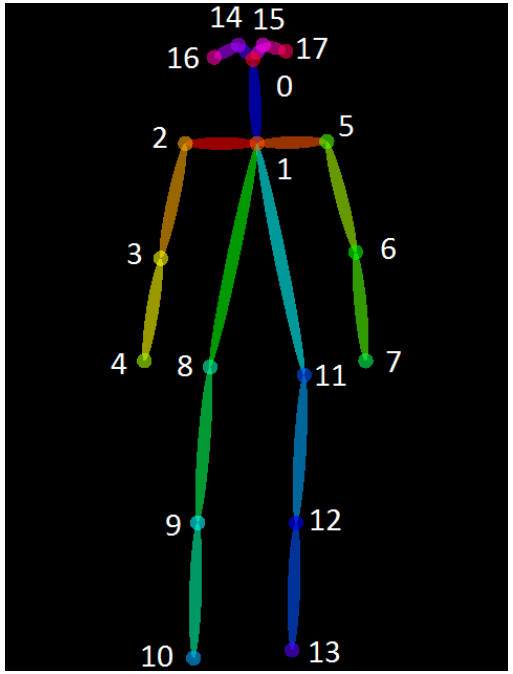
2D Keypoints Detection – OpenPose + Grabcut

OpenPose – State of the Art 2D Keypoint Detection Algorithm

- **Only detect 2D pixel locations of keypoints**
- **Accurately locate key points even though occlusion occurs**



Demo of OpenPose Body Estimation



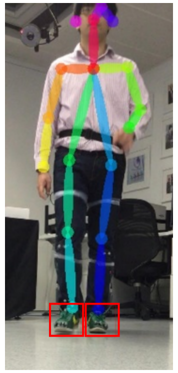
Body Keypoints Index

0, "Nose"
1, "Neck"
2, "RShoulder"
3, "RElbow"
4, "RWrist"
5, "LShoulder"
6, "LElbow"
7, "LWrist"
8, "RHip"
9, "RKnee"
10, "RAnkle"
11, "LHip"
12, "LKnee"
13, "LAnkle"
14, "REye"
15, "LEye"
16, "REar"
17, "LEar"
18, "Bkg"

Cao et al., IEEE CVPR, 2016
OpenPose code: github.com/CMU-Perceptual-Computing-Lab/openpose

2D Keypoints Detection – OpenPose + Grabcut

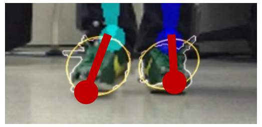
GrabCut – Mixture-models Foreground Segmentation Algorithm



