




Augmenting Access to Embodied Knowledge Archives: A Computational Framework

Giacomo Alliaata <giacomo_dot_alliaata_at_epfl_dot_ch>, Laboratory for Experimental Museology, EPFL, Switzerland 
<https://orcid.org/0000-0001-8281-4451>

Yumeng Hou <yumeng_dot_hou_at_epfl_dot_ch>, Laboratory for Experimental Museology, EPFL, Switzerland 
<https://orcid.org/0000-0002-7908-0693>

Sarah Kenderdine <sarah_dot_kenderdine_at_epfl_dot_ch>, Laboratory for Experimental Museology, EPFL, Switzerland 
<https://orcid.org/0000-0002-7190-9946>

Abstract

With the burgeoning use of digital technologies in safeguarding intangible and living heritage, memory institutions have produced a significant body of material yet to be made accessible for public transmission. The quest for new ways to unlock these massive collections has intensified, especially in the case of embodied knowledge embedded in complex formats, such as audiovisual and motion capture data.

This study examines a computational workflow that combines posture recognition and movement computing to bridge the gap in accessing digital archives that capture living knowledge and embodied experiences. By reflecting on how embodied knowledge is sustained and potentially valorised through human interaction, we devise a series of methods utilising vision-based feature extraction, pose estimation, movement analysis, and machine learning. The goal is to augment the archival experience with new modes of exploration, representation, and embodiment.

This article reports the computational procedures and algorithmic tools inspected through two use cases. In the first example, we visualise the archives of the Prix de Lausanne, a collection of 50 years of video recordings of dance performances, for archival exploration through the dancers' poses. In another experiment, movement encoding is employed to allow multimodal data search via embodied cues in the Hong Kong Martial Arts Living Archive, a comprehensive documentation of the living heritage of martial arts that is chiefly comprised of motion-captured performances by masters.

Though holding different application purposes, both projects operate on the proposed framework and extract archive-specific features to create a meaningful representation of human bodies, which reveals the versatile applications that computational capacities can achieve for embodied knowledge archives. The practices also represent a model of interdisciplinary involvement where the archivists, computists, artists, and knowledge holders join hands to renew strategies for archival exploration and heritage interpretation in a new light.

1 Introduction

Living heritage, also known as intangible culture, encompasses the practices, expressions, spaces, and knowledge that safeguard the rich diversity of human creativity and collective heritage, as defined by UNESCO. The preservation and exploration of living heritage has been of paramount importance in understanding and celebrating the cultural fabric of our society. To further enhance our engagement with this heritage, advancements in digital technologies have opened new possibilities for capturing and experiencing embodied knowledge. 1

In recent years, there has been a growing trend in the digital capturing of living heritage and embodied knowledge [Hou 2

et al. 2022]. These endeavours aim to document and preserve not only physical artefacts but also the intangible aspects of culture, such as rituals, performances, and traditional practices. However, GLAM sectors, encompassing galleries, libraries, archives, and museums, face significant challenges in effectively managing and providing access to these vast digital repositories [Jaillant 2022].

One of the key challenges faced by GLAM sectors is the sheer scale of digitised collections. As more and more heritage materials are digitised, the volume of data grows exponentially, making it increasingly difficult to navigate and explore these archives manually and essentially turning these collections into “dark archives” inaccessible to the public. Giovanna Fossati et al. further highlight how non-expert audiences are in need of more intuitive modes of access to discover these collections [Fossati et al. 2012]. To address this challenge, recent research has embarked on applying computational analysis, artificial intelligence (AI) in particular, to unlock various new modes of archival experience. Leveraging the transformative potential of AI^[1], we can enhance the retrieval, visualisation, and navigation of multimodal cultural data materials [Aske and Giardinetti 2023]. [Cameron, Franks, and Hamidzadeh 2023] [Hou, Seydou, and Kenderdine 2023]. Further, by involving the collaborative efforts of cultural domain experts in configuring AI tools, there is a promise to alleviate the gap between data science and GLAM sectors, highlighted through various uses cases on the implementation of AI strategies in Swedish heritage organizations [Griffin, Wennerström, and Foka 2023], and enable users to interact with digitised archives in meaningful yet trustworthy contexts [Jaillant and Rees 2023].

To tackle these challenges, we propose a computational framework that augments the archival experience of embodied knowledge. Our approach intends to extract human movements in the form of skeleton data and process it, showing its versatility through multiple use cases. By revalorizing and visualising knowledge, memory, and experience, we strive to enhance the cognitive reception and engagement of users, ultimately improving the overall archival user experience. Our experiments are conducted on two exemplar collections: the archives of the Prix de Lausanne, comprising an audiovisual dataset of dance performances, and the Hong Kong Martial Arts Living Archive (HKMALA), which contains both videos and motion capture recordings of Hong Kong martial arts performers. This work might be of interest to professionals of cultural heritage institutions wishing to explore new methods of access to their collections, as well as researchers of computational humanities with a focus on embodied knowledge archives.

By exploring this computational framework, we hope to contribute to the ongoing efforts to leverage visual AI technologies to transform the way we interact with and appreciate living heritage. Through the synthesis of cultural heritage and cutting-edge computational methods, we can bridge the gap between the past and the present, enabling a deeper understanding and appreciation of our shared human history.

2 Related Work

2.1 Embodied Knowledge as Part of Living Heritage

Though partly embedded in architectural sensorium and memory objects, intangible cultural heritage (ICH) is inherently living and manifests through bodily expressions, individual practices, and social enactment — on the multiple layers of embodiment. Recent research has revealed a stress shift in documenting, describing, and presenting the living nature of ICH by transforming intangible practices into tangible choreographic objects. For instance, the i-Treasures project (2013-2017) works to digitize the living heritage of folk dances, folk singing, craftsmanship, and contemporary music composition in an attempt to sustain traditional know-how via creative presentation [Dimitropoulos et al. 2014] [Grammalidis and Dimitropoulos 2015]. Taking a more public-facing approach, the Terpsichore project (2016 to 2020) integrates computational models, such as semantic web technologies and machine learning, with storytelling to facilitate data accessibility and affordance of digitised materials [Doulamis et al. 2017]. Famous choreographers have also been keen on documenting their practices with audiovisual recordings, resulting in extensive collections such as William Forsythe's Performative Archive, recently acquired by the ZKM Center for Art and Media Karlsruhe. While its focus differs from the performing arts, the Hong Kong Martial Arts Living Archive (HKMALA) project (2012-present) forges a multimodal digital archiving paradigm by motion-capturing martial art sequence performances of renowned masters. The motion-captured (MoCap) records are situated within a digital collation capturing different facets of Hong Kong's martial arts traditions, encompassing rituals, histories, re-enactments, and stories of practitioners [Chao et al. 2018].

The embodied facet of ICH not only refers to the knowledge of the body but also that which is dwelling in and enacted through the body [Craig et al. 2018]. Movements and gestures are considered typical mediums to express and process mindfulness. In various cases, they also mediate interactive processes, such as human-environment communication and knowledge formation. Movement data usually documents the active dimension of mankind's creativity over time. Hence, as Tim Ingold's delineates, instead of targeting solely result-driven examinations, the ideal usage of such data should facilitate a new mode of computation conveying the open-ended information embedded in bodily practices [Salazar 2018] [Ingold 2011].^[2] In accordance with this ideal, MoCap technologies have gained increasing popularity in ICH archiving. MoCap allows data collection to be neutral, simultaneous, “beyond the frame and within the volume” [Delbridge 2015], thereby fostering a perceivable human presence in the virtual CH environments, as surveyed by Alan Chalmers et al. [Chalmers et al. 2021].

7

2.2 Computational (Embodied) Archives

Digitization of archives allows researchers and practitioners to apply modern computational methods to these collections. A recent review by Colavizza et al. has outlined multiple axes on which AI can improve archival practices by automating recordkeeping processes and improving access to these collections as well as fostering new forms of digital archives [Colavizza et al. 2021]. Additionally, Lev Manovich has put forward a new way to analyse large datasets of cultural items through computational means, a method he has named “cultural analytics” [Manovich 2020]. Other researchers, such as Leonardo Impett, have focused specifically on computer vision models to better understand large visual cultural collections through “distant viewing” [Impett 2020b].

8

From the perspective of experimental museology, computational approaches can help improve the “civic value” of cultural archives [Edmondson 2004]. Through its operationalisation, the archive is augmented on three main levels [Kenderdine, Mason, and Hibberd 2021]. First, it enters the social and immersive dimension of situated museological experiences. Second, archivists engage in new interdisciplinary exchanges with media experts and the computer science community to solve many of the technical challenges cultural institutions face today. Third, new narratives can be imagined by extracting novel features and revealing “hidden structures” in the dataset [Olesen et al. 2016].

9

These hold especially true for audiovisual recordings, one of the primary examples of collections used to capture and document embodied practices. As the foremost mnemonic records of the 21st century, these moving image archives are part of our daily lives. In these last decades, major broadcasting institutions have digitised their entire collections. For instance, in Switzerland, the Radio Télévision Suisse (RTS) has more than 200,000 hours of footage [RTSArchives 2018], while in the United Kingdom, the British Broadcasting Corporation (BBC) preserves more than a million recorded hours [Wright 2017]. In parallel, we are observing many advances in the field of visual AI dedicated to human pose estimation with state-of-the-art models such as OpenPose [Cao et al. 2019], OpenPifPaf [Kreiss, Bertoni, and Alahi 2021], and BlazePose [Bavarevsky et al. 2020]. These AI algorithms can reliably extract human poses and movements from large moving image collections at scale, essentially creating computational embodied archives. The wealth of new data extracted by this process can then be further processed, unlocking new modes of access and novel ways of exploring these archives.

