

## Article

# Supporting Sign Language Narrations in the Museum

Nikolaos Partarakis <sup>1,\*</sup> , Xenophon Zabulis <sup>1</sup> , Michalis Foukarakis <sup>1</sup>, Mirodanthi Moutsaki <sup>1</sup>, Emmanouil Zidianakis <sup>1</sup> , Andreas Patakos <sup>1</sup>, Ilia Adami <sup>1</sup>, Danae Kaplanidi <sup>2</sup>, Christodoulos Ringas <sup>2</sup> and Eleana Tasiopoulou <sup>2</sup>

<sup>1</sup> Foundation for Research and Technology (ICS-FORTH), Institute of Computer Science, GR-70013 Heraklion, Crete, Greece; zabulis@ics.forth.gr (X.Z.); foukas@ics.forth.gr (M.F.); mmirto91@hotmail.com (M.M.); zidian@ics.forth.gr (E.Z.); patakos@ics.forth.gr (A.P.); iadami@ics.forth.gr (I.A.)

<sup>2</sup> Piraeus Bank Group Cultural Foundation, 6 Ang. Gerontas St., GR-10558 Athens, Greece; danae.kaplanidi@gmail.com (D.K.); RiggasCh@piraeusbank.gr (C.R.); TasiopoulouE@piraeusbank.gr (E.T.)

\* Correspondence: partarak@ics.forth.gr; Tel.: +30-28-1039-1754

**Abstract:** The accessibility of Cultural Heritage content for the diverse user population visiting Cultural Heritage Institutions and accessing content online has not been thoroughly discussed. Considering the penetration of new digital media in such physical and virtual spaces, lack of accessibility may result in the exclusion of a large user population. To overcome such emerging barriers, this paper proposes a cost-effective methodology for the implementation of Virtual Humans, which are capable of narrating content in a universally accessible form and acting as virtual storytellers in the context of online and on-site CH experiences. The methodology is rooted in advances in motion capture technologies and Virtual Human implementation, animation, and multi-device rendering. This methodology is employed in the context of a museum installation at the Chios Mastic Museum where VHs are presenting the industrial process of mastic processing for chewing gum production.

**Keywords:** virtual humans; sign language; augmented reality; museum narratives; storytelling



**Citation:** Partarakis, N.; Zabulis, X.; Foukarakis, M.; Moutsaki, M.; Zidianakis, E.; Patakos, A.; Adami, I.; Kaplanidi, D.; Ringas, C.; Tasiopoulou, E. Supporting Sign Language Narrations in the Museum. *Heritage* **2022**, *5*, 1–20. <https://doi.org/10.3390/heritage5010001>

Academic Editor: Nicola Masini

Received: 17 November 2021

Accepted: 17 December 2021

Published: 21 December 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The impact of the digital revolution has resulted in the emergence of many and diverse opportunities for people to also engage with culture through digital media. This was made more urgent by the COVID-19 pandemic and the resulting need for CHIs to provide more online experiences to the public. This resulted in a lot of Cultural Heritage (CH) content becoming available online.

In this work, we challenge the fact that opportunities to access digital media on-site and online are not granted in equal measure to all citizens. Although Cultural Heritage Institutions (CHIs) worldwide are developing strategies towards widening the public's interest through digitalization in order to address not only emerging requirements but also emerging crises, such as the COVID-19 pandemic, there has barely been any discussion in the cultural sector worldwide about the accessibility of interactive digital exhibits, digital media, and content for disabled people. An audit regarding the accessibility of museums has shown that disabled people face numerous potential stumbling blocks in the average CHI [1]. The idea that cultural venues, as a service to the public, have a responsibility to welcome everyone in inclusive settings is far from being universally embraced in the cultural sector.

To this end, the goal of providing accessible CH content depends on (a) raising the awareness of the public and the CH sector and (b) providing methodologies that could be easily applied by the majority of the Cultural and Creative Industries (CCIs) that produce and distribute such content. In this paper, a solution to support the second prerequisite is discussed, implemented, and validated in the museum context, employing widely adopted technologies and technical equipment.

## 2. Background and Related Work

### 2.1. Web Content Accessibility

Nowadays, web content is employed by a multitude of devices and applications, extending the traditional web-exclusive usage. Web 2.0 applications are compatible with more platforms and devices, thus making such content available in various non-desktop computing contexts. In this vein, accessibility of web content extends the barriers of browser-based interaction. The Web Content Accessibility Guidelines (WCAG) 2.0 [2] cover a wide range of recommendations for making Web content more accessible. At the same time, the User Agent Accessibility Guidelines (UAAG) explain how to make user agents accessible to people with disabilities [3]. Following these guidelines, designers and developers can make content and content-consuming applications accessible to a wider range of people with disabilities, including blindness. Although these guidelines provide a good starting point for achieving accessibility of web content, several limitations exist. First of all, the guidelines cannot always be evaluated by automated accessibility evaluators (e.g., [4]) as they contain several “manual checks”, which are things that should be checked individually by the developers or accessibility evaluation experts. Secondly, in many cases, accessibility conformance does not result in an accessible website. The development team should have access and expertise in assistive technologies to test how the web page is scanned (e.g., through binary switches) or “heard” through a screen reader. Finally, achieving maximum conformance requires the adaptation of content for people with hearing and cognitive impairments; this results in the need for the creation of alternative content for these individuals.

Over the years, many methodologies have been proposed for the development of universally accessible applications (e.g., [5]). Among these, the most promising ones are the ones that combine accessibility guidelines with adaptation to user profiles. Such approaches depend on the initial level of accessibility, supported through conformance with the accessibility guidelines and reinforced by content and UI adaptation based on the specificities of each user as defined by their user profile. Alternative approaches aim, e.g., towards the transformation of web content into audible information [6] for users to be able to listen to the web content, following a similar approach to the daisy format for accessible multimedia books [7].

The evolution of 3D user interfaces and 3D content provision has provided new content and new opportunities but has also posed new barriers in terms of accessibility. The approach followed in this research work is to employ well-known solutions brought by web content accessibility, in terms of making textual content accessible, and to focus on the new aspects, such as the interactions with VH that require new radical solutions.

### 2.2. Accessibility of CH Resources in the Museum

The mainstream accessibility approaches presented above fail to provide sufficient tools for supporting the diverse content provided by museums, both online and on-site. Despite the progress to date, the CH fruition enabled by interactive technologies still presents considerable limitations: (i) the accessibility of existing interactive systems has not yet been considered by application providers; (ii) current systems offer limited interactivity, personalization, and contextual grounding of the fruition experience; (iii) very few efforts are focused on exploiting the wealth of available digital content and specialized knowledge for the benefit of the public at large in a way that would help further capitalize on significant investments in this area; and (iv) there are no systematic technological solutions available for supporting museums and cultural heritage institutions in more effectively satisfying visitors’ expectations [8].

A potential solution could be the design and implementation of interactive technologies with adaptation features that could enable content and UI adaptation to the needs of each user. Some initial approaches have been proposed in this area, focusing on the visitor’s profile to adapt information and content [9,10].

### 2.3. *New Media in the Museum*

The evolution of digital technologies has brought new media technologies to the museum. In this work, a particular interest is VHS acting as museum narrators. The usage of VHS in Digital Cultural Heritage (DCH) environments has recently been studied [11–13]. For example, in ref. [14] the persuasiveness and overall emotional impact of VHS with different professional and social characteristics (a curator, a museum guard, and a visitor) in an immersive VM environment have been studied. In the study, persuasiveness relates to the VHS's capacity to engage, affect, and stimulate emotional and cognitive responses by employing different narration styles. In ref. [15], the authors underline the importance of aligning and fine-tuning the narrative styles and contents of the VHS, which should correspond in terms of appearance to their roles, and they highlight the importance of affective components in their storytelling approach. Under these conditions, the virtual experience stimulates attention and involvement [16–18] and thus can make the stories presented more credible and thereby influence the users positively and constructively. Furthermore, VHS contribute to the suspension of disbelief, which enables the user to become immersed and follow their story and turn of events.

### 2.4. *Sign Language and the Museum*

New media pose new interaction requirements and new considerations in terms of accessibility. The need to address the requirements of a diverse population in the museum in terms of accessibility is extremely challenging. More specifically, people with hearing disabilities are considered among the ones that face barriers to understanding both written and oral information. In this section, an analysis of the state-of-the-art signing VHS and sign language production is provided.