Rectangle of Interest

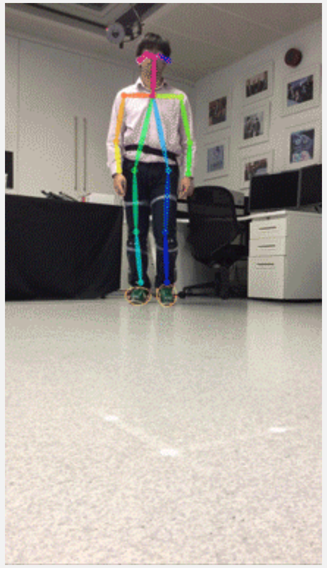
GrabCut + Morphology Operations

Fit Ellipse

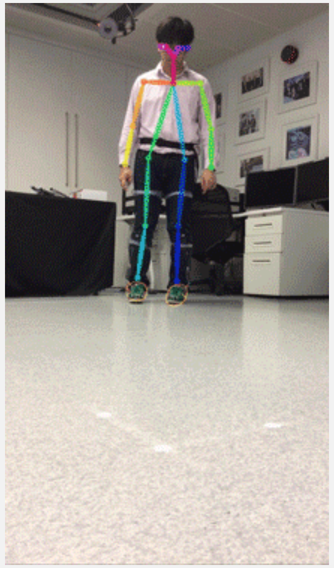


Connect ankle and ellipse centre

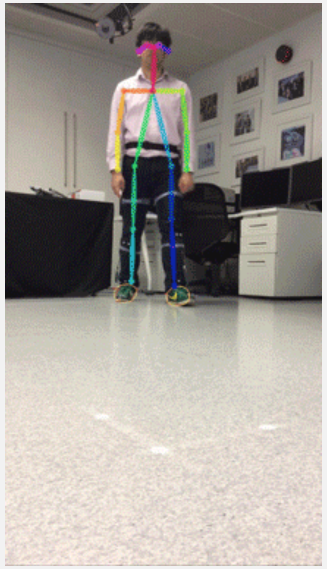
Normal



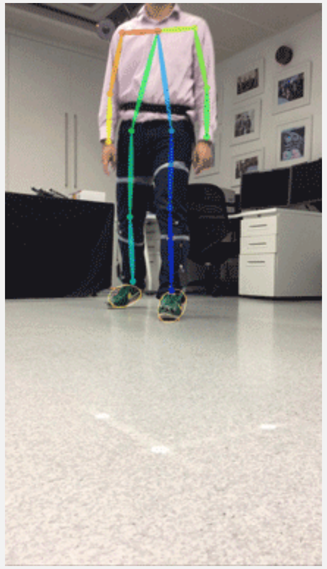
Supination



Pronation

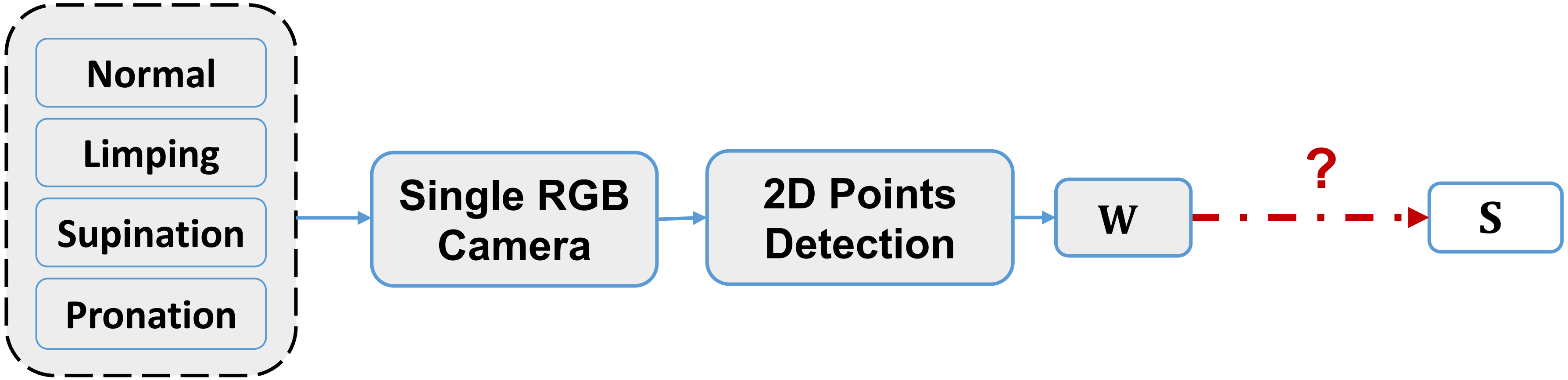


Limping



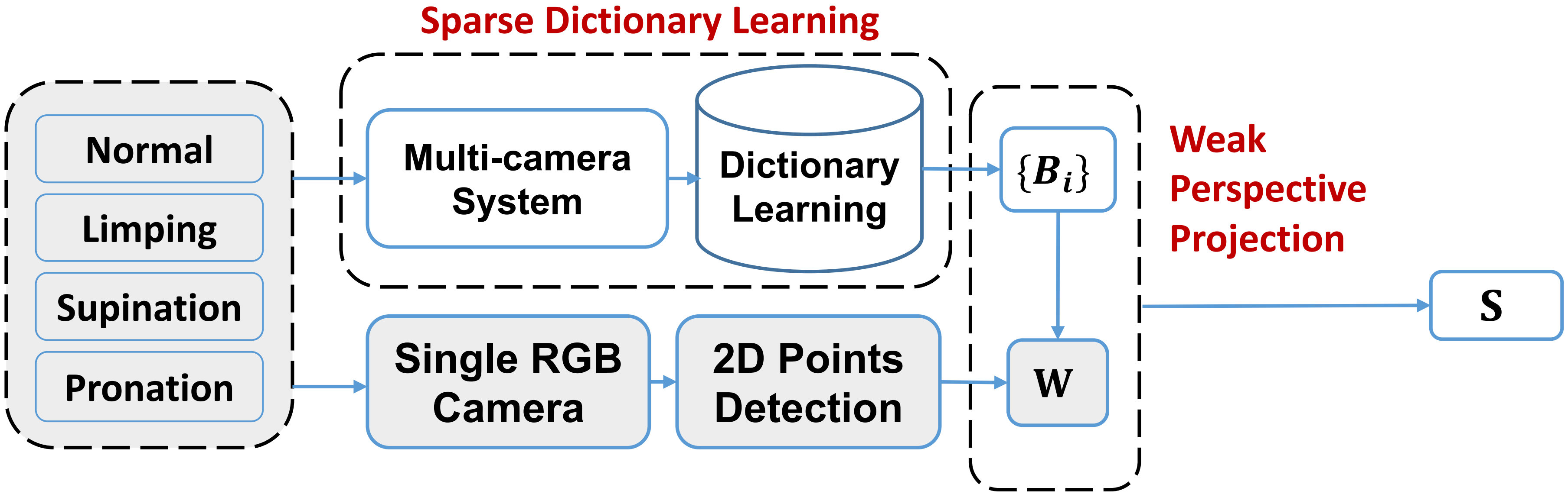
Rother, ACM SIGGRAPH 2004 Papers, 2004

3D Keypoints Reconstruction



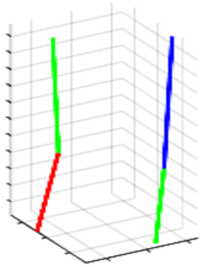
W 2D points
S 3D points

3D Keypoints Reconstrcution — Weak Perspective Projection + Sparse Dictionary Learning



