ShapeFindAR: Exploring In-Situ Spatial Search for Physical Artifact Retrieval using Mixed Reality

Evgeny Stemasov  
Tobias Wagner  
evgeny.stemasov@uni-ulm.de  
tobias.wagner@uni-ulm.de  
Institute of Media Informatics,  
Ulm University  
Ulm, Germany

Jan Gugenheimer  
Télécom Paris – LTCI, Institut  
Polytechnique de Paris  
Paris, France  
jan.gugenheimer@telecom-paris.fr

Enrico Rukzio  
enrico.rukzio@uni-ulm.de  
Institute of Media Informatics,  
Ulm University  
Ulm, Germany

Figure 1: ShapeFindAR is a proof-of-concept implementation of in-situ spatial search. It enables users to search for future physical artifacts through spatial (a) or textual (b) queries, while coupling the search and previewing process to the users’ physical context. Spatial search queries (a) rely on in-situ sketches, which may be drawn mid-air (1), be tracings of objects/features in the physical environment, or combine prior search results with sketched features (2). Textual search queries (b) rely on written or spoken terms known by the user. Alternatively, a user may use a photo-based label extraction to frame objects of interest (3) and receive labels to potentially use for searching (4).

ABSTRACT

Personal fabrication is made more accessible through repositories like Thingiverse, as they replace modeling with retrieval. However, they require users to translate spatial requirements to keywords, which paints an incomplete picture of physical artifacts: proportions or morphology are non-trivially encoded through text only. We explore a vision of in-situ spatial search for (future) physical artifacts, and present ShapeFindAR, a mixed-reality tool to search for 3D models using in-situ sketches blended with textual queries. With ShapeFindAR, users search for geometry, and not necessarily precise labels, while coupling the search process to the physical environment (e.g., by sketching in-situ, extracting search terms from objects present, or tracing them). We developed ShapeFindAR for HoloLens 2, connected to a database of 3D-printable artifacts. We specify in-situ spatial search, describe its advantages, and present walkthroughs using ShapeFindAR, which highlight novel ways for users to articulate their wishes, without requiring complex modeling tools or profound domain knowledge.

CCS CONCEPTS

• Human-centered computing → Human computer interaction (HCI); Mixed / augmented reality.  
• Information systems → Information retrieval.

KEYWORDS

Personal Fabrication, Spatial Search, In-Situ Search, Mixed Reality, Physical Artifact Retrieval, Model Repositories, 3D-Printing

ACM Reference Format:

Evgeny Stemasov, Tobias Wagner, Jan Gugenheimer, and Enrico Rukzio. 2022. ShapeFindAR: Exploring In-Situ Spatial Search for Physical Artifact Retrieval using Mixed Reality. In CHI Conference on Human Factors in Computing Systems (CHI ‘22), April 29-May 5, 2022, New Orleans, LA, USA. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3491102.3517682

This work is licensed under a Creative Commons Attribution International 4.0 License.
1 INTRODUCTION

Personal Fabrication is a powerful opportunity for technology enthusiasts and consumers alike: It empowers them to design and fabricate artifacts that are unrivaled in their degree of personalization and precision. While technology enthusiasts may be compelled to invest time in the process of personal fabrication, consumers may generally care more about the result [21] and expect a low-friction workflow [52]. To be able to include the whole spectrum of users to benefit from the progress of personal fabrication, future creation tools are increasingly addressing the needs of non-experts and laypeople [14, 46, 55], which are often not able or not willing to express their needs with the precision an expert would. This likely applies to the process of modeling artifacts, but also to the process of searching and retrieving the correct ones [50]. Through more affordable and robust hardware (e.g., 3D printers, laser cutters) and improved software workflows (e.g., modeling tools, slicers), personal fabrication is now in reach for a wider user base than ever before.

However, most established workflows used to define artifacts to fabricate (e.g., 3D modeling ⇒ slicing ⇒ 3D printing), require both learning and usage effort. Such challenges, in part, explain the popularity and importance of open, crowd-based model repositories, such as Thingiverse or MyMiniFactory [2, 13, 21]. Model repositories provide users ready-to-print artifacts, making the interaction with the repository more related to shopping and less like established notions of modeling or designing artifacts [52]. Model repositories are an appropriate alternative to modeling artifacts from the ground up, and will become even more viable as they grow in size and artifact diversity. However, model repositories require feasible ways to be searched, to be a viable asset for users. Text-based search was adopted from established domains (e.g., file search), but largely misses the spatial nature and physical context of searching for (future) physical artifacts. Similarly, query formulation and artifact previewing are disconnected from the physical context they are meant to be placed in later on. When facing a requirement or challenge, users then are either required to invest effort in modeling artifacts or are required to translate their requirements to search terms and subsequently refine them.

