Show, Don't Tell: Retrieving Cultural Assets Via Gestures

Sven HELMER [a,1], Gerhard GLÜHER [b] and Christian UPMEIER [b]
[a] *Free University of Bozen-Bolzano, Faculty of Computer Science, 39100 Bolzano, Italy*
[b] *Free University of Bozen-Bolzano, Faculty of Design and Art, 39100 Bolzano, Italy*

Abstract. Currently, almost all information retrieval systems utilize a textual interface for users to express their information needs. While this works perfectly well for searching text document collections, there are domains in which textualizing a query or context is much harder. One of these domains is the preservation of cultural heritage practices in the form of artifact usage, which involves the capture and representation of intangible assets. We have started work on a project in which we want to focus on the domain of (hand-held) tools and their use, going beyond a description in the form of text or 2D images. Our aim is to record the gestures necessary to operate the tools and to make a collection of tools accessible with this information, providing a much richer context.

Keywords. cultural heritage, information retrieval, search via gestures, human computer interaction, hand-based interfaces

1. Introduction

The importance of preserving cultural heritage has been steadily on the rise over the past few decades. Computer science is contributing to this effort by providing platforms for digitizing, modeling, and visualizing artifacts or archaeological sites and making them accessible to a wider audience [11]. While this is interesting for imageable objects such as pieces of art, researchers have started looking beyond this: for instance, by developing haptic devices allowing users to feel the forms of sculptures [2]. Rosner et al. sketch a vision that goes even further: the digitization of cultural practices [19]. The significance of this has also been acknowledged by UNESCO, who have published a "List of Intangible Cultural Heritage in Need of Urgent Safeguarding" on their web site [23]. One example is the textile art on the island of Taquile, which revolves around not only preserving the textiles themselves, but also the manual weaving techniques used to create them.

We plan to concentrate our efforts on hand-operated tools, which are so common in our daily lives that we rarely pay attention to them. Nevertheless, they are intriguing for they perfectly combine appearance with dense, sometimes complex inherent organization. Some of these objects present a basic form that developed over decades or even centuries with little variations. The grasping and handling of these tools for proper use – to make them work – add dimension: for example, André Leroi-Gourhan talks of the

[1]Corresponding Author: Sven Helmer, Free University of Bozen-Bolzano, Faculty of Computer Science, 39100 Bolzano, Italy; E-mail: shelmer@inf.unibz.it

intermeshing of tools and motive gestures [13]. The focus on hand-operated tools, especially on those from the pre-industrial age, also makes sense from the point of view of our regional context. We have access to local museums with excellent collections of tools from the period of the 18th to the beginning of the 20th century. Industrialization in Austria, which included South Tyrol in this period, occurred quite late, due to a reluctance of the aristocracy to modernization. In rural areas, the production of goods in artisan workshops or small manufactories continued up to World War I.

Similar to Rosner et al. [19] we would also like to go beyond the modeling and mere visualization of artifacts. Searching through collections is currently done mainly via a keyword search, meaning that every object has to have a textual description of each of its features to make it possible to use a text-based retrieval system to access a collection. However, words are not always the ideal way to describe the information needs of a user, especially when it comes to activities related to objects such as hand-operated tools. For example, trying to describe to someone in words how to uncork a bottle using a corkscrew is rather difficult, while demonstrating it is much easier. Also, as David Anderson notes in [1]: "... sometimes the most significant features of an object, and those that we would be most concerned to preserve, are not tangible features at all, but rather lie outside of the materiality of the object."

We want to make collections of objects accessible by describing the way they are handled. A user describes with gestures how an item is used or moved around, which can be accomplished with the help of data gloves, other haptic input devices, kinetic user interfaces, or hand tracking technology such as a Leap Motion controller [24]. In addition, we envision the presentation of search results to include more context than just diagrams, pictures, or textual descriptions of objects. Ideally, there would be a 3D animation of the mechanism, a video clip of actual usage, and maybe even a haptic experience through motion feedback.

While computer science can provide the tools to enable the digitization of cultural assets and practices, it cannot give answers to questions concerning the overall philosophy of pursuing an approach like this. Involving the field of cultural and art history, as well as the field of design history and applied design, we focus on two main questions. First, how can digitized collections of objects be enhanced by visual and tactile, even interactive features for augmented information conveyance on cultural heritage? Second, how can historic product development be transposed in a modern discourse of knowledge transfer and thus be of interest even for contemporary product design?

2. State of the Art

Current digital collections of cultural objects (either operating offline or online) generally follow similar principles. The objects are categorized by basic physical features, such as material, form, and size, and by descriptions, such as texts and images, and the system offers structured search functionality, e.g. with keywords or simple image matching. Although this kind of interface works well with, e.g. books or two-dimensional artwork, it is less suitable when dealing with objects having specific modes of operation in addition to their three-dimensional appearance. For instance, (hand-operated) tools and their use are an interesting example of cultural assets whose functional aspects are either completely missing or at least underrepresented in current digital collections.