$\{B_i\}$ Dictionary
W 2D points
S 3D points

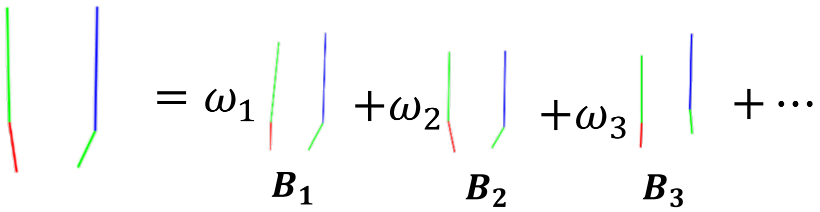
$W \longleftrightarrow S$

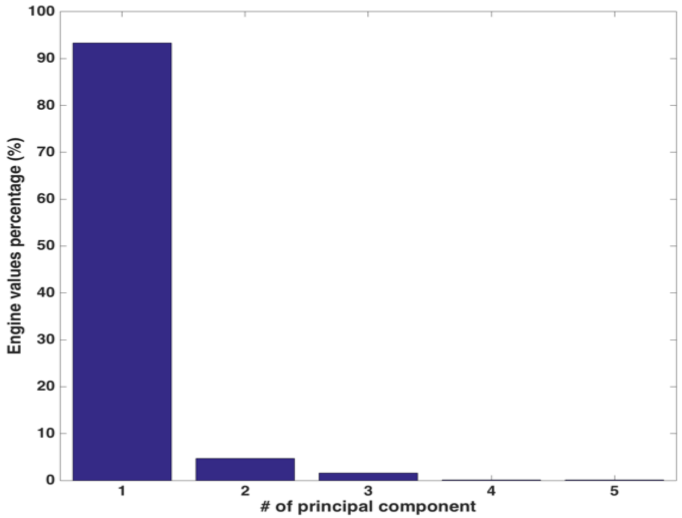


3D Keypoints Reconstruction — Weak Perspective Projection + Sparse Dictionary Learning

Sparse representation of lower limb 3D position

$$\boldsymbol{S} = \sum_{i=1}^{k} \omega_i \boldsymbol{B_i}$$





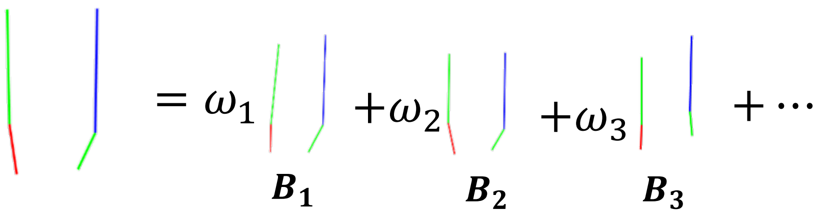
PCA Analysis

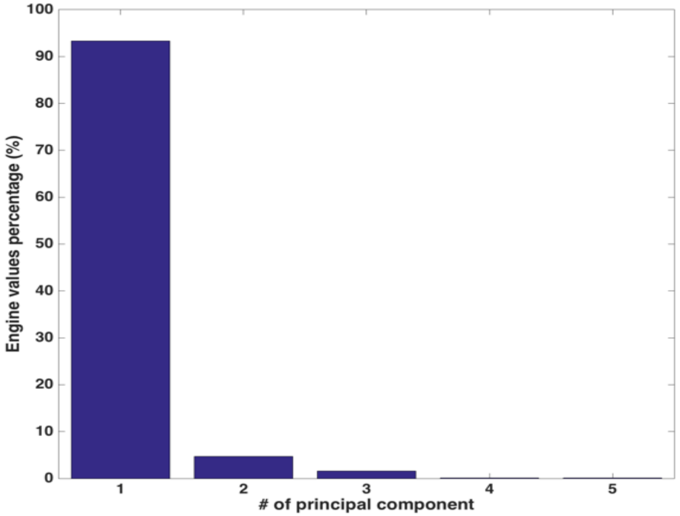
Mairal et al., J Mach Learn Res, 2010
Zhou et al., IEEE CVPR, 2015

3D Keypoints Reconstruction — Weak Perspective Projection + Sparse Dictionary Learning