#### 2.4.1. *Sign Language Translators Captured on Video*

A very common approach is that of the use of pre-recorded videos of human signers. Although this solution generally produces natural results, it is not viable as these videos cannot be easily updated and enriched with new information and thus become obsolete. Re-recording human signers eventually ends up being a tedious, time-consuming task and is also an expensive process, whereas “stitching” together clips of signs to produce words and sentences often leads to non-natural results due to discontinuities that certainly appear in the video editing [19,20].

#### 2.4.2. *Signing Virtual Humans*

To overcome the limitations of pre-recorded videos, animated characters performing sign language are used instead, providing more flexibility through the control of the 3D movement of their parts [21]. “Scripting” systems have been developed which allow users with familiarity with sign language to produce animations by combining signs from a lexicon with facial expressions. An example of such a tool is eSIGN, which allows developers to create sign databases, whereas Vcom3D, a character animation system, supports facial expressions, allowing ASL scripting. EMBR [22] and JASigning [23] are similar systems that automate, and thus ease, aspects of animations for animators, such as the transition movements between signs. On the other hand, other researchers follow a different path, focusing on animation production by generating/translating the spoken or written word to sign language [24–27].

Signing VHS are a relatively new research area with two decades of active research and some significant results. However, creating signing VHS involves multiple challenges, ranging from content representation, as a universal writing system for sign language does not exist, to realizing a comprehensible animation. Sign language is a highly multi-channel/multimodal language where hands/arms, the face, and the whole body must be synchronized on various levels. Therefore, state-of-the-art VHS reach rather low comprehensibility levels of 58–62%, with a single study reporting 71% [28].

Despite the challenges involved, great efforts are being made for the accurate reproduction of not only manual signs (hand signs, gestures) but non-manual (facial expressions, head tilting, mouthing, shoulder raising, etc.) too. Specifically, for the case of ASL it has been proven that when manual signs are combined with facial expressions, the comprehension levels of the animations improve greatly [29]. The SignCom project [30] retrieved motion-captured data from different parts of the body (e.g., head, hands, and torso) and combined them to build an animation system for the French Sign Language. The combination of both manual and non-manual signals in a coordinated manner is a critical parameter and a crucial step that needed to be taken to significantly affect and increase the comprehension of the animations by deaf users.

Examples of sign language animation accomplished through motion capture are Paula of DePaul University and TESSA, a demonstration of the European Visicast Project. They produce excellent animations that are lifelike and intelligible. However, their inventory of signs is limited by the need to acquire, store, and retrieve motion capture data from experienced human signers [31,32]. An early deviation from motion capture is SignSynth from the University of New Mexico, which employs a simple VRML VH to synthesize signs from stored Stokoe-like parameters [33]. The e-SIGN project, the successor to Visicast, has introduced a sign editor based on the Hamburg Notation System (HAMNOSYS) [34]. DePaul University has been modifying its animation software to allow the inverse kinematics that are necessary for synthesis [35].

When the VH is generated as part of a translation system [36], an initial translation step converts spoken/written language into a symbolic representation of the sign language (as described in the previous section). Whether human-authored or automatically translated, a symbolic plan is needed for the sign-language message. While multiple representations have been proposed [37–39], there is no universal standard. Beginning with this symbolic plan, pipelines generating VHS typically involve a series of complex steps. Animations for individual signs are often pulled from lexicons of individual signs. These motion plans for the individual signs are produced in one of several ways: key-frame animations [40], symbolic encoding of subsign elements [41], or motion capture recordings [42,43]. Similarly, non-manual signals are pulled from complementary datasets [44] or synthesized from models [45]. These elements are combined to create an initial motion script of the content. Next, various parameters (e.g., speed, timing) are set by a human, set by a rule-based approach [46], or predicted via a trained machine learning model [21,47]. Finally, computer animation software renders the animation based on this detailed movement plan [48].

#### 2.4.3. Sign Language Production through Video Synthesis

Sign language synthesis can be also performed by assembling individual previously filmed video clips of sign demonstrations [49]. In this case, further optimization of the passages between individual video clips is needed to enable smooth joining of the video clips in real-time, offering a high-quality video.

#### 2.5. Human Motion Digitisation for Sign Language Production

Regarding the digitization of human motion, the current state of the art includes the usage of marker-based and IMU-based systems. Marker-based systems have multiple cameras that encircle a specific volume. Retroreflective (i.e., with minimum scattering reflection) spherical markers are placed on the subject at a set pattern. The cameras emit IR light that is reflected directly back to its source with virtually no scattering by the markers. The marker-based system provides a complete 3D area where only the marked objects are tracked while the rest of the scenery is ignored. With enough cameras and an appropriate setup, it is possible to have an unobstructed view of a large area. Another benefit is that the system records only the markers; therefore, background items are not included in the final recording, making the output efficient. However, marker-based systems require more post-processing to extract joint angles from the cluster of markers, and they are not portable by their nature.

IMU systems are different in the sense that they do not measure displacement but acceleration. Each IMU is comprised of an accelerometer, a magnetometer, and a gyroscope [50,51], which provide measurements in three dimensions in relation to the earth's magnetic field. The output of one IMU is relative to a global coordinate frame. The biggest benefit of this method is that there is no need to generate a reference signal, such as the IR light in optical-based systems. However, they are sensitive to magnetic noise. Sources of magnetic noise can be electrical appliances, metal furniture, and metallic structures within a building [50]. There are commercially available suits that have embedded several IMUs to be worn by the subjects, and their output is very streamlined without the need for significant post-processing of the data.

To sum up, IMU-based suits are portable but have high specificity. However, electromagnetic disturbances can cause a high error in the azimuth angle (perpendicular to the north–south pole of the earth's magnetic field) for as long as the magnetic distortion is present [50]. Assuming the magnetic interference has some consistency (static magnetic field or distortion with a known frequency), it is possible to create filters to correct the signal to a reasonable degree [50,52]. Regardless of the technology used to acquire those recordings, the resulting data are always a chain of coordinate frames and the difference in position and orientation between them.

In the domain of sign languages, research has employed various types of MoCap systems such as:

- Collecting a Motion Capture Corpus of American Sign Language: two Immersion CyberGloves®, an applied Science Labs H6 eye-tracker, InterSense IS-900 acoustical/inertial motion capture system, and an Animazoo IGS-190 spandex bodysuit [53].
- Building French Sign Language Motion Capture Corpora: marker-based systems [54].
- VICON MoCap cameras and the use of VICON's MX system [55].
- Computer-vision-based optical approaches: (a) The HANDYCap system [56] uses three cameras, which operate at a high frame rate of 120 Hz (120 frames/second). Two cameras are focused on the whole body, while the third one is focused only on the face and provides data for further processing (for sensing facial expressions); (b) ten cameras are used for the optical motion capture, the Tobii eye tracker, and the markers [57].

## 2.6. Sign Language Datasets

Currently, several alternative sign language datasets have been generated and used. A classification of such datasets can be found below:

- Image datasets: (a) 2D static hand gesture color image dataset for ASL gestures [58] and (b) the dataset for the Irish sign language [59].
- Video and multi-angle video datasets: (a) the RWTH-PHOENIX-Weather corpus, a video-based, large vocabulary corpus of German sign language suitable for statistical sign language recognition and translation [60]; (b) the ASL Lexicon Video Dataset, a large and expanding public dataset containing video sequences of thousands of distinct ASL signs, as well as annotations of those sequences, including start/end frames and class labels of every sign [61]; and (c) SIGNUM: the database for Signer-Independent Continuous Sign Language Recognition [62].
- RGB video and RGB-D video synchronized datasets: (a) the Polish sign language dataset [63] and (b) the 3D body-part detection video dataset [64].
- Video and ToF camera datasets: PSL ToF 84: gestures as well as hand postures acquired by a time-of-flight (ToF) camera [65].

These datasets are extremely helpful for research in the context of sign-language-detection algorithms using a variety of technologies from CV to deep neural networks, but they are not sufficient to support the objectives of this research work due to their nature but also the lack of expressive power to support CH applications.



## 2.7. Discussion

The multitude of research outcomes and accessibility solutions presented in this section do not yet provide an adequate solution that could be exploited by the Cultural and Creative Sector (CCS) to offer accessible solutions in the museum context. Most of these remain research outcomes and thus cannot be employed in a production context. Thus, by critically exploring what was proposed by past research work, in this paper we adopt the solution proposed by the SignCom project [30]. Rooted in this fact, and with the support of past research in universal access and adaptive user interfaces, a methodology is proposed that can be adapted by the Cultural and Creative Industries (CCIs) to support the provision of accessible VHS in the museum. The emphasis of this work is on people with hearing disabilities, but other functional limitations are also addressed through the setup of the installed prototype.