We propose ShapeFindAR, a mixed-reality tool to search model repositories for 3D printing in-situ. It combines two types of search queries: textual and spatial. Users may sketch coarse 3D shapes that are then used for geometry-based (i.e., spatial) searches. They may likewise enter known search terms for textual searches. To leverage the user’s physical context, ShapeFindAR provides ways to support users in the task of defining and combining these two query types (Figure 1): users are able to trace features of the environment for spatial searches, or may retrieve labels for objects in the environment for textual searches. This bridges the disconnect between the users’ physical environment and the search process. ShapeFindAR was developed as an application for the Microsoft HoloLens 2\(^1\) and is connected to a custom database of 3D-printable artifacts. By leveraging additional, spatial modalities (i.e., sketches), ShapeFindAR allows users to omit the task of precisely defining search terms for artifact retrieval. While some objects have clear and established terminology, more niche artifacts may require knowledge of the terminology used to describe them (i.e., a certain degree of domain knowledge). An example can be seen in Figure 1 a-1: a ring meant to be worn on two fingers simultaneously. Users may not necessarily know the term “double ring”, but likely have an intuitive understanding of the geometry they desire. Hence, they would be able to sketch the shape in a coarse fashion and use this sketch, combined with imprecise search terms (e.g., “ring”), to find the desired artifact. Similarly, users may have an intuitive understanding how an elongated vase with handles looks like, but may not know the term “amphora” used to precisely denote it. By searching for geometry, users may receive results that fulfill their intended (functional) needs, without limiting themselves to objects originally meant for their goal. By conducting the sketching procedure in-situ (i.e., at the location of the future artifact), users benefit from previewing and referencing real-world features (e.g., by tracing them). ShapeFindAR also enables users to get machine-generated labels and guesses after framing an object of interest (Figure 1 b). This enables them to discover potentially unknown terms. In turn, this helps finding similar objects, or objects meant to interact with the existing object. By providing multiple ways (i.e., modalities) to express and refine searches, ShapeFindAR aims to allow users to choose the path to a desired artifact that fits the task and the user best. Users are also encouraged to iteratively refine their queries. A fitting 3D model could be used as a base for a new spatial search query, by sketching new features onto it, as seen in Figure 1 a-2. With ShapeFindAR, we take the perspective of 3D printing as a domain for physical artifact acquisition. However, we believe that in-situ spatial search, and, by extension, ShapeFindAR, are applicable to the search for and acquisition of physical artifacts in general, regardless of the actual method to fabricate or acquire it.

The contributions of this work are as follows:

- The concept of in-situ spatial search for (physical) artifact retrieval. It actively embeds both spatial input, textual input and the user’s physical context in the search procedure.
- Proof-of-concept implementation of ShapeFindAR. A Mixed-Reality application enabling in-situ spatial search of a model repository for 3D printing.
- Exploration of the concept of in-situ spatial search, using walkthroughs enacted with the ShapeFindAR prototype. The walkthroughs highlight novel ways for users to formulate and refine queries for future physical artifacts.

2 CONCEPT: IN-SITU SPATIAL SEARCH

With this section, we want to elaborate on the concept behind ShapeFindAR. In-situ spatial search supports users with the transfer of features in their physical context to search queries. This currently happens through users’ mental efforts, domain knowledge, and through textual queries only. In-situ spatial search also embraces the spatial nature of physical artifacts, not only for result-presentation, but, where beneficial, also during the search process. We argue that text-based search adds an unnecessary abstraction from a three-dimensional, spatial problem to a one-dimensional search term. Only in the final step, the retrieved artifact is transformed back to

\(^1\)https://www.microsoft.com/en-us/hololens, retrieved on 12.12.2021
3D by presenting a 3D object (either a digital, or a physical one) to the user. This abstraction towards a 1D space (i.e., text) is helpful and powerful for expert users that know the (domain-)specific language and terminology. However, it creates an unnecessary burden for laypeople and novices who may be able to coarsely describe their 3D problem in-situ without necessarily knowing the appropriate and precise terms.

The in-situ aspect of in-situ spatial search supports the transfer of features in the user’s physical context to search queries. Physical objects may be traced with (spatial) sketches, scanned (if the headset is capable of doing so), or be photographed to extract labels. These approaches transfer and translate context features (e.g., geometry) to search queries. With in-situ spatial search, more abstract aspects of context could also be available to use and embed in queries. Context may mean the room a user is in (e.g., the kitchen), the task the user is attempting to solve (e.g., attach an object to another), the general design of the room (e.g., modern), or the color climate of the interior (e.g., earth tones to be matched by the future artifact). ShapeFindAR enables users to trace objects and apply photo-based label extraction. Apart from the use of the context’s features for search queries, it is likewise used for previewing artifacts in context. This enables the user to estimate the form and function of the future artifact, while also allowing interactions like positioning and scaling to evaluate and improve the artifact prior to fabrication or acquisition in general.

The spatial aspect of in-situ spatial search embraces the spatial nature of physical artifacts during the search process. By embedding geometry and sketches in the search procedure, geometry receives an emphasis over terminology. If systems sidestep text and labels, functions and functionality ascribed to objects by contributors to an artifact database (e.g., users on Thingiverse) can be ignored. The general topology of cups, pencil holders, vases, pots, and other cylindrical containers is comparable. However, this geometry of a bin is labeled differently by contributors, while all can be used for “holding” tasks (e.g., holding pencils). Not all artifacts suitable for holding pencils can be found through the original label (e.g., “pencil” or “pen” holder). While geometry does not necessarily strictly define function, it often is tightly coupled to it, particularly for static objects. Furthermore, we argue that users are likely to have at least a coarse understanding of their desired geometry or shape. Approximating it through a spatial sketch to find existing artifacts (to use or remix) may be a feasible procedure to retrieve a fitting one. This avoids the enforced transfer of spatial requirements (3D to precise, written keywords (1D), which is particularly relevant if a fitting terminology is unknown or b) if the actual geometry of a future artifact is the most relevant aspect.