Sparse representation of lower limb 3D positions

$$\boldsymbol{S} = \sum_{i=1}^{k} \omega_i \boldsymbol{B_i}$$



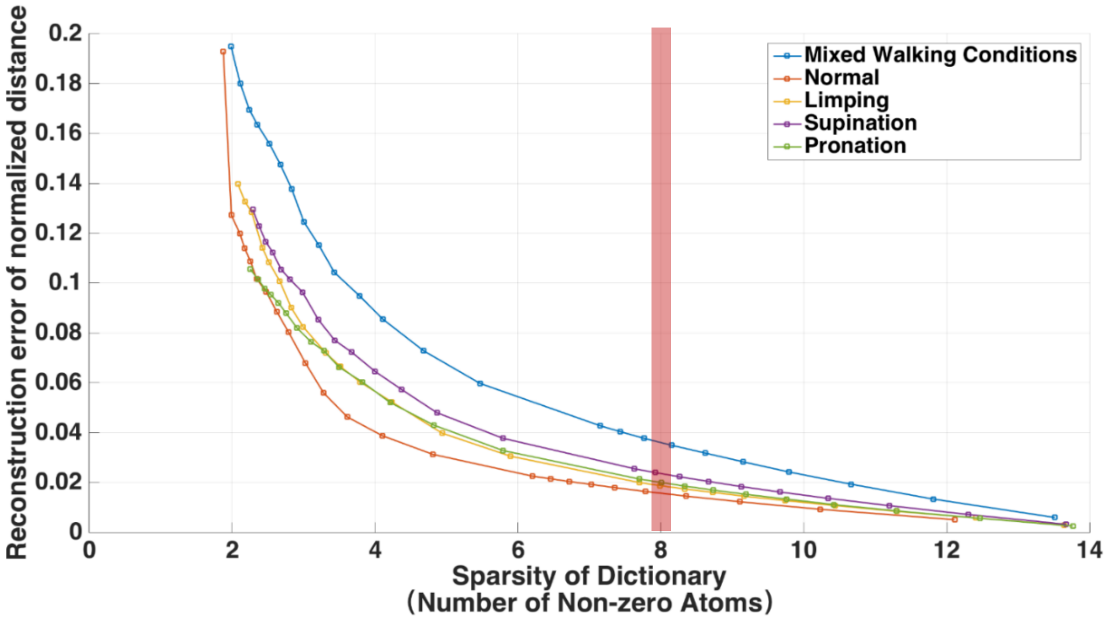


PCA Analysis

Sparse dictionary learning

$$\min_{\{\boldsymbol{B_i}\},\{\omega_{ij}\}} \sum_{j=1}^{n} \frac{1}{2} \left\| \boldsymbol{S_j} - \sum_{i=1}^{k} \omega_{ij} \boldsymbol{B_i} \right\|_F^2 + \lambda_1 \sum_{i,j} \omega_{ij}$$

$$s.t. \, \omega_{ij} \geq 0, \|\boldsymbol{B_i}\|_F \leq 1, \forall i \in [i, k], j \in [1, n]$$



Reconstruction Error of Dictionary Learning

Mairal et al., J Mach Learn Res, 2010
Zhou et al., IEEE CVPR, 2015

3D Keypoints Reconstruction — Weak Perspective Projection + Sparse Dictionary Learning

$$\Pi = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \end{pmatrix}$$ scaling matrix;

$$R \in SO(3) = \{ R \in \mathbb{R}^{3\times 3} \mid R^T R = I_3, det\, R = 1\}$$ rotation matrix;

T translation vector.

$$W = \Pi(RS + T)$$

2D $W \longleftarrow S$ 3D

2D $W \longrightarrow S$ 3D

Zhou et al., IEEE CVPR, 2015
Boyd et al., Found Trends Mach Learn, 2011

3D Keypoints Reconstruction — Weak Perspective Projection + Sparse Dictionary Learning

$\boldsymbol{\Pi} = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \end{pmatrix}$ scaling matrix;

$\boldsymbol{R} \in \boldsymbol{SO}(3) = \{ \boldsymbol{R} \in \mathbb{R}^{3\times3} \mid \boldsymbol{R}^T\boldsymbol{R} = \boldsymbol{I}_3, det\,\boldsymbol{R} = 1 \}$ rotation matrix;

T translation vector.

$$\boldsymbol{W} = \boldsymbol{\Pi}(\boldsymbol{RS} + T)$$

2D $\boldsymbol{W} \longleftarrow \boldsymbol{S}$ 3D

$$\boldsymbol{W} = \boldsymbol{\Pi RS} = \boldsymbol{\Pi R} \sum_{i=1}^{k} \omega_i \boldsymbol{B_i} = \sum_{i=1}^{k} \boldsymbol{M_i B_i}$$

$$\boldsymbol{S} = \sum_{i=1}^{k} \omega_i \boldsymbol{B_i}$$

$$\min_{\boldsymbol{M}_1,\ldots,\boldsymbol{M}_k} \frac{1}{2} \left\| \boldsymbol{W} - \sum_{i=1}^{k} \boldsymbol{M_i B_i} \right\|_F^2 + \lambda_2 \sum_i^k \|\boldsymbol{M_i}\|_2$$

2D $\boldsymbol{W} \longrightarrow \boldsymbol{S}$ 3D

Zhou et al., IEEE CVPR, 2015
Boyd et al., Found Trends Mach Learn, 2011

Experiment Settings

- **Device**
  - Wearable Mobile Phone
    - 30Hz
    - put in front of the subjects
  - Smart DX, BTS Bioengineering
    - 200Hz
    - Multi-camera motion capture system
    - Reflective ball put in knee, ankle, toe

