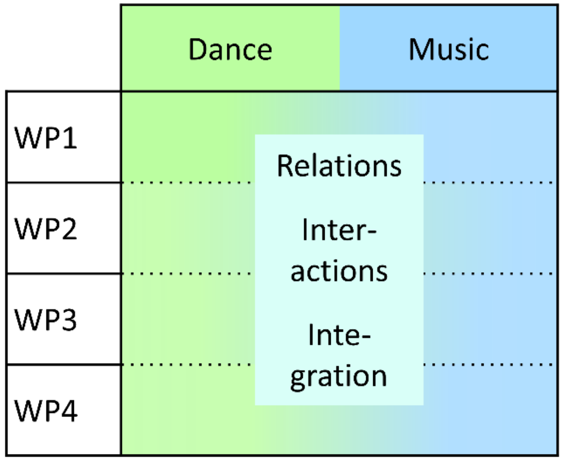
DjembeDance @ MSP

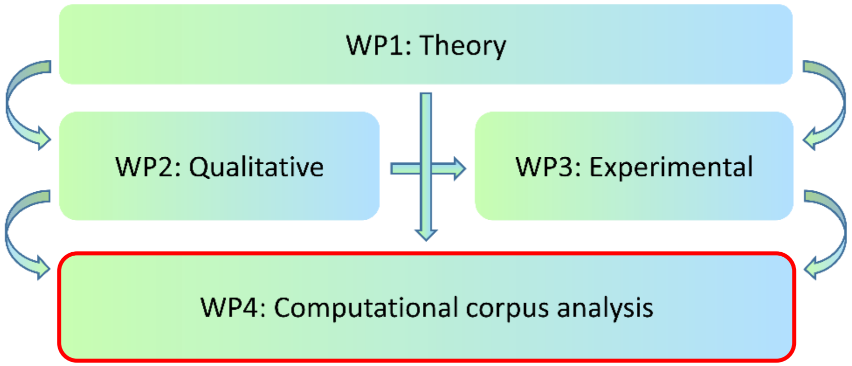
Sagar Dutta and Rainer Polak

TOC

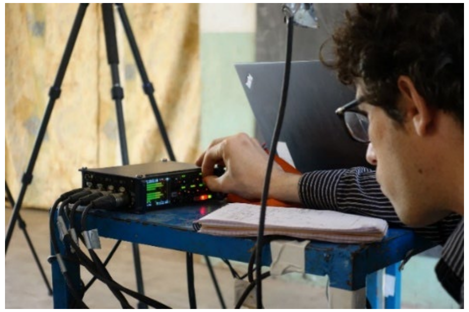
- Brief outline of DjembeDance
- Current work-task: synching mocap to audio
- Future work within DjembeDance