## 3. Proposed Method

The proposed methodology is comprised of six sequential steps, as shown in Figure 1. The methodology has been developed in the context of the Mingei project, targeting the representation and presentation of Heritage Crafts [66]. In the first step, the narrative is authored by CH professionals based on several research methods that could include archival research, ethnography, interviews, etc. The script is reviewed by the sign language translators and is optimized for oral sign language presentation. Then, the VHS are implemented considering the characteristics of the personality narrating the story (e.g., age, gender, occupation, historical clothing, etc.). Then, the narration is recorded using a motion capture suit and gloves, both in sign language and orally. In the oral recording, the narration audio is also recorded. Then, the segmentation occurs to identify possible reusable parts of the recording to be integrated into the sign language vocabulary. The resulting animations are imported into a 3D game engine and retargeted to animate the virtual narrator, with the animation sequence defined by the recorded animation file. In this stage, the resulting retargeted animations must be validated by sign language translators to assess their accuracy and readability. Finally, as the resulting animations can be exported again in fbx format, these can be employed to augment the content of various 3D applications delivered through a standalone application, the web, mobile devices, and VR.

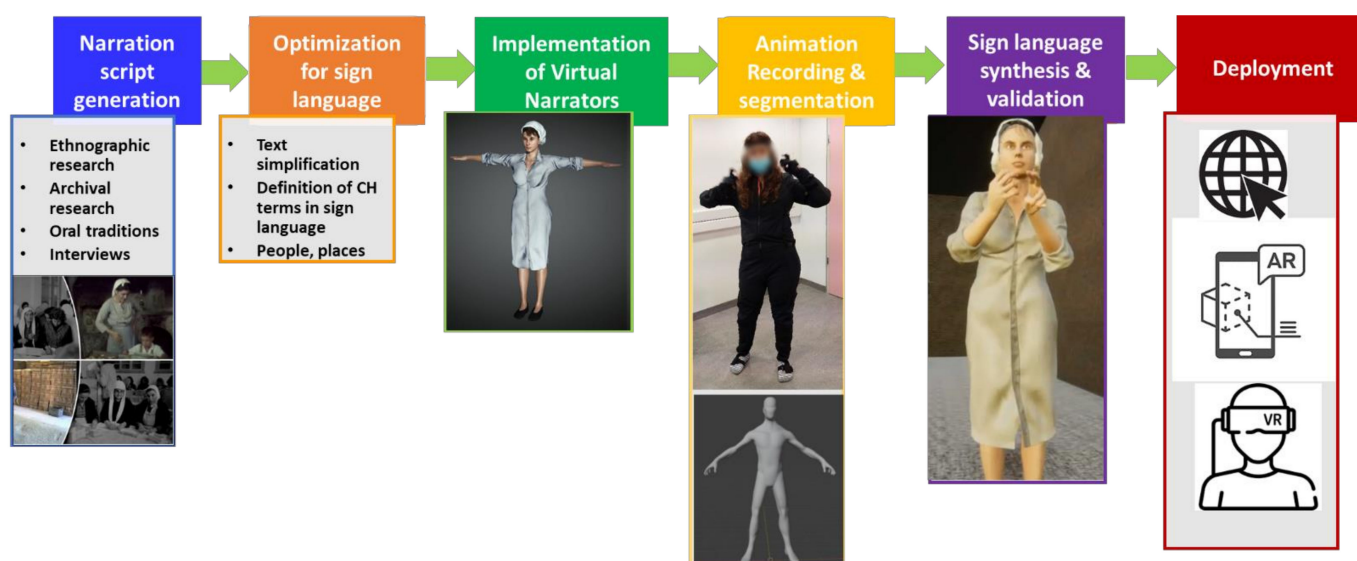


Figure 1. An example of a female VH (Stamatia).

The rest of this chapter dives deeper into the proposed methodology.

### 3.1. Narration Script Generation

The first step of the proposed method regards the research needed to create a narrative that could relate to various aspects of CH, such as an object, a collection, an historic event, a personal story, etc. In this sense, the research may include any scientific process that could include archival research, a study of resources, ethnographic research, oral traditions and testimonies, interviews, curatorial work, etc. The result is the collection of sufficient information for the creation of a narrative. The outcomes, as envisioned by this research work, are the text of the narration and the characteristics of the person narrating the story. Such characteristics may also include gender information, age, clothing, etc., and are useful for the implementation of the narrator VH.

### 3.2. Optimisation for Sign Language

In this step, the initial narration text that is meant to be used for oral narration must be reviewed and optimized by the sign language translators. This includes, if needed, text simplification which replaces text with simpler equivalents to simplify the sign language translation and enhance understanding. Although automated methods have been proposed (e.g., [67,68]), in this paper we propose that as sign language translators are part of the methodology they are also involved in the simplification of the text. This is also important as content for CH applications is more challenging because it usually contains specialized CH terms that may pose requirements in terms of translation. For such terms, in many cases, no translations exist so a more descriptive sign language translation is needed.

### 3.3. Implementation of Virtual Narrators

This step regards the implementation of the VH that will perform both the oral and the sign language narration. In this paper, no emphasis is put on the creation of VH as any form will suit the needs of this research work as long as it is compatible with several basic characteristics and as such could, for example, be represented by a humanoid, rigged, and skinned model. Any free or commercially available software can be employed for this reason. Of course, the software toolchain used will affect the realism of the final model as more advanced software solutions come with more features in terms of skinning, clothing, texturing, etc.

To this end, several guidelines can be provided. The VHs' bodies and clothes should be created to obtain one unified and optimized model, enhancing the visual impact of the characters with texture mapping and material editing. The 3D generation of the virtual bodies also has to take into consideration the total number of polygons used to create the meshes to keep a balance between the 3D real-time simulation restrictions and the skin deformation accuracy of the models. For VHs acting as conversational agents, an extra requirement involves the use of a blend-shape system for the facial animation, which is meant for working with software that supports on one hand the external BVH files for the animation of the body and, on the other hand, gives tools for controlling the facial animation.

### 3.4. Animation Recording and Segmentation

For animation recording, a Human Motion Digitization technique should be used that can capture motion from the entire body of the captured person but at the same time can capture hand gestures with high accuracy. Achieving the capture of accurate motion data is extremely important for ensuring a "readable" final result. Based on the authors' understanding and experience, a motion capture suit and smart gloves are a particularly satisfactory solution as they have the capacity to capture full-body motion, while, at the same time, the individual recording of the hand provides high-quality hand-movement recordings.

Once the narration animations were recorded, their segmentation was conducted [48] and the segments were exported in fbx format, using the HumanIK skeleton. This action creates a series of bones, body joints, and muscles and defines their rotations in the 3D

space over time. Segmentation is important for several reasons. First, it creates several isolated animations that can be validated individually. Second, in the case of an error, this is easily located and can be fixed by manually editing the animation. Third, if the error cannot be easily fixed, only the specific animation segment needs to be recaptured.

The segmentation process is conducted in collaboration with the sign language narrators and involves the isolation of key phrases for the recording that can be reused in the future and in the definition of the core narration part. The first part of the isolated “sentences” is used to create a vocabulary of sentences that can be reused in the future and thus do not have to be recorded for each narration (e.g., the introduction of the person acting as the narrator that could occur in different narration texts). Regarding the narration part, the output of the segmentation process is a collection of sentences in the form of motion data.

### 3.5. Sign Language Synthesis and Validation

The segmented animations from the previous step are used to create animation sequences combined with a VH model.

To do so in this research work, the import of the narration animations and the VHs into the Unity game engine is proposed. After that, add the VH to a scene and define an animator component to control their animations. The controller defines which animations the VH can perform, as well as when to perform them. Essentially, the controller is a diagram, which defines the animation states and the transitions among them. In Figure 2, an animation controller is shown where the VH initially performs an idle animation; it can then transit (arrows) to a state where the character introduces herself (“Self-introduction”) or to a state where she narrates a specific process (“Narration about Sifting Process”). At the end of this step, a VH with all the integrated animations of the narrative is available. In the case where the oral narration is also recorded, this is also integrated into the final model. For ease, a script in JSON format is created for the orchestration of animation sequences.



