Transforming Intangible Folkloric Performing Arts into Tangible Choreographic Digital Objects: The Terpsichore Approach

Anastasios Doulamis, Athanasios Voulodimos, Nikolaos Doulamis, Sofia Soile and Anastasios Lampropoulos

Photogrammetry and Computer Vision Lab, National Technical University of Athens, Zografou Campus, 15780, Athens, Greece

Keywords: Intangible Cultural Heritage, Performing Arts, Folkloric Dances, 3D Modelling, Cultural Heritage Digitisation.

Abstract: Intangible Cultural Heritage is a mainspring of cultural diversity and as such it should be a focal point in cultural heritage preservation and safeguarding endeavours. Nevertheless, although significant progress has been made in digitization technology as regards tangible cultural assets and especially in the area of 3D reconstruction, the e-documentation of intangible cultural heritage has not seen comparable progress. One of the main reasons associated lies in the significant challenges involved in the systematic e-digitisation of intangible cultural assets, such as performing arts. In this paper, we present at a high-level an approach for transforming intangible cultural assets, namely folk dances, into tangible choreographic digital objects. The approach is being implemented in the context of the H2020 European project "Terpsichore".

1 INTRODUCTION

Intangible Cultural Heritage (ICH) content means "the practices, representations, expressions, knowledge, skills – as well as the instruments, objects, artefacts and cultural spaces associated therewith". Intangible Cultural Heritage (ICH) is a very important mainspring of cultural diversity and a guarantee of sustainable development, as underscored in the UNESCO Recommendation on the safeguarding of Traditional Culture and folklore of 1989, in the UNESCO Universal Declaration on Cultural Diversity of 2001 and in the Convention for the Safeguarding of the Intangible Cultural Heritage (Kyriakaki, 2014). Improving the digitization technology regarding capturing, modelling and mathematical representation of performance arts and especially folklore dances is critical in: (i) promoting cultural diversity to the children and the youth through the safeguard of traditional performance arts; (ii) making local communities and especially indigenous people aware of the richness of their intangible heritage; (iii) strengthening cooperation and intercultural dialogue between people, different cultures and countries.

Although ICH content, especially traditional folklore performing arts, is commonly deemed worthy of preservation by UNESCO and the EU Treaty, most of the current research efforts focus on tangible cultural assets, while the ICH content has been given less emphasis. The primary difficulty stems by the complex structure of ICH, its dynamic nature, the interaction among the objects and the environment, as well as emotional elements (e.g., the way of expression and dancers' style). Of course there have been some notable efforts such as the i-Treasures project which provides a platform to access ICH resources and contribute to the transmission of rare know-how from Living Human Treasures to apprentices (Dimitropoulos, 2016) and the RePlay project, whose goal is to understand, preserve, protect and promote traditional sports (Linaza, 2013).

Towards this direction, the Terpsichore project aims to study, analyse, design, research, train, implement and validate an innovative framework for affordable digitization, modelling, archiving, e-preservation and presentation of ICH content related to folk dances, in a wide range of users (dance professionals, dance teachers, creative industries and general public).

Exploring the digitization technology regarding folklore performances constitutes a significant impact

Doulamis A., Voulodimos A., Doulamis N., Soile S. and Lampropoulos A.
Transforming Intangible Folkloric Performing Arts into Tangible Choreographic Digital Objects: The Terpsichore Approach.
DOI: 10.5220/0006347304510460

at European level. On one hand, the multi-cultural intangible dimension of Europe is documented, preserved, made available to everybody on Internet. On the other hand, the multifaceted value to the ICH content for usage in education, tourism, art, media, science and leisure settings is added.

Currently, digital technology has been widely adopted, which greatly accelerates efforts and efficiency of Cultural Heritage (CH) preservation and protection. At the same time, it enhances CH in the digital era, creating enriched virtual surrogates. Many research works have been proposed in the literature on archiving tangible cultural assets in the form of digital content (Li, 2010). Although the aforementioned significant achievements for improving the digitization technology towards a more cost-effective automated and semantically enriched representation, protection, presentation and re-use of the CH via the European Digital Library EUROPEANA, very few efforts exist in creating breakthrough digitization technology, improving the e-documentation (3D modelling enriched with multimedia metadata and ontologies), the e-preservation and re-use of ICH traditional music and fashion, folklore, handicraft, etc.

Terpsichore targets at integrating the latest innovative results of photogrammetry, computer vision, semantic technologies, time evolved modeling, combined with the story telling and folklore choreography. An important output of the project will be a Web based cultural server/viewer with the purpose to allow user's interaction, visualization, interface with existing cultural libraries and enrichment functionalities to result in virtual surrogates and media application scenarios that release the potential economic impact of ICH. The final product will support a set of services such as virtual/augmented reality, social media, interactive maps, presentation and learning of European Folk dances with significant impact on the European society, culture and tourism.

The remainder of this paper is structured as follows: In Section 2 we review the state of the art in the fields pertaining to the Terpsichore approach, which is presented in Section 3. Finally, Section 4 concludes the paper.