We developed ShapeFindAR as a proof-of-concept prototype embodying our vision of in-situ spatial search. With the means of personal fabrication, the process of artifact acquisition can be approached from distinct directions: 1) design (e.g., 3D modeling or sculpting an artifact from the ground up) or 2) retrieval (e.g., searching for an artifact on Thingiverse to fabricate). Both directions are equally viable to users focusing on the goal of “attaining an artifact”. Usually, the design or modeling approach exhibits the highest complexity, but also enables sufficiently proficient users to achieve any requirement they have [52]. However, designing artifacts requires knowledge and training, and therefore has a higher entry barrier. In contrast, retrieving an artifact generally requires knowledge of the correct terminology to pass to a search system. Alternatively, users have to rely on recommender systems, or invest time in elaborate iterations and refinement. When dealing with high-level concepts (e.g., “chair”, “vase”), terms are generally established and there is little to no mismatch between the users’ intentions and the understanding of the term by the artifact repository. The more specific and personal the requirements become, the more detailed and precise the search terms have to be to yield satisfying results. This makes the combination of multiple search modalities an appealing outlook.

Figure 2: Conceptual process of in-situ spatial search, as demonstrated with ShapeFindAR. The users reside in the same physical context in which a future artifact may be fabricated for (a). They may search the repository with 3D sketches or known words, composing a search query (b). As the search happens in-situ, users may also extract labels from the environment or trace existing objects. While refining their queries, users may learn new terms from the search results for textual searches and sketch new features onto the results. The query yields a result set (c) generated by the database server. Both query types (spatial, textual) can be combined, enabling multimodal search, while allowing users to be ambiguous in one or both modalities.
ShapeFindAR provides two main ways for searching: spatial queries (i.e., through sketching) and textual queries. A diagram of the search process is depicted in figure 2. Users are supported in the task of embedding their physical context in the queries: by tracing existing features for a spatial query, or by using label extraction to generate suggested search terms for a textual query. This is crucial, as any artifact users retrieve is meant to be fabricated or acquired otherwise – ultimately interacting with the physical context [4]. In-situ search lowers the effort required to define searches and preview results in-context, as it omits transfers between a location of (future) use (e.g., the table a vase will rest on) and a location of searching/design (e.g., a computer). Both query types may also be created without involving the users’ physical context (e.g., by relying on known terminology or known geometry). This allows users to freely encode their existing wishes and knowledge, but demands a certain level of domain knowledge (for search terms) or at least a coarse understanding of the desired geometry. Through the use of in-situ spatial search procedures, a set of more specific advantages emerges for users: users may circumvent established terminology, ignore objects’ intended function and use a sketch’s scale as a filter or a target scale. In general, in-situ spatial search is meant to enable users to formulate and refine ambiguous search queries in multiple modalities to ideally retrieve finished, fitting designs from a model repository.

All aforementioned tools can be seen as approaches to reduce effort, learning, and overcome challenges in personal fabrication. With in-situ spatial search and ShapeFindAR, we similarly aim to address novices to personal fabrication and design. However, we focus on outsourced design artifacts (i.e., 3D-printable models from experienced users uploaded to Thingiverse or MyMiniFactory). This is a contrast to procedures inferring or generating geometry [35, 45] or requiring users to model artifacts by aggregating primitives [6, 64].

3.2 In-situ Tools for Personal Fabrication
There is a disconnect between the location of artifact design and the artifact’s future location [4, 37]. This requires either assessment and transfer of requirements, or methods to compensate measurement errors [27], which is a challenge that is not exclusive to novices. In-situ methods for modeling and fabrication enable users to preview designs within their future context. This was demonstrated by Peng et al. [39], where design overlapped fabrication, or Yung et al. [66], where the target audience was children. In-situ tools relying on augmented or mixed reality also enable the augmentation and guidance of previously manual and analog fabrication approaches, such as 3D pen sculptures [65], carved models [17, 18], or linkages [23] by providing feedback during the process itself. They may also support remote collaboration and learning, as shown by Villanueva et al. [60]. Mixed reality approaches, such as the works by Weichel et al. [62] or Jeong et al. [23] enable situated design approaches, while other tools focus on fabrication [39, 69], or remixing [53].

Personal fabrication and any process that generates personal physical artifacts require grounding in the users’ physical context. This is achieved through situating design or retrieval processes in this particular context. Mixed reality is an outstanding opportunity to achieve this type of interaction. ShapeFindAR aims to provide a similar benefit, and, for instance, enables users to trace real-world artifacts to use these features for their searches. However, our focus lies on search and retrieval, instead of modeling or design.