**Figure 2.** Simple animator controller for a narrator VH.

What is important, in this step, is the validation of synthetic sign language narrations by VHs before any deployment of a signing solution to a CHI. To this end, the sign language translators must preview the signed narration and perform a validation of the readability of the outcome. This is crucial considering that minor errors in the retargeting process may result in an unnatural experience for end users. For example, a false thumb position is something that may be detected by a sign language translator but not by a developer that does not have an adequate understanding of sign language gestures.

### 3.6. Deployment

Regarding deployment, the proposed methodology does not pose any constraints on the development platform/technology. The created signer VHs can be integrated into any 3D-enabled software technology compatible with the fbx format. Furthermore, the output can be easily ported to support WebGL technologies for web-based integration. Finally, VH animations can be rendered to video to be hosted by any other app that supports video rendering.

In the case where VHs are to be presented in a physical space, several issues should be considered. The first is the registration of the device in the space. This regards the possibility of the device to understand its location within the space. This can be conducted



by localization algorithms and the detection of image features that are located in fixed locations in the space. Upon localization calibration, the AR device can be moved to whichever location within the space without restrictions regarding its localization. The second is the correct placement of the VH in the physical space. This can be achieved through plane detection provided by SLAM algorithms and appropriate scaling of the VH.

For AR deployment another consideration that should be taken into account is the device compatibility with ARCore for android-based devices [69] and ARKit for iOS-based devices [70].

#### 4. Use Case

For the validation of the proposed methodology, we used one of the pilots of the Mingei project, Mastic. It considers the cultivation of mastic in the island of Chios and its usage as part of a production process.

##### 4.1. Use Case Context

In this context, the historical information on the use case involves the Chios Gum Mastic Growers Association which was an agricultural cooperation established in 1937 in Chora of the island of Chios.

VHs have been created in the context of Mingei and represents the actual workers of the association. They narrate stories from their personal lives, their work life, and their duties at the factory. Through them, the museum visitors can virtually travel back in time to that era and learn how people lived but also get to know about the mastic processing stages, the machines' functionality, and more. VHs are integrated into an AR application that involves the augmentation of the physical museum exhibits with the narrator VHs. As part of this research work, the proposed methodology was followed to enrich the VHs with sign language narrations.

##### 4.2. Technical Setup and Accessibility Requirements

The AR app is available to museum visitors through AR-enabled tablet devices (Samsung Galaxy Tab 5) that are mounted on floor bases, as shown in Figure 3. These allow the tablet to be moved in the following degrees of freedom.



**Figure 3.** An example of a tablet device mounted on the floor base at the museum.

- Vertical movement: This allows the adjustment of the height of the tablet to support interaction by visitors of varying ages and heights and to support people that are using a wheelchair.
- Tilt: Regardless of the height of the tablet's base, the tilt supports the adjustment of the field of view of the device's camera.
- Rotation: This allows the rotation of the tablet to target different exhibits in the museum in the case where the same device is used to augment a larger space than the one covered by its camera field of view.

The main scenario of the app includes narrations that are orally presented by the VHs, which can be seen through the camera of the device in the exhibition space.

#### *4.3. Narration Script and Optimisation for Sign language*

Basic research was conducted, taking into account archives, photographic and audio-visual documentation, and interviews. More specifically, the archival research focused on the vast archive that PIOP holds on mastic, its cultivation, the production processes of mastic products, and the historical and social material. Part of the PIOP archive is archival material and the machines that were acquired by the Chios Mastic Growers Association. Furthermore, literature regarding the mastic tree, mastic production, and the historical and social facts was acquired through essays developed by academics for PIOP for the purposes of the Chios Mastic Museum.

The audio-visual material studied included: (a) documentaries produced by television channels and cinematographers; (b) documentation videos of history researchers and ethnographers that cooperated with PIOP for the production and collection of material for the Chios Mastic Museum (videos include interviews with mastic producers, men and women that used to work at the Chios Mastic Growers Association); and (c) advertisement clips of the ELMA chewing gum.

Finally, the audio material studied included: (a) interviews with former employers at the Chios Mastic Growers Association on topics including chewing gum production, distillation, and the description of the factory and the machines; (b) radio advertisements of the ELMA chewing gum; (c) women singing traditional songs regarding mastic cultivation; and (d) recordings from the traditional feast of 'Agha'.

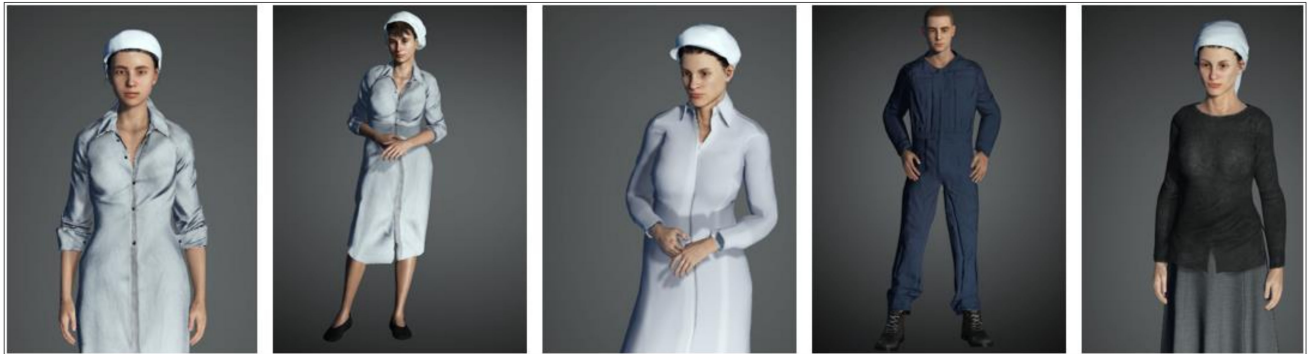
The profiles and stories of the VHs are a mix-and-match of the studied material. In creating the content of the profiles and the stories of the VH workers, the aim was a representation how life at the villages was, how the workers grew up in the village (i.e., education, agricultural life, leisure time, adolescence, and married life), what led them to seek work at the Association in Chora of Chios, how their working life in the Association was, and in which process(es) they worked in. All this information is divided into sections according to (a) family background and early and adult years of life; (b) work-life in the Association; and (c) an explanation of the processes.

Following the proposed methodology, in this step the narrations were reviewed by sign language translators and the required simplifications were made to the narration text to enhance understanding and improve the accuracy and efficiency of the sign language capturing.

#### *4.4. Implementation of Virtual Narrators and Optimisation*

The virtual human bodies and clothes were created to obtain one unified and optimized model, enhancing the visual impact of the characters with texture mapping and material editing. The 3D generation of the virtual bodies also has to take into consideration the total number of polygons used to create the meshes to keep a balance between the 3D real-time simulation restrictions and the skin deformation accuracy of the models. In this use case, avatars were created with a combination of different software: Adobe Fuse CC/Mixamo [71] was used for creating the body character, the clothes, the hair, and the rigging. The generated model was then imported into Autodesk 3DS max [72] for mesh geometry optimization.

Eight VH were created, seven of them representing people coming from different villages of southern Chios (also known as mastic villages), and one VH represents a woman coming from Sidirounta (a northern village). The age of the VHs was also defined as the middle age of the participants coming from villages of the same area. Figure 4 shows an example of the implemented VHs.



**Figure 4.** An example of VHs implemented for the use case.

For optimization purposes and in order to ensure smooth operation in variant device characteristics, the VH were created both in low and high quality. To this end, Autodesk 3DS max [72] was used for mesh geometry optimization. Manual methods, by using the editable poly tools, were preferred as they allowed keeping the regularity of the topology while the automatic methods generate a mess geometry which is not suitable for skin deformation or for a regular UV texture map generation. Using this approach, the high-resolution VHs have about ~40,000 triangles, while the low-quality VHs have about ~15,000 triangles. Furthermore, adaptations were made to the image textures and mapping methods to increase performance.

#### 4.5. Animation Recording and Segmentation

In this stage, the final narrations were available, and the sign language translators were requested to prepare one narration per recording session. For the acquisition of motion, the Rococo motion capture suit and smart gloves were used. The sign language translator was requested to wear the suit and then narrate in sign language the simplified narrative. Simultaneously, the recording was carried out using the equipment. Each session concluded with the narrator previewing the acquired narrations to ensure that the raw capture data previewed through a simplified skeleton were accurate. To make sure that minimum re-capturing occurred, the narrations were segmented into parts (paragraphs) and each part was captured and reviewed individually. Figure 5 shows examples of the sign language narrator in action.