Project ambitions

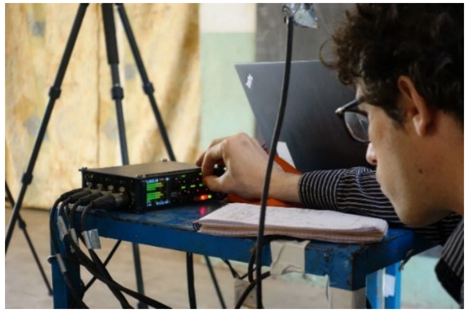




Audio



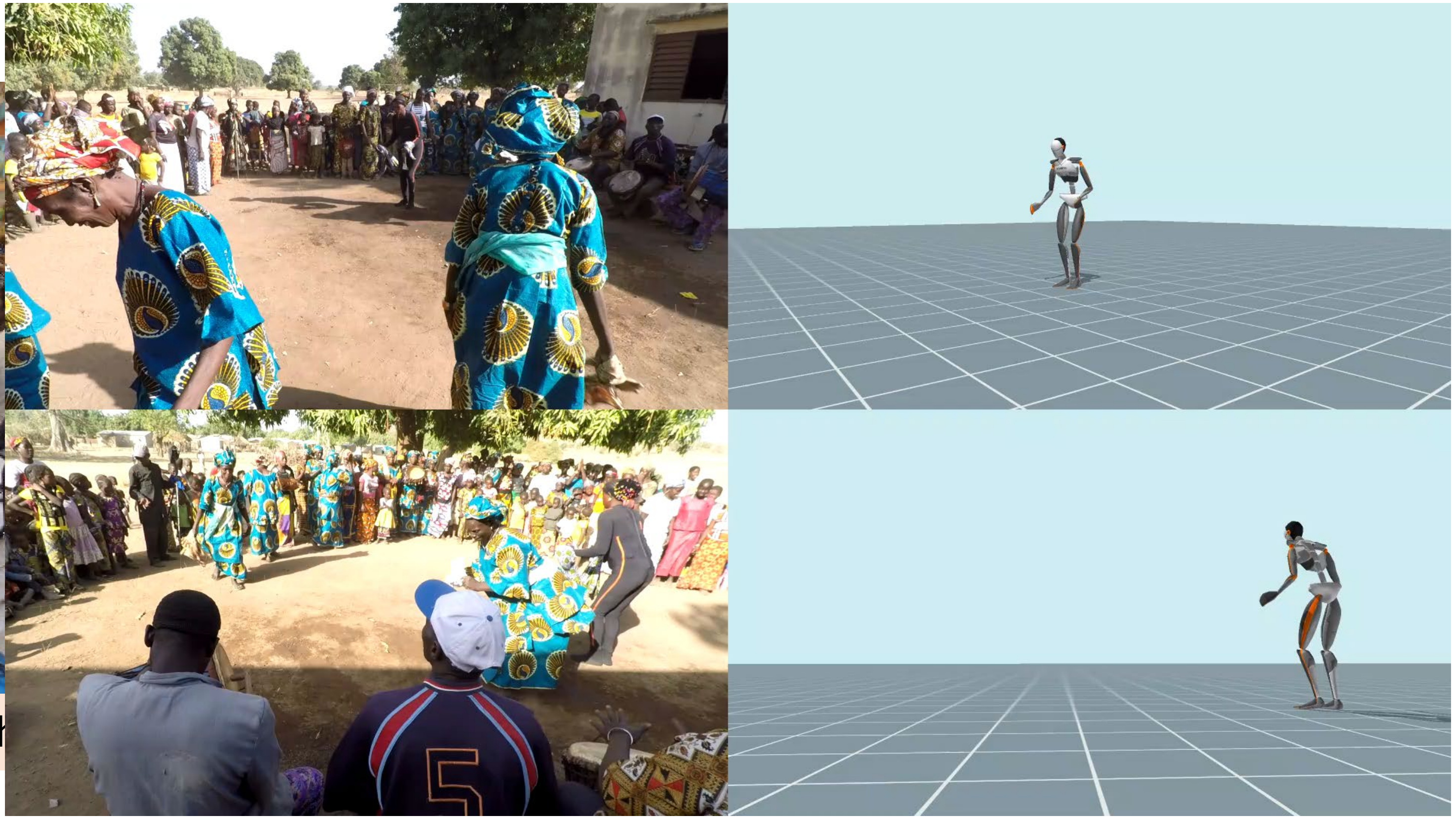
Audio and Video



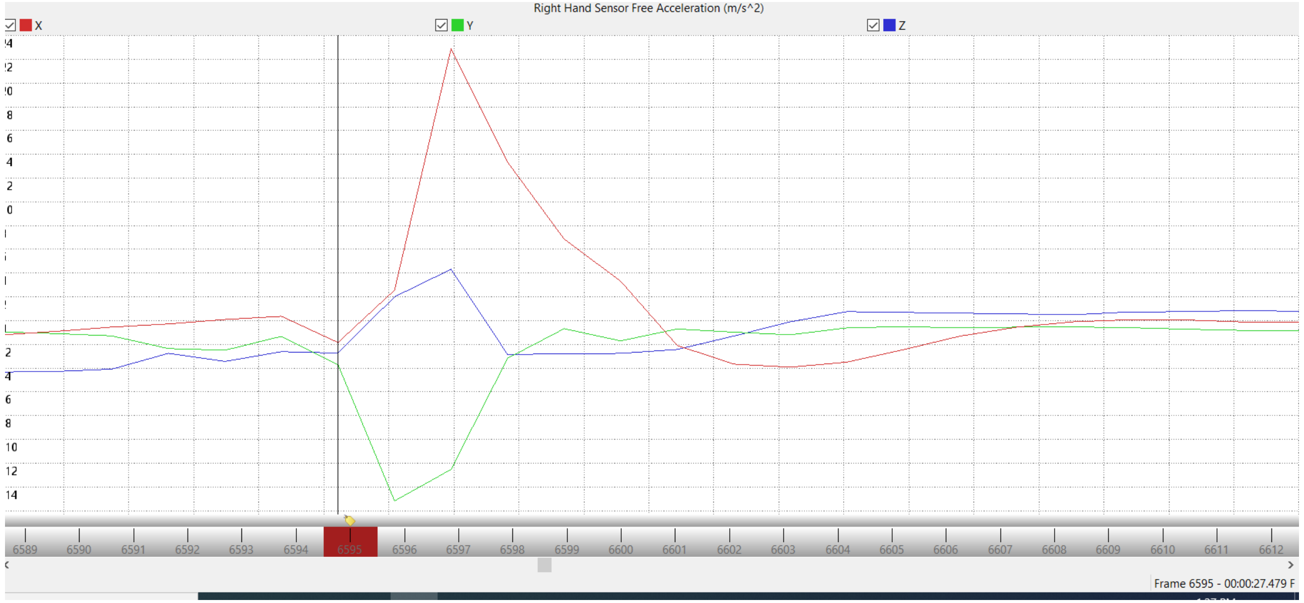
Motion Capture

Recording situations

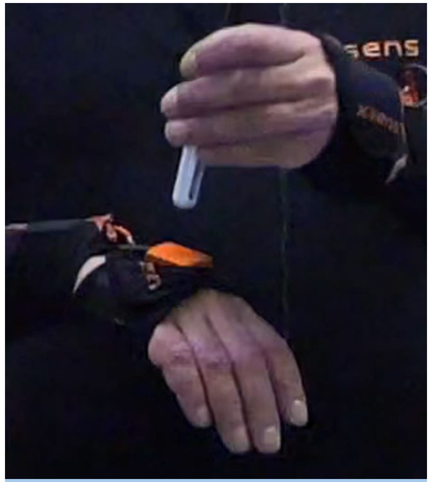
Urban professionals, 6 sessions, 6+ h



Synchronizing mocap to audio



Calibration procedure



1 frame = 4.16 msec

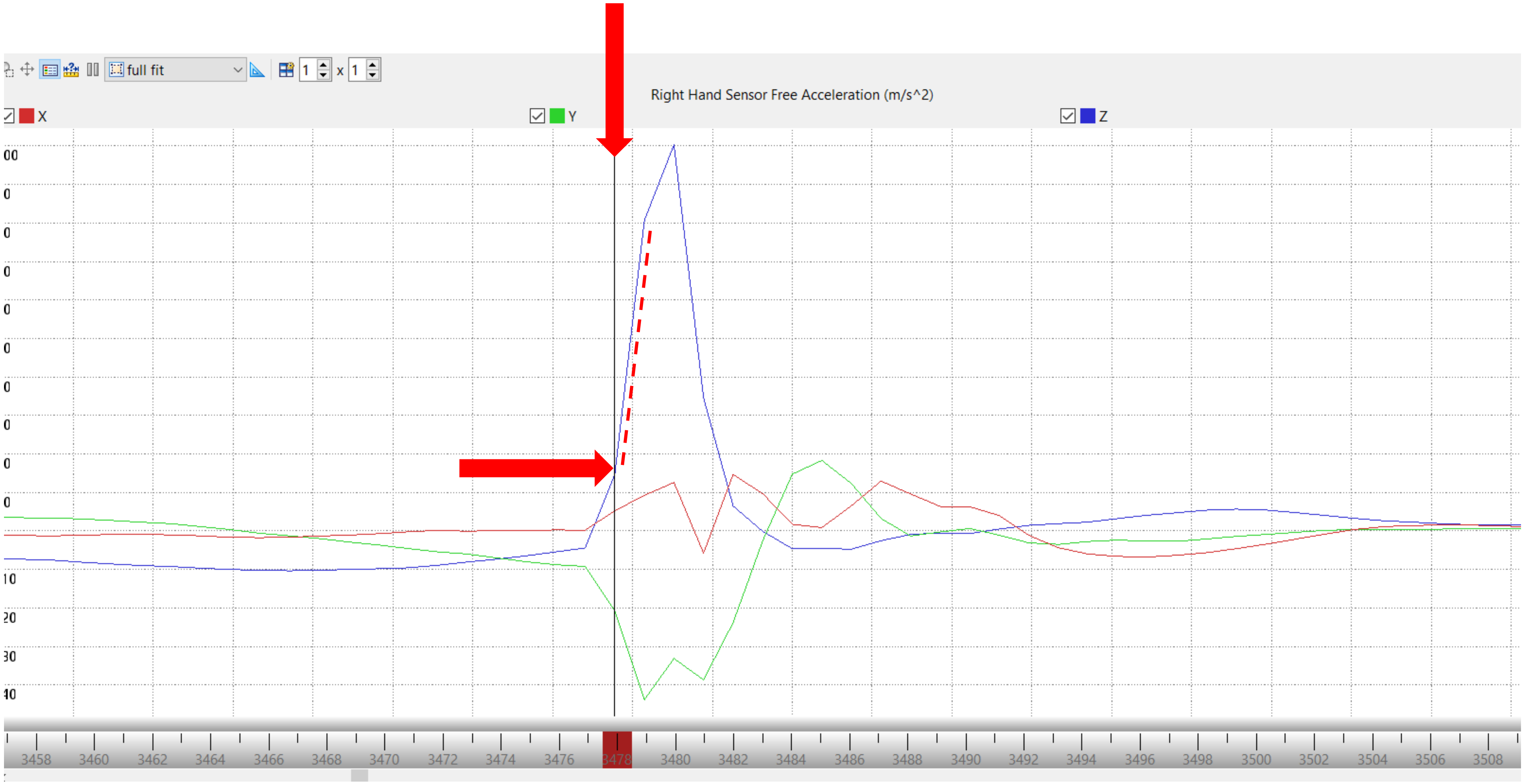
- Include in a single recording
  - calibration event: tick, rise-time:
    - <1 msec audio
    - <1frame (4.16msec) mocap
  - test event: clapperboard slate
- Calculate expectation for mocap slate from audio timings
- Interpret mocap data plot at expected time-point
- Result: Impact occurs
  - On the frame showing the steepest slope in resultant acceleration (XYZ)
  - 1-3 frames before first peak