3.3 Alternative-modality Interfaces for Fabrication or Search
The combination of different input modalities enables users to bypass established metaphors for CAD (i.e., for the design of artifacts) or information retrieval (i.e., for the search for artifacts). This includes uses of sketches as a coarse input [31, 34, 45], but also gestures [10, 26, 32] or speech [5]. Alternatively, tangible manipulation may be employed for input and provide appropriate feedback [48, 57] not found in most industrial CAD tools. Such ways to express and define geometry may, depending on the user, reduce the effort to design geometry [51]. In contrast to the interfaces that focus on the design and generation of artifacts, retrieval interfaces focus on the search for finished designs or parts [20]. Fraser et al. presented ReMap, which embodies a prototype of multimodal search [15], an aspect we aimed to embrace with ShapeFindAR. On a more technical level, various approaches for the comparison of 3D
models are present in the literature. These include approaches to align point sets [3, 44], computation of object similarity [8, 28, 30], and associated retrieval methods based on geometry [56]. In particular, gesture-based systems, such as DataMiming by Holz and Wilson [20] or sketch-based systems, such as the ones introduced by Pu et al. [41] Ebenhofer [12] or Eitz et al. [11] inspired ShapeFindAR. While DataMiming omitted visual feedback for the users and focused on gestures as input, it is conceptually similar to ShapeFindAR, as it uses a fairly natural way to describe artifacts, and offloads the precise definition or matching to other parties [20]. ShapeFindAR does not omit visual feedback but considers it to be a crucial part of the search process. It also benefits from context features (tracing, label extraction) and embraces multiple modalities to give users more freedom and ambiguity. With platforms like Thangs [40] offering 3D-model based search for 3D models, the notion of searching with shapes over labels is already in reach for consumers, albeit in an ex-situ fashion. Giunchi et al presented a combination of sketching and speech for the retrieval of chair models in VR [16]. While conceptually similar, our focus lies on users’ physical environments, instead of virtual ones. This requires ways to transfer spatial requirements to searches.

Such novel user interfaces for search or design are an opportunity to enable widespread access to personalized artifacts, by lowering the skill floor while retaining high expressivity for more proficient users. Novel modalities are applicable both to searching for existing artifacts [20, 49] and generating new ones [5, 32, 63]. ShapeFindAR can be considered to be a comparable approach to artifact retrieval, with an emphasis on coarse inputs that are – ideally – transferred to high-fidelity artifacts from a model repository. ShapeFindAR similarly aims to enable and support iterative search procedures, where users leverage their physical context and retrieved artifacts to formulate and refine searches.

### 3.4 Usage of Model Repositories

Model repositories (e.g., Thingiverse) and their associated communities of makers are an important component of today’s personal fabrication landscape. Repositories have been leveraged as a model source for remixing tools, such as Grafter by Roumen et al. [43], or Mix&Match by Stemkoski et al. [53]. They have also been the subject of studies and investigations, generating novel frameworks like PARTs by Hofmann et al. [19] that aim to increase re-use of design effort. Likewise, systematic analyses of repositories like Thingiverse and their users unveiled challenges in terms of artifact use and customization by novices [2, 21]. Works by Flath et al. [13] or Kyriakou et al. [29] evaluated patterns and degrees of knowledge and design reuse. Domain-specific usage, such as the exchange of assistive technologies [7] is likewise a field benefitting from a combination of design sharing, re-use, and customization.

We agree with and embrace the notion that crowdsourced models or artifacts are crucial for personal fabrication. Notably, any storefront – on- and offline – provides users a similar experience, with varying degrees of potential personalization. When considering stores to be an alternative to personal fabrication (currently chosen by a majority of the population [52]), approaches relying on searching and customizing existing designs can be considered to be a viable path to take. We approach this aspect with an emphasis on search and how it is applied to the task of finding future artifacts for one’s own, personal, physical context.

### 4 PROOF OF CONCEPT IMPLEMENTATION

The following sections describe the development of ShapeFindAR as proof-of-concept implementation. The system consists of 3 components: the application for Microsoft’s HoloLens 2, a server component connected to the database, and a data-gathering tool which was used to fill the database with models gathered from Thingiverse and MyMiniFactory. We report our approach and development process openly for replicability.

#### 4.1 HoloLens 2 Application

We implemented the user-facing part of ShapeFindAR with Microsoft’s HoloLens 2 using Unity2 2019.4.6f1. Version 2.4 of the Mixed Reality Toolkit (MRTK3) was used to implement the user interface and most of the interactions with it. Networking was implemented with the RestClient4 library for Unity.

The search interface is initially attached to the user’s right hand. If needed, the user can grab and position the panel statically in space (Fig. 3a). The search interface can be used to enter search terms by text or speech. After submitting a query, the interface loads a scrollable grid of results (Fig. 3b). The user can select results that appear promising, which are then downloaded in the background. As soon as the download and instantiation are completed, the model appears attached to the user’s left hand, forming a palette where the user may gather up to 5 models to compare and evaluate (Fig. 3b). Each artifact loaded into the palette can be grabbed and positioned in the space around the user. The objects can then be repositioned, rotated, and scaled freely, which enables an in-situ preview of them. The sketch search component can be activated by tapping on a button below the search interface. After activating the sketch mode, the users can draw in the space around them by pinching their thumb and index finger (Fig. 4a). The sketch can then be submitted as a query. To ensure appropriate results, sketch-based search currently requires the users to enter at least a coarse search term. This is mainly used to refine the search to receive faster results and can be removed in the future with either better (backend) hardware or a more optimized implementation of the matching algorithm.