Segmentation is the process where the recorded segments are further segmented into parts that can be translated to an instance of an animation. This process was used to create a set of exported motion segments that were then exported and used to implement the actual animations on each VH. Reusable phrases were also kept in the sign language dictionary for future use.



**Figure 5.** An example of a female VH (Stamatia).

#### 4.6. Sign Language Synthesis and Validation

In the use case, VHs employed already contained animation recordings for the oral presentation of the narrative. Thus, the process involved the addition of new animation and their grouping under the sign language narrative group. Furthermore, an extra animation sequence was implemented to support the sign language narrative.

The result was then used to implement a demonstrator application running in standalone mode through a desktop computer. This demo was used for the validation of the sign language synthesis by the sign language translators. Sessions were organized with their participation. The objectives of these sessions were to preview the resulting animations and locate any issues that reduced the understanding or even altered the meaning of the sign language translation. Any errors were corrected either by altering the animation itself or by retaking the entire animation if needed. In our case, only minor adjustments were needed to the animations, so no recapturing was performed.

#### 4.7. Mobile-AR App

VH narrators are hosted by a mobile app implemented in Unity3D. The application uses the google ARCore library for localizing the tablet device within the exhibition with the help of markers located in the exhibition (see Figure 6). Using the results of this localization, the unity app is capable of placing virtual markers in the exhibitions that upon activation present the VH associated with the marker. The positioning and scale of the VH are made using the plane detection capabilities provided by the implementation platform.





**Figure 6.** The localization process.

#### 4.8. Deployment and Preliminary Evaluation

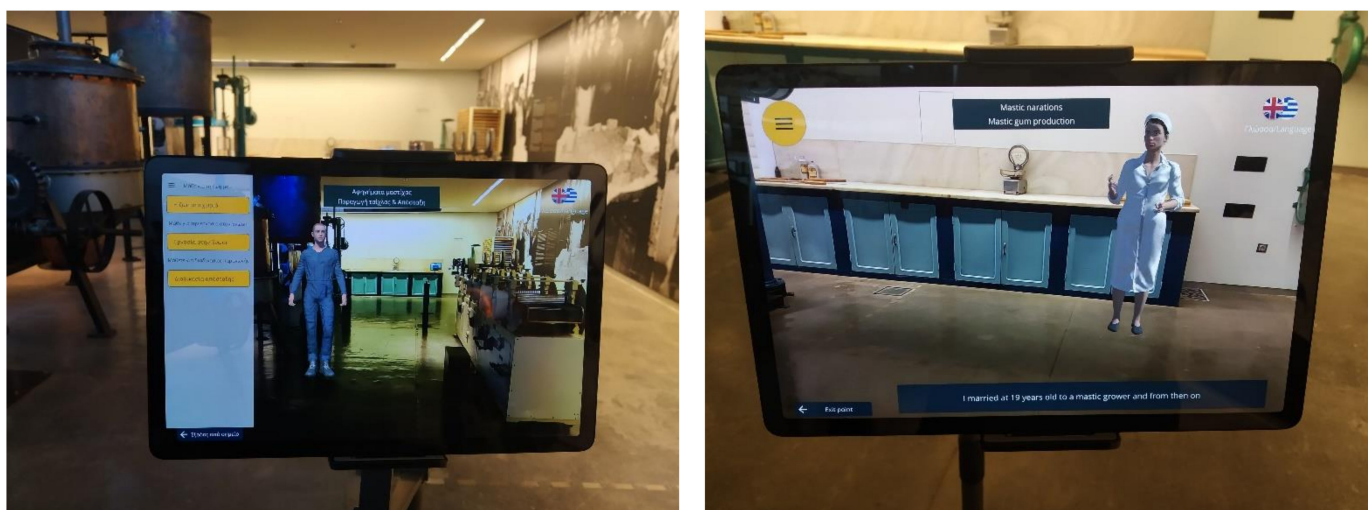
The deployment of the result was conducted in the context of an installation at the factory exhibition of the Chios mastic museum. The installation comprises four tablet devices mounted on four floor-mounted bases, each one at a different location within the museum. Each of the tablets is configured to recognize different spots in the exhibition (see Figure 7). Each spot is typically a machine used in the production process and the place of the appearance of a VH worker used to work with the specific machine.



**Figure 7.** View of the factory exhibition with the active spots as previewed through the tablet's screen.

Upon activation of a spot, the VH appears and provides different narration options. All narrations are available in plain language and sign language. Examples are presented in Figure 8.





**Figure 8.** Examples of VHs as seen from within the tablet seen inside the installation at the Mastic Museum of Chios.

The installation at the Mastic Museum of Chios was combined with school visits from two technical high schools of the island. During these visits, preliminary observations were made regarding the user experience by a usability expert. These observations regarded the overall usage of the system by users with no disability and were focused on user understanding of the provided UI and of the VH dialogues. Furthermore, the general satisfaction from using this form of storytelling in the museums was observed. From such observations, valuable feedback was received, and many issues were encountered and fixed on site. For example, the kids had the tendency to press multiple points of interest at once, thus masking the system lag while loading multiple animations. Regarding the value of the produced prototypes for the targeted population of users with disabilities, a separate evaluation is planned in the future which will guide us in exploring the future possibilities of extending and improving this research work.

## 5. Discussion

In this research work, a cost-effective methodology for the implementation of accessible VHs to act as storytellers in the context of online and on-site CH experiences is presented. The proposed methodology is rooted in advances in motion capture technologies and Virtual Human implementation, animation, and multi-device rendering. The proposed methodology could ultimately support a wide variety of users as (a) users using a wheelchair are supported through the height adaptation of the floor-mounted base; (b) users with cognitive impairments are supported through text simplification and animation recording of the simplified narration; (c) users with hearing impairments are supported through sign language recording; and (d) blind users will be supported through the standard narration option with the prerequisite that a screen reader is available to allow the user to initiate interaction with the system. For this purpose, the narration script also contains non-visual information meant to be orally presented by the screen reader to give localization information to the users of the application. An example of such a script is presented in Appendix A of this paper.

Overall, the advantages of the proposed methodology can be analyzed in the following main axis: (a) technological needs and the needs of acquiring specialized resources and (b) impact in terms of cultural appreciation by people with disabilities.

Advances relevant to the first axis and concerning the state of the art can be summarized as follows. Initially, this is a cost-effective methodology in terms of technical equipment and resources. Only a motion capture suit and gloves are required, with a cost of less than 3k. In terms of resources, a 3D development and animation team have only

the need for the addition of a sign language translator, who can be employed only for the needs of each project. Furthermore, the proposed methodology allows the creation of a reusable sign language phrases vocabulary that can be reused across projects and thus reduce the cost of acquiring sign language recordings. Additionally, the retargeting of sign language recording to VHS does not limit the usage of a specific sign language translator to being visually compatible with previous recordings (e.g., people are used to seeing the same person as the sign language translator in the context of, e.g., tv news). A mixture of animation recorded by different sign language translators is possible and thus more cases can be supported than the ones initially captured.

Advances relevant to the second axis can be summarized as follows. The proposed solution can greatly enhance the visiting experience for people with hearing disabilities by providing narrators that have a dual meaning. They can support the suspension of disbelief by supporting mental travels through storytelling, and at the same time, the storytelling happens in a language that people with hearing disabilities can understand, thus increasing the emotional impact of the storytelling for this target group.

Regarding the limitations of this research work, there are some non-technical and some technical. First, the proposed solution requires that the script is very well prepared and optimized from the beginning as changes may require recapturing part of the script. Second, and this is the same for all sign language solutions, there is no universal sign language, which means that the same translation requirements are needed for sign languages too. Finally, from a technical perspective, the proposed approach does not yet cover facial expressions complementing sign languages.

Regarding future improvements, several directions can be followed. First, for completeness of the proposed approach, we intend to integrate a face-capturing solution in our toolchain to support face morphing in sign language animations. Second, although the app is accessible for visually impaired users through audio narrations, some experimentation is needed with the screen-reading applications for mobile devices to ensure seamless interaction with our application. Finally, the entire solution could be packaged in the future in an all-in-one solution for narrations by VHS where the VH and the animation are configured in a visual editor; thus, there would be no need to write dedicated software for each usage scenario.