2 RELATED WORK

Although significant progress has been made in digitization technology as regards tangible cultural assets and especially in the area of 3D reconstruction – see the research achievements of the projects 3D-COFORM (www.3d-coform.eu), EPOCH (www.epoch-net.org), IMPACT (www.impact-project.eu), PRESTOSPACE (prestospace.org) – the e-documentation of intangible cultural assets is not yet evident, especially in the case of folklore performing arts. This is mainly due to the complex multi-disciplinarity of the folklore performances which presents a series of challenges ranging from the choreography, the folk music, the –uniforms, -music and from the digitization and computer vision to spatio-temporal (4D) dynamic modelling and virtual scene generation as discussed above. It is important to mention that this is the first time in the entire European Union that a Knowledge Alliance of this calibre is trying to undertake such innovative project, which aims to act as a pioneering mechanism for unifying ICH content with already existing digitized CH content from digital libraries (such as the folklore stories documented in the different e-records in EUROPEANA), and whose outcomes will not only lead to advanced scientific publications, but also patents, which will boost EU economic growth. In the following, we present the current e-documentation technologies and the respective limitations that the Terpsichore project aims to address.

2.1 3D Data Acquisition and Processing

The most popular methods for 3D capturing are divided into two main categories: active methods (laser scanners, range finders, structured light projectors) and passive methods (stereo vision and visual hulls). The most common used passive method is to attach distinctive markers to the body of a human and track these markers in images acquired by multiple calibrated cameras (www.vicon.com). In this case, method's accuracy is limited by the number of markers available. Markerless capture methods (Carceroni, 2001; Li, 2008; Pons, 2006) based on computer vision technology overcome this problem. However, these approaches do not fully exploit global spatiotemporal consistency constraints and are susceptible to error approximation.

The method of (Furukawa, 2007) addresses these limitations, however, it relies on PVMS software efficiency. In (Yamasaki, 2010) a system using 22 cameras is proposed. 3D modeling is based on the combination of volume intersection and stereo matching. However, the original shape distribution cannot be generated stably. Active methods offer higher stability and accuracy compared to passive methods. In (Cui, 2010) a 3D shape scanning system based on the ToF camera is presented. Although, these cameras are of low cost, they present limitations, such as very low X/Y resolution and random noise behavior. In (Sakashita, 2011) a system for capturing textured 3D shapes is presented, by using a multi-band camera in combination with an infrared structured light projector. However, it requires no other illumination in the environment. FusionKinect (Izadi, 2011) uses a Kinect camera to generate real-time depth maps containing discrete range measurements of the physical scene. However, depth data are inherently noisy and depth maps contain "holes" where no readings were obtained. After capturing, computer vision techniques are necessary for data reduction, filtering, optical flow and disparity estimation. The techniques used to solve correspondence problems are similar and can be categorized as energy-based and feature-based. Energy-based methods (Alvarez, 2002) minimize a cost function plus a regularization term, in the framework of (Horn, 1981), to solve for the 2D displacements and yield very accurate, dense flow fields. However, they fail as displacements get too large. Feature-based methods (Pons, 2006; Furukawa, 2011; Shrivastava, 2011; Liu, 2011) match features in different images. This kind of methods are able to overcome the problem of large displacements by using the concept of coarse-to-fine image warping, however, this downsampling removes information that may be vital for establishing correct matches. Data processing techniques will be used to refine acquired data (remove noises, remove "holes", accelerate 3D registration). The method of (Mitra, 2004) based on local least square fitting for estimating the normals at all sample points of a point cloud data set, in the presence of noise. The work of (Ruhnke, 2012) proposes an approach to obtain highly accurate 3D models from range data by jointly optimize the poses of the sensor and the positions of the surface points measured with a range scanning device. The work of (Rusu, 2009) uses Point Feature Histograms for accelerating 3D registration problem.

2.2 3D Modelling and Rendering

Modelling is a process to create a model, which by definition is an abstract representation that reflects the characteristics of a given entity either physical or conceptual. 3D modelling relies on computational geometry techniques such as skeleton extraction, division of space into subspaces and mesh reconstruction. The authors of (Menier, 2006) present an approach to recover body motions from multiple views using a 3D skeletal model, which is an a priori articulated model consists in kinematic chain of segments representing a body pose. In (Chen, 2009) an approach for simultaneously reconstructing 3D human motion and full-body skeletal size from a small set of 2D image features is presented. It resolves the ambiguity for skeleton reconstruction using pre-recorded human skeleton data. The approach of (Gall, 2009) recovers the movement of the skeleton, as well as, the possibly non-rigid temporal deformation of the 3D surface by using an articulated template model and silhouettes from a multi-view image sequence. Another volumetric approach is this of (Matsuyama, 2004) that uses silhouette volume intersection to generate the 3D voxel representation of the object shape and uses a high fidelity texture mapping algorithm to convert the 3D object shape into a triangular patch representation.