Example Tick 1

Right hand, Raw sensor data, Free acceleration

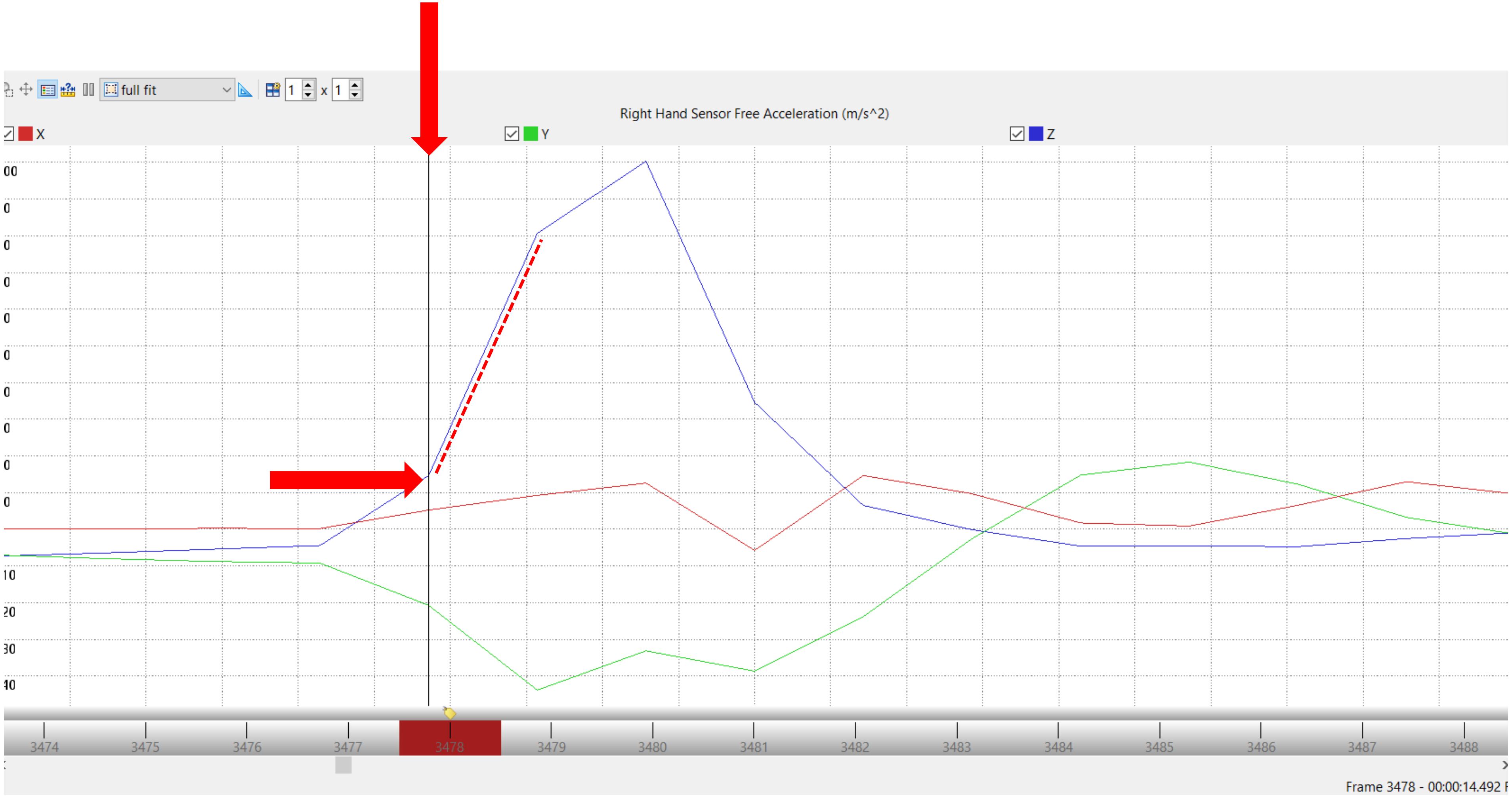
Tick 1 (zoom-in)

Example Slate 1 (Test event impact)



Impact (timing calculated from audio) - - - - - - - - - - Steepest slope

Slate 1 (zoom-in)



Impact ---------------- Steepest slope

Slate 2 (zoom-in)