10

2.3 Human Bodies as a Way of Access

In the field of digital humanities, various projects have relied on such approaches to analyse, explore, and better understand visual collections through human bodies. In digital art history, for instance, researchers have undertaken a comprehensive analysis of large art corpora through human gestures [Bernasconi, Cetinić, and Impett 2023] [Impett 2020a]. Similarly, in collaboration with choreographer Wayne McGregor, the Google Arts & Culture Lab has developed the Living Archive, a web-based interface to navigate the collection of postures in McGregor's choreographies and create new movements in the choreographer's style [McGregor and Lab 2019]. Furthermore, movement computing has gained popularity as an intelligent approach to transform the performed cultures into “choreographic objects”, and it is used for analysing, visualising, and interacting with datasets of dance heritage [Doulamis et al. 2017]. Andreas Aristidou et al. developed a Labanotation [Guest 1977] based framework for transforming movements in Cypriot dances to a high-dimensional feature model and constructing a deep-learned *motion signature* for similarity analysis [Aristidou et al.

11

2018] [Aristidou, Shamir, and Chrysanthou 2019]. Improved from *deep signature* encoding [Aristidou, Shamir, and Chrysanthou 2019], Jan Sedmidubsky et al. invented a text-like representation of 3D skeleton sequences and employed the benefits of text-learning models in a more complicated context [Sedmidubsky et al. 2020]. Likewise, Katerina El Raheb et al. combines posture recognition and Benesh movement notation [Benesh and Benesh 1977] to assist with multimedia annotation and interactive learning. In addition to a solely choreographic focus, the Nrityakosha project synthesises a marriage of detection algorithms and semantic models. The researchers related embodied attributes to concepts of Indian classical dances and, correspondingly, created a specialised ontology for describing knowledge in the multimedia archive [Mallik and Chaudhury 2012] [Mallik, Chaudhury, and Ghosh 2011]. By applying the methodology of “distant viewing” [Arnold and Tilton 2019] to embodied knowledge archives, Peter Broadwell and Timothy R. Tangherlini propose a computational analysis of K-pop dance, leveraging human-pose estimation algorithms applied to audiovisual recordings of famous K-pop groups and idols performances [Broadwell and Tangherlini 2021]. Finally, readers will find additional interesting use cases in the comprehensive review provided by Clarisse Bardiot in *Performing Arts and Digital Humanities: From Traces to Data* [Bardiot 2021].

3 Our Computation Framework

3.1 Rationale

The high-level goal of our computational framework is to enrich an embodied knowledge archive by extracting human skeleton data. This wealth of new data, in the form of sets of keypoints, captures both static postures (when looking at individual frames) as well as dynamic motions (when adding the temporal sequences of skeletons from successive frames). By applying motion extraction algorithms, we can therefore operationalise the abstract features of human poses and movements, essentially translating them into vectors that can then be further processed through other computational methods. Such a process augments the archive and unlocks a multitude of new scenarios, examples of which will be discussed through our two use cases. In particular, we investigate visualisations of the whole archive through human poses as well as motion-based retrieval.

12

3.2 Methodology

Based on the relevant literature reviewed above, we propose the following general framework to augment embodied knowledge archives and create new modes of experiencing them. This method applies to both moving image and MoCap archives, with the main difference in the first step: extracting human poses. Indeed, when working with videos, human pose estimation models such as OpenPose [Cao et al. 2019] are required to extract human postures in the form of a “skeleton”, or a list of keypoint edge pairs, as defined in the COCO dataset [Lin et al. 2014]. Depending on the model, these can be in two or three dimensions and handle single or multiple persons at the same time. Furthermore, we stress that these models can sometimes fail to accurately detect full bodies, especially in situations where part of the body is occluded. The number of keypoints detected can also impact how much these models are able to capture features specific to a certain discipline. While MoCap data generally comes with skeleton information, post-processing, normalisation, and format conversion are often necessary to produce clean and operable skeleton datasets.

13

Once skeleton data is extracted, poses need to be normalised so that they can be meaningfully compared. This involves scaling the skeleton data with respect to the image sizes, in the case of video datasets, as well as centring and scaling with respect to pose size. Subsequently, feature vectors can be computed based on this normalised set of keypoints. We note that these features can vary a lot from one application to another, especially if one is working with static poses (as in our first use case) or with motions (as in our second use case).

14

These steps result in a computational embodied archive, in which each element is a human pose, as a set of normalised skeleton keypoints, linked to its corresponding frame or timestamp in the recorded item and potentially augmented with additional computed features. The wealth of data extracted can then be further processed for a variety of scenarios, of which we present two examples. The first use case of this paper, which addresses a collection of dance performance video recordings, will explore how the whole dataset can be mapped in two dimensions through the poses dancers take in their performances. Afterwards, we present a second scenario on a dataset of martial arts

15

performances, extending to motion-based similarity retrieval.

4 Use Case 1: Accessing an Audiovisual Dance Performance Archive via Human Poses

4.1 The Prix de Lausanne Audiovisual Archive

The Prix de Lausanne is a competition for young dancers held yearly in Lausanne, Switzerland since 1973^[3]. Nearly each year of the competition has been recorded and digitised, resulting in a rich dataset of 1,445 mp4 videos of individual dance performances across forty-two years, along with information on the dancers and their chosen performances. This metadata, although valuable, does not document the embodied elements of dance performances, limiting how one can access such an archive. With our solution, we augmented the available data using computational methods and employed the extracted features to map the whole dataset in two dimensions through the human poses embedded in the material, revealing its hidden structures and creating new connections between similar poses.

16

4.2 Building Feature Vectors for the Dancers' Poses

We extracted human skeleton data using the MediaPipe BlazePose GHUM 3D model [Bavarevsky et al. 2020], which was mainly chosen for its out-of-the-box robustness and rapidity of inference for single human estimation. In the case of the Prix de Lausanne, since the videos record one participant at a time on stage, we found this model to be the most suitable. We also observed only occasional cases of occlusions when the dancers are wearing particularly large costumes. The BlazePose algorithm outputs skeleton data as a set of 33 keypoints in three-dimensional space, normalised to the image size. Compared to models trained on the COCO dataset, MediaPipe algorithm detects more keypoints, including the index fingers and index toes of the two hands and the two feet, allowing us to compute the angles of the wrists and ankles. Although the algorithm was not specifically trained on dance recordings, we believe these additional features help to better capture specific poses of dance performances, where even the positions of the hands and feet are crucial. Videos were processed to extract one pose every five seconds, a cautious design decision to avoid over-collecting similar poses merely due to temporal closeness. We do note however this is potentially not the best solution to ensure a full coverage of the diversity of poses, and future research in this direction could improve the results presented even further. It is worth noting that not all frames necessarily had a detected pose, since in some cases the dancer was too far away from the camera, or because some videos were not properly cut to the exact dance performance and thus had irrelevant frames at the beginning and/or the end. We measured the rate of detection by taking the mean of the ratio of poses detected over the number of frames per video, obtaining a mean detection rate of 76.80% across all the videos. The pose extraction process resulted in a total of 27,672 poses.

17

Once these poses had been extracted, we normalised them following a two-step procedure. We first subtracted the pose centre, defined as the mean point between the left and right hips. We then scaled the pose by its size, defined as the maximum between the torso size and the maximum distance between the pose centre and any keypoint, where the torso size was computed as the distance between the centre of the shoulders and the pose centre (the centre of the hips). The procedure ensures meaningful comparison between the poses by normalising them. An intuitive sample of normalised poses with their corresponding frames is provided in Figure 1.