3D rendering is a necessary process for visualizing modelled content. However, real-time rendering of detailed animated characters, especially in crowded simulations like dance, is a challenging problem in computer graphics. Textured polygonal meshes provide high-quality representation at the expense of a high rendering cost. To overcome this problem, several techniques focusing on providing level-of-detail representations have been proposed. Image-based pre-computed impostors (Tecchia, 2002) render distant characters as a textured polygon to accelerate rendering of animated characters. A much more memory-efficient but view-dependent approach is to subdivide each animated character into a collection of pieces, in order to use separate impostors for different body parts (Kavan, 2008). In (Pettré, 2006) a three-level-of-detail approach is described, combining the animation quality of dynamic meshes with the high performance offered by static meshes and impostors. The technique in (Andújar, 2007) adopts a relief mapping approach to encode details in arbitrary 3D models with minimal supporting geometry.

2.3 Symbolic Representation, Ontologies and Harvesting

During dancing performances, motion gestures are used to communicate a storyline in an aesthetically pleasing manner. Although, humans automatically perceive and understand such gestures, from the point of view of computer science these gestures have to be analyzed under an appropriate framework with appropriate features, such as repetitive patterns, motion trajectories and motion inclusions, in order to extract their semantics. In the work of (Moon, 2008) a generative statistical approach to human motion modeling and tracking that utilizes probabilistic latent semantic analysis to describe the mapping of image features to 3D human pose estimate, is presented. The latent variables describe intrinsic motion semantics linking human figure to 3D pose estimates. In (Yang, 2010) the dance motion is analyzed to extract the repetitive patterns and compute prerequisite relations among them. These relations are illustrated by a concept map that is constructed automatically. The authors of (Kahol, 2004) based on human anatomy propose a method for deriving choreographer segmentation profiles of dance motion capture sequences. A more stable and descriptive framework for human motion analysis is the Laban Movement Analysis (LMA) framework. LMA has been proposed and used from the viewpoint of the analysis of body motions and it can be used for not only motion analysis but also extracting and generating expression of movements in general (Aristidou, 2014a; Aristidou, 2014b). The method of (Bouchard, 2007) uses LMA Effort component as a basis for motion capture segmentation, which is more meaningful than kinematic features, and it is easier to compute for general motions than semantic features. In (Woo, 2000) LMA has been used to obtain dancer's intention and estimate emotional or sensitivity information of the dance performance. In (Santos, 2011) the performance of different signal features regarding the qualitative meaning of LMA semantics is examined. Especially, this work is based on the study of body part trajectories and the objective is to apply multiple feature generation algorithms to segment the signals according to LMA theory, in order to find patterns and define the most prominent features in each of the descriptors defined on Labanotation (Chi, 2000). “Synchronous Objects” project (synchronousobjects.osu.edu) is focusing on the topic of transforming choreographies into symbolic representations. However, the dancers have been interactively marked by animators frame by frame with the aim to track their movements. Finally, within the project “MotionBank” (motionbank.org) the synchronous object is followed up with the aim to create “online scores” visualizing and illustrating the choreograph’s intention.

3 4D MODELLING OF ICH: THE TERPSICHORE APPROACH

To achieve a reliable 4D modelling of intangible cultural heritage (ICH) assets, such as dances, a new pioneer framework should be adopted. The research tools needed to be applied are depicted in Figure 1. The goals are to: (a) explore a scalable capturing framework as regards 3D reconstruction of ICH in terms of accuracy and cost-effectiveness, (b) develop the captured visual 3D signals into symbolic data that represent the overall human creativity, (c) define an interoperable Intangible Cultural Metadata Interface (ICMI) for folklore performing arts, and (d) nominate the appropriate codification of the extracted symbolic data structures in an interoperable form that permits interconnection with existing specifications of national digital cultural repositories or international like the EU digital library EUROPEANA (www.europeana.eu/portal/) and the UNESCO Memory of the World (Figure 2).

During the 20th century there have been several attempts to model human creativity in performing arts. In 1920s Rudolf Laban developed a system of movement notation that eventually evolved into modern-day Laban Movement Analysis (LMA) (Pforsich, 1977)0, which provides a language for describing, visualizing, interpreting, and documenting all varieties of human movement, in an attempt to preserve classic choreographies. In the early 2000s LMA extensively used for analyzing dance performances (Davis, 2001) and create digital archives of dancing in the area of education and research.

The main goal of the Terpsichore project is to leverage such past efforts towards the transformation of intangible cultural heritage content to 3D virtual content, through the development and exploitation of affordable digitization technology. Towards this direction, fusion of different scientific and technological fields, such as capturing technology, computer vision and learning, 3D modeling and reconstruction, virtual reality, computer graphics and data aggregation for metadata extraction, is necessary. There are four main distinct research directions in the context of the project, which will be presented in the following.

3.1 Choreographic Analysis, Design and Modelling

The main research objective in this section is to analyze the spatial specifications, the attributes and the properties of folklore traditional performing arts.