- **Four Subjects**
  - 1 female, 3 males
  - walking straight to the camera

- **Four Walking Conditions**
  - Normal
  - Supination
  - Pronation
  - Limping

Experiment Settings

- **Device**
  - Wearable Mobile Phone
    - 30Hz
    - put in front of the subjects
  - Smart DX, BTS Bioengineering
    - 200Hz
    - Multi-camera motion capture system
    - Reflective ball put in knee, ankle, toe

- **Four Subjects**
  - 1 female, 3 males
  - walking straight to the camera

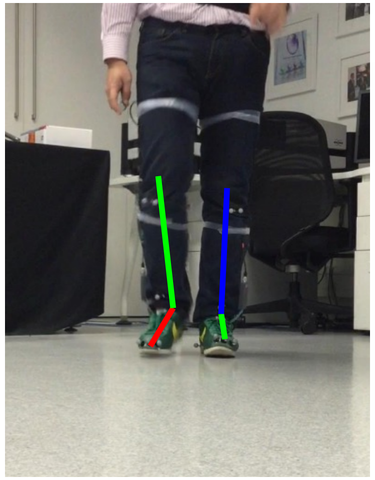
- **Four Walking Conditions**
  - Normal
  - Supination
  - Pronation
  - Limping

- **Leave-one-out Cross Validation**

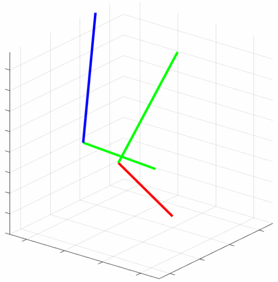
- **For each walking condition of each subject**
  - Extract frames of all other walking conditions (3 conditions of current subject + 4 conditions X 3 other subjects)
  - Train the dictionary
  - Use the trained dictionary for weak perspective projection
  - Validate based on extracted angular features