**Author Contributions:** Conceptualization, X.Z. and N.P.; data curation, D.K., E.T., M.M. and C.R.; funding acquisition, N.P. and X.Z.; investigation, D.K.; methodology, N.P. and X.Z.; project administration, N.P. and X.Z.; resources, D.K. and M.M.; software, A.P., M.F. and E.Z.; supervision, N.P., X.Z. and E.Z.; validation, D.K. and I.A.; visualization, A.P., N.P. and X.Z.; writing—original draft, N.P.; writing—review and editing, N.P., X.Z., M.F., D.K., E.T. and C.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work has been conducted in the context of the Mingei project that has received funding from the European Union’s Horizon 2020 research and innovation program under grant agreement No 822336.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data available upon request.

**Acknowledgments:** The authors would like to thank the Chios Mastic Museum of the Piraeus Bank Group Cultural Foundation for its contribution to the preparation of the exhibition and MiraLab SRL for implementing the VH used in the use case.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

In this annex we present an example of a narration script together with embedded information for the developers of the AR application. The information includes visual

content that should be combined with the presentation and non-visual content that should be narrated when presenting the application to users with visual impairments.

Virtual Human: Woman in her late 40s, wearing her clothes (early 1980s).

Location of the Virtual Human: Cleaning table: standing beside the table for cleaning mastic.

Non-visual information: torso shot of the woman seated with hands on the table working and the mastic on the table near her; she turns her head towards you to greet you.

Text: Hello! I am Irini.

Visual and non-visual information: Select: Learn about Irini's life → Select: Life 1 or Life 2 or Life 3/Present process/Learn about work life in the Association → Select: Work 1 or Work 2 or Work 3

Learn about Irini's life [Life 1]: I come from the village of Nenita. Both my parents were migrants from Asia Minor after the catastrophe of Smyrna. My father was a fisherman, and my mother did the household and cultivated mastic trees. My brother and I never had the chance to properly go to school because firstly, we had to help our mother in the fields, and later, the Second World War started. During the War, it was also hard. Not for everyone, of course. For example, the wealthier families of the village were even allowed by the Germans to leave their lights on until late at night when they had a feast. For our family, I thank God that my father could at least exchange fish for seeds in other villages so that we could cultivate more things to eat.

Learn about Irini's life [Life 2]: After the War, my brother decided to become a sailor because our income from mastic was not enough. Many young men did that back then because they earned a lot and thus were able to send money to their families. As for myself, I decided to learn sewing in Chora. When I finished my courses, I became a dressmaker in our village and mainly made clothes, such as trousers, dresses, and shirts from linen and silk. But in the meantime, I also helped my mother in the mastic fields.

Learn about Irini's life [Life 3]: At 21 years old, I got married to the son of the coffee shop owner who also cultivated mastic and olive trees. We met in the fields while collecting mastic because my mother and I were going as daneikes to help in their fields; they were our neighbors. We started flirting in the fields, and after a year, we got married and had two children. I continued to be a part-time dressmaker and also helped my husband in the fields when I could. But unfortunately, my husband died after ten years. Such a tragedy... It was difficult to make it on my own at the village with two young children, so I asked my cousin, who was already working at the Chios Gum Mastic Growers Association at Chora, if they would also take me as a worker. The warehouse manager understood my situation and, as the Association always wants to help people from the mastic villages, he hired me. I then took the children with me, and we stayed at the house of my aunt in Chora which was just fifteen minutes walk from the Association. My mother stayed in the village and took care of our fields along with the help of the daneikes, but on the weekends and my days off, I am still going back to the village to help.

Present process: [Non-visual information: You are in a room consisting of two big tables with mastic quantities on them as exhibited in the museum;] This is the room for cleaning mastic after it has been washed. Cleaning is about removing any attached foreign matter from the mastic tear with the help of a knife with a sharp edge. With this tool, we can superficially dig the tear in a precise manner.

Learn about work life in the Association [Work 1]: About a hundred women work here as temporary workers; that means, we work for two consecutive weeks and then we stay at home for one week. I started from this task when I first came here because many new workers start from here, but as the years went by, I remained because I clean mastic well and rather fast. In the beginning, my shift was from 6 a.m. until 2 p.m. because of the children and their school schedule, but after they finished, I also worked on the afternoon shift from 2 p.m. to 10 p.m.

Learn about work life in the Association [Work 2]: [Non-visual information: she also sits on the table, an elder woman working beside her and then falling asleep on the table]

I am not the older one here. Many older women remain at this task because you can be seated and there is no load to carry. But, of course, because of age, they also get tired quicker, and sometimes, one or two have been spotted taking a nap on the table during midday. The younger ones take care of them and watch out if the supervisor is coming to wake them up [Non-visual information: she nudges her to wake up]. Not only women from mastic villages work here. There are also some coming from northern villages where mastic is not cultivated. We (women from mastic villages) already know how to clean mastic because most of us are also mastic growers, and we also do it before delivering our production at the Association. But the other women have no previous contact with mastic. The supervisor shows them the process, and that is why they usually clean small tears which is easier. Women from mastic villages usually clean the large tears, ‘pittas’, because they are more difficult and precious.

Learn about work life in the Association [Work 3]: The supervisor arranges the quantities of mastic every morning according to the orders that the Association has, and at the end of the shift, she weights our production and writes it on our record. As you can imagine, work can become very competitive. When George Stagkoulis was the president of the Association until 1978, he and the supervisors were unofficially urging the workers to work on a bet on who would produce the most. It was fun but also tiring sometimes. I remember when Stagkoulis was coming to the room to check on us, the supervisor warned us so that we would stop talking; sometimes, we were also singing or saying poems. But Stagkoulis should have never witnessed that. He was strict as a boss, but he was fair and a good man. Everyone knew him and the work he has done for the mastic villages.

## References

1. Weisen, M. How accessible are museums today? In *Touch in Museums*; Routledge: Oxfordshire, UK, 2020; pp. 243–252.
2. Caldwell, B.; Cooper, M.; Reid, L.G.; Vanderheiden, G.; Chisholm, W.; Slatin, J.; White, J. Web content accessibility guidelines (WCAG) 2.0. *WWW Consort. (W3C)* **2008**, *290*, 1–34.
3. Gunderson, J. W3C user agent accessibility guidelines 1.0 for graphical Web browsers. *Univ. Access Inf. Soc.* **2004**, *3*, 38–47. [[CrossRef](#)]
4. Oikonomou, T.; Kaklanis, N.; Votis, K.; Kastori, G.E.; Partarakis, N.; Tzovaras, D. Waat: Personalised web accessibility evaluation tool. In Proceedings of the International Cross-Disciplinary Conference on Web Accessibility, Hyderabad Andhra Pradesh, India, 28–29 March 2011; pp. 1–2.
5. Doulgeraki, C.; Partarakis, N.; Mourouzis, A.; Stephanidis, C. A development toolkit for unified web-based user interfaces. In Proceedings of the International Conference on Computers for Handicapped Persons, Linz, Austria, 9–11 July 2008; Springer: Berlin/Heidelberg, Germany, 2008; pp. 346–353.
6. Mourouzis, A.; Partarakis, N.; Doulgeraki, C.; Galanakis, C.; Stephanidis, C. An accessible media player as a user agent for the web. In Proceedings of the International Conference on Computers for Handicapped Persons, Linz, Austria, 9–11 July 2008; Springer: Berlin/Heidelberg, Germany, 2008; pp. 474–481.
7. Leas, D.; Persoon, E.; Soiffer, N.; Zacherle, M. Daisy 3: A Standard for Accessible Multimedia Books. *IEEE MultiMedia* **2008**, *15*, 28–37. [[CrossRef](#)]
8. Partarakis, N.; Klironomos, I.; Antona, M.; Margetis, G.; Grammenos, D.; Stephanidis, C. Accessibility of cultural heritage exhibits. In Proceedings of the International Conference on Universal Access in Human-Computer Interaction, Toronto, ON, Canada, 17 July 2016; Springer: Cham, Switzerland, 2016; pp. 444–455.
9. Partarakis, N.; Antona, M.; Zidianakis, E.; Stephanidis, C. Adaptation and Content Personalization in the Context of Multi User Museum Exhibits. In Proceedings of the 1st Workshop on Advanced Visual Interfaces for Cultural Heritage co-located with the International Working Conference on Advanced Visual Interfaces (AVI\* CH), Bari, Italy, 7–10 June 2016; pp. 5–10.
10. Partarakis, N.; Antona, M.; Stephanidis, C. Adaptable, personalizable and multi user museum exhibits. In *Curating the Digital*; England, D., Schiphorst, T., Bryan-Kinns, N., Eds.; Springer: Cham, Germany, 2016; pp. 167–179.
11. Machidon, O.M.; Duguleana, M.; Carrozzino, M. Virtual humans in cultural heritage ICT applications: A review. *J. Cult. Heritage* **2018**, *33*, 249–260. [[CrossRef](#)]
12. Addison, A. Emerging trends in virtual heritage. *IEEE MultiMedia* **2000**, *7*, 22–25. [[CrossRef](#)]
13. Karuzaki, E.; Partarakis, N.; Patsiouras, N.; Zidianakis, E.; Katzourakis, A.; Pattakos, A.; Kaplanidi, D.; Baka, E.; Cadi, N.; Magnenat-Thalmann, N.; et al. Realistic Virtual Humans for Cultural Heritage Applications. *Heritage* **2021**, *4*, 4148–4171. [[CrossRef](#)]
14. Sylaiou, S.; Kasapakis, V.; Gavalas, D.; Dzardanova, E. Avatars as storytellers: Affective narratives in virtual museums. *Pers. Ubiquitous Comput.* **2020**, *24*, 829–841. [[CrossRef](#)]