Cross Modal Synchronization

Step 1: Select segment with clap onsets

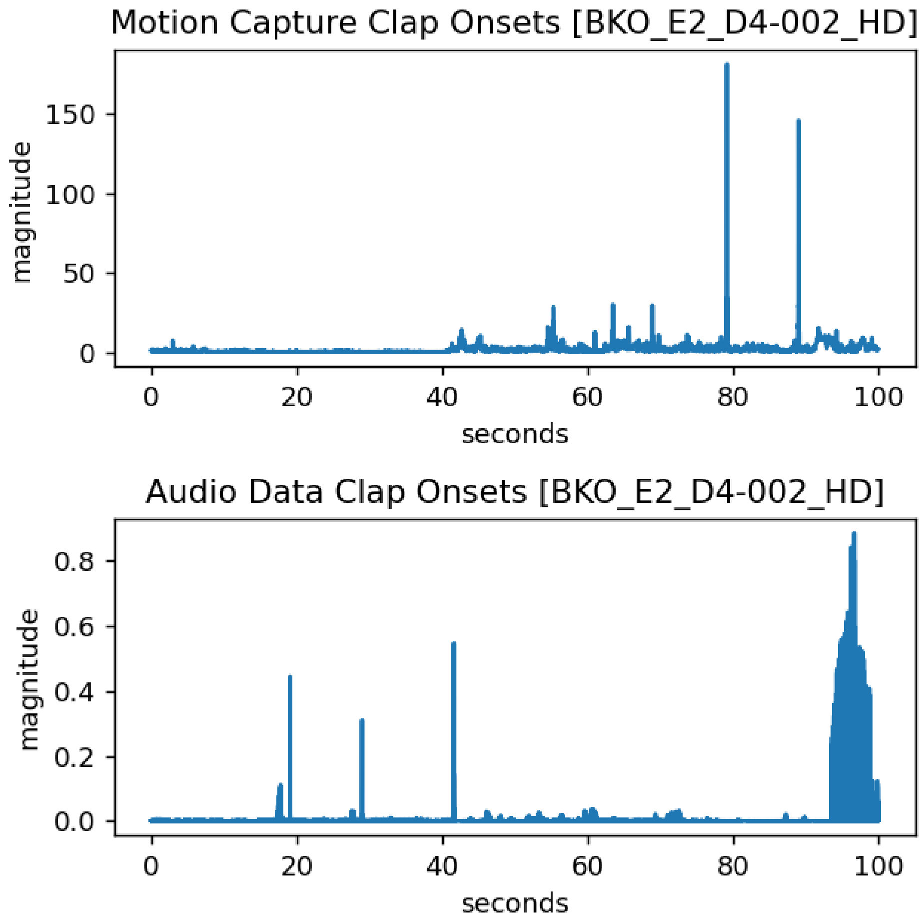


Fig. 1: Visualization of first 100 seconds

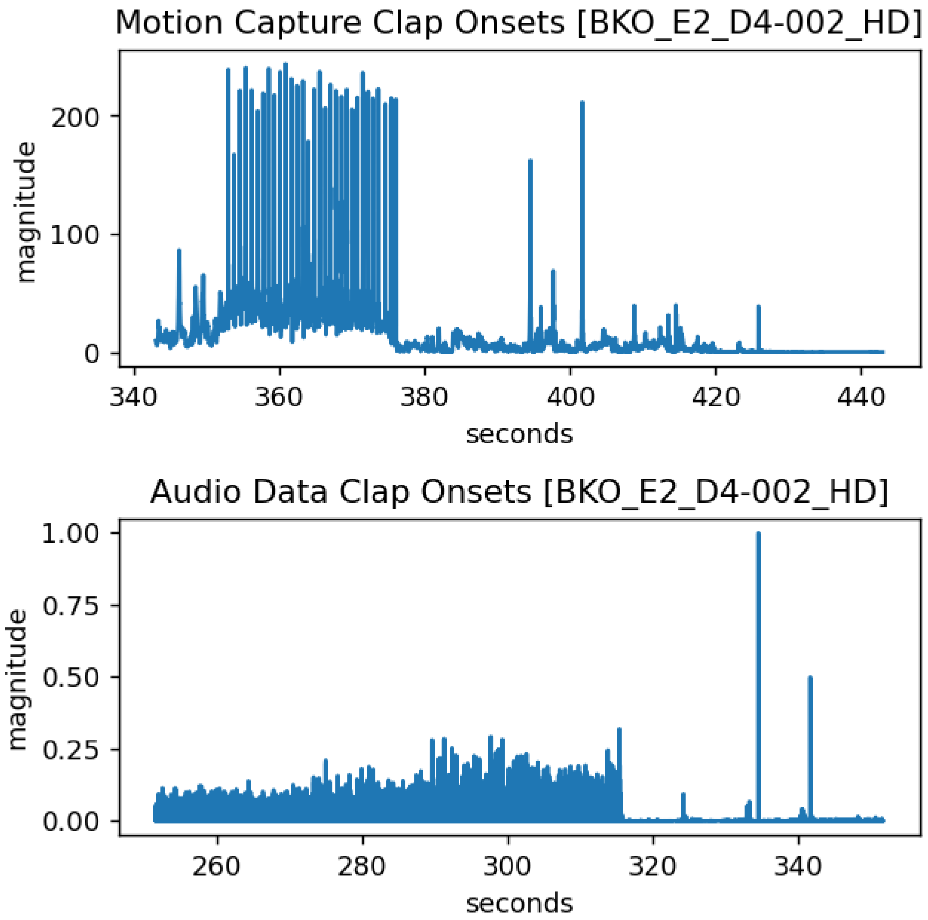


Fig. 2: Visualization of last 100 seconds

Step 2: Set threshold for peak detection

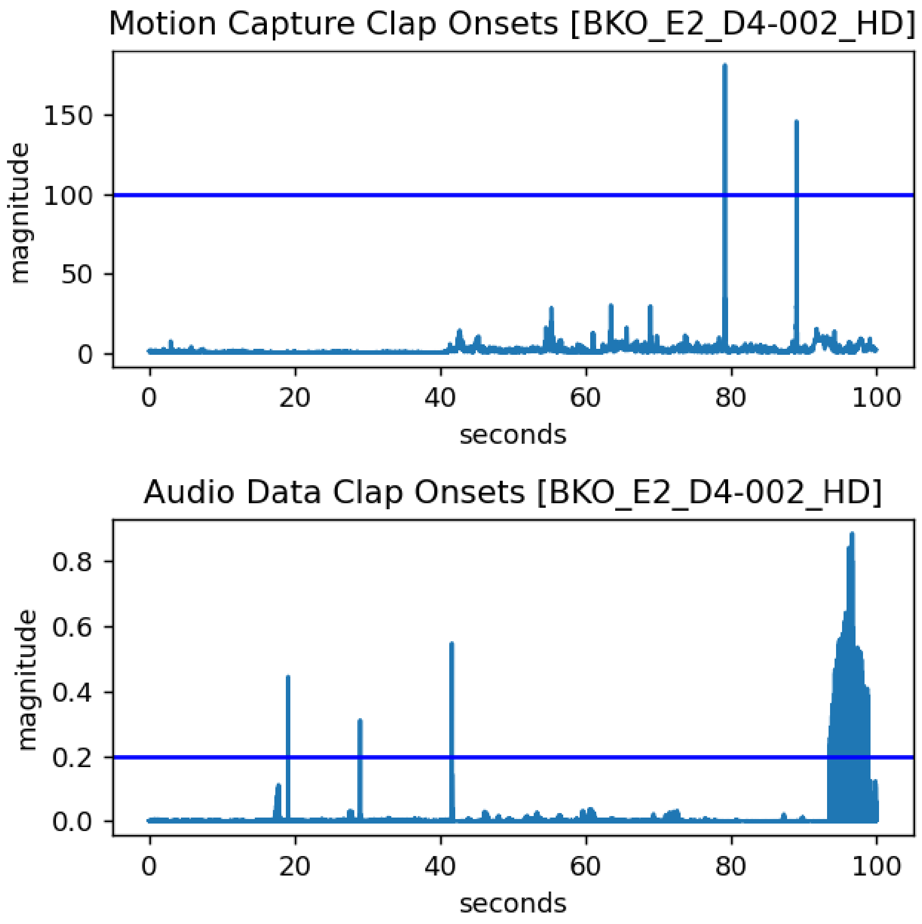