Qualitative Results

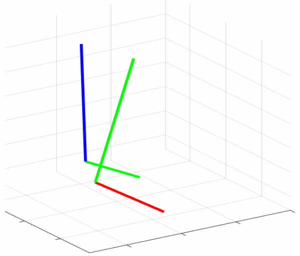
- **Normal**



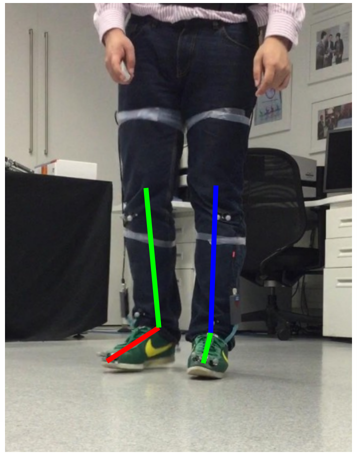
Ground truth



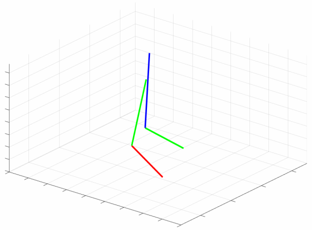
Estimation



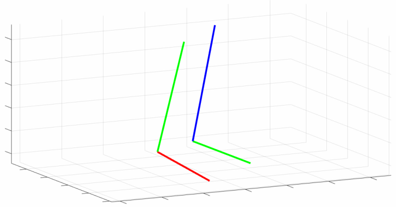
- **Limping**



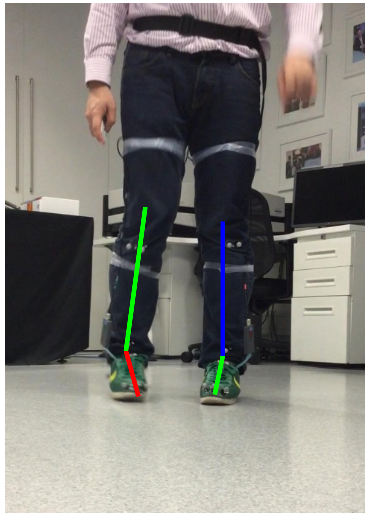
Ground truth



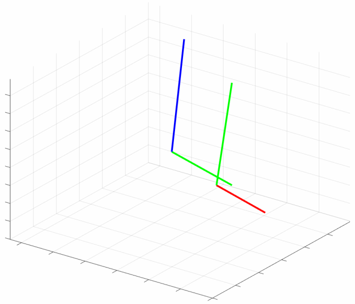
Estimation



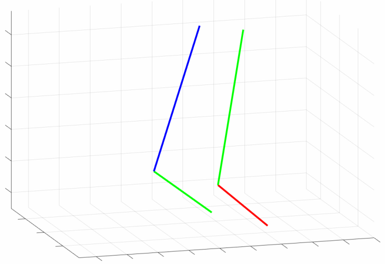
- **Supination**