18

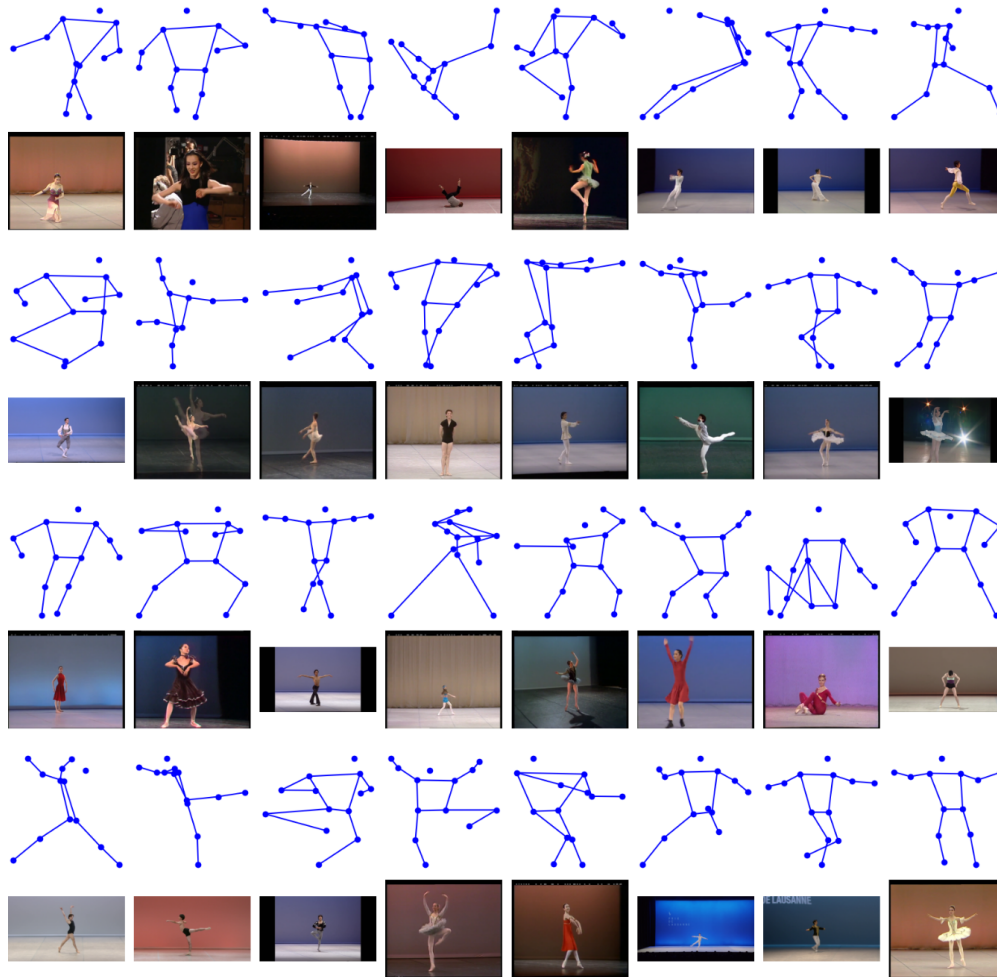


Figure 1. Sample of normalised poses with corresponding frames. Only the main keypoints are displayed for ease of visualization (the nose, the shoulders, the elbows, the wrists, the hips, the knees, and the ankles).

Subsequently, two feature vectors were built and tested. For the first, we simply took the vector of the keypoints and flattened it, resulting in a 99-dimensional vector. For the second, we built a custom feature vector, taking a combination of pose lengths (distances between two keypoints) and pose angles (angles defined by three keypoints). Table 1 shows the features computed.

Category	Feature
Pose Lengths (Distances Between Joints)	Left elbow — Right elbow Left wrist — Right wrist Left knee — Right knee Left ankle — Right ankle Left hip — Left wrist Right hip — Right wrist Left hip — Left ankle Right hip — Right ankle Left ankle — Left wrist Right ankle — Right wrist Left shoulder — Right ankle Right shoulder — Left ankle Left wrist — Right ankle Right wrist — Left ankle
Pose Angles (Defined by Three keypoints)	Left hip — Left knee — Left ankle Right hip — Right knee — Right ankle Left shoulder — Left elbow — Left wrist Right shoulder — Right elbow — Right wrist Left hip — Left shoulder — Left elbow Right hip — Right shoulder — Right elbow Left shoulder — Left hip — Left knee Right shoulder — Right hip — Right knee Left elbow — Left wrist — Left index finger Right elbow — Right wrist — Right index finger Left knee — Left ankle — Left index toe Right knee — Right ankle — Right index toe

Table 1. A list of custom features used to represent dance poses in the Prix de Lausanne archive.

4.3 Mapping the Prix de Lausanne Archive

One of the fundamental applications unlocked by computational archives is the possibility to visualise whole collections. To this end, we apply dimensionality reduction (DR) algorithms to embed the high-dimensional feature spaces in two or three dimensions, in order to create a novel way of visualising and navigating the whole archive. We tested both standard DR methods, such as Principal Component Analysis (PCA), and t-distributed Stochastic Neighbour Embedding (tSNE), as well as a more recent approach, Uniform Manifold Approximation and Projection (UMAP). We leveraged scikit-learn implementations for the two first algorithms and the official UMAP Python implementation for the latter [Pedregosa et al. 2011] [Halko, Martinsson, and Tropp 2011] [Martinsson, Rokhlin, and Tvgert 2011] [van der Maaten and Hinton 2008] [McInnes et al. 2018]. For tSNE and UMAP we compared using cosine and Euclidean metrics as the distance measure to compute the projection. We therefore tested five options on 27,672 poses, with both sets of feature vectors. Figure 2 shows the resulting embeddings in two dimensions.

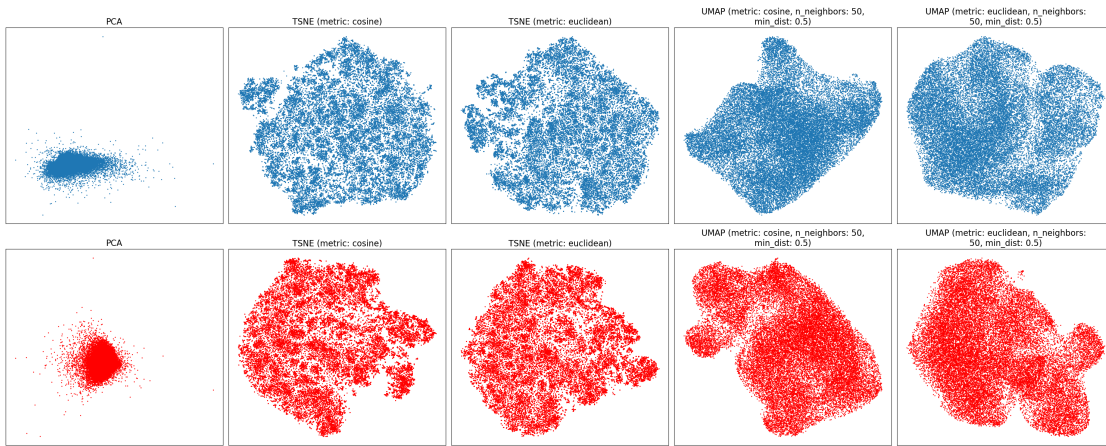


Figure 2. Embeddings of 27,672 poses in two dimensions. The top row (blue) uses the flattened vectors of keypoints, while the bottom row (red) uses the custom features computed. From left to right, the algorithms used are PCA, tSNE (cosine metric), tSNE (Euclidean metric), UMAP (cosine metric), and UMAP (Euclidean metric).

We can immediately observe that PCA performs poorly on this dataset, as expected due to its linear nature that cannot preserve much variance with just two dimensions (only 49% with the flattened vectors of keypoints and 45% with the custom features). Both tSNE and UMAP, however, give much more interesting results, where structures in the dataset can be observed. Since we did not observe meaningful differences between the two sets of feature vectors used, we pursued further examinations with the flattened vectors of keypoints to exploit the heuristic quality of the algorithms independent of the choice of features. Therefore, we set out to investigate tSNE and UMAP behaviours in more depth by studying the effects of their hyperparameters. For tSNE, we observed how the embedding in two dimensions behaves for different perplexities (ϵ), the main parameter for this algorithm, related to the number of neighbours considered for any given point during the computation. For UMAP, we investigated the number of nearest neighbours (n), controlling how UMAP balances local versus global structure in the data, as well as the minimum distance (d_{min}) the algorithm is allowed to pack points together in the low-dimensional space. We used the Euclidean distance as it seemed to produce slightly clearer results. Figure 3 shows how tSNE behaves for different values of ϵ , while Figure 4 displays the results of UMAP with varying n and d_{min} values.

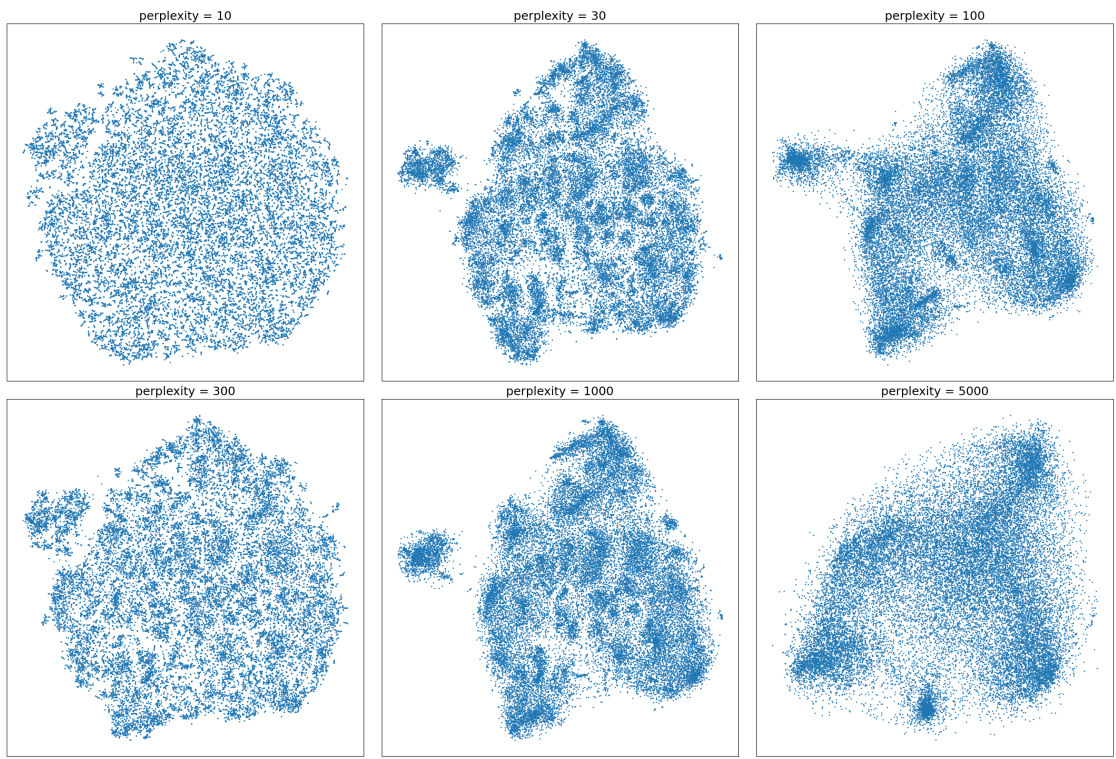


Figure 3. Effect of the perplexity (ϵ) on the tSNE embeddings of 27,672 poses (flattened vectors of keypoints, Euclidean distance). Lower values of ϵ result in a more granular embedding.