The analysis describes all the aspects that are needed for recording human creativity based on tradition; thus apart from the human movement, the way of expression, the associated emotional characteristics, the style as well as additional contextual data, such as climate conditions, social-cultural factors, stylistic variations, the accompanied untold scenarios and stories should be recorded. All

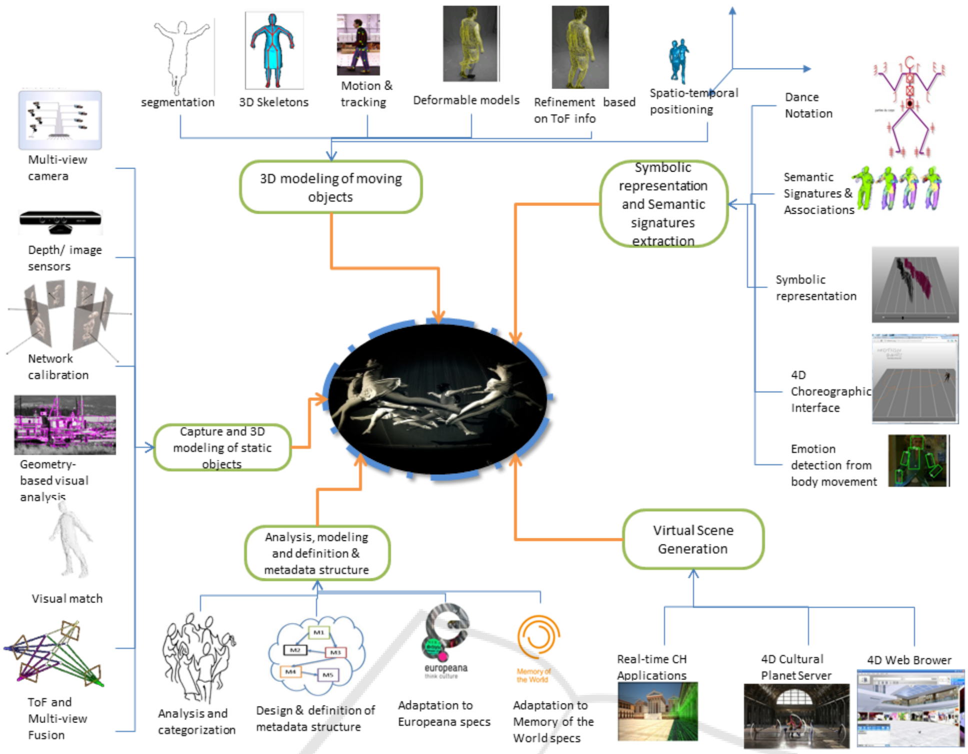

Figure 1: The main research components towards an intangible cultural heritage content (ICH) digitalization.

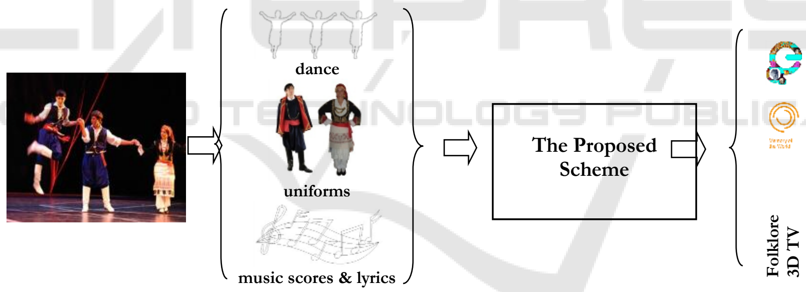

Figure 2: Framework of the proposed intangible cultural content digitalization.

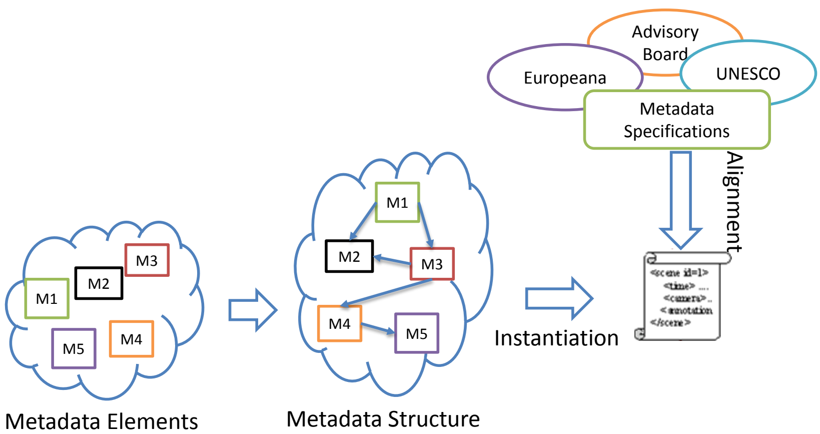

Figure 3: Description of the basic metadata elements used for representing the folklore performing arts.

the aforementioned specifications will be surveyed in such a way to derive an interoperable description framework based on which we are able to design and define the Intangible Cultural Metadata Interface (ICMI). ICMI aims at specifying a set of metadata that are necessary for representing the rich intangible cultural heritage information and especially in case of folklore traditional performing arts. In addition, ICMI introduces an interoperable description scheme framework able to specify the metadata structure and the relations between the extracted metadata. Thus, ICMI specifies not only the appropriate metadata for representing human creativity but also the structure and semantics of relations between its components. Figure 3 presents an indicative description framework of the basic metadata elements used for the description of folklore performing arts. As is observed, the metadata of the ICMI are discriminated into four main categories; the low level feature metadata, the contextual and environmental metadata, the socio-cultural factors as well as the emotional attributes.