---

2https://unity.com/products/unity-platform, retrieved on 16.12.2021  
3https://microsoft.github.io/MixedRealityToolkit-Unity, retrieved on 26.8.2021  
4https://unity.com/proyectos/RestClient, retrieved on 15.12.2021
users may also sketch onto downloaded and positioned meshes, to extend them and use them as a new query (Fig. 4b). This feature also enables users to position multiple models they retrieved from the repository and connect them through their sketch, which is then treated as a single spatial query by the server. Thereby, users do not necessarily have to sketch entire objects, but may start with an object and add features they deem missing from it (e.g., adding handles to a vase).

The image-based label retrieval extracts possible terms for textual queries based on an image captured by the headset. This component of ShapeFindAR is fundamentally a mapping from a 2D image to a search term (1D). It is enabled by connecting and holding the index fingers and thumbs of both hands for one second (dwell time) in the field of view of the device. This activates a "framing mode" (Fig. 5a). Afterwards, the users can drag their hands apart, to form a frame around the object they want to use as a query (Fig. 5a). The captured image is used as a query to the Google Cloud Vision API, which retrieves suggested labels based on the image’s contents.

4.2 Data Collection

While Thingiverse and MyMiniFactory provide crucial infrastructure for makers, they are – by design – focused on established ways to be searched: textual queries along with sorting and filtering options. Notably, most filters focus on tags and categories (i.e., more abstract terminology). Sorting similarly focuses on non-spatial dimensions, like popularity or (textual) relevance. We chose Thingiverse and MyMiniFactory as data sources for ShapeFindAR. Thingiverse can be considered the de-facto standard for model repositories for 3D printing, while MyMiniFactory aims to curate the models offered on the platform. However, neither of these platforms offers ways for geometry-based searches. Thangs\(^6\) allows users to upload 3D models as queries, but offers no way to define them (either coarse sketches or precise models), while also functioning ex-situ only. To comply with the platforms’ API guidelines and to enable our novel functionalities without straining public infrastructure, we chose to duplicate a subset of their libraries. We chose high waiting times between all requests, to avoid putting unnecessary strain on the servers of Thingiverse and MyMiniFactory. This led to timespans of multiple days to gather batches of 500 models. The keywords used were basic mechanical artifacts popular in 3D printing communities like hooks, and decorative artifacts like vases or figures. Furthermore, the most popular artifacts on each of the repositories were gathered. Data like the original source, the designer, the licenses used were all transferred to our database subset.

The entire dataset we gathered consists of 4118 objects in total. 2717 objects consist of one part only, while the remaining ones have 2 or more files (e.g., .stl) associated with them. 2356 objects out of 4118 provide mesh metadata suitable for our implementation of spatial search queries.

4.3 Postprocessing

Postprocessing of the metadata and the meshes themselves was necessary after each iteration of data gathering. The tool used for mesh handling was the trimesh library for Python by Dawson–Haggerty et al. \(^9\). Some meshes were corrupted, and were for instance missing correctly calculated normals or did not describe a volume. We did not apply any "opinionated" automated mesh repair methods, due to the risk of corrupting the original geometry. The data sources Thingiverse and MyMiniFactory both employ the concepts of categories and tags. While tags are largely unmoderated and can be freely chosen by the users, available categories are specified by the repository administration. The categories used by Thingiverse and MyMiniFactory are comparable, but not identical. Categories were therefore grouped and matched into new categories used by our database.

To improve the text- and voice-based search, fields like the artifact name, description, or tags were subject to postprocessing with NLTK\(^9\) (stemming, lemmatization, and stopword removal). To enable geometry-based searches, spatial metadata of the meshes was also established during this step. As detailed registrations and comparisons of meshes are computationally expensive, we implemented methods to pre-filter the result. We calculated two ratios based on the ordered dimensions of the bounding box to filter meshes based on their proportions. The voxel representation used for aligning and matching sketches to objects was computed at this step and

\(^6\)https://cloud.google.com/vision/, retrieved on 11.12.2021
\(^7\)https://www.nltk.org/, retrieved on 1.12.2021
\(^9\)https://thangs.com/, retrieved on 16.12.2021

---

Figure 4: The sketching mode of ShapeFindAR. The users are free to sketch queries in the space around them with a pinch gesture (a). To iteratively refine spatial search queries, users may also sketch additions onto an object retrieved from the repository (b) and submit this combination as a new query.

