Artificial Intelligence for Engineering Design, Analysis and Manufacturing

cambridge.org/aie

Enabling multi-modal search for inspirational design stimuli using deep learning

Elisa Kwon[1], Forrest Huang[2] and Kosa Goucher-Lambert[1]

[1]Department of Mechanical Engineering, University of California, Berkeley, CA, USA and [2]Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA, USA

Research Article

Cite this article: Kwon E, Huang F, Goucher-Lambert K (2022). Enabling multi-modal search for inspirational design stimuli using deep learning. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* **36**, e22, 1–18. https://doi.org/10.1017/S0890060422000130

Received: 21 December 2021
Revised: 27 May 2022
Accepted: 3 June 2022

Key words:
Deep learning; design creativity; inspirational design stimuli; multi-modal search

Author for correspondence:
Kosa Goucher-Lambert,
E-mail: kosa@berkeley.edu

Abstract

Inspirational stimuli are known to be effective in supporting ideation during early-stage design. However, prior work has predominantly constrained designers to using text-only queries when searching for stimuli, which is not consistent with real-world design behavior where fluidity across modalities (e.g., visual, semantic, etc.) is standard practice. In the current work, we introduce a multi-modal search platform that retrieves inspirational stimuli in the form of 3D-model parts using text, appearance, and function-based search inputs. Computational methods leveraging a deep-learning approach are presented for designing and supporting this platform, which relies on deep-neural networks trained on a large dataset of 3D-model parts. This work further presents the results of a cognitive study ($n = 21$) where the aforementioned search platform was used to find parts to inspire solutions to a design challenge. Participants engaged with three different search modalities: by keywords, 3D parts, and user-assembled 3D parts in their workspace. When searching by parts that are selected or in their workspace, participants had additional control over the similarity of appearance and function of results relative to the input. The results of this study demonstrate that the modality used impacts search behavior, such as in search frequency, how retrieved search results are engaged with, and how broadly the search space is covered. Specific results link interactions with the interface to search strategies participants may have used during the task. Findings suggest that when searching for inspirational stimuli, desired results can be achieved both by direct search inputs (e.g., by keyword) as well as by more randomly discovered examples, where a specific goal was not defined. Both search processes are found to be important to enable when designing search platforms for inspirational stimuli retrieval.

Introduction

During idea generation, designers are known to benefit from external inspirational stimuli toward achieving desirable design outcomes such as greater novelty, feasibility, or innovativeness (Chan *et al.*, 2011; Fu *et al.*, 2013b; Goucher-Lambert *et al.*, 2020). The efficacy of a stimulus in providing inspiration during the design process can depend on a variety of features. For example, the modality of stimulus representation, analogical distance of the example to the design problem, and the timing of example delivery have all been shown to impact the way in which designers utilize stimuli (Linsey *et al.*, 2008; Tseng *et al.*, 2008; Chan *et al.*, 2011). Inspirational stimuli may vary with respect to modality of presentation. Different uses of visual stimuli to support design ideation have been explored, such as when combined with text (Borgianni *et al.*, 2017), other images (Hua *et al.*, 2019), or in contrast to interactions with physical products (Toh and Miller, 2014). Representing stimuli visually compared to physically, or when combined with textual examples, has been shown to increase idea novelty (Linsey *et al.*, 2008; Toh and Miller, 2014). The impact of analogical distance of stimuli from the design problem is also important to consider. Relative to the designer's approach to a design problem, far-field examples have been found to contribute to idea novelty (Chan *et al.*, 2011; Goucher-Lambert and Cagan, 2019). However, near-field examples can also lead to design creativity as well as greater feasibility, relevance, and quantity of ideas (Chan *et al.*, 2015; Goucher-Lambert *et al.*, 2019, 2020). A given stimulus may be more useful depending on when it is accessed during the design process. Inspirational stimuli provided after ideation on a design task has begun have been found to be more effective than when provided before ideation (Tseng *et al.*, 2008). During ideation, designers who receive stimuli when stuck produce more ideas than those who receive them at predefined intervals, indicating the importance of timing of example delivery (Siangliulue *et al.*, 2015). The level of abstraction of inspirational examples can also impact their influence on the design process. Design stimuli at the concept level may provide more rapid inspiration but miss the richer design details available in more comprehensive documents like patents (Luo *et al.*, 2021). Examples can differ further by being provided with descriptions that are more general versus domain-specific (Linsey *et al.*, 2008) or constitute concrete design examples versus abstract system properties (Vasconcelos *et al.*, 2017).

© The Author(s), 2022. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

While extensive prior research, as highlighted above, has uncovered the characteristics of inspirational stimuli that contribute to their usefulness to designers, less is known regarding how designers naturally discover them. Prior researchers have mostly provided carefully curated examples to designers in controlled studies to study how specific independent variables of inspirational stimuli affect design outcomes. To better understand the process of searching organically for inspiration during design, a creativity-support platform is developed that allows designers to search flexibly in realistic contexts and researchers to collect data through custom instrumentation. Toward these research goals, the core contribution of this work is two-fold:

(1) The development of a platform enabling search for inspirational stimuli. This platform provides designers with the ability to search with multi-modal inputs and control the degree of similarity between retrieved results and input queries.
(2) An investigation of search processes in design employed when using this platform during a cognitive study. This study compares the impact of using the afforded modalities on overall search outcomes and behaviors.

The platform developed in this work involves the computational extraction of features of inspirational stimuli, and the subsequent ability to search based on these features using multiple modalities. Semantic, visual, and function-based features are specifically explored in this work, following past studies on design by analogy, as introduced in the section "Computational methods for inspirational stimuli retrieval". By providing this creativity-support platform (described in the section "Platform development") to designers during a cognitive study (described in the section "Cognitive study design"), important insights regarding designers' processes of searching for inspirational stimuli can be uncovered. The findings of the cognitive study reveal how designers search by different modalities and the effect of using these modalities on the inspirational stimuli designers engage with and discover. These insights into designers' search behaviors and strategies, and how they are differently influenced by search modality, can be helpful to future researchers when further developing computational retrieval-based systems best fitted to the engineering design process.

Related works

The first objective of this work is to leverage computational methods to develop a platform that enables the retrieval of inspirational stimuli based on a given search input. This section thus firstly presents a brief review of computational techniques used to derive relationships between design ideas and inspirational stimuli. The proposed platform also aims to support flexible search for inspirational stimuli. Therefore, existing work on design-support tools that specifically allow multi-modal interactions, for example, visual sketch-based inputs, is also surveyed. A second objective of this work is to conduct a cognitive study to investigate how designers use the developed platform to search for inspiration. Search processes from a cognitive perspective are thus discussed to gain insight into designers' search behavior.

Computational methods for inspirational stimuli retrieval

In order to extract meaningful stimuli relevant to a given design problem, design idea, or search query, computational methods and tools are needed to derive similarity relationships between inspirational stimuli in a given dataset and a designer's input. Research in design by analogy offers insight into different techniques used to establish relationships between examples in the design space, which may be based on, for example, semantic, functional, or visual information.

A variety of sources from which these stimuli may be derived have been explored. Information-rich repositories such as patent databases or biology textbooks are expansive sources of examples that are commonly used to provide relevant design information in both textual and pictorial representations (Chan *et al.*, 2011; Cheong *et al.*, 2011; Fu *et al.*, 2013b). Examples from these and other sources are often used as functionally related inspirational stimuli to support design by analogy. Using patent databases as sources of design examples, function-based relationships between these designs can be defined. Murphy *et al.* built a functional vector space model to quantify functional similarity of a design problem to designs described in patents (2014). This approach forms a functional vocabulary through text-based processing of patent documents, resulting in a vector representation of the patent database. Latent semantic analysis (LSA) is another method for defining text-based contextual similarity between patents, used by Fu *et al.* (2013a, 2013b). VISION is an exploration-based design-by-analogy tool developed by Song and Fu that uses nonnegative matrix factorization to assign topics to patents based on different concepts, including function (Song and Fu, 2022). Patent data has also been used to train semantic network databases to support engineering design activities. While some semantic networks such as WordNet or ConceptNet contain common words (Han *et al.*, 2022), the Technology Semantic Network (TechNet) was developed using patent data to formulate a semantic network database specialized in technology-based knowledge (Sarica *et al.*, 2020, 2021). Beyond patents, crowd-sourced design solutions and ratings have also been provided to designers as sources of inspiration (Goucher-Lambert and Cagan, 2019; Kittur *et al.*, 2019). Goucher-Lambert and Cagan used natural language processing approaches to categorize near and far inspirational stimuli based on the frequency of terms appearing in crowd-sourced responses (2019).

Another category of approaches utilizing text-based functional relationships include those that facilitate search for analogies in biologically inspired design. Goel *et al.* used a structure–behavior–function knowledge representation (2009) to represent biological models and provide biological inspiration in multiple modalities, for example, text and visually represented behavior and structure models (Vattam *et al.*, 2011; Goel *et al.*, 2012). Representations of and relationships between biological analogies have been differently approached by Chakrabarti *et al.* to emphasize behavior of natural systems (e.g., motion) (2005). To implement keyword search for relevant biological analogies, Cheong *et al.* extracted a set of biologically meaningful keywords corresponding to functional terms in engineering (2011). Nagel and Stone further contributed a computational method that presents relevant biological concepts based on desired functionality, as searched for by the designer (2012). Object functionality can be differently defined based on the interaction context in which an object is used, which Hu *et al.* explored with a functional similarity network, a generative network, and a segmentation network (2018).