The pool of the basic metadata used for describing the intangible cultural assets of traditional folklore performing arts are framed with a metadata structure format able to interoperably represent the semantics relations between the metadata components of Figure 3. Special emphasis should be given in order to align the ICMI format with existing specifications of digital cultural libraries, such as EUROPEANA and UNESCO the Memory of World, since this will allow the easy archiving, usage and re-usage of the digitized intangible cultural content with the content of large digital cultural repositories.

3.2 Capturing and 3D Modelling of Static Objects

Another important research issue deals with technologies able to capture and virtually 3D reconstruct traditional performing arts under an affordable, cost-effective and accurate digitization framework. To do this, an innovative technological framework able to combine advanced technologies in the area of photogrammetry and computer vision should be introduced. A scalable capturing framework that combines 3D modelling technology coming by a set of multi-view stereo camera architecture and a low cost depth sensors (Kinect) is necessary to be adopted. The balance of using multi-view stereo imaging (high resolution cameras plus dense image matching techniques) and low cost depth sensors able to generate depth information in real time are obtained in terms of accuracy and cost-effectiveness. In other words, the fusion of the information from different sensors in order to increase the accuracy or reduce the number of cameras necessary for data capturing remains a challenge.

As regards the multi-view imaging architecture, the 3D information is extracted by applying dense image matching techniques. For each stereo model observing common content a correspondence is determined for each pixel individually. By using these correspondences between all stereo models, all 3D Points can be triangulated based on their viewing rays at once. This leads to a very dense and accurate point cloud. In order to acquire the movements in time, this step is performed for each frame in time for all synchronized cameras – leading to 3D point clouds for each time stamp. However, this is a computational intensive process that increases the total cost of digitization. 3D modelling tools appropriately designed for time varying shapes should be used. Such methodologies exploit motion information as well as tracking methods Subsequently, a volumetric integration of this depth information not only enables the extraction of a volumetric representation, but also to fill small gaps and reduce noise.

The fusion process should be performed under a calibrated network of cameras or depth sensors. Network calibration is very important since it allows the implementation of super-resolution methods and detection of confidence data as obtained either by the depth sensors or the multi-view imaging. The use of self-calibration methodologies able to automatically calibrate the network of cameras and depth sensors using computer vision tools is a major research issue. Figure 4 presents the aforementioned methodology.

3.3 3D Modelling of Moving Objects

Another research field includes imaging methodologies based on the combination of photogrammetric, computer vision and computer graphics techniques able to automate the 3D modelling procedure of moving objects. Towards this direction, initially, segmentation algorithms able to isolate the foreground objects from the background is applied. This step is critical since it allows the human objects to be separated from the background content. The foreground objects are dynamically updated through the motion estimation captured by the capturing layer. On the contrary, background content is updated by the information provided by the depth sensor network. A set of deformable models is used to describe the human movement. These set of deformable models are provided dealing with the

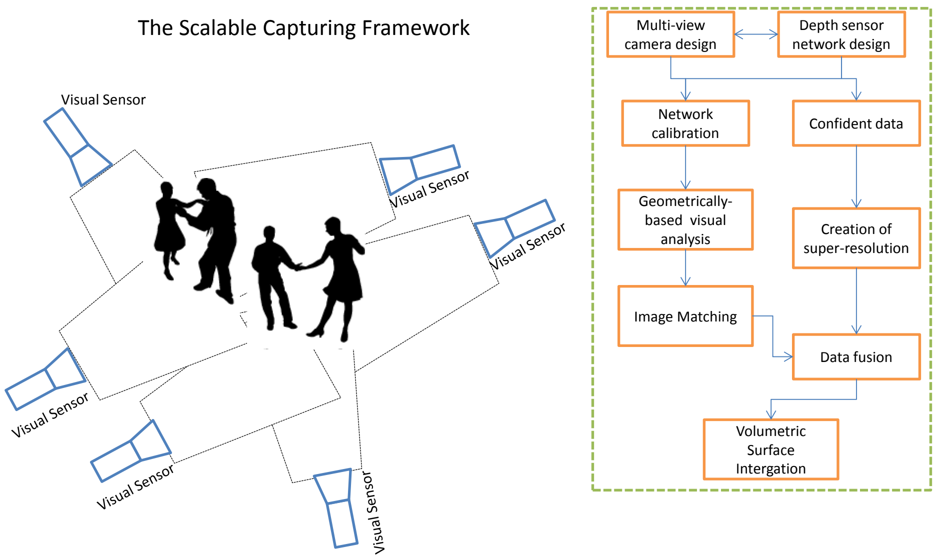

Figure 4: A representation of the methodology resulting in a scalable capturing framework in terms of accuracy and performance.

