AI-Assisted Performance Analysis

Deep Learning for Live and Archival Theater

Michael J. Rau,[1] Peter Broadwell,[2] Simon Wiles,[2] Vijoy Abraham[2]

[1] Theater and Performance Studies [2] Center for Interdisciplinary Digital Research [1,2] Stanford University

This project seeks to build on recent advances in human pose-estimation and object tracking technologies powered by emergent deep-learning methods to develop new tools and techniques for the analysis and augmentation of theatrical performances.

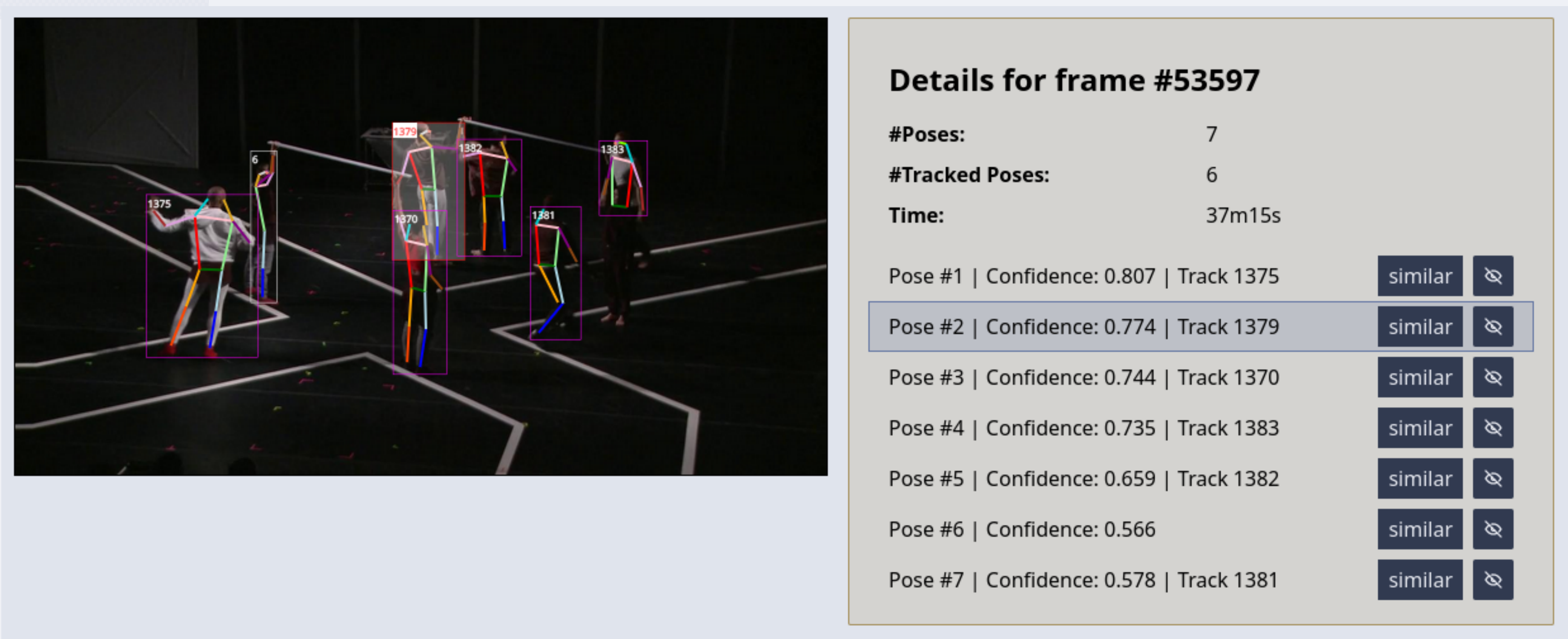
ONLINE

Modern pose-estimation and tracking methods are capable of performing inference of body positions and motions from images captured by standard visual-light cameras, unlocking for archival and live footage possibilities that previously only existed in constrained environments with expensive motion-capture suits and equipment.