Less frequently explored in prior research are computational methods to support visual analogy. Setchi and Bouchard devised a method to index images based on semantic information from image labels and textual descriptions as one method of providing

https://doi.org/10.1017/S0890060422000130 Published online by Cambridge University Press

images as design inspiration (2010). To establish relationships between examples based on non-textual information, emerging methods using visual analogy in design have considered image-based search. Recent work by Zhang and Jin has demonstrated how visual analogy can be supported by sketch-based retrieval of visually similar examples (2020, 2021). Specifically, they used a deep-learning model to construct a latent space for a dataset of sketches and computationally determined visual similarities within this space (Zhang and Jin, 2020). Short and long-distanced visual analogies can then be identified based on the level of visual similarity shared between sketches (Zhang and Jin, 2021). Kwon *et al.* explored the use of image-based search to find visually similar examples to aid alternative-use concept generation (2019). Visual information, along with topic-level international patent classification (IPC) labels, have also been used in the retrieval of images from patent documents (Jiang *et al.*, 2020, 2021). Jiang *et al.* used a convolutional neural network-based method to perform image-based search using visual similarity and shared domain knowledge.

The methods and systems explored here are used to define text and visual-based relationships within various design stimuli repositories (e.g., patent images, design concepts, sketches, etc.). In prior research, these derived relationships have been used to identify stimuli related to a specified input, such as a design problem or search term. The current work also relies upon computational methods to extract semantic, functional, and visual information from potentially inspirational examples as well as designers' search inputs, as expressed through multiple modalities.

Motivating multi-modal search for inspirational stimuli

A second consideration for the platform developed in this work is the interface through which designers explore and discover inspirational stimuli. The role of non-text-based modalities in providing flexible modes of interaction and expressing search is examined. In general, expressing design ideas with visual attributes importantly supports cognitive processes of emergence and reinterpretation. Shape emergence is a process where designers perceive emergent patterns not initially intended in a visual stimulus (Soufi and Edmonds, 1996). Reinterpretation of visual stimuli is a process that leads to the formation of alternate interpretations and restructuring of design problems (Gross, 2001). During design exploration, these processes can importantly trigger new mental images and thus new ideas for design (Menezes and Lawson, 2006). Designers can benefit from interacting with a system through sketch-based inputs specifically, since in early-stage idea exploration, the act of sketching itself can assist with idea formation (Botella *et al.*, 2018). The ability for a creativity-support tool to uncover meaning from a designer's developing sketch, intent, and task context can be valuable for activating appropriate computational aid at the right time (Do, 2005). As an example, Kazi *et al.* developed DreamSketch, a sketch-based user interface that uses generative design methods, to provide designers with potential 3D-modeled design solutions based on early-stage 2D-sketch-based designs (2017). SketchSoup is another interface that inputs rough sketches and generates new sets of sketches, which may be explored and inspire further concept generation (Arora *et al.*, 2017). Interfaces that capture these sketch-based inputs can therefore be useful for supporting search and exploration of the design space.

In addition to 2D sketches, design ideas can be expressed in a 3D representation, for which creativity support is also possible. Through the InspireMe interface, Chaudhuri and Koltun provided data-driven suggestions for new components to add to a designer's initial 3D model (Chaudhuri and Koltun, 2010). Retrieval of inspiring examples based on 3D-represented design ideas can facilitate emergence and reinterpretation processes important for the design process. Conventionally, 3D-modeling environments recognize the unambiguous selection and placement of different elements to build a model, and thus provide limited support for new ideas to emerge or old ideas to be reinterpreted (Gross, 2001). It is also important to note that while CAD modeling enhances visualization and communication of ideas by providing a form to early design ideas, it may also cause premature design fixation and limit ideation (Robertson *et al.*, 2007). Systems capable of recognizing and reinterpreting conceptual or early-stage 3D design are valuable for overcoming limitations related to developing 3D models in a typical CAD environment.

Cognitive processes underlying search for inspiration

To implement useful features in the proposed search platform and link interactions with the platform to insights on search behavior, cognitive processes involved when searching for inspiration are reviewed. Early work on the role of search processes in design identified incidental experience and intentional learning as relevant sources of knowledge (Purcell and Gero, 1992). More recently, inspiration has been proposed as an iterative process that begins with an intention, is actualized by a search input, and ends when the problem has been solved (Goncalves *et al.*, 2016). In this process, active approaches to find specific stimuli more intentionally or passive approaches to randomly encounter relevant stimuli may take place (Herring *et al.*, 2009; Goncalves *et al.*, 2016). Active search refers to the deliberate search for a particular stimulus with a specific goal in mind (Eckert and Stacey, 2003). Alternatively, when what designers are searching for is unclear, they typically depend on randomly finding relevant stimuli. Randomness of web-based search, for example, has been found to be beneficial for inspiration due to the sometimes unexpectedness of results, related to more passive search strategies (Herring *et al.*, 2009). In information retrieval theory, search behavior has classically been categorized as exploratory versus specific (or lookup) (Sutcliffe and Ennis, 1998). Lookup search activities involve precise search goals whereas exploratory search is related to knowledge acquisition and evolving needs (Marchionini, 2006). Users have been found to examine more results in open-ended exploratory search tasks than during lookup tasks (Athukorala *et al.*, 2016). For computational tools to successfully support search for inspiration, user studies suggest that they should provide control and flexibility over the level of abstraction versus literalness of search terms (Mougenot *et al.*, 2008). To facilitate search for inspiration, it is important that active and passive search strategies are both supported. Designers should be able to express what they are looking for with a high level of agency and encounter inspirational stimuli more passively when what they are looking for is undetermined. Relevant to the current work, these insights into search for inspiration both guide the design of the search platform and provide a basis for interpreting the anticipated results of the cognitive study presented.

Platform development

To effectively support and subsequently study how inspirational stimuli are retrieved in the design process, a platform that enables

https://doi.org/10.1017/S0890060422000130 Published online by Cambridge University Press

similarity-based, multi-modal search for stimuli in the form of 3D-model parts was developed. This section describes in detail (1) the process of defining similarity among stimuli using deep-neural networks and (2) the development and design of the multi-modal search interface that participants interacted with in a cognitive study.

Neural network development enabling inspirational stimulus retrieval

A major component of the platform is a system that supports the search for and retrieval of inspirational design stimuli in large datasets using multi-modal inputs. Relying solely on text-based search using semantic relationships may limit discovery of inspirational stimuli to concepts that are well enough defined to express using words. As described previously, search processes using more passive or exploratory processes are sometimes preferred and may not be as well supported by tools requiring such direct input (Goncalves *et al.*, 2016). Introducing new modes of expressing search may be one approach to aid different search strategies when needed during the design process. As such, the proposed system is designed to support queries of 3D-model examples while also maintaining support for text-based queries. Beyond aiming to support additional query modalities, the platform provides a measure of similarity that allows users to control the similarity level between retrieved examples and their multi-modal queries. It is believed that having such agency in the system will allow researchers to better understand and analyze users' intentions in the retrieval process of inspirational examples.

Using 3D-model parts as inspirational stimuli

To support the research goals in the current work, a large-scale dataset of 3D models is used to train the deep-neural networks and provide users with relevant examples. Specifically, the PartNet dataset was used, which contains 26,671 unique 3D models (assemblies) in 24 object categories, each further splitting into trees of individually named parts within each assembly (e.g., cap as a child of bottle) (Mo *et al.*, 2018). Names for each part are assigned in the dataset through expert-defined annotations. In total, the dataset contains 573,585 part instances, across 24 object categories. Each object category contains varying numbers of part instances. For example, bags, hats, bowls, mugs, and scissors contain on the order of ∼1 K parts whereas vases, trash cans, lamps contain ∼10 K parts and chairs, storage furniture, and tables contain >100 K parts. The 24 object categories include everyday objects at various scales (e.g., microwave, scissors, tables). Since these categories represent only a small subset of possible objects that mechanical engineers might design, part-based data within these objects are instead used and presented in the proposed system. These parts (e.g., legs, cover, lid) may be present in object categories beyond those in the dataset. This allows the system to cover diverse design cases and to potentially provide inspiration between distant design goals. While the PartNet dataset was used in this work, alternative datasets could be leveraged that similarly contain large-scale, hierarchical, fine-grained annotations of data.

The use of such a large-scale 3D-model dataset also allows the system to leverage data-driven, deep-learning-based methods. These methods are used to *extract computationally derived similarities* between stimuli within the platform, based on their semantic, visual, and functional features. The platform uses deep-neural networks to contrastively model similarities between design examples (3D-model parts in this application) and natural-language-model keywords. This deep-learning approach directly consumes 2D snapshots of 3D-modeling parts and utilizes knowledge from large text corpuses, which subsequently enables the efficient retrieval of relevant examples in the large dataset used. Deep-neural networks are suitable candidates for this task because they are highly effective in understanding complex patterns in high-dimensional data, such as multi-perspective image snapshots of 3D models in the platform. In their review of data-driven methods to support design-by-analogy, Jiang *et al.* identify deep-learning models as an effective technique for learning complex features from datasets (Jiang *et al.*, 2022).

Computationally deriving similarity between 3D-model parts

Using the PartNet dataset, three neural networks were used with the intent to embed raw 3D-model data to high-level concepts and modeling parameters to be used in the platform. Each of these networks handles a unique modality or type of similarity. These networks are respectively (1) a text network that encodes similarity of design concepts in natural language; (2) an appearance network that encodes similarity of 3D models only by their appearance and geometric presence; and (3) a functionality network that extends beyond (2) to encode similarity of functions of 3D models based on their neighboring 3D parts.