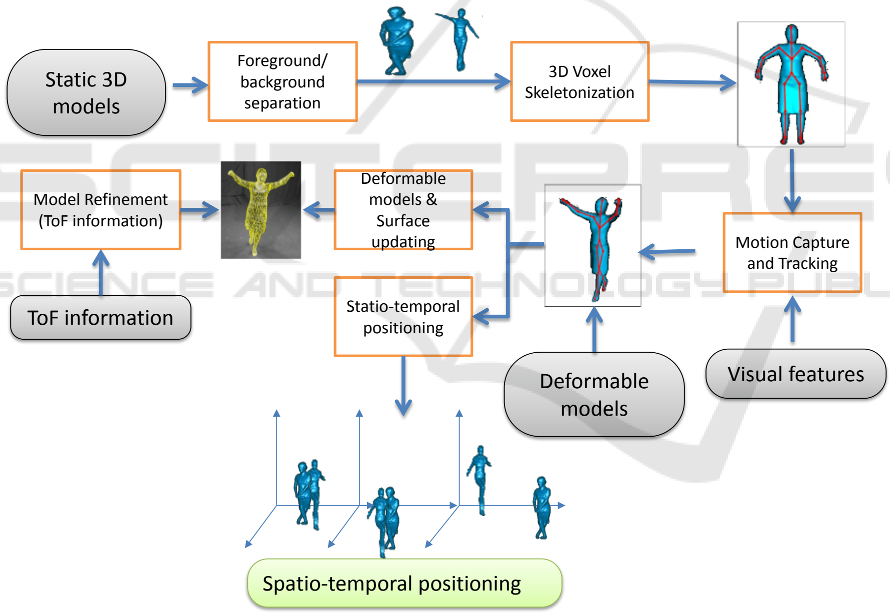

Figure 5: The 3D modelling methodology for moving objects.

analysis, design and modelling of folklore traditional performing arts.

In order to allow a cost effective 3D digitization process, a voxel 3D skeletonization is performed. In this way, we are able to reduce the amount of information needed to describe the human object; therefore, a reduction of the cost of processing resulting in more accurate and cost-effective 3D modelling of moving objects is possible. 3D skeletons are medial axes transform and encode mostly motion information. 3D skeletons separate motion estimation from shape and thus, they allow a more accurate 3D model updating through time.

The derived 3D skeletons are tracked in time using motion capture and tracking computer vision methodologies. Tracking process is assist through the selection of appropriate geometrically enriched data. By taking into consideration motion tracking, which is performed on the 3D skeletons to increase accuracy and computational efficiency, as well as the set of deformable models appropriate for a specific folklore performance, we are able to automatically update the detected 3D models to fit the properties of the current human movement and shape constraints as obtained by the static 3D models.

Despite the efficiency of the aforementioned methodology, possible errors (in the 3D skeletonization and tracking process) generate erroneous virtual 3D reconstructions. To address this difficult, we enhance the results of the 3D modelling of moving objects using information derived from the depth sensor network. In this case, we are able to exploit the depth information to improve 3D modelling accuracy. Figure 5 depicts this architecture.

3.4 Symbolic Representation and Extraction of Semantic Signatures

3D modelling is critical for encoding the complex 3D reconstructions of performing arts into a set of compact semantic signatures in a similar way that a music song is encoded using a music score. For this reason, computational geometry algorithms are used to decode the spatial-temporal trajectory of the performances. This includes methodologies for positioning both in 3D and temporal space. Then, semantic signatures and respective spatial-temporal associations are extracted to represent the performances with high level concepts. It is clear that semantic analysis aids the digitization and computer vision process and vice versa.

The visualization of scores fosters the understanding of the dance and it helps visitors, dancers and choreographs to comprehend the structure and the intension of the dance. A more formalized documentation of the dance will be supported with an automated mapping of capturing data to formal and abstract chorographic notations ("Symbolic Representations"). Hereby, the captured dance could be coded using traditional approaches like the Laban Motion Analysis or modern ballet notation forms like the "peacemaker" from William Forsythe and David Kern.

3.5 Virtual Scene Generation

Finally, the complete reconstructed 4D-scenario can be visualized within an interactive Web-Viewer. Virtual scene generation exploits both the produced 3D models, but also the codified symbolic representations, generating an innovative and unique research framework able to allow for manipulation, usage and re-usage of the cultural objects. This interactive Web-Viewer links the extracted Symbolic Representations to the 4D-reconstruction of the choreography. Within this Web-Viewer beside different viewpoints also the perspective of a specific dance can be chosen.