2.1. Computer Science

While there is interesting research work in the area of computer science trying to go beyond keyword or feature search, we are not aware of any involving gestures to describe the handling of artifacts that go further than moving or placing them in a virtual space. In the following we give a brief overview of existing techniques. Ichida et al. present a method for interactive retrieval of virtual 3D shapes using physical objects [8]. Using a set of interlocking cubes, a user can build physical objects describing the shapes they are searching for. Although this certainly is a novel way to describe queries, it lacks descriptions of handling the objects. Henriques et al. create an immersive environment to visualize 3D object retrieval [7]. However, their technique is mainly about displaying the retrieved objects by distributing them in a virtual space according to their similarity. Moustakas et al. allow a user to sketch primitive objects which are then used as an input parameter for the search [15]. An interesting feature of this system is that a user sketches the object in 3D, rather than producing a 2D drawing. The authors propose different input devices for describing 3D objects, such as a camera for real-time tracking, an air mouse, or even a data glove. However, the system is only interested in the position of a hand, not in gestures or more subtle movements. "Masterpiece" is a sophisticated multi-modal application framework allowing a user to search for and manipulate virtual 3D objects using input as diverse as speech, gesture, and sketch recognition [17]. While this system includes gesture recognition, it is not used in the actual search for the objects, but in the manipulation and design of virtual 3D objects. In a framework for museums, Moustakas and Tzovaras present a technique for simulating ancient technology via virtual reality and haptic rendering [16]. With the help of a force feedback mechanism, museum visitors can manipulate virtual objects; a search functionality is not available, though.

2.2. History of Art and Design

In an art-historic and design context, research and analysis of functional objects do exist, but the results of this analysis are usually embedded in a discourse defined by text. Therefore, many aspects of how an object actually functions or why it was designed in a specific way depend on text comprehension and context awareness, whereas visual (or even experienced) conveyance of such information is mostly missing. In particular, the efforts that go into trying to establish design research as a scientific discipline in its own right rely on (textual) descriptions, semantics, aesthetics, and cognition theories [3, 9, 12, 18]. This does not just hold true for the general case. Even in the rare cases in which the specific use of everyday objects is investigated, the observations are translated into descriptive texts [5, 6, 21].

3. Our Approach

We are about to embark on a project about building a framework for the retrieval of information from a digitized collection of hand-operated tools via gestures with the plan of implementing a prototype of the framework, demonstrating the technical feasibility. We have identified several important steps on the way to the finished prototype. First, we need to lay the theoretical foundations by analyzing a specific collection of hand-

operated tools and identifying their fundamental properties in terms of usage. This means categorizing and describing the objects, not only in terms of their appearance, but also in terms of their mode of operation, going beyond mere technical descriptions, adding cultural and historical aspects. For example, scissors are used in a very typical fashion (similar to shears, clippers, or pliers) by repeatedly moving the thumb and one or several fingers towards and away from each other.

Second, we need to investigate models for representing hand-operated tools according to their modes of operation. This not only involves the movement of a hand through space, but also relative positions of fingers to each other and the positions of joints and other parts of the hand. A gesture would be described as a spatio-temporal sequence of basic events or movement primitives. For our scissor example this means identifying the repeated opening and closing movements of the digits of a hand. We also need to investigate which motion-tracking technology is suitable for this task. Currently, we are looking into data gloves, Leap Motion controllers [24], and, if and when it becomes available, Google's new wave radar technology [14]. However, again, this is not just about technical details, we are also interested in uncovering cultural references.

The insights gained during the first two steps will be applied when it comes to actually digitizing the artifacts. This involves creating virtual representations of the objects (e.g. via 3D scanning) and digitizing gestures and movements. Finally, for querying the digitized collection a measure determining the similarity of gestures has to be defined, as this will be crucial for the retrieval, comparison, and ranking of the returned objects. The importance of different features considered for the similarity measure depend heavily on the cultural aspects of the results obtained earlier. So, when querying the system with the gestures for using scissors, the answer would contain different types of scissors, but also include similar tools such as the aforementioned shears, clippers, and pliers (although these tools would be ranked lower).

4. Conclusion and Outlook

From the point of view of computer science, our goal is to extend the boundaries of today's information retrieval systems by developing a novel, gesture-based query interface that is much better suited for describing the information need of a user searching for manually operated artifacts. Particular hand movements used for operating tools are usually hard to translate into text, but easy to demonstrate. Consequently, it makes much more sense to capture these movements directly, transferring them into an internal representation that can be utilized for retrieval without taking a detour via text. For the moment, the main aim of our approach is to support professionals and researchers, such as designers, archaeologists, anthropologists, and historians, conducting searches in artifact collections. Once the technology becomes more inexpensive, it could also be used by museums to develop interactive displays.

However, our work goes beyond the technological dimension, we also want to generate new impulses in a cultural setting: with the help of our project, the investigation into the interface between man and object can be simulated, manipulated, improved, and visualized. We want to contribute to the current philosophical discourse about the so-called material culture by introducing the new dimension of direct experience of a forgotten or lost object, thereby bridging the gap between oral history and the reconstruction of

knowledge by heuristic narrative methods. This may even spark a new debate about the actor network theory by Bruno Latour, stating that objects are actors on an equal footing with other participants in social systems, contributing substantially to civilization by being able to create and preserve human communication [4, 10, 20, 22].