The text network in the platform relies on the Universal Sentence Encoder (Cer *et al.*, 2018) pre-trained on web text to find parts with names similar to the keyword queries provided by users. The Universal Sentence Encoder is trained on nontechnical text to solve general text-understanding tasks such as sentiment analysis and question classification. As a result, the model should be able to obtain a general semantic understanding of English words and thus be able to identify synonyms (e.g., "box" should be semantically similar to "container" in the embedding space). Alternative semantic networks exist beyond the Universal Sentence Encoder, such as TechNet (Sarica *et al.*, 2020), which consists of technology-related terms. However, since the PartNet dataset contains everyday objects that are not highly technical, the use of the Universal Sentence Encoder to understand common words is sufficient for this work. The Universal Sentence Encoder is also effective for working with, not only sentences, but short phrases, which other semantic embedding methods, for example, BERT (Bidirectional Encoder Representations from Transformers), are not trained on (Devlin *et al.*, 2019).

The appearance network was trained by embedding knowledge from 2D snapshots of 3D-model parts. This network is trained to consider snapshots of the same 3D model as "similar" to each other, and snapshots from different 3D models as "dissimilar" from each other in the embedding space. This leads to the model learning the general physical form and presence of the 3D model by visually analyzing it from different angles. More concretely, consider a training example (x) as a 3D-modeling part (e.g., a leg of a chair). Eight 2D snapshots (images) $[S(x)_i, i \in 1, \ldots, 8]$ of the part are first taken by rendering the part in Blender. Snapshots are normalized to the size of the image, meaning that the whole part takes the size of the entire image and the relative scale of the part is not considered. After obtaining these screenshots, each snapshot is passed through a neural network f to get a single n-dimensional real-valued embedding $f(S(x)_i) \in R^n$. These embeddings for other examples in the dataset were similarly gathered.

https://doi.org/10.1017/S0890060422000130 Published online by Cambridge University Press

To train this model toward the goal of considering snapshots of the same 3D model as similar, other examples are also needed to allow the model to contrast the dissimilar embeddings with the similar embeddings. As such, to obtain a loss function for each embedding in the dataset, real-value embeddings of multiple other 3D models in the dataset were also gathered. Without loss of generality, another randomly sampled but different 3D model (a) and its embeddings $f(S(a)_i)$ were considered in the following formulation. The model was then trained with sampled *positive* pairs that consist of snapshots that come from the same 3D model,

$$p^+ = (f(S(x)_i), f(S(x)_j)),\ i \neq j, \tag{1}$$

and negative pairs:

$$p^- = (f(S(x)_i), f(S(a)_j)). \tag{2}$$

The following training objective L [in Eq. (3)] was used to minimize the distance [measured by the distance function (D)] of positive pairs and maximize the distance of negative pairs (up to the margin m):

$$L = D(p^+)^2 - (\max(m - D(p^-), 0))^2, \tag{3}$$

$$\mathrm{D}(a, b) = ||a - b||_1. \tag{4}$$

On a high level, this model considers these snapshots as similar among themselves and dissimilar to snapshots of other 3D models in the latent space. Such similarity is considered primarily by the overall appearance and geometric presence of the 3D-model parts.

The functionality network was built to learn a slightly different notion of similarity than the appearance network. While considering the exact functions of different 3D models could be difficult and greatly depend on context, as a first step toward this goal, 3D models are considered to be similar if they have similar neighboring parts within their respect assemblies. Hu *et al.* demonstrate the effectiveness of this approach in capturing the function of 3D models through the usage contexts of the models (2018). Using this method means that 3D-model parts that perform a certain function should have similar neighbors in their respective assemblies (e.g., different styles of chair legs, despite having different appearances, are considered similar since they share "chair seat" as a neighbor). The functionality network builds upon the appearance network such that it takes the appearance embeddings and transforms them into function-aware embeddings. The functionality network is trained with a very similar paradigm as the appearance network, with an almost identical loss function to Eq. (3). The only difference is that the functionality network (g) is now used to obtain a transformation of the appearance embeddings [$g(f(S(x)_i))$], and the group of similar parts extend beyond the snapshots of a single 3D-model part itself to neighboring parts. For instance, given a chair leg x and a chair seat z, and an irrelevant lamp cover b, positive pairs are formed

$$p^+ = (g(f(S(x)_i)), g(f(S(z)_j))), \tag{5}$$

as well as negative pairs:

$$p^- = (g(f(S(x)_i)), g(f(S(b)_j))). \tag{6}$$

These pairs are then trained using the same loss function [Eq. (3)]. Figure 1 displays how the functional embeddings are derived from appearance embeddings using the described networks.

The appearance network consists of five stacked groups of a convolution layer with kernel sizes of 5×5 or 3×3 followed by a 4×4 or 2×2 max pooling layer (see Fig. 1 for arrangement). This network also consists of a final 4×4 convolution layer that flattens the output to 128-dimensional appearance embeddings. The functional network then takes these appearance embeddings and passes them through its four stacked 128-dimensional fully connected layers and one 64-dimensional fully connected layer to produce 64-dimensional functional embeddings. On a high level, these embeddings encode the context of the usage of the 3D-model parts and consider 3D-model parts that are used along with other parts as similar by assuming that they have similar functions.

Implementation and training of neural networks

The appearance and functionality models are implemented with Tensorflow Keras. An Adam Optimizer with a learning rate of 0.001 was used to train each model until the validation loss plateaued. The appearance model took 26 h to train on a machine with two GPUs (a NVIDIA GeForce 1080 Ti and a Titan X Pascal), while the functional network took 10 h to train on the same machine. The text network did not involve any training as it is directly taken from the pre-trained Universal Sentence Encoder provided in Tensorflow Hub.

Front-end user interface for multi-modal search

Leveraging the underlying platform for inspirational stimuli retrieval described in the previous section , functionality for multi-modal search was subsequently enabled for use in a cognitive study. This was achieved by comparing the semantic, visual, and functional features of the participant's input to the parts in the dataset populating the platform. The modalities of input available and additional features of the search interface are discussed below.

Search modalities: keyword, part, and workspace-based

Using the search interface that relies on the neural networks described above, participants were able to search for parts in the dataset using three types of input. The first search type is keyword-based, where text input by the participant is embedded using the text network, as described in the section "Computationally deriving similarity between 3D-model parts". Embedding values are then compared against those of the dataset's part names and the nearest neighbors from the dataset are retrieved. The results from a keyword search for the term "container" is shown in Figure 2a. The second and third search types are part-based and workspace-based, where new parts are retrieved using visual snapshots taken of a selected 3D-model part or the participant's current workspace (composed of 3D-model parts), respectively. For workspace searches, snapshots of the whole workspace are taken, which may include multiple parts. These snapshots are passed through the appearance and functional networks and the resulting appearance and functional embedding values are compared with those of other parts in the dataset. The same computational approach used to derive similarities between 3D-model parts, as described in the section "Computationally deriving similarity between 3D-model parts",

https://doi.org/10.1017/S0890060422000130 Published online by Cambridge University Press

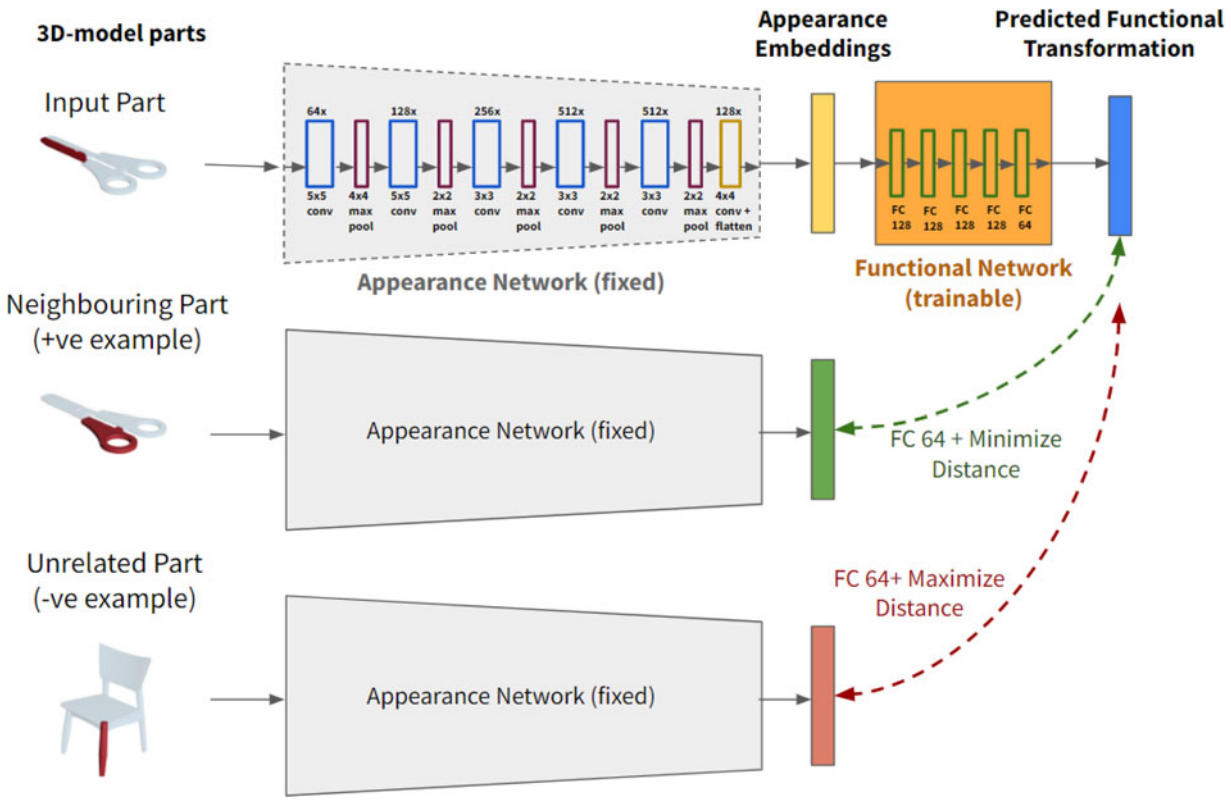


Fig. 1. Overview of process of transformation of embeddings from appearance network to functional embeddings. Appearance embeddings of **input part** (scissor blade) were used to generate a predicted functional transformation using the functional network. Functional network was then trained by considering this prediction as similar to the appearance embedding of a **neighboring part** (scissor handle) and dissimilar to the appearance embedding of an **unrelated part** (chair leg). Intermediate representation within the functional network was used as the functional embedding of each model part in the dataset.