4 CONCLUSIONS

In this paper we have presented the concept of the Terpsichore project. The research directions of the project will span across various fields, including choreographic analysis, 3D data capturing, 3D modelling of static and moving objects and symbolic representations, among others. Through the described approach Terpsichore aims to study, analyse, design, research, train, implement and validate an innovative framework for affordable digitization, modelling, archiving, e-preservation and presentation of Intangible Cultural Heritage content related to folk dances, in a wide range of users, including dance professionals, dance teachers, creative industries and the general public. Future

ACKNOWLEDGEMENTS

This work has been supported by the H2020-MSCA-RISE project "Transforming Intangible Folkloric Performing Arts into Tangible Choreographic Digital Objects (Terpsichore)" funded by the European Commission under grant agreement no 691218. The authors would like to help all partners for their contribution and collaboration.

REFERENCES

Alvarez, L., R. Deriche, J. Sánchez, and J. Weickert, 2002. "Dense Disparity Map Estimation Respecting Image Discontinuities: A PDE and Scale-Space Based Approach," Journal of Visual Communication and Image Representation, vol. 13, no. 1–2, pp. 3–21, 2002.

Andújar, C., J. Boo, P. Brunet, M. Fairén, I. Navazo, P. Vázquez, and À. Vinacua, 2007. "Omni-directional Relief Impostors," Computer Graphics Forum, vol. 26, no. 3, pp. 553–560, 2007.

Aristidou A. & Y. Chrysanthou, "Feature extraction for human motion indexing of acted dance performances." In Proceedings of the 9th International Conference on Computer Graphics Theory and Applications, 2014 (pp. 277-287).

Aristidou, A., E. Stavrakis, & Y. Chrysanthou, "LMA-Based Motion Retrieval for Folk Dance Cultural Heritage." In Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection, 2014, (pp. 207-216). Springer International Pub

Bouchard D. and N. Badler, “Semantic Segmentation of Motion Capture Using Laban Movement Analysis,” in Proceedings of the 7th international conference on Intelligent Virtual Agents, Berlin, Heidelberg, 2007, pp. 37–44.

Carceroni, R.L. and K. N. Kutalakos, 2001. “Multi-view scene capture by surfel sampling: from video streams to non-rigid 3D motion, shape and reflectance,” in Eighth IEEE International Conference on Computer Vision, 2001. ICCV 2001. Proceedings, 2001, vol. 2, pp. 60 – 67 vol.2.

Chen, Y.-L. and J. Chai, 2010. “3D Reconstruction of Human Motion and Skeleton from Uncalibrated Monocular Video,” in Computer Vision – ACCV 2009, H. Zha, R. Taniguchi, and S. Maybank, Eds. Springer Berlin Heidelberg, pp. 71–82.

Chi, D. M. Costa, L. Zhao, and N. Badler, “The EMOTE model for effort and shape,” in Proceedings of the 27th annual conference on Computer graphics and interactive techniques, New York, NY, USA, 2000, pp. 173–182.

Cui, Y., S. Schuon, D. Chan, S. Thrun, and C. Theobalt, 2010. “3D shape scanning with a time-of-flight camera,” in 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2010, pp. 1173 –1180.

Davis, E., 2001. “BEYOND DANCE Laban's Legacy of Movement Analysis,” Brechin Books, 2001

Dimitropoulos, K., P. Barmpoutis, A. Kitsikidis, and N. Grammalidis, 2016. "Extracting dynamics from multi-dimensional time-evolving data using a bag of higher-order Linear Dynamical Systems", in Proc. 11th International Conference on Computer Vision Theory and Applications (VISAPP 2016), Rome, Italy.

Dobbyn, S., J. Hamill, K. O’Conor, and C. O’Sullivan, 2005. “Geopostors: a real-time geometry / impostor crowd rendering system,” in Proceedings of the 2005 symposium on Interactive 3D graphics and games, New York, NY, USA, 2005, pp. 95–102.

Furukawa, Y. and J. Ponce, 2010. “Dense 3D Motion Capture from Synchronized Video Streams,” in Image and Geometry Processing for 3-D Cinematography, R. Ronfard and G. Taubin, Eds. Springer Berlin Heidelberg, 2010, pp. 193–211.

Gall, J, C. Stoll, E. de Aguiar, C. Theobalt, B. Rosenhahn, and H.-P. Seidel, 2009. “Motion capture using joint skeleton tracking and surface estimation,” in IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2009, 2009, pp. 1746 –1753.

Horn B. K. P. and B. G. Schunck, 1981. “Determining optical flow,” *Artificial Intelligence*, vol. 17, no. 1–3, pp. 185–203.

Izadi, S., D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison, and A. Fitzgibbon, 2011. “KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera,” in Proceedings of the 24th annual ACM symposium on User interface software and technology, New York, NY, USA, 2011, pp. 559–568.

Kahol, K., P. Tripathi, and S. Panchanathan, “Automated gesture segmentation from dance sequences,” in Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. Proceedings, 2004, pp. 883 – 888.

Kavan, L., S. Dobbyn, S. Collins, J. \vZára, and C. O’Sullivan, 2008. “Polypostors: 2D polygonal impostors for 3D crowds,” in Proceedings of the 2008 symposium on Interactive 3D graphics and games, New York, NY, USA, pp. 149–155.