Fig. 3: Threshold selection for peak detection

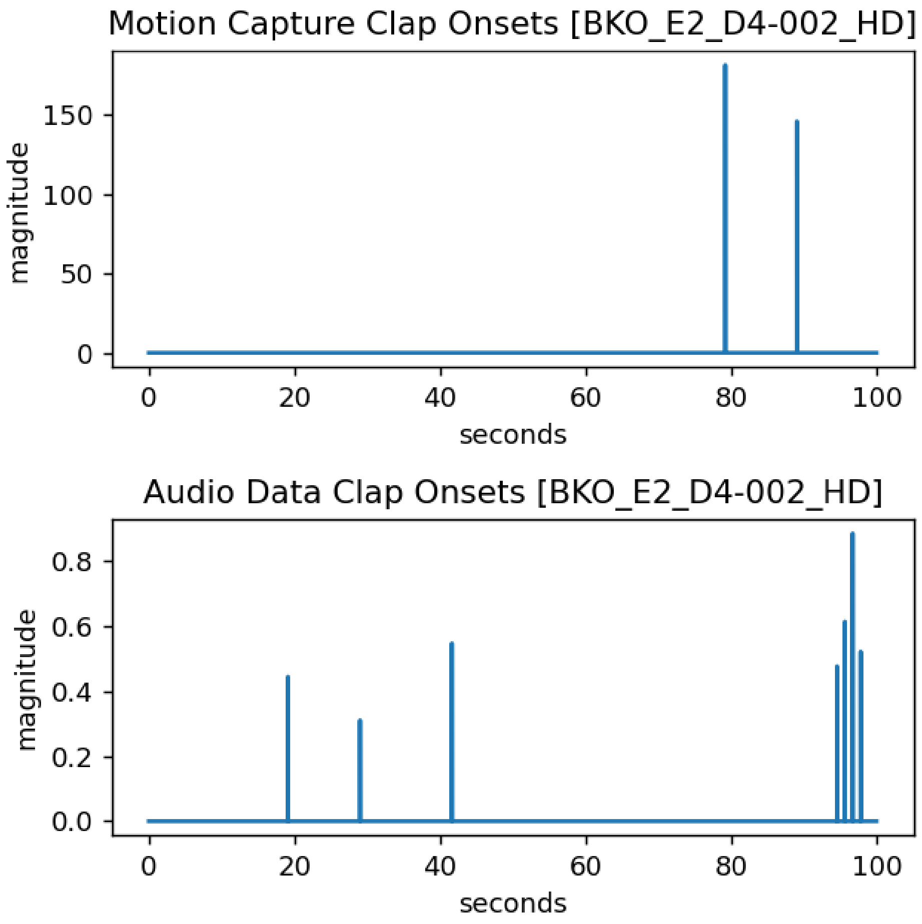


Fig. 4: Extracted Peaks

Step 3: Normalizing Peak Value to Unity and Mocap Up-sampling

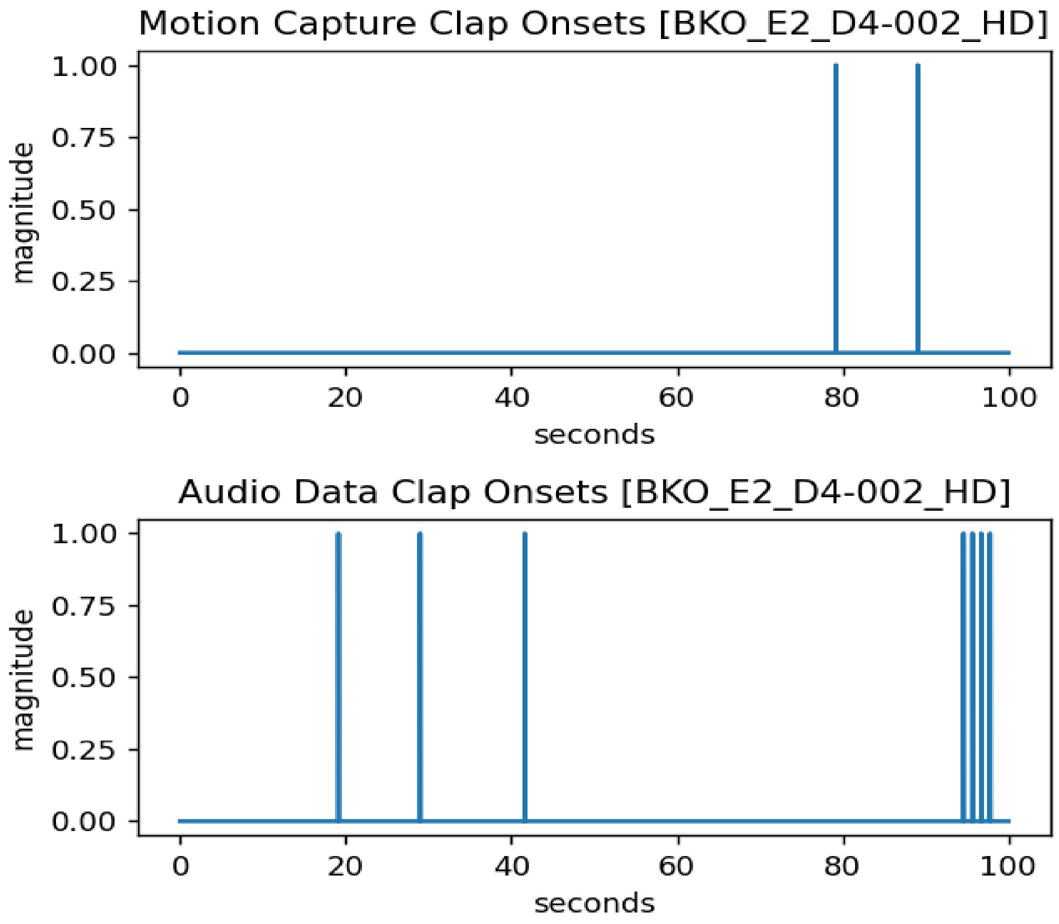


Fig. 5: Normalization of peaks

@240Hz
mocap onsets: [79.175, 89.0375]
mocap peak frame: [19002 21369]