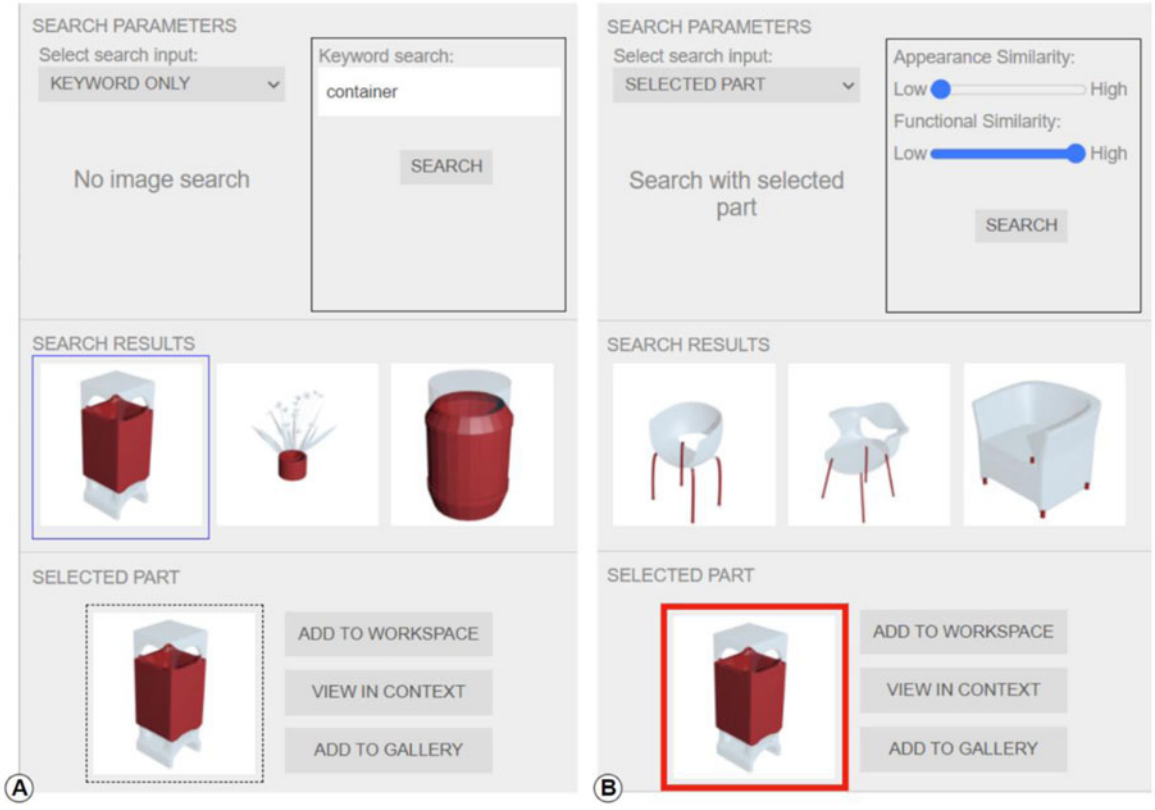


Fig. 2. (a) Search results for a **keyword search** of the term "container"; (b) search results for a **part search** of a result from keyword search for "container".

https://doi.org/10.1017/S0890060422000130 Published online by Cambridge University Press

is used to produce embedding values for search inputs from the relevant neural networks.

Part and workspace-based searches are made using two additional user-specified parameters, appearance similarity and functional similarity, which participants can specify in the platform interface with sliders. The closest neighbors are retrieved for the participants according to the weighted sum of the distances specified by the appearance and functional sliders in the user interface. Figure 2b shows the use of similarity sliders and the search results for a part search of the first keyword search result for "container". Sliders controlling similarity in appearance and function allow participants to conduct multiple searches using the same part or workspace input with increased agency. In the example shown in Figure 2b, parts are searched for with low similarity in appearance but high similarity in function to the selected container. As represented in Figure 1, neighboring parts of visually similar parts to the input part are considered functionally similar to the input. In this example, the shared visual characteristics between the chair seats and container have caused chair legs to be considered functionally similar to the container. Based on the results retrieved, participants are then able to modify these inputs to continue to search for new results.

Interactions with parts retrieved from search

After relevant 3D-model parts are retrieved from the dataset, the model pushes the images of the 3D models, as well as their associated STL files, to the web front-end of the platform, which is based on the editor code of the open-source three.js library. Participants can thus preview three of the retrieved 3D models in the "Search Results" panel of the interface (Fig. 2). An example in Figure 3a shows how parts can also be added to and modified in the user's 3D workspace using the "Add to Workspace" button. Workspace-based searches are made with snapshots of the entire workspace with parts added by the participant using this action. Moreover, since all results are retrieved from the PartNet dataset, which contains information on neighboring parts in the assembly of the results, participants may view this information (Fig. 3b) using the "View in Context" button. For a selected part, this action allows further understanding of the retrieved parts' utility in their original context. Finally, participants have the ability to use the "Add to gallery" button to save a part to a gallery of collected 3D parts (Fig. 3c). The gallery is accessible to the participant to access and select parts from at any point during the design task. For any given search result, participants could perform none to all actions, in any order.

These search modalities and part interactions are envisioned to enable search for inspiration during early-stage design. Keyword and part searches may provide initial, rapid inspiration by retrieving results based on the designer's text-based query or based on similarity to a previously discovered part. Workspace search, more similar to 2D or 3D sketch-based retrieval platforms introduced in the section "Motivating multi-modal search for inspirational stimuli" (e.g., SketchSoup, InspireMe, DreamSketch), can further support the discovery of inspiration during later stages of design, based on the designer's developing 3D-model. In general, the various representations of inspiration provided by the platform (i.e., 2D representation, 3D representation, text label) make it suitable for aiding various stages and forms of early-stage design, such as in generating conceptual design ideas, 2D sketches, or 3D sketches.

Cognitive study design

To understand the processes and behaviors associated with searching for and exploring design examples, a cognitive study was conducted using the platform. During the study, participants searched with different modalities available in the platform to find and select relevant 3D parts that could help inspire solutions to a design challenge. The main approach taken in this work was to analyze participants' interactions in the platform and relate these actions to strategies involved in searching for inspirational examples. A 30-minute study was administered to understand how participants engaged with the three search types available in the platform. Participants searched for parts using each search modality in three separate subtasks and worked toward collecting inspirational stimuli for a given design challenge.

Participants

Participants were recruited from announcement emails sent to undergraduate and graduate mechanical engineering students at the University of California, Berkeley. Twenty-three participants (15 males and 8 females) with varying levels of design experience, ranging from less than 1 year to 9 years, volunteered for the study. Participants were offered $10 compensation for their participation in the 30-minute study. Due to data collection errors, data from two participants were excluded from the analysis. All participants

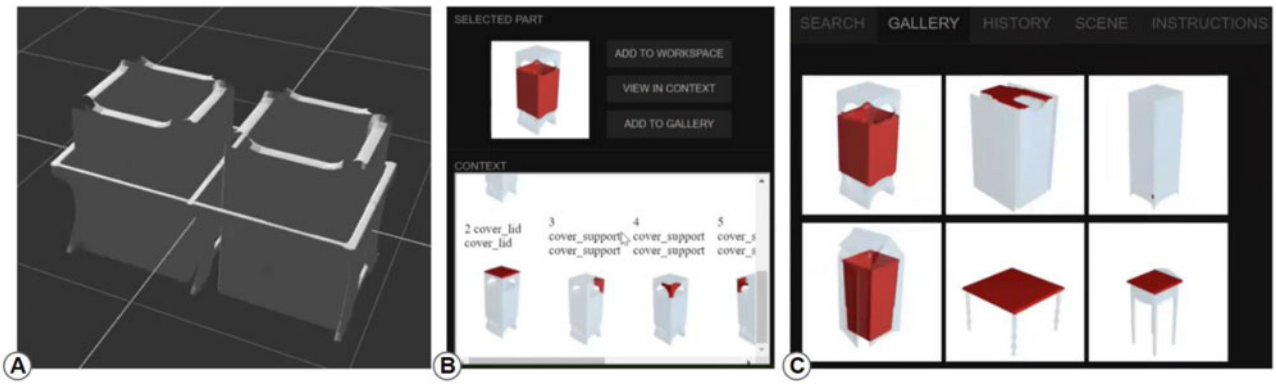


Fig. 3. Interactions with selected part in Figure 2 – (a) adding part to the workspace; (b) viewing part in context by seeing related parts with text labels in the same object assembly; and (c) adding part to a gallery of saved 3D parts.

https://doi.org/10.1017/S0890060422000130 Published online by Cambridge University Press

completed the study while connected virtually with the experimenter over a Zoom meeting, where all participants consented to sharing their screens for the duration of the task. Any issues completing the task or clarifications needed could thus be addressed in real time.