Kyriakaki, G., Doulamis, A., Doulamis, N., Ioannides, M., Makantasis, K., Protopapadakis, E., Hadjiprocopis, A., Wenzel, K., Fritsch, D., Klein, M., Weinlinger, G., 2014. "4D Reconstruction of Tangible Cultural Heritage Objects from web-retrieved Images," International Journal of Heritage in Digital Era, vol. 3, no. 2, pp. 431-452.

Laumond, and D. Thalmann, 2006. “Real-time navigating crowds: scalable simulation and rendering,” Computer Animation and Virtual Worlds, vol. 17, no. 3–4, pp. 445–455, 2006.

Li, R., T. Luo, and H. Zha, 2010. “3D Digitization and Its Applications in Cultural Heritage,” in *Digital Heritage*, M. Ioannides, D. Fellner, A. Georgopoulos, and D. G. Hadjimitsis, Eds. Springer Berlin Heidelberg, pp. 381–388.

Li, R. and S. Sclaroff, 2008. “Multi-scale 3D scene flow from binocular stereo sequences,” Computer Vision and Image Understanding, vol. 110, no. 1, pp. 75–90, 2008.

Linaza, M, Kieran Moran, Noel E. O’Connor, 2013. "Traditional Sports and Games: A New Opportunity for Personalized Access to Cultural Heritage", 6th International Workshop on Personalized Access to Cultural Heritage (PATCH 2013), Rome, Italy.

Matsuyama, T., X. Wu, T. Takai, and S. Nobuhara, 2004. “Real-time 3D shape reconstruction, dynamic 3D mesh deformation, and high fidelity visualization for 3D video,” Computer Vision and Image Understanding, vol. 96, no. 3, pp. 393–434.

Menier, C., E. Boyer, and B. Raffin, 2006. “3D Skeleton-Based Body Pose Recovery,” in Proceedings of the Third International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT’06), Washington, DC, USA, 2006, pp. 389–396.

Mitra, N.J, A. Nguyen, and L. Guibas, 2014. “Estimating surface normal in noisy point cloud data,” *International Journal of Computational Geometry & Applications*, vol. 14, no. 04n05, pp. 261–276.

Moon, K. and V. Pavlovic, “3D Human Motion Tracking Using Dynamic Probabilistic Latent Semantic Analysis,” in Canadian Conference on Computer and Robot Vision, 2008. CRV ’08, 2008, pp. 155 –162.

Pettré, P. de H. Ciechomski, J. Maïm, B. Yersin, J.-P. Laumond, and D. Thalmann, 2006. “Real-time navigating crowds: scalable simulation and rendering,” Computer Animation and Virtual Worlds, vol. 17, no. 3–4, pp. 445–455.

Pforsich, J., 1977. “Handbook for Laban Movement Analysis", New York: Janis Pforsich, 1977.

Pons, J.-P., R. Keriven, and O. Faugeras, 2006. "Multi-View Stereo Reconstruction and Scene Flow Estimation with a Global Image-Based Matching Score," International Journal of Computer Vision, vol. 72, no. 2, pp. 179–193, Jul. 2006.

Ruhnke, M., R. Kummerle, G. Grisetti, and W. Burgard, 2012. "Highly accurate 3D surface models by sparse surface adjustment," in *2012 IEEE International Conference on Robotics and Automation (ICRA)*, 2012, pp. 751 –757.

Rusu, R. B., N. Blodow, and M. Beetz, 2009. "Fast Point Feature Histograms (FPFH) for 3D registration," in *IEEE International Conference on Robotics and Automation, 2009. ICRA '09*, 2009, pp. 3212 –3217.

Santos L. and J. Dias, "Motion Patterns: Signal Interpretation towards the Laban Movement Analysis Semantics," in Technological Innovation for Sustainability, vol. 349, L. M. Camarinha-Matos, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 333–340.

Sakashita, K., Y. Yagi, R. Sagawa, R. Furukawa, and H. Kawasaki, 2011. "A System for Capturing Textured 3D Shapes Based on One-Shot Grid Pattern with Multi-band Camera and Infrared Projector," in 2011 International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), 2011, pp. 49 –56.

Tecchia, F., C. Loscos, and Y. Chrysanthou, "Image-based crowd rendering," IEEE Computer Graphics and Applications, vol. 22, no. 2, pp. 36 –43, Apr. 2002.

Woo, W., J. Park, and Y. Iwadate, "Emotion analysis from dance performance using time delay neural networks," in proc. of the JCIS-CVPRIP, 2000, pp. 374–377.

Yamasaki, T. and K. Aizawa, 2007. "Motion segmentation and retrieval for 3D video based on modified shape distribution," EURASIP J. Appl. Signal Process., vol. 2007, no. 1, pp. 211–211, 2007.

Yang, Y., H. Leung, L. Yue, and L. Deng, 2010. "Automatically Constructing a Compact Concept Map of Dance Motion with Motion Captured Data," in Advances in Web-Based Learning – ICWL 2010, vol. 6483.