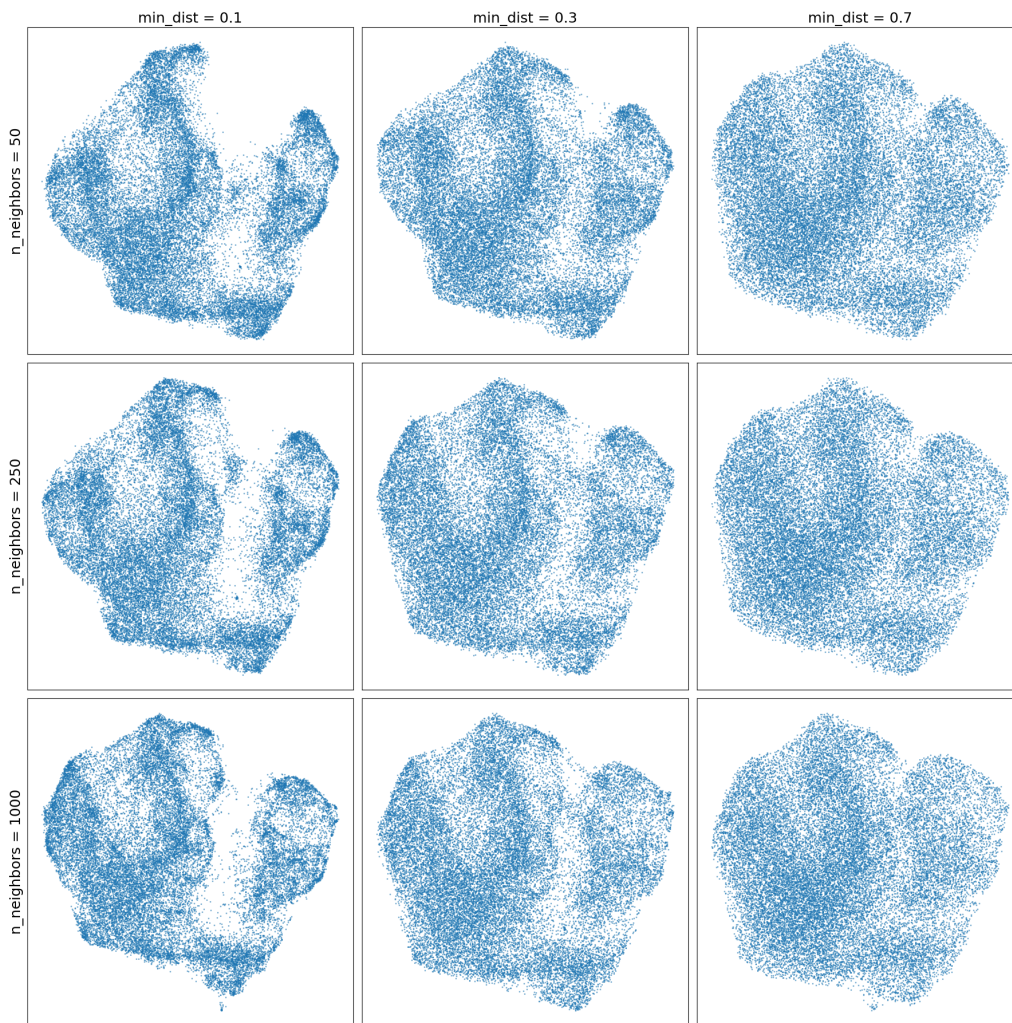


Figure 4. Effect of the number of neighbours n and the minimum distance d_{min} on the UMAP embeddings of 27,672 poses (flattened vectors of keypoints, Euclidean distance). Lower values of d_{min} result in a more packed embedding, but global structures appear to be quite stable.

To conclude our exploration of the Prix de Lausanne moving image collection, we set out to produce a map of the human poses embedded in the archive. Following the outlined procedure, we extracted the keypoints, computed the corresponding feature vectors, and applied UMAP (with Euclidean distance, $n = 50$, $d_{min} = 0.5$) on the flattened vectors of keypoints to create Figure 5, where we display the corresponding pose at each point. For the sake of visibility, we only show a portion of all the poses extracted, resulting in a map of 2,768 poses. Through this visualisation, we can verify the effectiveness of our approach in grouping together similar poses, thus unlocking a new way of seeing an embodied knowledge archive.

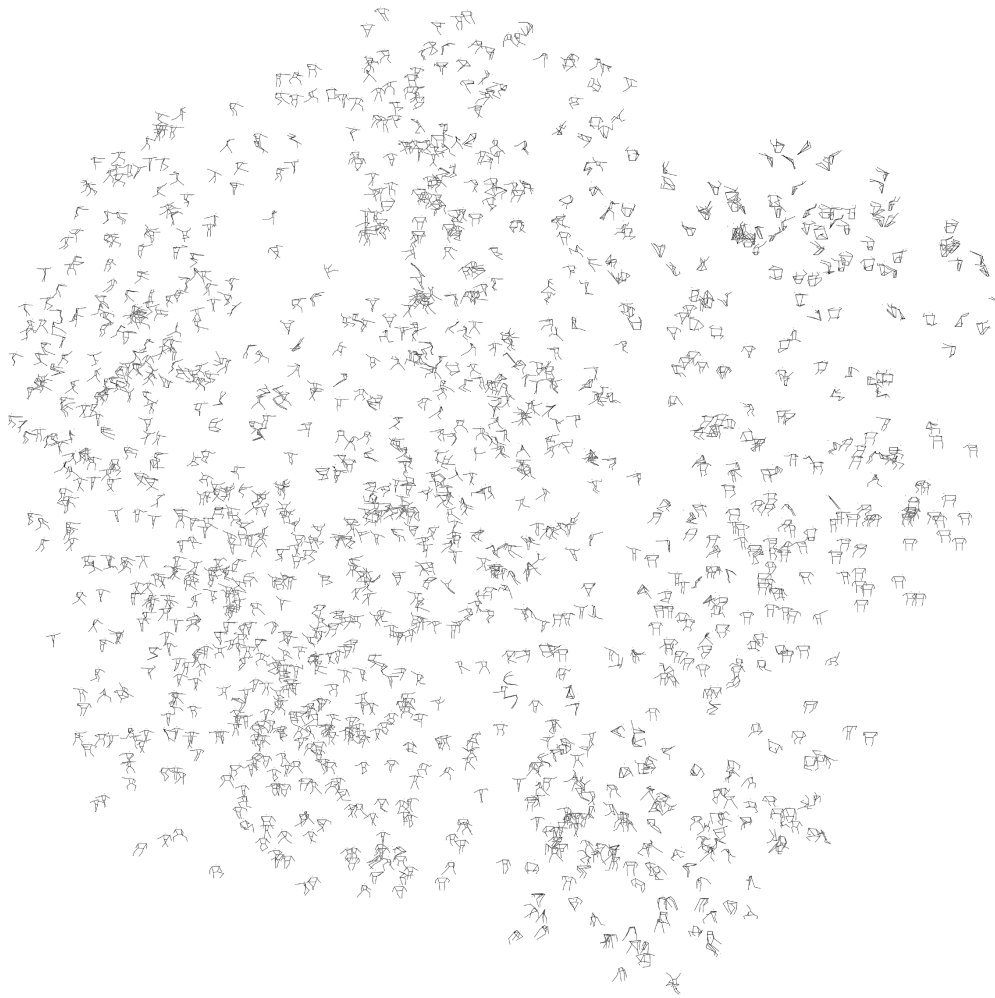


Figure 5. Map of 2,768 dancers' poses, based on the 2D embeddings of the vectors of flattened keypoints with UMAP (Euclidean distance, $n = 50$, $d_{min} = 0.5$).

5 Use Case 2: Accessing Martial Art Recordings through Motion Traits

In the first use case, the extracted human skeletons were been treated as static postures, thus operating on individual frames of the moving image recordings. This is suitable for the purpose of visualising the whole archive in two dimensions because we can draw these skeletons, using them as glyphs representing frames in the collection to create compelling images, as in Figure 5. However, one might argue that embodied knowledge archives are not only about human postures but also, if not more so, about human motions. This second use case thus investigates such an approach, proposing an innovative motion-based retrieval mode. In the context of the Hong Kong Martial Arts Living Archive (HKMALA) project, a retrieval system allows users to select a specific motion within the collection (by choosing the start and end frames of a video sequence, for instance) and returns similar motions in the archive.