@48KHz
mocap onsets: [79.175, 89.0375]
mocap peak frame: [3800400, 4273800]

@48KHz
audio onsets: [19.1212917, 28.9886042, 41.612375, 94.571375, 95.6385833, 96.7002083, 97.7877708]

audio peak frame: [917822 1391453 1997394 4539426 4590652 4641610 4693813]

Step 4: Apply correlation and find offset

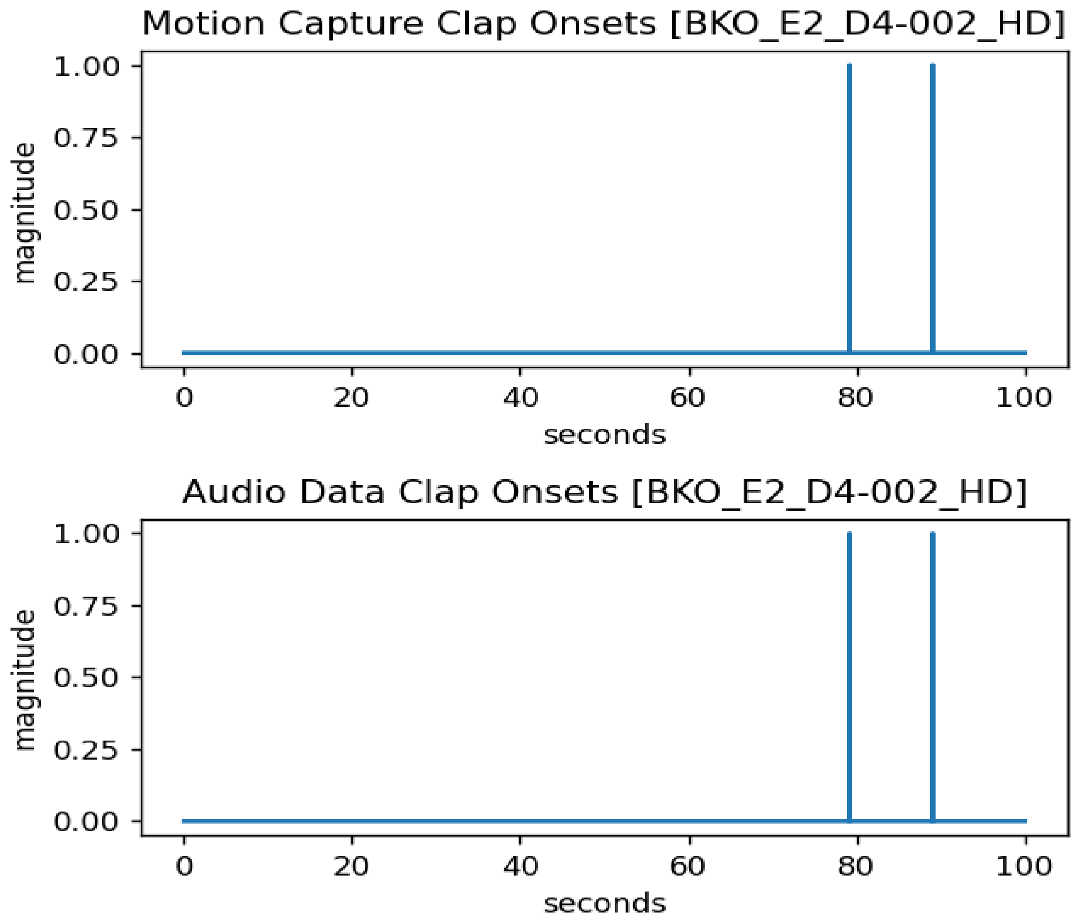


Fig. 6: Synchronization of mocap and audio

Audio Frames to shift for sync. : 2882462
Time shift: 60.051 seconds

Synchronized Peaks: [79.175, 89.0375]