Figure 5: The image-based label extraction of ShapeFindAR: The user activates a framing mode through a gesture, encloses the artifact to be classified in the frame, and confirms the selection via dwell-time (a). The "best guess" is pre-filled into the search field and additional suggested keywords are listed below (b).
4.4 Server Implementation

All server-side software was written in Python. The server component was written using Flask.¹⁰ The Flask application provides a REST (representational state transfer) API for the HoloLens client: searches by text, searches by sketch, download of metadata, files, and thumbnails. For the database component, MongoDB¹¹ 4.4.0 was used, due to its flexible data model and its document-based approach. To accelerate queries, our database instance employs multiple indexes. Apart from indexing the ID fields, which accelerates retrieval of object details and files, a textual index and a 2D spatial index are used. This ideally accelerates searches submitted through their respective index. For queries ultimately converted to text, the database uses a text index¹², based on the title, description, tag, and category fields. The fields are weighted based on their specificity (with categories being the least and names being the most specific). Searches passed to this index omit stopwords and employ stemming and lemmatization for the English language.¹³ A second index configured for the database is a 2D spatial index.¹⁴ We leverage the accelerated query time for 2D-data we derive from the meshes’ object-aligned bounding box (OABB). The ratios between height/width and width/depth are used to pre-filter models based on their proportions.

Each search request is initially treated as a textual search, but gets by highly imprecise and general terms (e.g., “object”). This enables pre-filtering the results to amounts manageable by the server for computationally expensive operations and manageable by the user to get an appropriate overview. After having reduced the potential search result set, the server filters the results based on the previously introduced metric of OABB ratios. This set of results that match based on the OABB ratio is intersected with the set retrieved by the textual search. With this selection, it is now possible to calculate per-voxel overlap, which is the server’s ranking/scoring function. The stored voxel representations are already size-normalized (i.e., their largest extent is set to 100 and the other extents are scaled proportionally to match). The incoming sketch is similarly normalized in scale. However, the original extents are retained, if the resulting model is meant to be scaled to match it. The sketch is then voxelized. The resolution of the voxelization is likewise defined by the model’s largest dimension, as the voxelization pitch was chosen to yield 20 voxels across the largest dimension. The voxelized meshes are then used to run the ICP algorithm (iterative closest point).¹⁵ The ICP procedure aligns the sketch to the repository model, as the latter usually has an orientation that follows the principal axes of inertia. The starting parameters for the ICP procedure are chosen based on the inertia of the mesh. The voxelized sketch is then rotated and moved with the calculated transformation. Due to the previously conducted size-normalization, no scaling is involved in this step. With the 2 models aligned and with maximal overlap, the amount (i.e., the number of voxels) of overlap is counted. Based on the resulting overlap, we calculate multiple similarity metrics: normalized by the number of voxels in the sketch, normalized by the number of voxels in the repository model, and an average of these two normalized values. To accommodate for different volumes (e.g., sketches being entirely wrapped by the repository model), the average of the two normalized values is used as the core ranking for the search results.

5 APPLICATION WALKTHROUGHS

The following scenarios were enacted with the ShapeFindAR prototype and present brief, self-contained examples of actual interactions with the system. The search results seen in the examples are based on our custom dataset and are therefore a subset of models available on Thingiverse or MyMiniFactory. However, due to a sizeable number of different objects of different categories, the user interacts with the dataset and the search results in a similar fashion to the potential interaction happening with a more extensive database. Furthermore, the sketch search is not ideally calculating similarities, but fails in specific edge cases (e.g., completely flat sketches, or when the user’s sketch has entirely different proportions than any fitting object in the database). It does, however, generally retrieve a set of results where fitting or similar artifacts can be found among the first ones and iterated with. With the walkthroughs, we want to emphasize different ways in which in-situ spatial search enables users to ambiguously and iteratively approach physical artifact retrieval.

5.1 Scenario 1: Bypassing Terminology

ShapeFindAR supports users by affording them a degree of Term-Abstraction: not having to know or choose the terminology of a domain. This is enabled through the use of photo-based label extraction (2D → 1D) and sketch-based search (3D → 3D).

The first example can be seen in Figure 6. A microcontroller (Arduino Uno), is in a user’s vicinity. The name “Arduino” does not appear on the device, as it is made by an alternative manufacturer.

![Figure 6: Term-Abstraction: Finding artifacts when terminology is unknown. 1) Framing the object to be searched for; 2) using one of the classified labels; 3) search results and scaling of a selected one; 4) previewing the result in-situ.](https://example.com/shapefindar.jpg)
of domain knowledge. ShapeFindAR enables both a textual and a spatial query refinement process.

An example can be seen in Figure 8. The user initially searches for bowls to use for decoration (Fig. 8-1). As she notices that she would prefer ones that are generally taller, she can either take one of the pots retrieved from the repository and scale it in the Y direction, or she can enable the sketching mode and add onto the downloaded pot to elongate it (Fig. 8-2). After having changed the pot’s proportions this way, she may submit the combination of previous result and sketch, along with a coarse search term (Fig. 8-3) as a new query. Lastly, she may preview the results in-situ and apply other manipulations such as scaling to the (3D) results (Fig. 8-4). While the initial query is formulated without leveraging in-situ features, the query refinement may happen through referencing the physical context (e.g., comparing proportions of the results with existing decoration and altering it accordingly).