Study objective

The study objective presented to participants was to use the platform to search for, and save, 3D parts that inspired solutions to the following design challenge: "design a multi-compartment disposal unit for household waste". Participants were told that parts inspiring solutions to the design challenge could include those they might want to directly incorporate into potential solutions. The design challenge presented to participants was developed to fit the context of the search platform, which is populated with parts related to household objects. Pilot testing revealed that this design prompt engaged several object categories in the PartNet dataset, including some that are highly relevant to the task (e.g., trash can, storage furniture).

Study overview

The study was divided into three subtasks (A, B, C), as summarized in Figure 4, where each task involved the use of a different search type (keyword, part, workspace), but worked toward the same design challenge. The study objective and task instructions were embedded in a Qualtrics survey link sent to participants at the start of the experiment. For each subtask, participants read the associated training and instructions, and then completed the task in an external link. At the end of the experiment, participants responded to a series of open-ended and multiple-choice questions about their experience using the search platform. Table 1 additionally summarizes the search types, inputs, and requirements of each subtask of the study.

Task A: In Task A, all participants were instructed to first search by keyword beginning with the term "container" (Fig. 2a). They were instructed to make four additional keyword searches (min.) and to save min. three parts to their galleries.

Task B: Participants then continued with their progress from Task A in Task B by conducting a part search with a part saved to their gallery during Task A. As before, the instructions were to conduct min. four additional part searches and save min. three more parts. Participants were also instructed to not make any additional keyword searches.

Table 1. Overview of search types and inputs specified for each subtask of the cognitive study

Subtask	Search type	Initial search input
A	Keyword	"Container"
B	Part	Any previous keyword search result
C	Workspace	Workspace consisting of parts retrieved in Subtasks A and B

Task C: Finally, in Task C, participants conducted workspace searches and made their first search consisting of parts either previously added to the workspace, or newly added from parts saved during Tasks A and B. A min. of four additional workspace searches were made and a min. of three parts were saved, without making any new keyword or part searches.

This study design, while constrained, ensured that participants used each search modality for a comparable portion of the design study, enabling an investigation into the use of the search platform's modalities and features. Without prescribing these constraints, for example, a minimum number of searches, sufficient interaction with each search modality and feature may not have been observed, given participants' lack of familiarity with the novel search inputs introduced.

The motivation for the ordering and division of tasks was to easily teach participants how to engage with the search platform. The order was selected since parts need to be discovered initially through keyword search to subsequently conduct part and workspace searches. Tasks were ordered to first use the most intuitive search mode (keyword) and to last introduce the least familiar and most difficult mode (workspace). Pilot testing revealed that learning about each search type at study onset overloaded participants with too much information to effectively engage with each search type, therefore each search type was introduced and used in separate tasks.

After completing the study, participants were asked to provide open-ended descriptions of any strategies used when conducting each type of search. Participants also evaluated the intuitiveness and usefulness of different features in the platform on five-point Likert scales. These features included searching for new parts and gaining more information about parts. Finally, participants self-evaluated the broadness of their exploration of the part repository and of their final gallery of saved parts on five-point Likert scales.

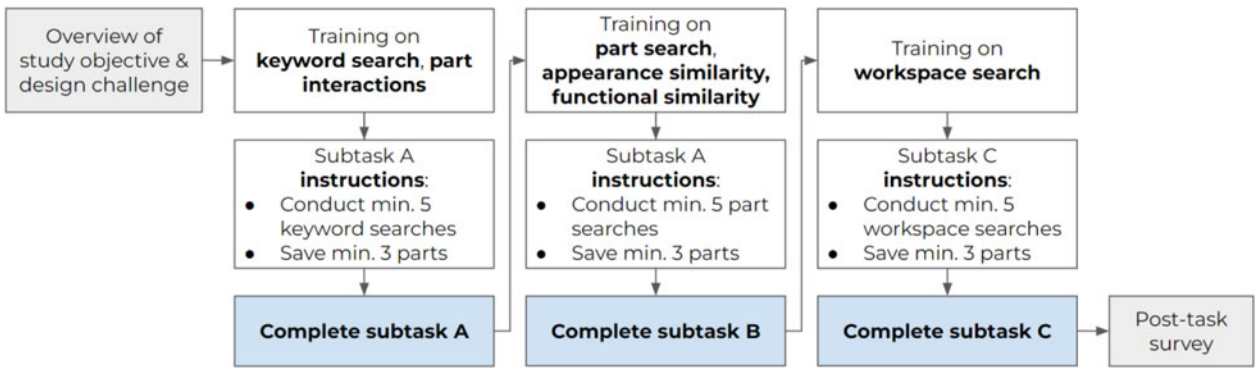


Fig. 4. Overview of flow between subtasks: training on search types and features in the interface preceded presentation of instructions and completion of each subtask.

https://doi.org/10.1017/S0890060422000130 Published online by Cambridge University Press

Results

The two main outcomes of this work are next presented. The first significant outcome is the development and illustration of the search platform defined in the section "Platform development". A quantitative illustration of the retrieval behavior with reference to the related definitions of similarities used in the neural networks underlying the search platform are described. The second outcome consists of the insights gained into search behavior when using this platform during a cognitive study. The study's findings were examined from both the perspectives of what participants discovered and how participants searched for inspirational stimuli through the use of different search modalities.

Quantitative retrieval behavior of neural networks

Prior to studying how participants interacted with the examples using the built system, it is important to understand the intrinsic ability of the system and its networks in accomplishing the feature that it is designed for, that is, providing users agency to control similarity of the retrieved examples to their input queries. A few quantitative measures are presented in this subsection based on several definitions of similarity that can be computed automatically based on existing datasets. This provides a partial, but objective, understanding of the system's networks' ability to retrieve similar visual stimuli that allow us to understand the characteristics of the models in relation to other definitions of similarity. These similarity definitions are also directly relevant to the development of the platform outlined in the section "Computationally deriving similarity between 3D-model parts". This subsection first outlines the procedure taken to compute overall retrieval behavior of the proposed networks quantitatively using any similarity definition. Three definitions of relevance for the stimuli retrieval task are then described and the measures specific to each of these definitions are then calculated and presented using the same procedure. While these measures provide a comparative and quantitative view of the model's behavior, it is important to note that these definitions of similarity are **not** ground-truths of our task (which does not currently exist) and are intended to serve as a reference to help readers better contextualize the behavior of the models supporting our platform. As Jiang *et al.* note in their review of data-driven design-by-analogy methods, there is a lack of gold-standard benchmarking for tools supporting visual and multi-modal analogies (Jiang *et al.*, 2022).

Ranked-based accuracy computation

A ranked-based measure common in retrieval system research is used to measure the proposed networks' retrieval performance given a binary definition of similarity (i.e., given two examples, this is the ground-truth label of whether the models are similar or not). In the context of this work, each example is a 2D snapshot of a 3D model taken from a random angle. All 2D snapshots of other models in the dataset were then embedded. Then, for each embedding that corresponds to a snapshot, the most similar other embeddings ranked using a common distance metric are then queried. This process is analogous to the querying process where the initial snapshot acts as the search query: For each querying 3D model, we compute its nearest neighbors using neural networks. Then, top-1 accuracy is the percentage of these queries where a "similar" (as defined in the sections "Definition and results of self-similarity" and "Definition and results of concept-based similarity") 3D model is ranked as the top nearest neighbor, and top-10 accuracy is the percentage of these queries where a similar 3D model is ranked as one of the top-10 nearest neighbors. For instance, top-1 accuracy refers to the percentage of 3D models that rank a similar model (according to the ground truth) as the first most similar embedding (according to the proposed networks). Top-10 accuracy was similarly reported in the results by relaxing the embedding similarity rank to any of the top-10 ranked similar embeddings. Note that the snapshots of the query snapshots themselves were excluded (but not other snapshots of the same 3D model) from the ranking procedure as the same snapshots would produce identical embeddings in the network, hence degenerating the measurements.

It is important to note that the accuracies were computed individually on three splits of the dataset. The networks were trained on 70% of the data and the rank-based accuracy measures were computed on a held-out test set (20% of the data) and a validation set (10% of the data). It is most important to consider the accuracies on the held-out test set, a set of 3D models that were unseen to the networks and the authors prior to the computation of the accuracies. These accuracies best reflect the generalization capability of the networks to new, unseen data of 3D models.

Definition and results of self-similarity

The first and most direct measure of relevance is only considering the model itself as relevant. As the networks were trained to encode different snapshots of the same 3D model to be similar (outlined in the section "Computationally deriving similarity between 3D-model parts"), the network's behavior could be more intuitively considered by its ability to consider different snapshots of an unseen 3D model (in the validation set) as similar.