23

5.1 Hong Kong Martial Arts Living Archive

Traditional martial arts are considered knowledge treasures of humanities sustained through generations by diverse ethnic groups. Among the genres, Southern Chinese martial arts (SCMA) is arguably one of the most prolonged systems embodying Chinese mind-body ideologies, yet it is now facing challenges in knowledge transmission and the risk of being lost. In preserving the living heritage of SCMA, the Hong Kong Martial Arts Living Archive inspects a comprehensive set of digitisation tools to capture the martial practices, with a chief focus on motion capturing form sequence performances, or *taolu*. Since its origination in 2012, HKMALA has built a 4D motion archive spanning over 20 styles and 130 sets of empty-hand and weapon sequences.^[4] The archive also collates various historical and

24

reconstituted materials documenting martial cultures, encompassing rituals, traditions, armaments, and objects preserved in Hong Kong [Chao et al. 2018].

The HKMALA's exceptional volumes hold the promise to enable various scholarly inquiries. However, its current archival organisation is based on a manual catalogue that merely indicates how the content has been curated for exhibitions. Additionally, the data consists of multiple modalities, such as texts, images, videos, and MoCap data, thereby impeding public access and dissemination at scale. There is a clear need for an effective yet meaningful way of data access.

Addressing the aforementioned challenges, we propose devising a motion-based retrieval framework that leverages machine learning to encode motion-wise information in multimodal recordings. The approach facilitates content retrieval through embodied cues, operating seamlessly across MoCap and moving image data formats. In the following paragraphs, we aim to provide an overview of the technical procedures. For a more comprehensive understanding of the implementation process, readers are invited to refer to the descriptions in [Hou, Seydou, and Kenderdine 2023].

5.2 Encoding and Retrieving Martial Arts Movements

Chinese martial arts are renowned for their strong emphasis on form training, which utilizes codified movement sequences to practise fundamental techniques and technique combinations. Less known, yet equally important, is the stress on combative tactics that involve mindful operation of whole-body movement. Thus, training is imparted through a combination of freeform combat and integration into *taolu* practices [Ma 2003]. Given that assessment methods in Chinese martial arts typically consider visually discernible characteristics and mechanical parameters, we applied the same configurations to model martial arts movements. Designed in compatibility with Laban metrics for movement analysis [Guest 1977], the approach aims to articulate both qualitative and quantitative qualities regarding the technical stances, ballistic activities, and reactions to the opponent or surroundings, as shown in Table 1 in [Hou, Seydou, and Kenderdine 2023].

The process began with the extraction of 3D human skeletons from MoCap data, or the extraction of 2D poses from video recordings, followed by depth estimation to reconstruct the 3D structure. The coordinates and joint angles were then computed from raw skeletons into basic kinematic metrics, which were normalized for further processing. Feature augmentation was employed to enrich the kinematic feature set, incorporating metrics that reflect linear and angular motion characteristics, along with the transition probabilities of motion time series. Finally, the top 50 features with the highest variance were selected using the variance threshold approach and used to train a discriminative neural network.

Holding a hypothesis that the characteristics of a long motion sequence can be represented by its segmentations, our encoding method was aimed to pinpoint representative motion units within the sequence. Firstly, we gathered high-dimensional features from equally segmented media data and employed deep learning to train a latent space, embedding them into a vector space where semantically similar dimensions are clustered together while dissimilar ones are set apart. Subsequently, K-means clustering was applied to identify cluster centres as representatives. To create a descriptive and compact scheme for efficient retrieval, we encoded the inverse document frequency (IDF) of each representative into a normalised histogram, or *metamotion*, to represent the distribution of archetypal qualities within a motion sequence.

5.3 A Motion-Based Retrieval Engine

To facilitate the envisioned scenario of exploratory archival browsing, a retrieval system was deployed to allow querying archival sequences containing similar movements. The system integrates different retrieval algorithms, including locality-sensitive hashing (LSH) and Ball-tree methods, which prove effective in achieving optimal retrieval efficiency, outperforming the baseline approach in existing research.^[5]

On the user end, an interactive search engine was deployed to enable motion-based data exploration across formats. An intuitive search case is illustrated in Figure 6, showcasing a cross-model retrieval result from a hybrid of videos and MoCap data. Additionally, Figure 7 demonstrates the initial development of the archival browser, presenting the retrieved items with ontological descriptions for the conceptual entities associated with the query item. This design aims

to improve data explainability and interoperability for future reuse, with ontology annotations sourced from the scholarly work of The Martial Art Ontology (MAon).^[6]

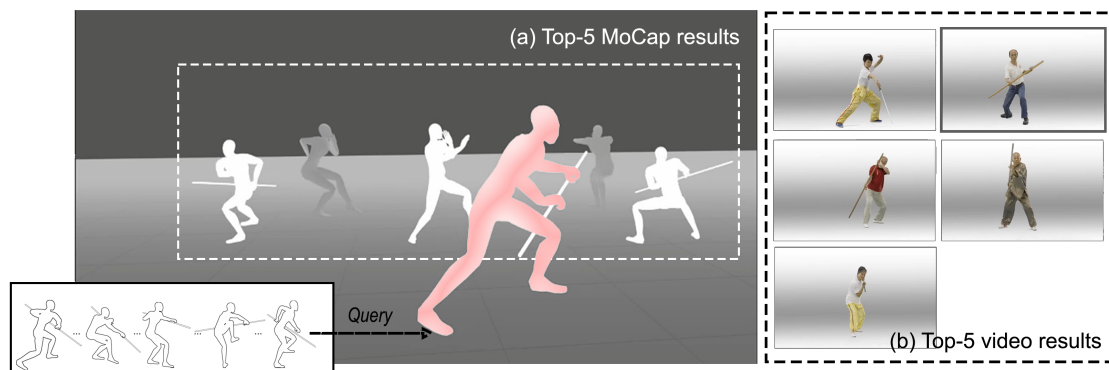


Figure 6. An example of the top-5 similar sequences retrieved by a query using a MoCap dataset (left) and a video dataset (right) [Hou, Seydou, and Kenderine 2023].

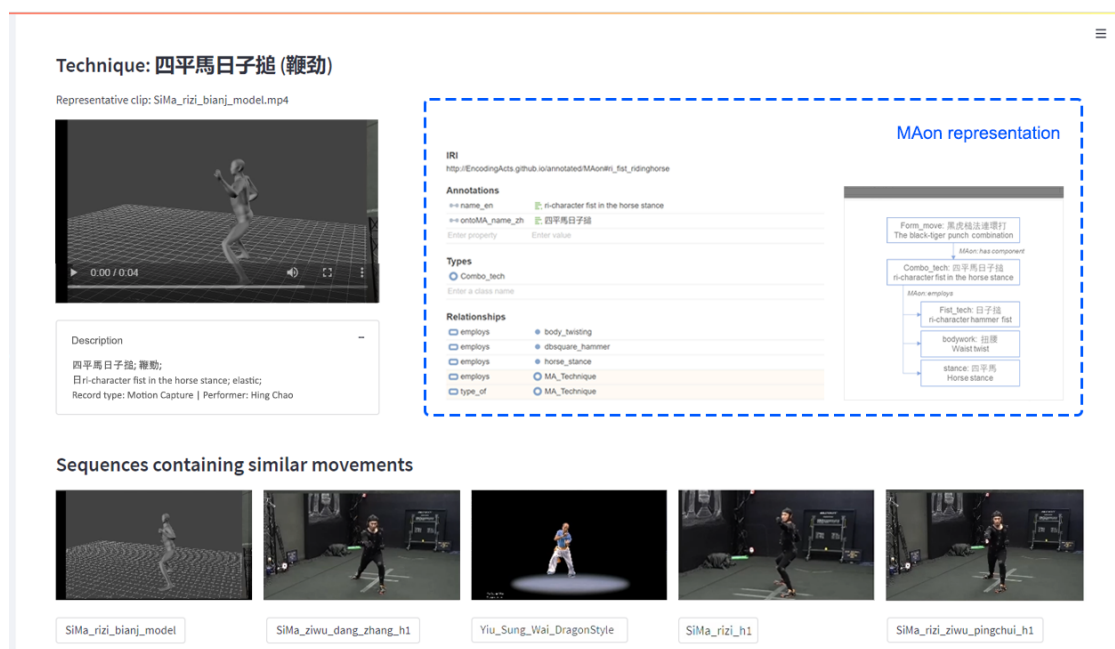


Figure 7. Illustration of a motion search example supplemented with ontological representations of the concepts of techniques.