### 5.3 Scenario 3: Inspiration

Photo-based label extraction can be employed to provide inspiration and related items to the user. By applying this label extraction, users transfer aspects of their physical context to textual queries. This is similar to features found on platforms like Pinterest [61] or shopping interfaces like IKEA’s smartphone app [54]. The items retrieved do not necessarily have to be the same artifact, but be similar to it, in terms of use-cases or domain. For a phone, the search may retrieve phone mounts or cases, which do not exhibit an identical geometry, but are meant to complement the geometry of the artifact.

Figure 9 depicts such a process. A user frames a watch (Fig. 9-1) and the resulting best guess is filled into the search field of ShapeFindAR (Fig. 9-2). Additional guesses are listed below and can likewise be used for a textual query (Fig. 9-2). While this is comparable to a coarse text search, it enables users to benefit from the machine’s perception of the object or the scenery, ideally expanding the users’ existing domain knowledge ad-hoc. This may
include more precise terms (e.g., wristwatch instead of watch) or more general ones, in comparison to the user’s wording. We consider this approach to be a feasible way to initiate a search with the goal of browsing a model repository. The results range from objects that are meant to complement the watch (e.g., a watch-stand – Fig. 9-3) or replacement parts for the watch (e.g., a watch strap – Fig. 9-4), all of which can be previewed in-situ.

5.4 Scenario 4: In-situ Features
With ShapeFindAR, users may also trace features available in the vicinity. This enables users to transfer aspects of their physical context to spatial queries. The features can be either traced coarsely and serve as a basis for a sketch extending upon them, or be traced completely to search for (geometrically) similar artifacts.

A user may trace an existing flowerpot he wishes to replace or improve its design (e.g., by adding a decorative cachetop or planter to it). The ratio of the existing pot is important for this task, so the user tries to approximate the contours of the pot as closely as possible (Fig. 10-1). This does not necessarily leave a complete sketch of the pot. Afterwards, he fills some of the gaps of the outline to provide ShapeFindAR with volumetric features to work with (Fig. 10-2). Along with the search term “planter”, the sketch search is submitted and yields decorative pots with comparable proportions to the sketch (and therefore to the original pot – Fig. 10-3). As the sketch was not entirely gap-free, a paper bin with sparse lines is among the first results (Fig. 10-3). Lastly, the user may move the 3D preview of the future artifact to cover the physical pot and evaluate aesthetics or apply modifications such as scaling (Fig. 10-4).

6 LIMITATIONS AND DISCUSSION
In this chapter, we want to discuss the conceptual and technical limitations of ShapeFindAR and the presented concept of in-situ spatial search.

6.1 Conceptual Limitations
Conceptually, it is important to consider the skill requirements and gaps associated with spatial sketching. Transferring partial requirements (e.g., tracing a physical object) to a satisfying sketch may prove hard for some users. We emphasize coarse approximation of shapes, but this may not suffice for complex queries, where for instance precise measurements are required. This is, to a degree, compensated by other modalities available to the users (i.e., text). However, this could potentially be overcome in the future with even better matching algorithms, which may consider more aspects of the environment. Alternatively, this may require ways to compensate for the lack of precision through automated corrections, or novel fabrication workflows [42, 55].

6.2 Technical Limitations
ShapeFindAR is, as a proof-of-concept, not without technical limitations. The ranking computed by the server is not impeccable, but under the candidates the system retrieves, fitting results can usually be found. The spatial search component is arguably slow, with one sketch-to-model comparison taking up to 2 seconds. This leads to requests times of approximately up to 60 seconds. The more artifacts are considered to be feasible candidates, the longer the search may take. Accelerating the server would for instance require approaches such as shape histograms [28]. To scale ShapeFindAR to database sizes comparable to Thingiverse, far more optimizations and more aggressive filtering approaches are required. Unlike textual queries, sketch queries are far more individual and harder to cache or precompute. The ICP algorithm we use to align sketches and repository meshes, was originally meant to align point clouds originating from the same scene [67]. This is not the case with our combination, as a sketch merely approximates the target object. The approach of pre-filtering based on the OABB proved to be not ideal in some cases. If “flat” sketches are submitted (i.e., when the user merely traces an outline), similarly flat repository models are treated preferably. Automated extrusion would likely improve search performance and reduce the effort required from the users to fill volumes in their sketches. We argue that while there is room for improvement, coarse and ambiguous input, along
with similarly coarse (but appropriate) results, may lead to more productive, creative explorations of the potential solutions to users’ requirements.

6.3 Opportunities of In-Situ Spatial Search
Regardless of the aforementioned limitations, we strongly believe that ShapeFindAR is a proof-of-concept that demonstrates the opportunities of in-situ spatial search for future physical artifacts. By abstracting away from terminology, users may express their desired artifacts in a modality that is arguably more natural and fitting: in 3 dimensions and in-situ, instead of ex-situ and converted to text and labels. In-situ spatial search enables Term- and Function-Abstraction. By ascribing terminology to the artifacts, they are prevented from appearing in searches not following this particular terminology (i.e., if novices to a domain are meant to express search terms). In-situ spatial search may also enable scale-invariance in a search process. With ShapeFindAR, we embrace the aspect that scale can be freely chosen by the user. This freedom enables scale-invariant searches. By normalizing all spatial input and output (while also focusing on proportions), ShapeFindAR sidesteps established conventions of size. Future implementations may also automatically scale objects to match the user’s sketch, or use the sketch as a way to filter objects by size.