With this definition of similarity (all snapshots corresponding to the same 3D model in the respective sets are labeled as similar), top-1 and top-10 accuracies are computed using procedures outlined in the section "Ranked-based accuracy computation". For training/validation/test sets, the top-1 accuracy of the appearance network is 1.13%/3.53%/2.24% and the top-10 accuracy of this network is 5.78%/17.0%/11.3%, where eight 2D snapshots from each 3D model in the dataset are compared. For the functional network, the top-1 accuracy is 2.93%/8.42%/5.73% and the top-10 accuracy is 15.2%/32.0%/23.5%, with an additional relaxation of definition that all parts that belong to the same model are considered similar. When assessing self-similarity, training set accuracy is lower than validation and test set accuracies. The larger size of the training set and thus the density of subspaces of similar mechanical parts in the latent space of embeddings can account for the lower training set performance. The retrieval of similar parts that are not the same model, and thus not considered similar by this metric, is more likely in the training set. These metrics are highly conservative estimates of the network's actual task performance as there are far more relevant screenshots of 3D models than those generated from the same models as the input query (i.e., there are many types of chair legs in the dataset). To illustrate such difficulty of our task, the accuracy of a naïve random sampler can be computed and compared to our task. A random sampler would have achieved a top-1 accuracy of 0.000416%/0.00273%/0.00137%,[1] which is three orders of magnitude lower than the network's measures. This shows that the networks used by our platform demonstrate reasonable behavior and

[1]This also explains the difference of accuracies between the three sets because the naïve probabilities of retrieving the ground-truth 3D model, which reflects the task accuracies, are different across sets from the beginning.

https://doi.org/10.1017/S0890060422000130 Published online by Cambridge University Press

leads to the development of other definitions of similarity described in the remainder of this section.

Definition and results of concept-based similarity

Beyond the 3D models themselves, there are many other models that can be considered similar semantically in the dataset. For instance, there are many chair legs that could be similar. To account for this relaxed definition of relevance, the text annotations of 3D models available in the dataset used (PartNet) were utilized to consider similarity. Two 3D models were considered to be relevant if both consist of exactly the same text label. These text labels represent larger clusters of 3D models in the dataset. The procedure outlined in the section "Ranked-based accuracy computation" is similarly used to compute top-1 and top-k similarities using this definition of text-concept-based similarity. For training/validation/test sets, the top-1 accuracy of the appearance network is 25.6%/26.8%/26.8% and the top-10 accuracy is 66.1%/69.1%/68.1%. For the functional network, the top-1 accuracy is 40.5%/39.9%/40.4% and the top-10 accuracy is 85.7%/85.7%/85.8%, with the further relaxation of definition that all text annotations of other parts that belong to the same model are considered similar.

Definition and results of physical form relevance

Besides calculating semantic relevance using text labels, relevance can similarly be computed by the physical forms of the models. The physical similarity of two models is computed by the three-dimensional intersection over union (IoU) of the models, such that the models are super-positioned to find the overlapping volume, which is then divided by the sum of the total volume of the models. To ensure the consistency of this measurement, six extra random orientations (in addition to the default position of the models) were taken between the models during super-positioning and the maximum value of the seven orientations was taken as the final IoU. Two models are considered as similar if the IoU between them is within the top 5% out of all other pairs for a particular model. This criterion controls the difficulty of our retrieval task such that a completely random retriever would get 5% top-1 accuracy in such a task.

The above process requires the models to be closed for the volume computation to be correct. Therefore, the convex hulls of both models and the intersected volumes are further computed to ensure correctness. These computational steps of convex hull, IoU, and volume of models are done with Blender 2.8.2. Moreover, since this process is computationally expensive (and scales quadratically with the size of the candidates), 200 random examples were sampled from the test set of the PartNet dataset and then manually reviewed to be approximately convex before being used as candidates of this experiment. The final top-1 accuracy for this similarity criteria on these candidates for the appearance network is 65.9% and the top-10 accuracy is 95.5%. We did not report this measure for the functional network due to the high number of parts required to relax this definition of physical form relevance to include all parts belonging to the models containing the sampled parts.

Overall, the ranked-based accuracy measures computed in this subsection provide insights for the retrieval behavior of the neural networks underlying the search platform used in the cognitive study. Different perspectives of similarity are considered including self, semantic (concept-based), and visual (physical form) similarity that allows us to further understand this neural network-based methods' retrieval characteristics. These measures are summarized in Table 2. The development and behavior of the platform are further discussed in the section "Discussion of multi-modal search platform development and behavior".

Results of cognitive study

The developed platform, which uses the similarity relationships described in the previous section, allows users to search for and interact with retrieved parts. In this section, the results of a cognitive study are presented, which demonstrate how participants search for stimuli in the platform and the content of these retrieved parts. In the cognitive study, participants searched for 3D-model parts using keyword searches in Task A, part searches in Task B, and workspace searches in Task C. Throughout the study, the following actions could be taken on any search result: adding it to the workspace, viewing it in context, or saving it to a gallery. This work considers how these interactions reveal the ways different modalities of expressing search affect and support the search process. The focus of the present study is on investigating the use of the described search platform to search for inspirational stimuli, and not necessarily the impact of these stimuli on design outcomes. Specific objectives of this study are to identify differences in search modality by how participants (1) search for new parts and (2) engage with the retrieved parts, as well as (3) what participants discovered. These results extend upon findings in prior work by Kwon *et al.* (2021).

Search for new parts

To understand how different search modalities are used, the search inputs defined by participants are first discussed. Specific inputs that participants tend to modify across searches are important to identify to support multi-modal search. Frequencies of each search modality used, and slider movements made, are analyzed to examine how participants used the platform to search for new parts. Differences between search types in the total number of searches made were compared with a Chi-square test. The number of searches made using each search type, including repeated searches, significantly differs [χ^2 (2, $N = 677$) = 9.8, $p < 0.01$], where the number of part searches (264) compared to keyword (207) and workspace (206) searches is the highest. These

Table 2. Summary of rank-based accuracies for similarity measures of test set data describing retrieval behavior of neural networks

	Appearance network		Functional network	
Similarity measures	Top-1 accuracy (%)	Top-10 accuracy (%)	Top-1 accuracy (%)	Top-10 accuracy (%)
Self-similarity	2.24	11.3	5.73	23.5
Concept-based similarity	26.8	68.1	40.4	85.8
Physical-form similarity	65.9	95.5	N/A	N/A

https://doi.org/10.1017/S0890060422000130 Published online by Cambridge University Press

implications of the differences in search frequency are further discussed in the section "Discussion of cognitive study results".

Total search counts include all keyword searches, and both new and modified part and workspace searches. New part searches are defined as those where a unique part is used as the search input. New workspace searches are made when the workspace input contains a newly added part. Participants could also make modified searches, where the same part or workspace from a previous search is selected and adjustments are made only to sliders specifying appearance and functional similarity. The numbers of new and modified part searches made across participants are summarized in Table 3. Also included are the numbers of modified part searches that increase (+) or decrease (−) appearance and/or functional similarity from a previous search.

As shown in Table 3, more total number of searches are conducted with the same part (131) than a different part (104) from the previous search. However, when examining the proportion of new and modified searches made by each participant, a repeated measures ANOVA did not reveal a significant difference [$F(1,20)$] = 0.55, $p = 0.5$). Modified search counts combined across participants vary significantly with respect to whether modifications are made in functional similarity (37), appearance similarity (60), or both (34) [χ^2 (2, $N = 131$) = 9.3, $p < 0.01$]. The proportion of these modified searches within participant does not differ across modification types [$F(2,40)$] = 0.03, $p = 0.97$). This result signifies that while some participants may have conducted many appearance-modified searches, this was not observed across all participants. Of the 21 participants, only 4 conducted more than 5 appearance-modified searches.

The same analysis performed for part searches was done to identify how workspace searches were made, as summarized in Table 4. The number of workspace searches made with modifications to functional (24), appearance (28), or both types of similarity (24), did not significantly differ. Different from part searches, more workspace searches are made with new search inputs (i.e., with an added part to the workspace) than with the same workspace configuration (105 vs. 76). A significant difference was observed in the proportion of new and modified workspace searches made by each participant, as revealed by a repeated measures ANOVA test [$F(1,20)$] = 7.43, $p < 0.05$). These combined results demonstrate how search inputs and desired similarity are differently defined when engaging with part versus workspace searches.

Engagement with parts retrieved from search

The search platform, beyond supporting retrieval of parts, allows participants to further engage with the shown parts. Participants can engage with a part through the interactions enabled in the platform, as outlined in the section "Interactions with parts retrieved from search", for example, by viewing it in context (to gain contextual information), adding it to the workspace (to see and manipulate its 3D representation), or saving it to their gallery. The number of times each interaction was made was counted to determine how results from each search type are engaged with differently. Frequencies of interactions with search results were compared across search modalities to assess differences in how participants engage with results retrieved from each search modality. There is a significant difference between search modalities in both the total number of search results that users engaged with [χ^2 (2, $N = 106$) = 18.6, $p < 0.001$] and did not engage with [χ^2 (2, $N = 581$) = 23.0, $p < 0.001$], as shown in Figure 5. This result suggests that, despite being instructed to conduct the same number of searches using each search modality, participants interacted differently with each search modality and the parts retrieved.