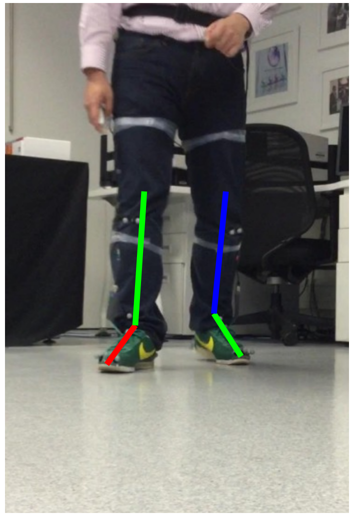
Ground truth



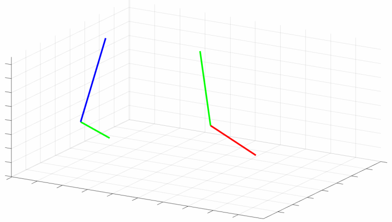
Estimation



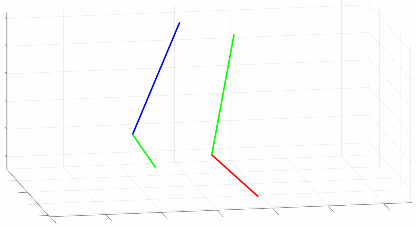
- **Pronation**



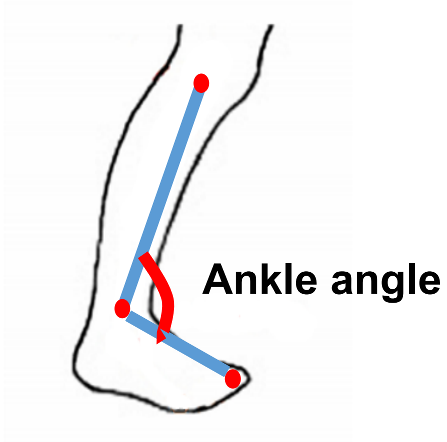
Ground truth

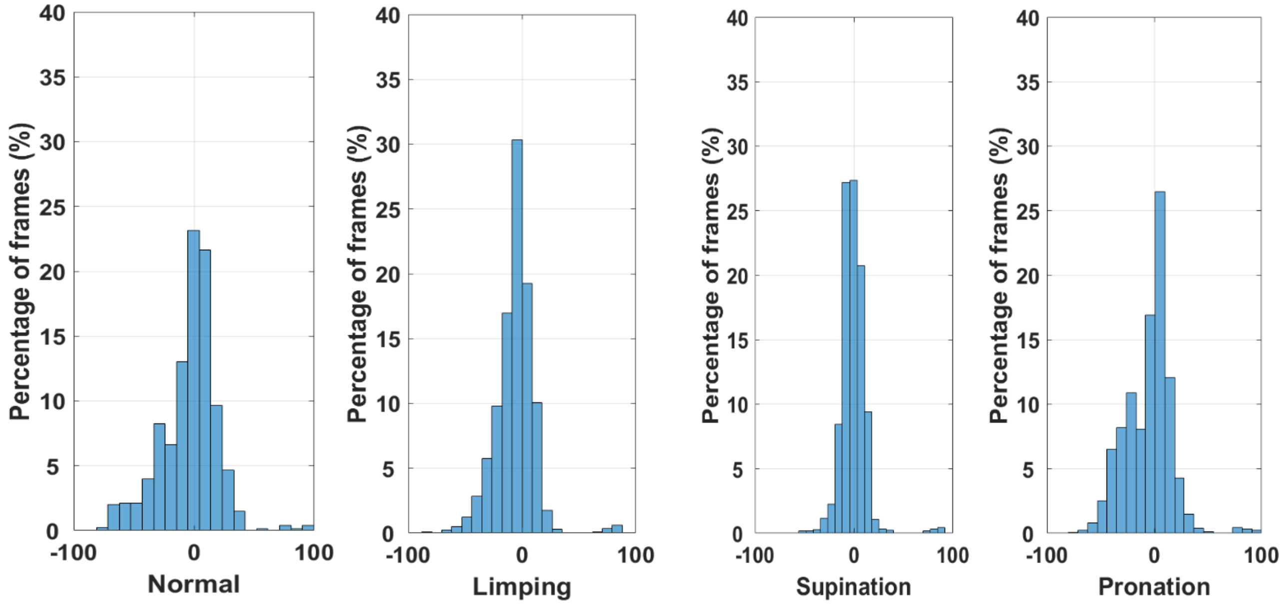


Estimation

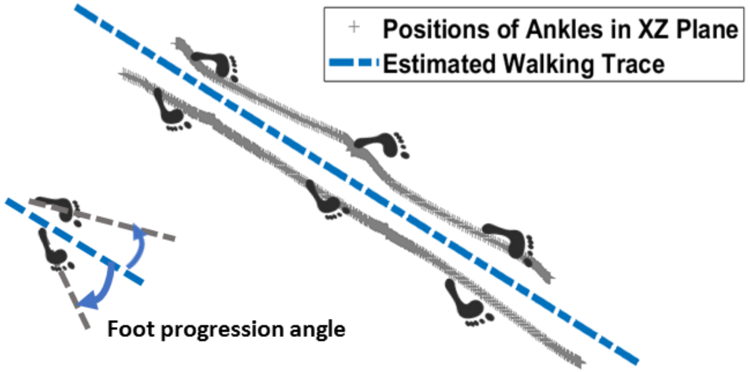


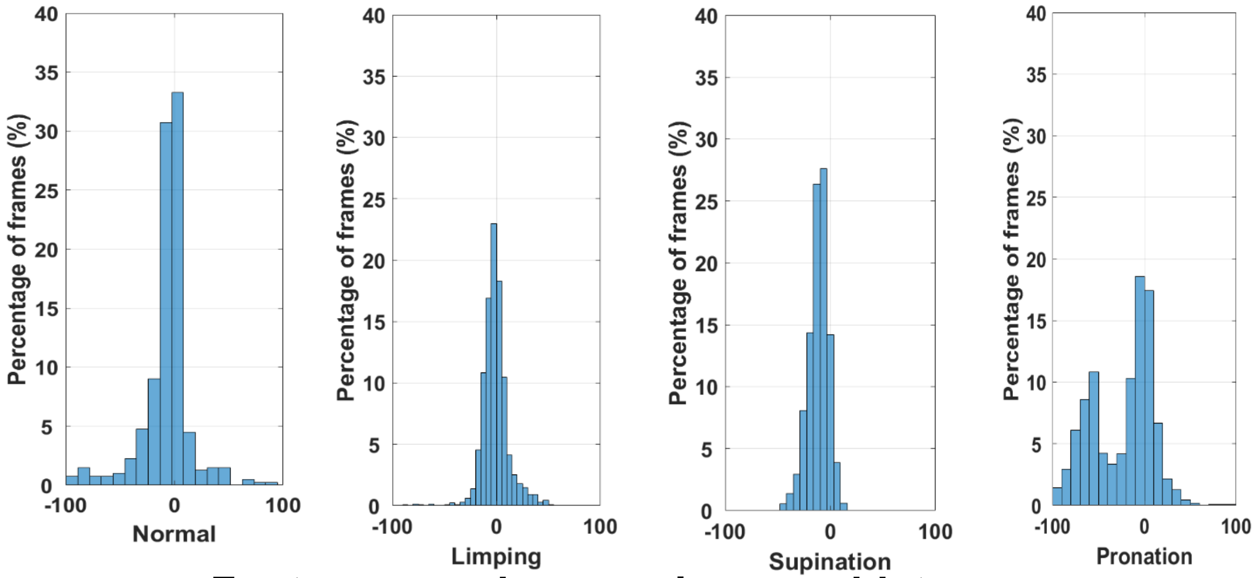
3D Angular Features Validation based on Multi-camera System