We argue that the “ideal search” for a physical artifact should happen in the most fitting (for the task) and easy (for the user) modality, and allow the combination of modalities to express partial requirements. There is no “one size fits all” approach to the search for physical artifacts, especially for personal use. A standardized artifact (e.g., an M3 screw) is easy to find, if the user knows the appropriate denomination (i.e., possesses the necessary domain knowledge). Likewise, precise measurements are easier to express through text (e.g., 4cm), compared to (spatial) sketches. As promising as a spatial-only search system may appear, it covers and excels at a subset of possible tasks. It also implicitly assumes that all users lack domain knowledge to express fitting search terms. If a user happens to possess the ability to express a precise search term in text, they may still achieve fitting results quickly. This requires systems to still offer textual means, which is what we did with ShapeFindAR, while also adding a transfer function from the physical context to textual queries (label extraction). However, proportions or geometry that are not standardized require other means to be encoded in a search, such as spatial sketches or 3D scans.

7 FUTURE WORK
We do not assume that 3D printing is the only domain ShapeFindAR may be applied to. Ideally, it may be a search and preview frontend to any database of physical artifacts. Databases that are domain-specific (e.g., furniture [36]) or general artifact databases (e.g., general shopping interfaces) could be addressed by ShapeFindAR and the concept of in-situ spatial search. With the capabilities of personal fabrication devices increasing, it may apply to any physical artifact in the (distant) future. Additionally, it is intriguing to explore users’ strategies for in-situ spatial search in a task-oriented user study. Technical improvements to ShapeFindAR are possible in terms of processing speed and the addition of further input modalities. With devices like the HoloLens being outfitted with a depth sensor, more detailed interaction with the environment is a promising direction. For instance, one may scan existing artifacts and use them as search input (3D → 3D). This is fairly similar to the label extraction approach (2D → 1D) present in ShapeFindAR, but abstracts even further from labels (Term-Abstraction) and emphasizes geometry instead (3D → 3D). This scan could also be altered with added sketch features or deformation. Silhouettes of artifacts are an intriguing aspect not fully considered in the ShapeFindAR prototype. Sketching outlines and applying basic operations like in revolving (for volumes with rotational symmetry), extrusion, or inflation (e.g., [22]) may be easier to execute than sketching the entire geometry (2D → 3D → 3D).

8 CONCLUSION
We presented the concept of in-situ spatial search and ShapeFindAR, a mixed reality search interface to a repository of 3D-printable models. The prototype system embeds multiple modalities for users to search for the physical artifact they desire: sketching shapes, photo-based search (i.e., label extraction), or textual search. The search itself, along with the object preview, happens in-situ, at the location of the future artifact, and happens iteratively, based on input gained from the search results and the physical environment. We specified this concept as in-situ spatial search, and described conceptual advantages, along with walkthroughs possible with ShapeFindAR. We consider this work to be a step towards lower-effort interfaces for personal fabrication, which may emerge more as sophisticated shopping interfaces [52], and less as simplified interfaces for (3D) modeling.

With this work, we argue for better ways for novices and even consumers (who are not involved in personal fabrication) to search and preview future physical artifacts. While users with intrinsic motivation for fabrication are willing to invest time in learning and designing, the majority of people can be considered “consumers” [52]. They are willing to benefit from unique, personal artifacts, but not willing to invest time in the process. For them, novel, low-effort, means for artifact retrieval, over artifact modeling, are required. The process of retrieval, in turn, should not enforce the use of a specific language to formulate queries and demand as few transfers between (future) artifact context and the search interface as possible.

ACKNOWLEDGMENTS
We thank Thingiverse, MyMiniFactory, and their contributors for providing access to such vast amounts of diverse models. We furthermore thank our anonymous reviewers, who have supported us in improving the manuscript.

REFERENCES
[1] Harshit Agrawal, Udayan Umapathi, Robert Kovacs, Johannes Frohnhofen, Hsiang-Ting Chen, Stefanie Mueller, and Patrick Baudisch. 2015. Prototyper: Physically Sketching Room-Sized Objects at Actual Scale. In Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology (UIST ’15). ACM, New York, NY, USA, 427–436. https://doi.org/10.1145/2807442.2807505
[2] Celena Alcock, Nathaniel Hudson, and Parmit K. Chilana. 2016. Barriers to Using, Customizing, and Printing 3D Designs on Thingiverse. In Proceedings of the 19th International Conference on Supporting Group Work (GROUP ’16). ACM, New York, NY, USA, 195–199. https://doi.org/10.1145/2957276.2957301
[3] K. S. Arun, T. S. Huang, and S. D. Blostein. 1987. Least-Squares Fitting of Two 3-D Point Sets. IEEE Transactions on Pattern Analysis and Machine Intelligence