15. Partarakis, N.; Doulgeraki, P.; Karuzaki, E.; Adami, I.; Ntoa, S.; Metilli, D.; Bartalesi, V.; Meghini, C.; Marketakis, Y.; Kaplanidi, D.; et al. Representation of socio-historical context to support the authoring and presentation of multimodal narratives: The Mingei Online Platform. *J. Comput. Cult. Herit.* **2022**, *15*, in press. [CrossRef]
16. Geigel, J.; Shitut, K.S.; Decker, J.; Doherty, A.; Jacobs, G. The digital docent: Xr storytelling for a living history museum. In Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology, Ottawa, ON, Canada, 1–4 November 2020; pp. 1–3.
17. Dzardanova, E.; Kasapakis, V.; Gavalas, D.; Sylaiou, S. Exploring aspects of obedience in VR-mediated communication. In Proceedings of the 2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX), Berlin, Germany, 5–7 June 2009; pp. 1–3.
18. Carrozzino, M.; Colombo, M.; Tecchia, F.; Evangelista, C.; Bergamasco, M. Comparing different storytelling approaches for virtual guides in digital immersive museums. In Proceedings of the International Conference on Augmented Reality, Virtual Reality and Computer Graphics, Otranto, Italy, 24–27 June 2018; Springer: Cham, Switzerland, 2018; pp. 292–302.
19. Kacorri, H. *TR-2015001: A Survey and Critique of Facial Expression Synthesis in Sign Language Animation*; CUNY Academic Works: Brooklyn, NY, USA, 2015.
20. Huenerfauth, M. *Learning to Generate Understandable Animations of American Sign Language*; Rochester Institute of Technology: Rochester, NY, USA, 2014.
21. Lu, P.; Huenerfauth, M. Collecting and evaluating the CUNY ASL corpus for research on American Sign Language animation. *Comput. Speech Lang.* **2014**, *28*, 812–831. [CrossRef]
22. Heloir, A.; Kipp, M. EMBR—A realtime animation engine for interactive embodied agents. In Proceedings of the 9th International Conference on Intelligent Agents, Amsterdam, Netherlands, 10–12 September 2009; pp. 393–404.
23. Jennings, V.; Elliott, R.; Kennaway, R.; Glauert, J. Requirements for a signing avatar. In Proceedings of the 4th Workshop on the Representation and Processing of Sign Languages: Corporal and Sign Language Technologies, Valletta, Malta, 17–23 May 2010; pp. 33–136.
24. Huenerfauth, M.; Zhao, L.; Gu, E.; Allbeck, J. Evaluation of American Sign Language Generation by Native ASL Signers. *ACM Trans. Access. Comput.* **2008**, *1*, 1–27. [CrossRef]
25. Elliott, R.; Glauert, J.R.W.; Kennaway, J.R.; Marshall, I.; Safar, E. Linguistic modelling and language-processing technologies for Avatar-based sign language presentation. *Univers. Access Inf. Soc.* **2007**, *6*, 375–391. [CrossRef]
26. Fotinea, S.E.; Efthimiou, E.; Caridakis, G.; Karpouzis, K. A knowledge-based sign synthesis architecture. *Univ. Access Inf. Soc.* **2008**, *6*, 405–418. [CrossRef]
27. Segundo, S. Design, development and field evaluation of a Spanish into sign language translation system. *Pattern Anal. Appl.* **2012**, *15*, 203–224. [CrossRef]
28. Kennaway, R.; Glauert, J.R.W.; Zwitserlood, I. Providing signed content on the Internet by synthesized animation. *ACM Trans. Comput. Interact.* **2007**, *14*, 15. [CrossRef]
29. Huenerfauth, M.; Lu, P.; Rosenberg, A. Evaluating importance of facial expression in american sign language and pidgin signed english animations. In Proceedings of the 13th International ACM SIGACCESS Conference on Computer and Accessibility, Dundee, Scotland, 24–26 October 2011; pp. 99–106.
30. Gibet, S.; Courty, N.; Duarte, K.; le Naour, T. The SignCom system for data-driven animation of interactive virtual signers: Methodology and Evaluation. *ACM Trans. Interact. Intell. Syst. (TiiS)* **2011**, *1*, 1–23. [CrossRef]
31. McDonald, J.; Alkoby, K.; Carter, R.; Christopher, J.; Davidson, M.; Ethridge, D.; Furst, J.; Hinkle, D.; Lancaster, G.; Smallwood, L.; et al. A Direct Method for Positioning the Arms of a Human Model. In Proceedings of the Graphics Interface 2002, Calgary, AB, Canada, 27–29 May 2002; pp. 99–106.
32. Signing Avatars. Available online: [http://www.bbcworld.com/content/clickonline\\_archive\\_35\\_2002.asp](http://www.bbcworld.com/content/clickonline_archive_35_2002.asp) (accessed on 10 August 2021).
33. Grieve-Smith, A. A Demonstration of Text-to-Sign Synthesis, Presented at the Fourth Workshop on Gesture and Human-Computer Interaction, London. 2001. Available online: [www.unm.edu/~grvsmth/signsynth/gw2001/](http://www.unm.edu/~grvsmth/signsynth/gw2001/) (accessed on 5 September 2021).
34. Zwiterslood, I.; Verlinden, M.; Ros, J.; van der Schoot, S. Synthetic Signing for the Deaf: eSIGN. In Proceedings of the Conference and Workshop on Assistive Technologies for Vision and Hearing Impairment, CVHI 2004, Granada, Spain, 29 June–2 July 2004.
35. McDonald, J.; Wolfe, R.; Schnepf, J.; Hochgesang, J.; Jamrozik, D.G.; Stumbo, M.; Berke, L.; Bialek, M.; Thomas, F. An automated technique for real-time production of lifelike animations of American Sign Language. *Univers. Access Inf. Soc.* **2015**, *15*, 551–566. [CrossRef]
36. Karpouzis, K.; Caridakis, G.; Fotinea, S.-E.; Efthimiou, E. Educational resources and implementation of a Greek sign language synthesis architecture. *Comput. Educ.* **2007**, *49*, 54–74. [CrossRef]
37. Elliott, R.; Glauert, J.R.W.; Kennaway, J.R.; Marshall, I. The development of language processing support for the ViSiCAST project. In Proceedings of the fourth international ACM conference on Assistive technologies, Arlington, VI, USA, 13–15 November 2000.
38. Braffort, A.; Filhol, M.; Delorme, M.; Bolot, L.; Choisier, A.; Verrecchia, C. KAZOO: A sign language generation platform based on production rules. *Univers. Access Inf. Soc.* **2015**, *15*, 541–550. [CrossRef]
39. Adamo-Villani, N.; Wilbur, R.B. ASL-Pro: American Sign Language Animation with Prosodic Elements. In *Universal Access in Human-Computer Interaction. Access to Interaction*; Antona, M., Stephanidis, C., Eds.; Springer International Publishing: Cham, Switzerland, 2015; pp. 307–318.