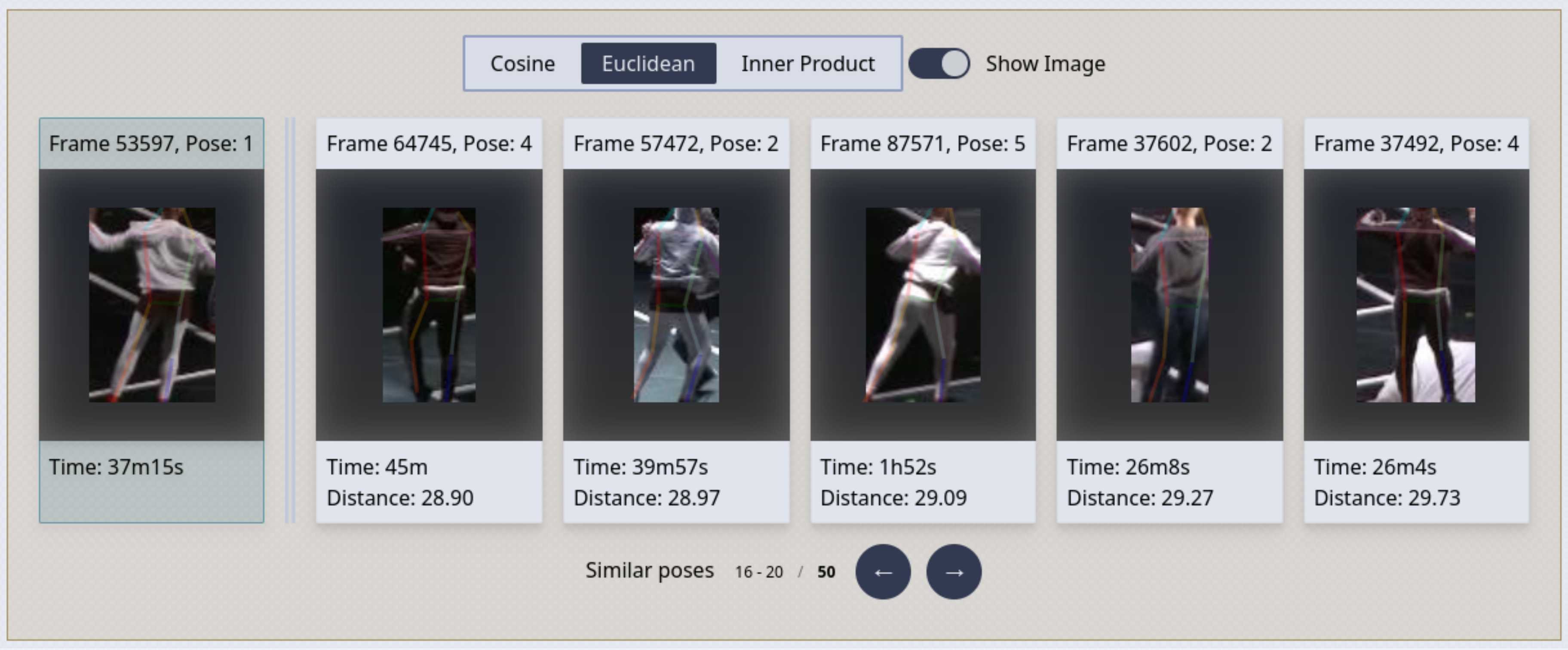
Leveraging these methods, we are developing software and methodologies that use pose data to aid in the investigation, analysis, elucidation, and quantification of insights into aspects of theatrical performance including directorial style and choreographic patterns.

- A window into a performance is provided by the overview which presents pose-estimation results on a timeline.
- Here it is possible to see how many poses have been identified on a frame-by-frame basis, and how many have been successfully tracked across frames.

- The Frame Viewer component visualizes an individual frame and associated pose data