6 Discussion

6.1 Visual AI Toward “Augmented Archives”

At cultural heritage archives, including born-digital archives, metadata is commonly organised according to an archive-specific curatorial structure with topical domain-language descriptions of archival items. For instance, both the Prix de Lausanne and HKMALA collections warrant a series of textual tags describing the contents of dance and martial art performances, as well as information about performers. Such documentation, although useful, entails a mode of interpretation and access rather limited to expert users. One needs to know what dance performances such as *The Giselle* or *The Nutcracker* are to look for them in the Prix de Lausanne archive, and one must understand that martial arts are comprised of a series of *taolu* in order to understand HKMALA collections. Furthermore, the visual and embodied features represented in these collections are difficult to capture verbally. Models that describe movement, such as those based on notation systems like Labanotation, have shown promise. However, such models typically

necessitate the alignment between the movement and the notation language, as well as an advanced understanding of the subject matter by the annotator, in order to effectively encode and decode the knowledge behind movements. As a result, it is very difficult, if not impossible, to capture embodied knowledge features using literal formats or in an easily accessible way. Therefore, this kind of metadata, although necessary and useful in many domains, is not suitable for access by the general public. It does not support the “casual modes of access” that are fundamental for museum-goers and general audiences [Whitelaw 2015].

In this article, we have introduced a computational framework aimed at enhancing cultural archives by enriching metadata with embodied knowledge descriptors derived from audiovisual and multimodal materials. Utilising computer vision models, we extract and transform human poses and skeleton features into meaningful data features, which potentially foster a range of usage scenarios. At the archival level, curators can operationalise their collections as “data sources” [Kenderdine, Mason, and Hibberd 2021], envisioning new ways for querying, annotating, and analysing whole datasets to forge novel narrative paradigms. In parallel, the enriched metadata improves data accessibility while the diversified query channels improve data findability, augmenting the archives’ compliance with FAIR (findability, accessibility, interoperability, and reusability) data principles, regardless of variations in language, topic, and data format. Furthermore, our approach demonstrates the potential for interpreting data traits within and across cultural contexts. This not only contributes to the reusability of archival content but also resonates with the findings of Colavizza et al., which highlight the growing role of AI in archival practices, particularly in the “automation of recordkeeping processes” and the improvement of procedures for “organising and accessing archives” [Colavizza et al. 2021]. Our experimentation, as exemplified through two distinctive heritage archives, represents an innovative step forward with a focus on embodied knowledge, allowing queries through more perceptive than quantitative channels, which we believe is more natural and humanistic.

33

6.2 Archive Navigation and Serendipitous Discovery

The operationalisation of embodied knowledge archives through visual AI, resulting in “augmented archives”, provides a vast array of data that can further be processed in order to create new modes of access adapted for general audiences. Following the principles of “generous interfaces” [Whitelaw 2015], our method supports explorative behaviours, encouraging a new paradigm of “seeking the unknown” [Winters and Prescott 2019]. By laying out the full archive through the lens of easily understood concepts, such as dancers’ postures, users can more readily comprehend the archive without specific knowledge about the topic. Users do not need to have a specific goal in mind or be looking for something in particular. Instead, they can wander around, browsing like an “information flaneur” and enjoying the Prix de Lausanne archive in a way that traditional modes of access, based on querying metadata or skimming grids and lists of information, could not offer [Dörk, Carpendale, and Williamson 2011]. Furthermore, each pose links to a timestamp in a dance performance, grouping together similar poses in the low-dimensional space to create serendipitous discoveries. Indeed, this mode of access rewards users for simply browsing the collection and stumbling upon new performances as they move from pose to pose on the map. Figure 8 explicates this process by showcasing poses similar to an input pose with the corresponding video frames.

34

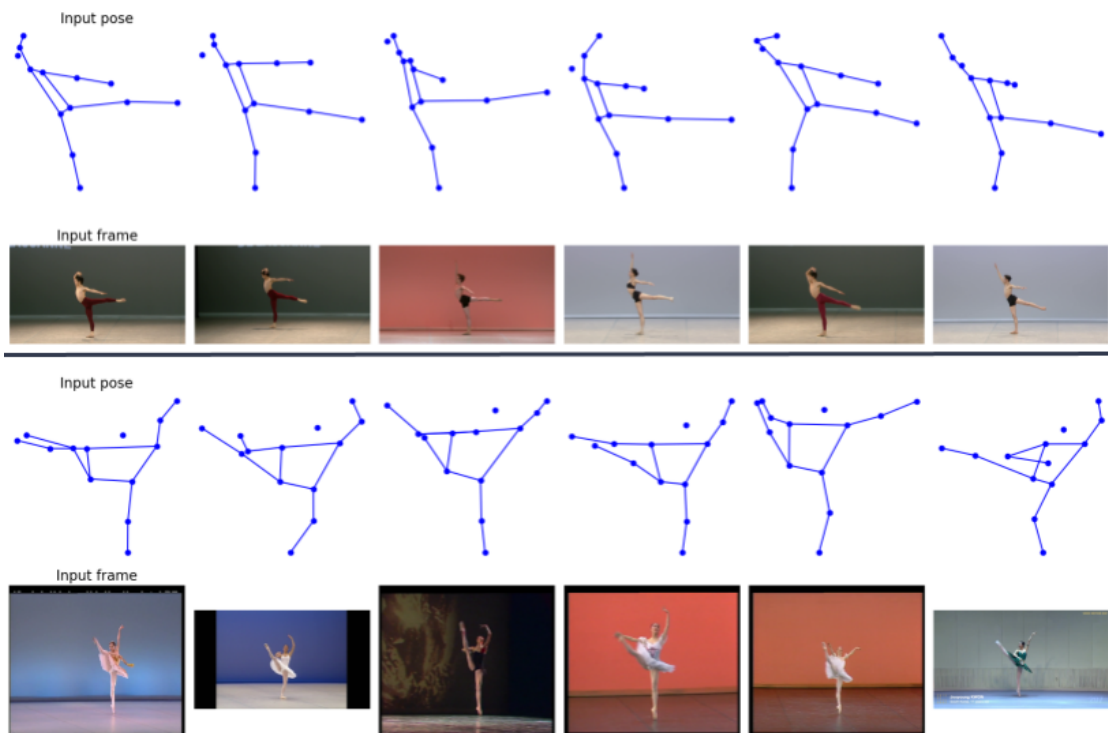


Figure 8. Examples of similar poses in the Prix de Lausanne archive (top 5 matches). Simplified poses and corresponding video frames are displayed. Notice how connections between different performers are discovered.

This new mode of access is enhanced by dimensionality reduction algorithms. However, one must take care in deciding which algorithm to employ and with what parameters. In this work, we have analysed Principal Component Analysis (PCA), t-distributed Stochastic Neighbour Embedding (tSNE), and Uniform Manifold Approximation and Projection (UMAP). PCA is a well-established dimensionality reduction technique, but, due to its linear nature, it often fails to properly capture a real-life dataset with only a handful of dimensions. This is clearly confirmed in Figure 2, where all embeddings generated with tSNE and UMAP show much more interesting structures without creating a large clump of points like PCA does. Regarding the choice of feature vectors used, as mentioned, we did not observe noticeable differences between the flattened vectors of keypoints and the custom features computed. We hypothesise that the custom features (reported in Table 1) are not more discriminative than the base keypoints. Although in subsequent analyses we decided to use the flattened vectors of keypoints, we believe this offers us an opportunity for more collaboration, since involving dance experts could help us craft a more precise and adequate set of features, based on Labanotation, for instance [Guest 1977]. To gain a deeper understanding of these algorithms, we have investigated the effect of parameters on tSNE and UMAP.

35

Figure 3 shows that increasing the perplexity yields very different layouts, with lower values displaying numerous small groupings of items while higher values reveal more large-scale structures. This is in accordance with expectations, since with higher perplexities tSNE considers more points when computing the vicinity of each item, thus better capturing global structures. Surprisingly, in Figure 4, when increasing the number of neighbours (n) considered with UMAP, the outputs appear to be much more stable, with global structures already visible with $n = 50$. It is in this case that the second parameter, d_{min} , affects more the results, yielding more sprayed-out mappings the higher it is (since the minimum distance directly controls how packed the points can be in the low dimensional space). These results indicate that when creating archive visualisations, Like the map of poses in Figure 5, for instance, higher d_{min} might be more suitable to avoid overlaps between similar poses. However, if one were to apply clustering algorithms on these embeddings to group together similar poses, the more fine-grained and packed structures obtained with lower d_{min} would potentially yield better results. Therefore, we argue that it is not a matter of which embedding is better overall but rather which embedding is better adapted to the specific narrative or mode of access sought.

36

6.3 Limitations and Future Work Directions

Although we believe our computational framework to be well thought out, there are still some limitations we would like to highlight, hopefully to later address them.

37

First, human pose extraction from monocular videos is never perfect. Human bodies can be incorrectly estimated or missed due to the influence of occlusion, monotonous colour patterns, or camera angles, to name a few. Indeed, through our “naive” approach to the Prix de Lausanne archive, we only achieved a detection rate of 76.80%.^[7] Furthermore, by checking a sample of skeletons extracted with the corresponding frames, as in Figure 8, we noticed that keypoints were not always correctly extracted. One possible reason is that pose estimation was done systematically every five seconds without a measure of quality or significance. Nevertheless, our approach proves sufficient for the use case described, as it still produces enough data to map the whole archive properly. Further developments could yield more interesting and precise results, for instance, by extracting skeleton data on a finer temporal granularity and then filtering only to keep the “better” poses.^[8]

38

Second, the feature vectors computed for the Prix de Lausanne archive are somewhat naive. Taking the lesson from the HKMALA feature computing, collaboration with dance experts could facilitate the design of more relevant, dance-specific features with precise measurements able to better capture and compare human bodies during dance performances. Nonetheless, our results demonstrate that even naive methods can produce new modes of access to embodied knowledge archives. Thus, we are confident that our method can be generalised for use by other archives and diverse embodied disciplines.