40. Huenerfauth, M.; Kacorri, H. Release of Experimental Stimuli and Questions for Evaluating Facial Expressions in Animations of American Sign Language. In Proceedings of the the 6th Workshop on the Representation and Processing of Sign Languages: Beyond the Manual Channel, The 9th International Conference on Language Resources and Evaluation (LREC), Reykjavik, Iceland, 31 May 2014.
41. Ebling, S.; Glauert, J. Building a Swiss German Sign Language avatar with JASigning and evaluating it among the Deaf community. *Univers. Access Inf. Soc.* **2015**, *15*, 577–587. [[CrossRef](#)]
42. Segouat, J.; Braffort, A. Toward the Study of Sign Language Coarticulation: Methodology Proposal. *IEEE* **2009**, 369–374. [[CrossRef](#)]
43. Duarte, K.; Gibet, S. Heterogeneous data sources for signed language analysis and synthesis: The signcom project. In Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC 2010) (Vol. 2, pp. 1–8), Valletta, Malta, 17–23 May 2010; European Language Resources Association: Luxembourg, 2010.
44. Huenerfauth, M.; Marcus, M.; Palmer, M. Generating American Sign Language Classifierpredicates for English-to-ASL Machine Translation. Ph.D. Thesis, University of Pennsylvania, Philadelphia, PA, USA, 2006.
45. Kacorri, H.; Huenerfauth, M. Continuous Profile Models in ASL Syntactic Facial Expression Synthesis. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Berlin, Germany, 7–12 August 2016. [[CrossRef](#)]
46. Ebling, S.; Glauert, J. Exploiting the full potential of JASigning to build an avatar signing train announcements. In Proceedings of the Third International Symposium on Sign Language Translation and Avatar Technology, Chicago, IL, USA, 18–19 October 2013. [[CrossRef](#)]
47. Al-khazraji, S.; Berke, L.; Kafle, S.; Yeung, P.; Huenerfauth, M. Modeling the Speed and Timing of American Sign Language to Generate Realistic Animations. In Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility, Galway, Ireland, 22–24 October 2018; pp. 259–270.
48. Partarakis, N.; Zabulis, X.; Chatziantoniou, A.; Patsiouras, N.; Adami, I. An Approach to the Creation and Presentation of Reference Gesture Datasets, for the Preservation of Traditional Crafts. *Appl. Sci.* **2020**, *10*, 7325. [[CrossRef](#)]
49. Solina, F.; Krapež, S. Synthesis of the sign language of the deaf from the sign video clips. *Electrotech. Rev.* **1999**, *66*, 260–265.
50. Bachmann, E.R.; Yun, X.; McGhee, R.B. Sourceless tracking of human posture using small inertial/magnetic sensors. In Proceedings of the 2003 IEEE International Symposium on Computational Intelligence in Robotics and Automation. Computational Intelligence in Robotics and Automation for the New Millennium (Cat. No. 03EX694), Kobe, Japan, 16–20 July 2003; IEEE: Manhattan, NY, USA, 2003; Volume 2, pp. 822–829.
51. Brigante, C.M.N.; Abbate, N.; Basile, A.; Faulisi, A.C.; Sessa, S. Towards Miniaturization of a MEMS-Based Wearable Motion Capture System. *IEEE Trans. Ind. Electron.* **2011**, *58*, 3234–3241. [[CrossRef](#)]
52. Madgwick, S. *An Efficient Orientation Filter for Inertial and Inertial/Magnetic Sensor Arrays*; Report x-io and University of Bristol (UK): Bristol, UK, 2010; Volume 25, pp. 113–118.
53. Lu, P.; Huenerfauth, M. Collecting a motion-capture corpus of American Sign Language for data-driven generation research. In Proceedings of the NAACL HLT 2010 Workshop on Speech and Language Processing for Assistive Technologies, Los Angeles, CA, USA, 5 June 2010; Association for Computational Linguistics: Stroudsburg, PA, USA, 2010; pp. 89–97.
54. Gibet, S. Building French Sign Language Motion Capture Corpora for Signing Avatars. In Proceedings of the Workshop on the Representation and Processing of Sign Languages: Involving the Language Community, LREC, Miyazaki, Japan, 30 May 2018.
55. Jedlička, P. Sign Language Motion Capture Database Recorded by One Device. In *Studentská Vědecká Konference: Magisterské a Doktorské Studijní Programy, Sborník Rozšířených Abstraktů, Květen 2018*; Západočeská univerzita v Plzni: Plzni, Czech Republic, 2019.
56. Havasi, L.; Szabó, H.M. A motion capture system for sign language synthesis: Overview and related issues. In Proceedings of the EUROCON 2005-The International Conference on “Computer as a Tool”, Belgrade, Serbia, 21–24 November 2005; IEEE: Manhattan, NY, USA, 2005; Volume 1, pp. 445–448.
57. Benchiheub, M.; Berret, B.; Braffort, A. Collecting and Analysing a Motion-Capture Corpus of French Sign Language. In Proceedings of the 7th LREC Workshop on the Representation and Processing of Sign Languages: Corpus Mining, Portorož, Slovenia, 28 May 2016.
58. Barczak, A.L.C.; Reyes, N.H.; Abastillas, M.; Piccio, A.; Susnjak, T. A new 2D static hand gesture colour image dataset for ASL gestures. *Res. Lett. Inf. Math. Sci.* **2011**, *15*, 12–20.
59. Oliveira, M.; Chatbri, H.; Ferstl, Y.; Farouk, M.; Little, S.; O’Connor, N.E.; Sutherland, A. *A dataset for irish sign language recognition*; Doras: Dublin, Ireland, 2017.
60. Forster, J.; Schmidt, C.; Hoyoux, T.; Koller, O.; Zelle, U.; Piater, J.H.; Ney, H. RWTH-PHOENIX-Weather: A Large Vocabulary Sign Language Recognition and Translation Corpus. In Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC’12), Istanbul, Turkey, 21–27 May 2012; pp. 3785–3789.
61. Athitsos, V.; Neidle, C.; Sclaroff, S.; Nash, J.; Stefan, A.; Yuan, Q.; Thangali, A. The american sign language lexicon video dataset. In Proceedings of the 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Anchorage, AK, USA, 23–28 June 2008; IEEE: Manhattan, NY, USA, 2008; pp. 1–8.
62. SIGNUM. Available online: <https://www.bas.uni-muenchen.de/Bas/SIGNUM/> (accessed on 2 September 2021).
63. Oszust, M.; Wysocki, M. Polish sign language words recognition with kinect. In Proceedings of the 2013 6th International Conference on Human System Interactions (HSI), Sopot, Poland, 6–8 June 2013; IEEE: Manhattan, NY, USA, 2013; pp. 219–226.

64. Conly, C.; Doliotis, P.; Jangyodsuk, P.; Alonzo, R.; Athitsos, V. Toward a 3D body part detection video dataset and hand tracking benchmark. In Proceedings of the 6th International Conference on Pervasive Technologies Related to Assistive Environments, Rhodes, Greece, 29–31 May 2013; pp. 1–6.
65. Kapuscinski, T.; Oszust, M.; Wysocki, M.; Warchol, D. Recognition of hand gestures observed by depth cameras. *Int. J. Adv. Robot. Syst.* **2015**, *12*, 36. [[CrossRef](#)]
66. Zabulis, X.; Meghini, C.; Partarakis, N.; Beisswenger, C.; Dubois, A.; Fasoula, M.; Nitti, V.; Ntoa, S.; Adami, I.; Chatziantoniou, A.; et al. Representation and preservation of Heritage Crafts. *Sustainability* **2020**, *12*, 1461. [[CrossRef](#)]
67. Aluísio, S.M.; Specia, L.; Pardo, T.A.; Maziero, E.G.; Fortes, R.P. Towards Brazilian Portuguese automatic text simplification systems. In Proceedings of the eighth ACM symposium on Document engineering, Sao Paulo, Brazil, 16–19 September 2008; pp. 240–248. [[CrossRef](#)]
68. Alonzo, O.; Seita, M.; Glasser, A.; Huenerfauth, M. Automatic Text Simplification Tools for Deaf and Hard of Hearing Adults: Benefits of Lexical Simplification and Providing Users with Autonomy. In Proceedings of the CHI '20: CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020. [[CrossRef](#)]
69. ARCore Supported Devices. Available online: <https://developers.google.com/ar/devices> (accessed on 9 December 2021).
70. ARKit Supported Devices. Available online: <https://developer.apple.com/library/archive/documentation/DeviceInformation/Reference/iOSDeviceCompatibility/DeviceCompatibilityMatrix/DeviceCompatibilityMatrix.html> (accessed on 9 December 2021).
71. Adobe Systems Incorporated. Mixamo. Available online: <https://www.mixamo.com/#/?page=1&query=Y-Bot&type=Character> (accessed on 30 August 2021).
72. ds-Max. Available online: <https://www.autodesk.fr/products/3ds-max> (accessed on 10 September 2021).