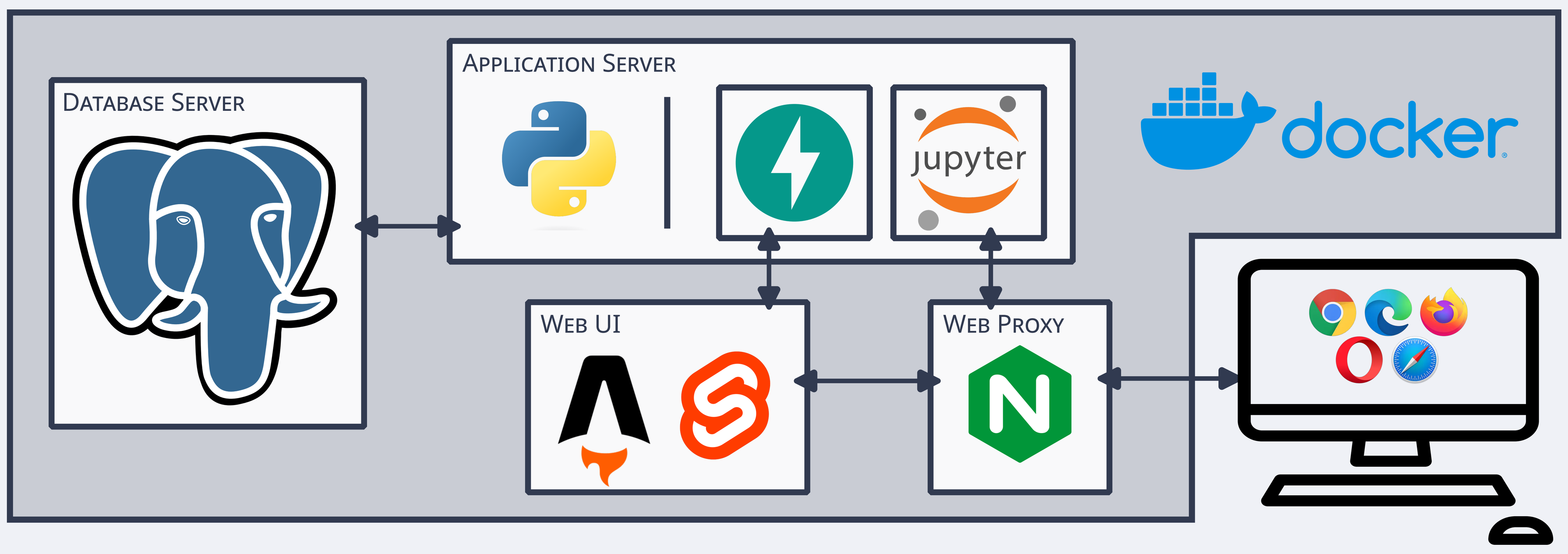


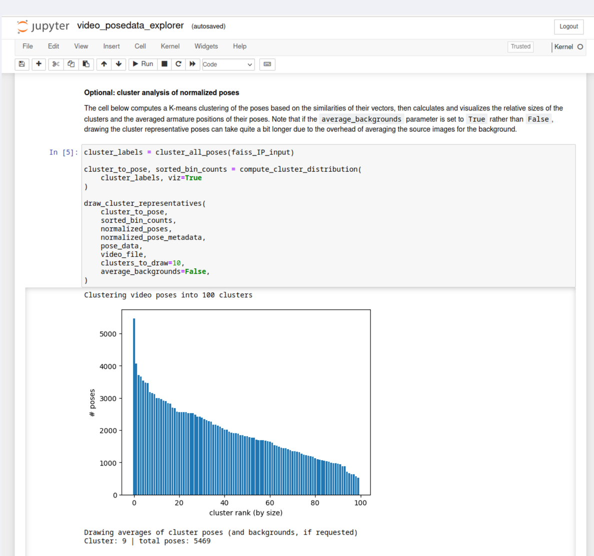
- Using vector similarity searching, the poses most similar to the target pose can be identified and located within the performance. Work continues to augment simple nearest-neighbors -style similarity across entire poses with other heuristics that help to focus in on the context-specific kinds of similarity that are meaningful to researchers.



Technical Stack

- At the present time the machine-learning stack is centered around pose-estimation provided by **OpenPifPaf**, augmented with object tracking across frames based on work by **ByteTrack**. Vector similarity is performed using algorithms from the **Faiss** library.
- The application stack is oriented around a **PostgreSQL** database augmented with **vector storage and search** capabilities to handle pose/armature vectors and embeddings.
- A **Python**-based backend handles ingestion and analysis of new video files, committing data to the database and surfacing it to the primary user interface via a **FastAPI**-based REST API.
- The web-based frontend is developed in **Astro** and **Svelte**.
- The application stack is developed and deployed as a (currently) **four-container docker orchestration** for convenience and portability.
- The Python backend also serves a **Jupyter** notebook server out of the same environment to facilitate further exploration and experimentation.





https://github.com/sul-cidr/mime

TAPS Stanford Department of Theater & Performance Studies

CIDR

Stanford LIBRARIES