39

Lastly, unlike standard computer vision challenges, it is difficult but necessary to quantify what makes a “good” estimation of poses or movement segments in the archival context, yet the standard varies across people and cultural themes. To this end, we resorted to evaluating the distribution and quality of embeddings, supplemented with an ad-hoc expert review of small sampling sets. Future improvement is suggested to integrate expert review dynamically along with the model training process, such as by intaking expert judgement as a score and feeding it back to the model, so as to enable a human-in-the-loop machine-learning process.

40

6.4 Outlook: Towards a New Archival Experience

The research presented in this paper was conducted within the larger context of “computational museology” [Kenderdine, Mason, and Hibberd 2021], a discipline aimed at developing new modes of access to cultural heritage through computational approaches, specifically designed for situated experiences in museum settings. To this end, further work will rely on leveraging the findings highlighted in this paper to create interactive installations to access these embodied knowledge archives. We contend that the implementations presented in this paper establish a robust foundation for developing interactive modes of access tailored to meet the requirements of casual users within situated experiences. In particular, two main directions will be pursued. First, based on the fundamental concept of placing visitors “inside” the archive [Shen et al. 2019], rather than looking at it on a plain screen, immersive environments (IEs) will be employed to create situated experiences in which users can navigate the whole archive. In practice, dimensionality reduction techniques will be employed to compute a mapping in two- or three-dimensions in order to generate a virtual world in which each point or glyph represents a pose (and its corresponding frame or clip). This virtual cloud of human bodies will then be displayed in large IEs, allowing visitors to freely navigate and discover the archive. Applying immersive technologies to the mediation of cultural heritage already has some interesting applications, particularly in the context of archaeological enquiry [Sciuto et al. 2023], and we believe such technologies can also serve a purpose for embodied knowledge archives.

41

Second, an interactive retrieval system can be developed, either based on pose or motion similarity. Users could strike a pose or perform a certain movement, detected in real-time with motion-tracking solutions, and the system would retrieve relevant clips from the collection. Such an experience would yield an interesting performative aspect that sees the visitor as a co-creator with the machine, essentially transforming them into a variable of the generative system, and other people around them as an audience witnessing the exchange in a “third-person’s perspective” [Mul and Masson

42

2018].

7 Conclusion

Embodied knowledge archives are an integral part of our living heritage, and they contain important aspects of our cultures. Yet, it is still difficult to explore embodied knowledge archives, especially for more casual audiences that lack specific knowledge on the topic of the collection. To answer this challenge, we have proposed in this work a computational framework that leverages motion extraction AI to augment these datasets with a wealth of rich data, enabling new modes of analysis and access.

43

The proposed method was applied to two embodied knowledge archives, containing dance and martial arts performances, respectively, which showcases its application to multimedia content and diverse access scenarios. In the former example, we devised a method to visualise a whole collection in two dimensions through the human poses embedded in the archival materials, revealing their structure and fostering serendipitous discoveries. In the latter, we extended our method to motion encoding from static poses and devised a motion-based query system, offering a new way to search an embodied knowledge archive. These scenarios showcase how our computational framework can operationalise this type of collection and unlock a variety of new modes of access suitable for non-expert audiences.

44

Acknowledgements

The authors are grateful to the Prix de Lausanne for the opportunity to work on their audiovisual archive, as part of the SNSF's Sinergia grant, *Narratives from the Long Tail: Transforming Access to Audiovisual Archives* (CRSII5_198632).

45

The Hong Kong Martial Arts Living Archive is a longitudinal research collaboration between the International Guoshu Association, the City University of Hong Kong, and the Laboratory for Experimental Museology (eM+), EPFL.

46

Notes

[1] By *transformative potential*, we mean the potential of AI to augment the data available by extracting new features that can open innovative venues for accessing the archive.

[2] Tim Ingold argues that some movement is “automatic and rhythmically responsive” to its surroundings and “along multiple pathways of sensory participation” [Ingold 2011].

[3] See <https://www.prixdelausanne.org>

[4] The concept of four-dimensional (4D) space denotes a dynamic 3D space moving through time.

[5] See [Hou, Seydou, and Kenderdine 2023, section 4].

[6] See The Martial Art Ontology (MAon), v1.1, <https://purl.org/maont/techCorpus>.

[7] A “naive” approach in pattern recognition implies a straightforward and easy-to-implement algorithm that finds all matching occurrences of a given input.

[8] One would first need to define what constitutes a “better” pose, however.

Works Cited

Aristidou et al. 2018 Aristidou, A. (2018) “Deep motifs and motion signatures”, *Transactions on Graphics*, 37(6), pp. 1–13.

Aristidou, Shamir, and Chrysanthou 2019 Aristidou, A., Shamir, A., and Chrysanthou, Y. (2019) “Digital dance ethnography: Organizing large dance collections”, *Journal on Computing and Cultural Heritage*, 12(4), pp. 1–27.

Arnold and Tilton 2019 Arnold, T. and Tilton, L. (2019) “Distant viewing: Analyzing large visual corpora”, *Digital Scholarship in the Humanities*, 34(Supplement 1), pp. i3-i16.

Aske and Giardinetti 2023 Aske, K. and Giardinetti, M. (2023) “(Mis)matching metadata: Improving accessibility in digital

- visual archives through the EyCon project”, *Journal on Computing and Cultural Heritage*, 16(4).
- Bardiot 2021** Bardiot, C. (2021) *Performing arts and digital humanities: From traces to data*. Hoboken, NJ: John Wiley & Sons.
- Bavarevsky et al. 2020** Bazarevsky, V. et al. (2020) “Blazepose: On-device real-time body pose tracking”, *arXiv*. <https://arxiv.org/abs/2006.10204>.
- Benesh and Benesh 1977** Benesh, R. and Benesh, J. (1977) *Reading dance: The birth of choreology*. London: Souvenir Press.
- Bernasconi, Cetinić, and Impett 2023** Bernasconi, V., Cetinić, E., and Impett, L. (2023) “A computational approach to hand pose recognition in early modern paintings”, *Journal of Imaging*, 9(6).
- Broadwell and Tangherlini 2021** Broadwell, P. and Tangherlini, T.R. (2021) “Comparative K-pop choreography analysis through deep-learning pose estimation across a large video corpus”, *DHQ: Digital Humanities Quarterly*, 15(1). Available at: <https://digitalhumanities.org/dhq/vol/15/1/000506/000506.html>
- Cameron, Franks, and Hamidzadeh 2023** Cameron, S., Franks, P., and Hamidzadeh, B. (2023) “Positioning paradata: A conceptual frame for AI processual documentation in archives and recordkeeping contexts”, *Journal on Computing and Cultural Heritage*, 16(4).
- Cao et al. 2019** Cao, Z. et al. (2019) “Openpose: Realtime multi-person 2D pose estimation using part affinity fields”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Available at: <https://arxiv.org/pdf/1812.08008.pdf>.
- Chalmers et al. 2021** Chalmers, A. et al. (2021) “Realistic humans in virtual cultural heritage”, *Proceedings of RISE IMET 2021*. Nicosia, Cyprus, 2-4 June 2021. New York: Springer, pp. 156-165.
- Chao et al. 2018** Chao, H. et al. (2018) “Kapturing kung fu: Future proofing the Hong Kong martial arts living archive”, in Whatley, S., Cisneros, R.K., and Sabiescu, A. (eds.) *Digital echoes: Spaces for intangible and performance-based cultural heritage*. New York: Springer, pp. 249-264.
- Colavizza et al. 2021** Colavizza, G. (2021) “Archives and AI: An overview of current debates and future perspectives”, *Journal on Computing and Cultural Heritage*, 15(1).
- Craig et al. 2018** Craig, C.J. (2018) “The embodied nature of narrative knowledge: A cross-study analysis of embodied knowledge in teaching, learning, and life”, *Teaching and Teacher Education*, 71, pp. 329-240.
- Delbridge 2015** Delbridge, M. (2015) *Motion capture in performance: An introduction*. New York: Springer.
- Dimitropoulos et al. 2014** Dimitropoulos, K. et al. (2014) “Capturing the intangible an introduction to the i-Treasures project”, *Proceedings of the 9th international conference on computer vision theory and applications*, vol. 3. Lisbon, Portugal, 5-8 January 2014. Setúbal, Portugal: SCITEPRESS, pp. 773-781. Available at: <https://ieeexplore.ieee.org/abstract/document/7295018>.
- Doulamis et al. 2017** Doulamis, A.D. (2017) “Transforming intangible folkloric performing arts into tangible choreographic digital objects: The terpsichore approach”, *Proceedings of the 12th international joint conference on computer vision, imaging and computer graphics, theory, and applications*. Porto, Portugal, 27 February-1 March 2017. Setúbal, Portugal: SCITEPRESS, pp. 451-460. <https://doi.org/10.3030/691218>.
- Dörk, Carpendale, and Williamson 2011** Dörk, M., Carpendale, S., and Williamson, C. (2011) “The information flaneur: A fresh look at information seeking”, *Proceedings of the SIGCHI conference on human factors in computing systems*. Vancouver, BC, Canada, 7-12 May 2011. New York: ACM, pp. 1215-1224. <https://doi.org/10.1145/1978942.1979124>.
- Edmondson 2004** Edmondson, R. (2004) *Audiovisual archiving: Philosophy and principles*. Paris: United Nations Educational, Scientific and Cultural Organization.
- El Raheb et al. 2018** El Raheb, K. (2018) “A web-based system for annotation of dance multimodal recordings by dance practitioners and experts”, *Proceedings of the 5th international conference on movement and computing*. Genoa, Italy, 28-30 June 2018. <https://doi.org/10.1145/3212721.3212722>.
- Fossati et al. 2012** Fossati, G. et al. (2012) “Found footage filmmaking, film archiving and new participatory platforms”, in Guldmond, J., Bloemheugel, M., and Fossati, G. (eds.) *Found footage: Cinema exposed*. Amsterdam: Amsterdam University Press, pp. 177-184.
- Grammalidis and Dimitropoulos 2015** Grammalidis, N. and Dimitropoulos, K. (2015) “Intangible treasures: Capturing the intangible cultural heritage and learning the rare know-how of living human treasures”, *Proceedings of the 2015 digital heritage international congress*, Granada, Spain, 28 September-2 October 2015. Available at:

<https://diglib.eg.org/handle/10.2312/14465>.