References

[1] David P. Anderson. Nailing smoke – curation at the bleeding edge of technology. *Communciations of the ACM*, 59(12):37–39, December 2016.

[2] M. Bergamasco, C. Avizzano, G. Di Pietro, F. Barbagli, and A. Frisoli. The museum of pure form: system architecture. In *10th Int. Workshop on Robot and Human Interactive Communication*, pages 112–117, Bordeaux, Paris, France, September 2001.

[3] Ute Brandes. *Designtheorie und Designforschung*. Fink Verlag, Munich, 2009.

[4] Fiona Candlin and Raiford Guins. *The Object Reader*. Routledge, London, 2009.

[5] Katharina Frerus and Dietmar Rübel. *Die Tücke des Objekts*. Reimer Verlag, Berlin, 2009.

[6] Moritz Grund. *Der Designer und die Dinge – ein Selbstversuch*. Niggli Verlag, Sulgen, 2012.

[7] Diogo Henriques, Daniel Mendes, Pedro B. Pascoal, Isabel Trancoso, and Alfredo Ferreira. Evaluation of immersive visualization techniques for 3d object retrieval. In *IEEE Symposium on 3D User Interfaces (3DUI'14)*, Minneapolis, Michigan, March 2014.

[8] Hiroyasu Ichida, Yuichi Itoh, Yoshifumi Kitamura, and Fumio Kishino. Interactive retrieval of 3d virtual shapes using physical objects. In *Virtual Reality (VR'04)*, pages 231–232, Chicago, Illinois, March 2004.

[9] Geesche Joost and Arne Scheuermann. *Design als Rhetorik*. Birkhäuser Verlag, Basel, 2012.

[10] Carl Knappett. *Thinking through Material Culture*. University of Pennsylvania Press, Philadelphia, 2005.

[11] D. Koller, B. Frischer, and G. Humphreys. Research challenges for digital archives of 3D cultural heritage models. *Journal of Comput. Cult. Herit.*, 2(3):7–17, 2010.

[12] Klaus Krippendorff. *Die semantische Wende – Eine neue Grundlage für Design*. Birkhäuser Verlag, Basel, 2013.

[13] A. Leroi-Gourhan. *Gesture and Speech*. MIT Press, Cambridge, Massachusetts, 1993.

[14] Jaime Lien, Nicholas Gillian, M. Emre Karagozler, Patrick Amihood, Carsten Schwesig, Erik Olson, Hakim Raja, and Ivan Poupyrev. Soli: Ubiquitous gesture sensing with millimeter wave radar. *ACM Trans. Graph.*, 35(4):142:1–142:19, July 2016.

[15] Konstantinos Moustakas, Georgios Nikolakis, Dimitrios Tzovaras, Sébastien Carbini, Olivier Bernier, and Jean-Emmanuel Viallet. 3D content-based search using sketches. *Personal and Ubiquitous Computing*, 13(1):59–67, 2009.

[16] Konstantinos Moustakas and Dimitrios Tzovaras. Virtual simulation of cultural heritage works using haptic interaction. In *6th Hellenic Conf. on AI (SETN'10)*, pages 389–394, Athens, Greece, May 2010.

[17] Konstantinos Moustakas, Dimitrios Tzovaras, Michael G. Strintzis, Sébastien Carbini, Olivier Bernier, Jean-Emmanuel Viallet, Stephan Raidt, Matei Mancas, Mariella Dimiccoli, Enver Yagci, Serdar Balci, and Eloisa Ibanez Leon. Masterpiece: Physical interaction and 3d content-based search in vr applications. *IEEE MultiMedia*, 13(3):92–100, 2006.

[18] Felicidad Romero-Tejedor. *Der denkende Designer*. Olms Verlag, Hildesheim, 2007.

[19] Daniela Rosner, Marco Roccetti, and Gustavo Marfia. The digitization of cultural practices. *Communications of the ACM*, 57(6):82–87, 2014.

[20] Stefanie Samida, Manfred Eggert, and Hans Peter Hahn. *Handbuch Materielle Kultur*. Metzler Verlag, Stuttgart, 2014.

[21] Hermann Sturm. *Die Tücke der Funktion*. Klartext Verlag, Essen, 2005.

[22] Elisabeth Tietmeyer and Claudia Hirschberger. *Die Sprache der Dinge*. Waxmann Verlag, Münster, 2010.

[23] UNESCO. Lists of intangible cultural heritage and the register of good safeguarding practices. http://www.unesco.org/culture/ich/en/lists, 2016. [Online; accessed May 2016].

[24] Frank Weichert, Daniel Bachmann, Bartholomäus Rudak, and Denis Fisseler. Analysis of the accuracy and robustness of the leap motion controller. *Sensors*, 13(5):6380–6393, 2013.