The differences in frequency between the expected and observed values for each set of results are plotted in Figure 5. The expected value is the total number of parts engaged with (106) or not (581), divided by 3 (the number of tasks). This value represents the number of parts expected to be engaged with or not in each task if no task differences exist. Parts that are engaged with include those viewed in context, added to the workspace, or saved to the gallery. Parts not engaged with are those retrieved from search and seen by the participant, with no further interaction made. The highest proportion of parts that were further engaged with were retrieved by keyword search, while results not engaged with were mostly those retrieved by part search. On average, participants spend 343 s, 195 s, and 451 s in subtasks A, B, and C, respectively. These results suggest that increased part engagement of keyword search results does not occur due to increased time spent at the beginning of the study. Reduced time spent on subtask B, despite high frequency of part searches, further demonstrates participants' lack of engagement with these search results.

To more closely consider how users engage with search results, the number of parts in each task that are viewed in context or added to the workspace are compared. The number of parts viewed in context significantly differs between tasks [χ^2 (2, $N = 104$) = 13.3, $p < 0.01$]. Displayed in Figure 6, results from keyword search are more frequently viewed in context than expected, and fewer results from workspace search are viewed than expected. As in Figure 5, expected values in Figure 6 refer to the total numbers of parts viewed in context (104) or added to the gallery (101),

Table 3. Frequencies of new and modified part searches with changes in functional and/or appearance similarity (+: increasing similarity, −: decreasing similarity)

Search input		Search counts
New search (different part from previous)		104
Modified search (same part as previous)	Change in functional similarity	37 (17+, 20−)
	Change in appearance similarity	60 (30+, 30−)
	Change in both similarity types	34
	Total modified searches	131

Table 4. Frequencies of new and modified workspace searches with changes in functional and/or appearance similarity (+: increasing similarity, −: decreasing similarity)

Search input		Search counts
New search (new part added to workspace)		105
Modified search (no new parts added to workspace)	Change in functional similarity	24 (14+, 10−)
	Change in appearance similarity	28 (19+, 9−)
	Change in both similarity types	24
	Total modified searches	76

https://doi.org/10.1017/S0890060422000130 Published online by Cambridge University Press

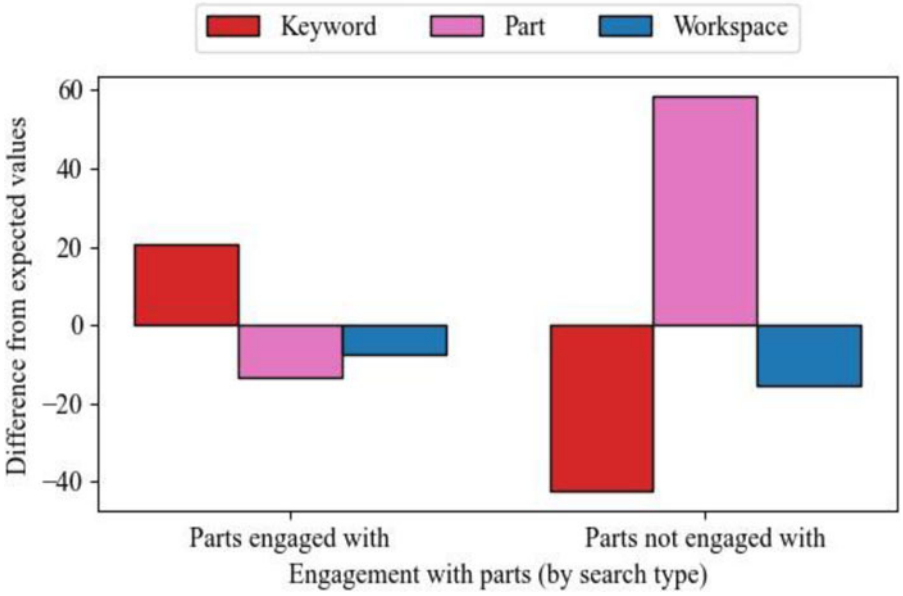


Fig. 5. Differences between observed and expected values of parts engaged with and not engaged with, shown by search type.

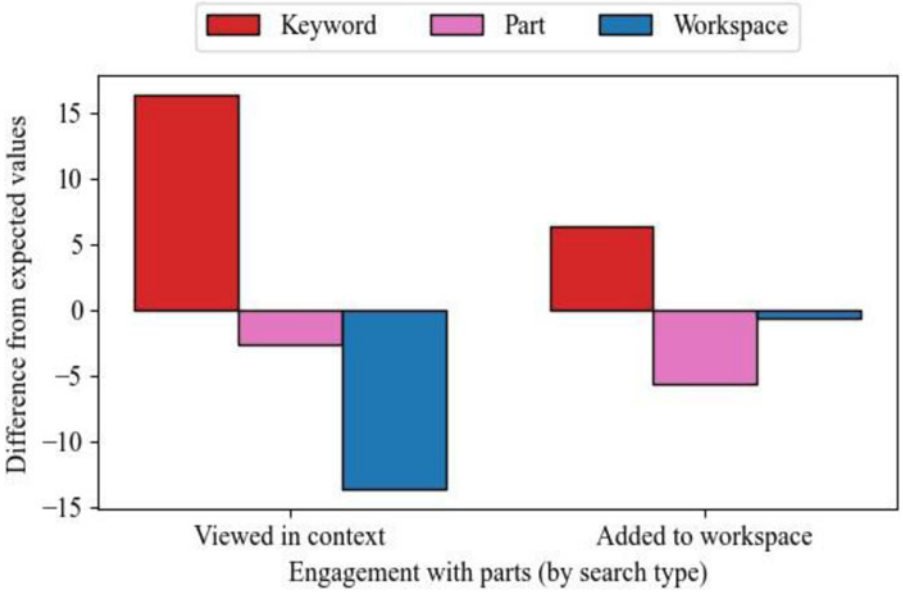


Fig. 6. Differences between the observed and expected values of parts viewed in context and added to the workspace, shown by search type.

divided evenly between tasks. Numbers of parts added to the workspace do not differ significantly between tasks. Combined, these results suggest that keyword search encourages increased engagement with individual results, while part-based search does not. A more detailed analysis of these results can be found in the section "Discussion of cognitive study results".

Coverage of design stimuli space by retrieved parts

Regarding participants' interactions with the interface, the role of search modality in the overall discovery of inspirational stimuli is investigated using a measure of coverage of the design stimuli space. By deriving a measure of coverage, the relative diversity of parts within the appearance and function-based embedding spaces discovered using each search modality can be compared. The parts retrieved by all participants throughout the study are first shown based on their representation within the appearance-based neural network in Figure 7. Parts are color coded based on the search type used when they were retrieved. The visualization represents the parts reduced from the 128-dimensional appearance-based neural network to a two-dimensional space using principal component analysis (PCA). The reduced space accounts for 72.8% of the total variance of the original space. As highlighted, examples of closely and distantly related parts in appearance are shown in the 2D projection of the embedding space. The pair of closely related parts displayed are also nearest neighbors in terms of Euclidean distance in the full 128-dimensional appearance embedding space.

It is important to note that "closeness" between parts in the full 128-dimensional embedding space may appear differently visually when projected into 2D. These differences provide insight into features learned by the neural network, which may be difficult to discern visually. As a representative example of this concept, Figure 8 displays a series of parts that are "close" to a reference part (in this case, a trash can lid). In Figure 8, parts

https://doi.org/10.1017/S0890060422000130 Published online by Cambridge University Press

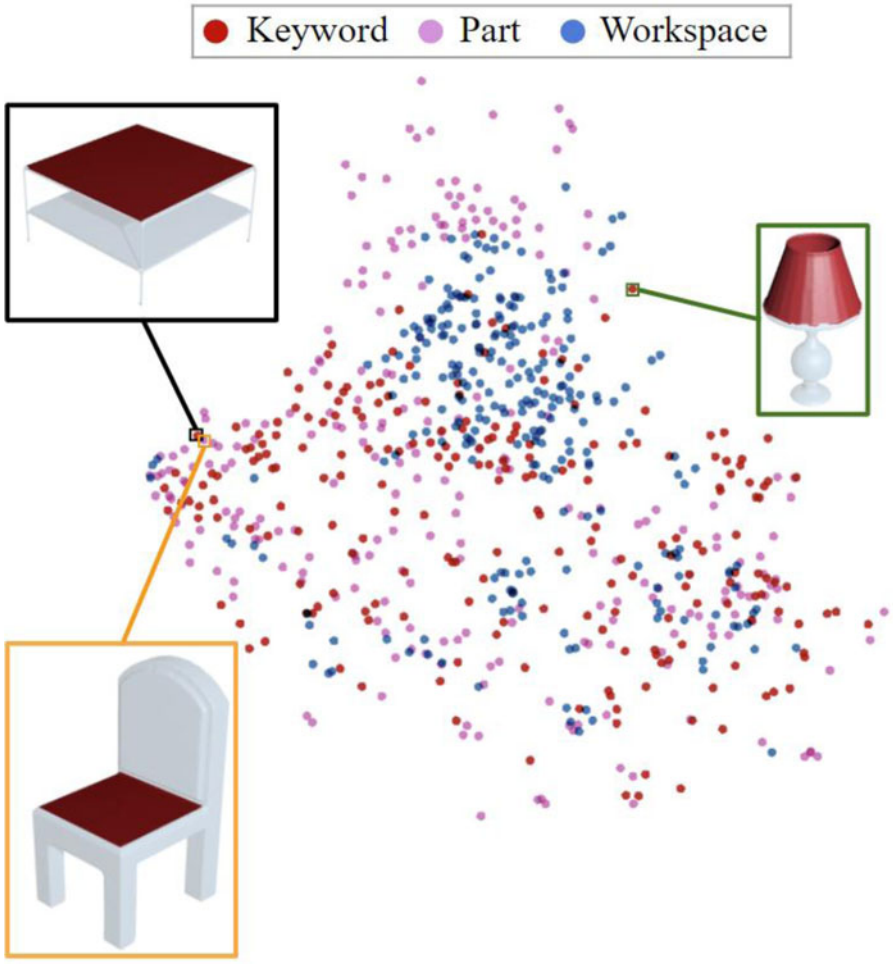


Fig. 7. 2D visualization of the 128-dimensional appearance-based neural network: parts retrieved during the study using each search modality are represented based on their appearance embeddings. Closely related (yellow, chair seat) and distantly related (green, lamp shade) examples in appearance to the reference part (black, tabletop) are shown.

labeled 1–4 are the top 4 nearest neighbors in Euclidean distance in the full embedding space to the reference trash can lid, while Part * appears close in distance in the reduced embedding space. As shown, parts with high appearance similarity, as determined by the appearance-based neural network, may not have the same relative distance in the 2D projection of the embedding space. In this example, Part *, though not a nearest neighbor in the full embedding space, does appear to share high visual similarity to the trash can lid, by human inspection. This discrepancy between human and model evaluation of appearance-based similarity can be explored in future work.