- Griffin, Wennerström, and Foka 2023** Griffin, G., Wennerström, E., and Foka, A. (2023) "AI and Swedish heritage organisations: Challenges and opportunities", *AI & Society*. <https://doi.org/10.1007/s00146-023-01689-y>.
- Guest 1977** Guest, A.H. (1977) *Labanotation: Or, kinetography laban: The system of analyzing and recording movement*. New York: Taylor & Francis.
- Halko, Martinsson, and Tropp 2011** Halko, N., Martinsson, P.G. and Tropp, J.A. (2011), "Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions", *SIAM Review*, 53(2), pp. 217–288. <https://doi.org/10.1137/090771806>.
- Hou et al. 2022** Hou, Y. et al. (2022) "Digitizing intangible cultural heritage embodied: State of the art", *Journal on Computing and Cultural Heritage*, 15(3), pp. 1–20.
- Hou, Seydou, and Kenderdine 2023** Hou, Y., Seydou, F.M., and Kenderdine S. (2023) "Unlocking a multimodal archive of southern chinese martial arts through embodied cues", *Journal of Documentation*. <https://doi.org/10.1108/JD-01-2022-0027>.
- Impett 2020a** Impett, L. (2020a) "Analyzing gesture in digital art history", in Brown, K. (ed.) *The routledge companion to digital humanities and art history*. New York: Routledge, pp. 386-407.
- Impett 2020b** Impett, L. (2020b) *Painting by numbers: Computational methods and the history of art*. EPFL.
- Ingold 2011** Ingold, T. (2011) *Being alive: Essays on movement, knowledge and description*. New York: Taylor & Francis.
- Jaillant 2022** Jaillant, L. (2022) *Archives, access and artificial intelligence: Working with born-digital and digitized archival collections*. Bielefeld, Germany: Bielefeld University Press.
- Jaillant and Rees 2023** Jaillant, L. and Rees, A. (2023) "Applying AI to digital archives: Trust, collaboration and shared professional ethics", *Digital Scholarship in the Humanities*, 38(2), pp. 571-585.
- Kenderdine, Mason, and Hibberd 2021** Kenderdine, S., Mason, I., and Hibberd L. (2021) "Computational archives for experimental museology", *Proceedings of RISE IMET 2021*. Nicosia, Cyprus, 2-4 June 2021. New York: Springer, pp. 3-18.
- Kreiss, Bertoni, and Alahi 2021** Kreiss, S., Bertoni, L., and Alahi, A. (2021) "Openpipaf: Composite fields for semantic keypoint detection and spatio-temporal association", *IEEE Transactions on Intelligent Transportation Systems*, 23(8), pp. 13498-13511.
- Król and Mynarski 2005** Król, H. and Mynarski, W. (2005) *Cechy ruchu-charakterystyka i możliwości parametryzacji [Features of movement-characteristics and capabilities of parametryzation]*. Katowice, Poland: Akademia Wychowania Fizycznego.
- Lin et al. 2014** Lin, T.Y. (2014) "Microsoft COCO: Common objects in context", *Proceedings of the 13th annual European conference of computer vision*. Zurich, Switzerland, 6-12 September 2014. New York: Springer, pp. 740-755.
- Ma 2003** Ma, M. (2003) *Wu xue tan zhen [Examination of truth in martial studies]*. Taipei, Taiwan: Lion Books.
- Mallik and Chaudhury 2012** Mallik, A. and Chaudhury, S. (2012) "Acquisition of multimedia ontology: An application in preservation of cultural heritage", *International Journal of Multimedia Information Retrieval*, 1(4), pp. 249–262.
- Mallik, Chaudhury, and Ghosh 2011** Mallik, A., Chaudhury, S., and Ghosh, H. (2011) "Nrityakosha: Preserving the intangible heritage of indian classical dance", *Journal on Computing and Cultural Heritage*, 4(3), pp. 1–25.
- Manovich 2020** Manovich, L. (2020) *Cultural analytics*. Cambridge, MA: The MIT Press.
- Martinsson, Rokhlin, and Tygert 2011** Martinsson, P.-G., Rokhlin, V., and Tygert, M. (2011) "A randomized algorithm for the decomposition of matrices", *Applied and Computational Harmonic Analysis*, 30(1), pp. 47–68. <https://www.sciencedirect.com/science/article/pii/S1063520310000242>.
- McGregor and Lab 2019** McGregor, W. and Lab, G.A.C. (2019) *Living archive*. Available at: <https://artsexperiments.withgoogle.com/living-archive>. (Accessed: 30 June 2023).
- McInnes et al. 2018** McInnes, L. et al. (2018) "Umap: Uniform manifold approximation and projection", *The Journal of Open Source Software*, 3(29). <https://doi.org/10.21105/joss.00861>.
- Mul and Masson 2018** Mul, G. and Masson, E. (2018) "Data-based art, algorithmic poetry: Geert Mul in conversation with

Eef Masson”, *TMG Journal for Media History*, 21(2).

- Olesen et al. 2016** Olesen, C.G. (2016) “Data-driven research for film history: Exploring the Jean Desmet collection”, *Moving Image: The Journal of the Association of Moving Image Archivists*, 16(1), pp. 82–105.
- Pedregosa et al. 2011** Pedregosa, F. et al. (2011) “Scikit-learn: Machine learning in Python”, *Journal of Machine Learning Research*, 12(85), pp. 2825–2830.
- RTSArchives 2018** RTSArchives (2018) *Le nouveau site RTSarchives*. Available at: <https://www.rts.ch/archives/5919889-le-nouveau-site-rtsarchives.html>. (Accessed: 30 June 2023).
- Salazar 2018** Salazar Sutil, N. (2018) “Section editorial: Human movement as critical creativity: Basic questions for movement computing”, *Computational Culture: a Journal of Software Studies*, 6.
- Sciuto et al. 2023** Sciuto, C. et al. (2023) “Exploring fragmented data: Environments, people and the senses in virtual reality”, in Landeschi, G. and Betts, E. (eds.) *Capturing the senses: Digital methods for sensory archaeologies*. New York: Springer, pp. 85-103.
- Sedmidubsky et al. 2020** Sedmidubsky, J. et al. (2020) “Motion words: A text-like representation of 3D skeleton sequences”, *Proceedings of the 42nd annual European conference on information retrieval*. Lisbon, Portugal, 14-17 April 2020. New York: Springer, pp. 527-241. https://doi.org/10.1007/978-3-030-45439-5_35.
- Shen et al. 2019** Shen, H. et al. (2019) “Information visualisation methods and techniques: State-of-the-art and future directions”, *Journal of Industrial Information Integration*, 16, pp. 100–102.
- Whitelaw 2015** Whitelaw, M. (2015) “Generous interfaces for digital cultural collections”, *DHQ: Digital Humanities Quarterly*, 9(1). Available at: <https://www.digitalhumanities.org/dhq/vol/9/1/000205/000205.html>.
- Winters and Prescott 2019** Winters, J. and Prescott, A. (2019) “Negotiating the born-digital: A problem of search”, *Archives and Manuscripts*, 47(3), pp. 391–403.
- Wright 2017** Wright, R. “The future of television archives”, *Digital Preservation Coalition*, 29 November. Available at: <https://www.dpconline.org/blog/wdpd/the-future-of-television-archives>.
- ZKM Center for Art and Media Karlsruhe 2023** ZKM Center for Art and Media Karlsruhe (2023) “William Forsythe: Improvisation technologies: The website project”. Available at: <https://zkm.de/en/project/william-forsythe-improvisation-technologies-the-website-project>. (Accessed 30 June 2023).
- van der Maaten and Hinton 2008** van der Maaten, L. and Hinton, G. (2008) “Visualizing data using t-sne”, *Journal of Machine Learning Research*, 9(86), pp. 2579–2605. Available at: <https://www.jmlr.org/papers/volume9/vandermaaten08a/vandermaaten08a.pdf>.



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.