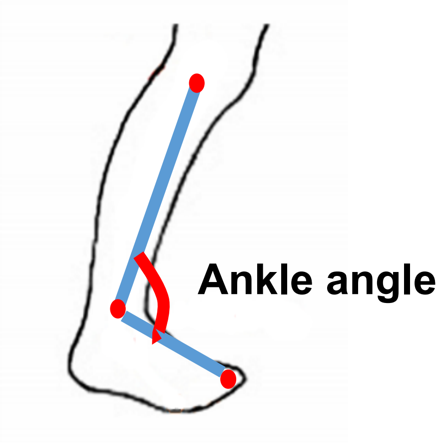
Ankle angular error histogram

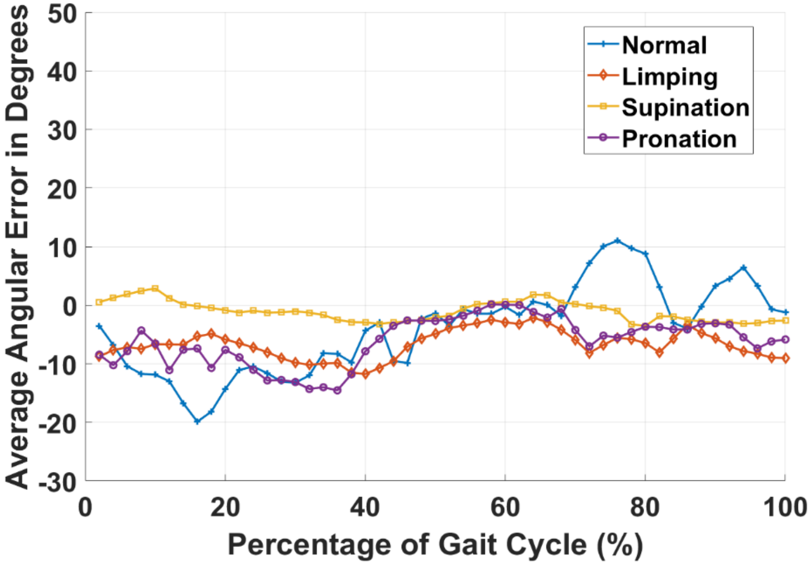




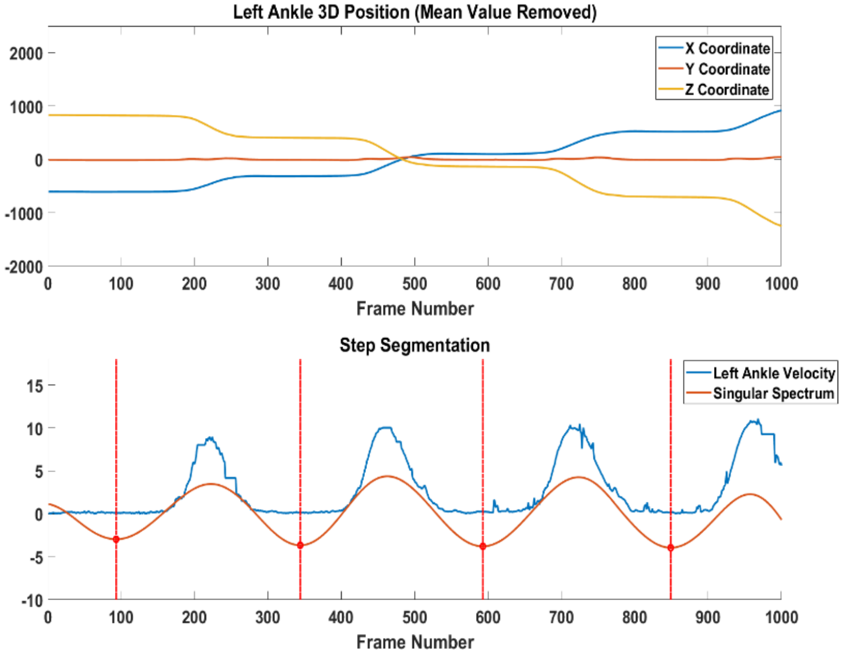
Foot progression angular error histogram

3D Angular Features Validation based on Multi-camera System

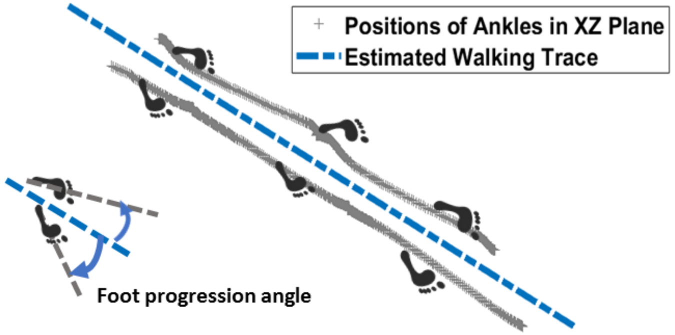


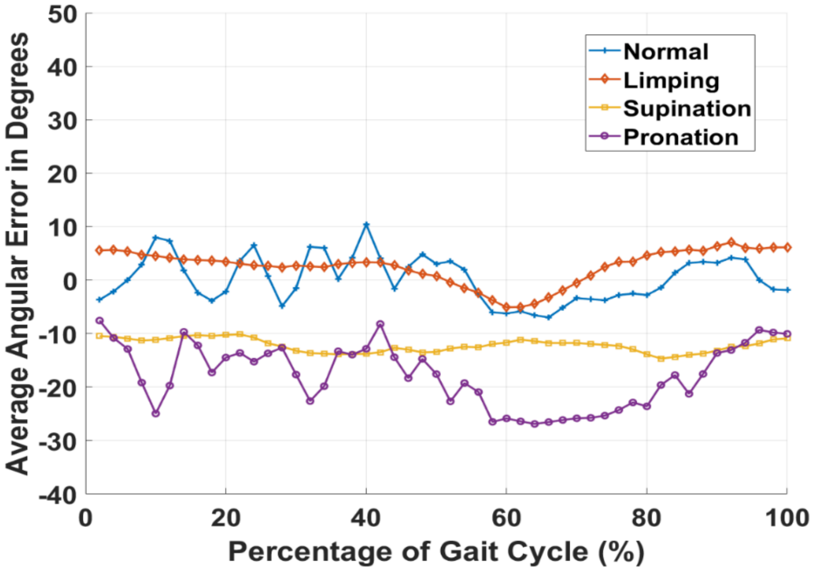


Ankle angular error of different walking conditions in a gait cycle



Singular Spectrum Analysis [1, 2]





Foot progression angular error of different walking conditions in a gait cycle

[1] Deligianni et al., Inf. Fusion, 2018
[2] Jarchi et al., IEEE TBME, 2014

Summary

- **Developed a novel framework to estimate 3D gait angular features of the lower limbs**

  - 2D joint detection + 3D coordinates reconstruction

- **Validated our result based on a state-of-the-art 3D multi-camera system**

- **Accuracy compares well with methods based on markers and depth information**

  - Within 10 degrees for estimation error in most frames

- **Future work**

  - Methodology Improvement

  - Clinical test + Larger datasets