Step 5: Mocap Frame Correction

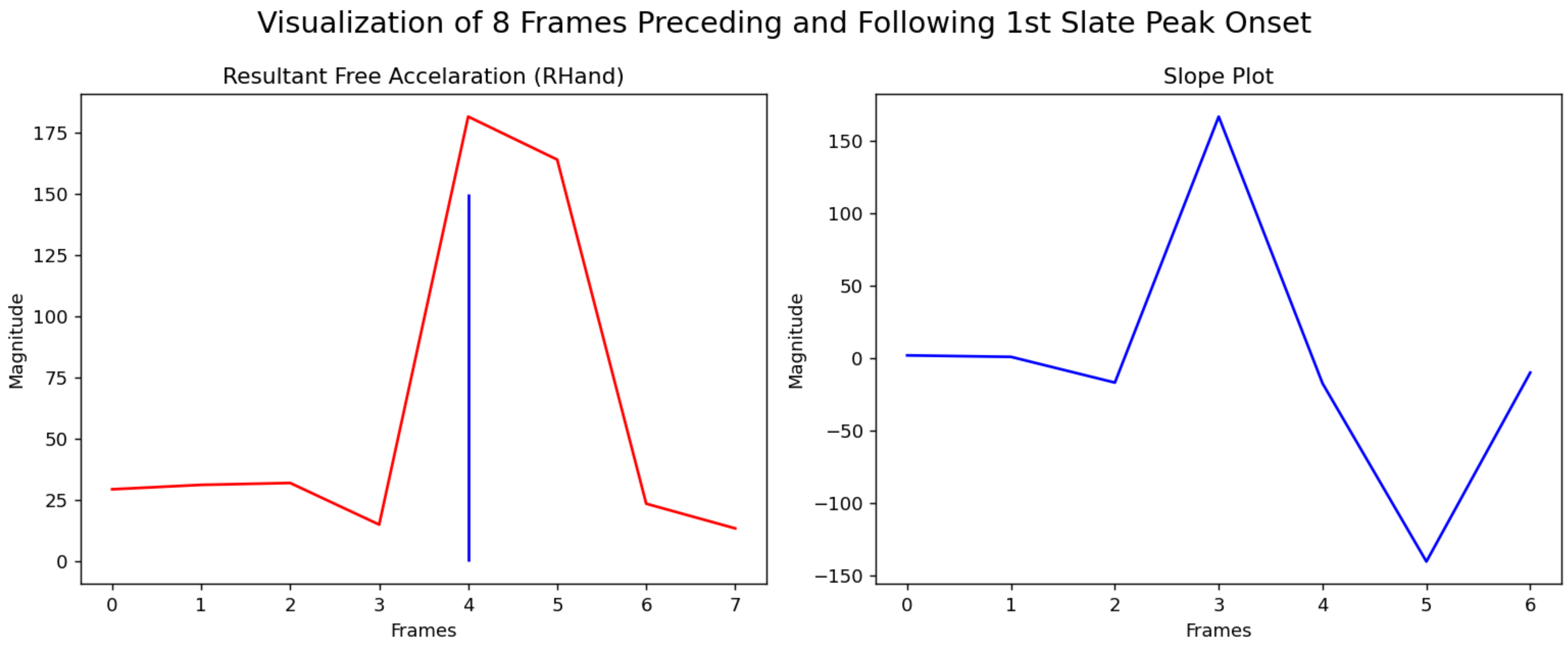


Fig. 7: Synchronization of mocap and audio

Future work within DjembeDance

- MIR: Extract from audio
  - Drum stroke onsets (timings)
  - Drum stroke sounds (pitch/timbre)
  - Vocals pitch over time (melody)
- Dance analysis in MoCap
  - Distribution of perceptually salient events (contrasts) across a reference periodicity (metric cycle)
    - Kinetic energy (aggregated velocity, acceleration)
    - Directional change
  - Spatiotemporal patterns:
    - Movement motifs per body-part (Basic gestures)
    - Full-body composite rhythms
- Music-Dance analysis: Correlational, Causal

Movement patterns and musical meter: periodicity coordination (Toiviainen et al. 2010)

- Objective: Analyze periodicity in music-induced movement across various tempi
- Methodology:
  - Focus on potential and kinetic energy of the body.
  - Utilize location markers for movement analysis.
  - Principal Component Analysis (PCA) of body segments to identify dominant movement patterns
- Key Findings:
  - Participants synchronized with periods of one, two, and four beats
  - Systematic relationship of body part size/mass and periodicity/tempo
    - Large size/mass body parts (e.g., torso) tend to use longer periodicities/slower tempos (e.g., 2/4-beat cycles)
    - Smaller size/mass body parts (e.g., arms) tend to use shorter periodicities/faster tempos (e.g., 1-beat cycle)

Toiviainen, P., Luck, G., & Thompson, M. R. (2010). Embodied Meter: Hierarchical Eigenmodes in Music-Induced Movement. Music Perception

Movement patterns and musical meter: periodicity coordination (Toiviainen & Carlson 2022)

- Objective: Investigates how eigen movements in music-induced dancing depends on musical content and genre.
- Methodology:
  - Use of trial-space-frequency tensor for capturing the complexity of movements
  - Decomposition of data using tensor decomposition techniques
  - Identification of 12 principal movement components
- Key Findings:
  - Time-frequency domain decomposition is better
  - Some eigen movements synchronized with music
  - Difference found in eigen movement amplitude between genres

Toiviainen, P., & Carlson, E. (2022). Embodied Meter Revisited. *Music Perception*

Limitations of the Toiviainen approache(s)

- Principal Components are noncorrelated thus movement patterns synchronized to the different metrical levels are noncorrelated as well.
- Assumes stationary data within the analysis window of PCA which may not capture certain dynamics of movement (e.g., translation and rotation).
- Their (2010) calculation of kinetic energy from speed excludes kinematic variables of velocity and acceleration from the analysis.
- Phase information is lost in the (2022) frequency-domain analysis, which is good for the analysis of frequency-locked body parts, but not good for phase-dependent analysis, e.g.,
  - the synchronization quality (accuracy and precision) between music and dance;
  - composite rhythm patterns emerging from coordinated but phase-shifted body-parts.

Further analysis

- Empirical Mode Decomposition
- Independent Component Analysis
- Denoising Source Separation
- Multidimensional Scaling

Further ideas (collaboration? qualification?)

- Use video
  - Use video as additional information to improve beat tracking algorithms
  - Compare mocap:video (1 dancer)
    - Use MoCap as ground-truth/training data for 3D←2D/video estimation
  - Compare lead dancer (mocap/video) with other dancers (video)
  - Acknowledge audience member behaviours
    - Compare audience:dancers
    - Estimate audience contribution to music:dance correlation