As Figure 7 helps to visualize, the parts discovered by participants during the study appear to vary in their overall coverage of the two embedding spaces by the search modality (keyword vs. part vs. workspace) used to search for them. Quantitatively, this result can be demonstrated by comparing the total variation of each set of parts in the original embedding spaces (represented as each set of colored points in the 2D visualizations). In the approach taken, three 128×128 covariance–variance matrices are first computed for the keyword, part, and workspace search results based on their definitions within the 128-dimensional appearance-based neural network. Diagonal elements of each matrix represent variances in each dimension of the embedding space. A significant Levene's test determined that the variances across the diagonal elements of the three matrices were not equal ($F = 20.9$, $p < 0.001$), signifying a difference between search types in the coverage by parts of the appearance embedding space.

Similarly, the parts retrieved in the study are also represented based on their embeddings in the function-based neural network, shown in Figure 9. The 2D visualization, reduced from the 64-dimensional functional embedding space using PCA, accounts for 89.7% of total variance. Figure 9 displays a cabinet door and its closest neighboring part by Euclidean distance in the full functional embedding space, a sink drawer face. Also shown is a set of chair legs, distantly related in function to the cabinet doors. These parts exemplify how functional relationships are represented in the neural network, as detailed in the section "Computationally deriving similarity between 3D-model parts". As intended in the design of the functional network, two types of doors that are used in different contexts are functionally similar based on their shared relation to box structures. A difference was found between how parts retrieved using each search type covered the functional embedding space ($F = 6.77$, $p < 0.01$).

In addition to comparing the variances across diagonal elements of each covariance–variance matrix using Levene's tests, a more intuitive representation of this measure is the trace of the matrix, that is, sum of the diagonal elements. The trace equals the sum of variances of each dimension of the original neural networks and represents total variation in the respective embedding spaces. Total variation provides a metric for comparing how parts accessed by each search type differently cover the search space. Table 5 summarizes the differences between parts retrieved using each search modality, with respect to the total variation and highest variance of a single variable in both embedding

https://doi.org/10.1017/S0890060422000130 Published online by Cambridge University Press

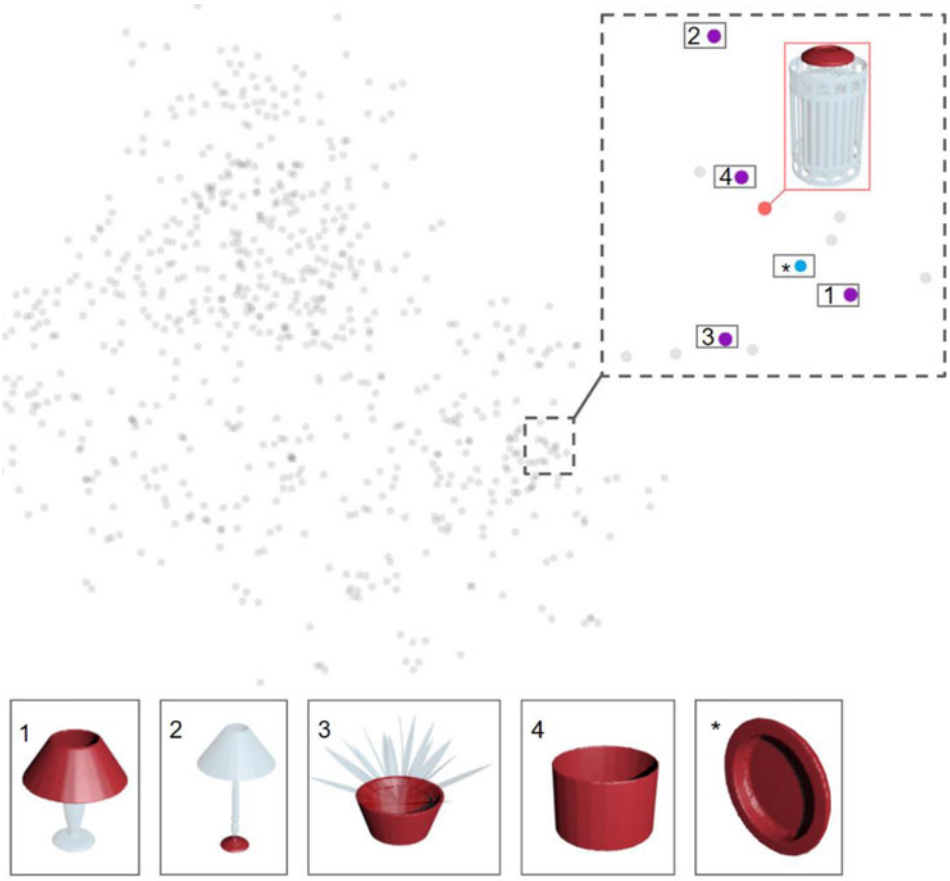


Fig. 8. Expanded view of a cluster of parts in the 2D visualization of the appearance-based neural network. Parts 1–4 are the top 4 nearest neighbors in Euclidean distance to a trash can lid in the full 128-dimensional embedding space. Part * is close based on visual inspection to the trash can lid in the 2D projection.

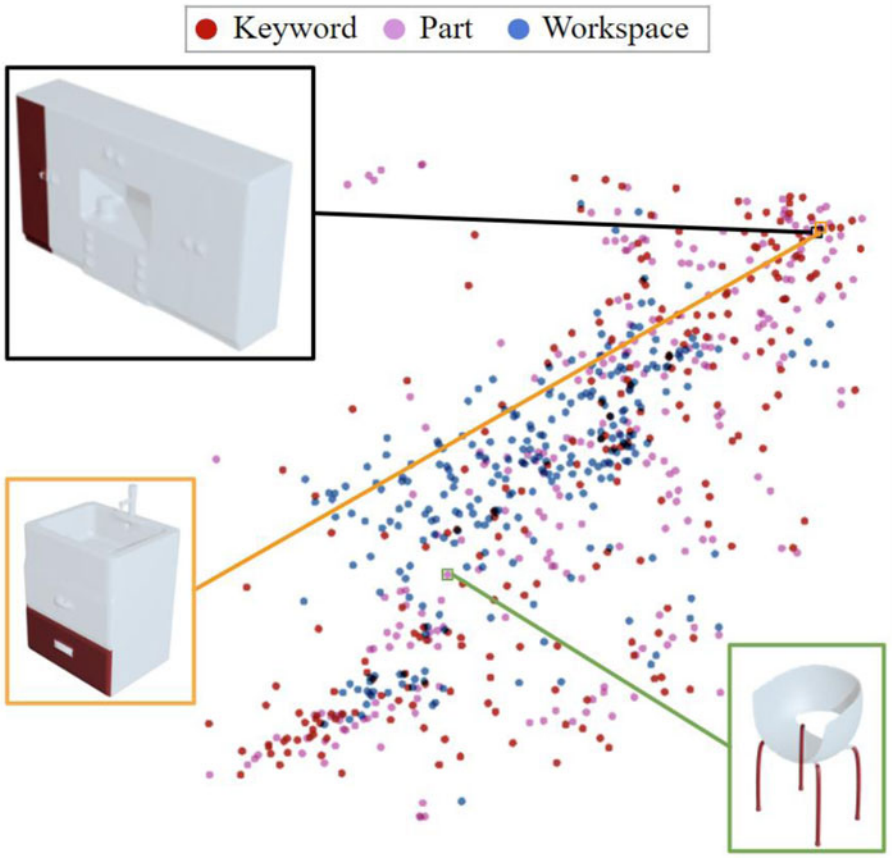


Fig. 9. 2D visualization of the 64-dimensional function-based neural network: parts retrieved during the study using each search modality are represented based on their functional embeddings. Closely related (yellow, sink drawer face) and distantly related (green, chair legs) examples in function to the reference part (black, cabinet door) are shown.

https://doi.org/10.1017/S0890060422000130 Published online by Cambridge University Press

Table 5. Total variation and highest variance of a single variable in appearance (128-dimensional) and functional (64-dimensional) embedding spaces by search type

	Appearance embedding space		Functional embedding space	
Search type	Total variation	Highest variance	Total variation	Highest variance
Keyword	6.0×10^{-3}	1.4×10^{-4}	7.8×10^{-4}	7.9×10^{-5}
Part	7.2×10^{-3}	1.4×10^{-4}	7.8×10^{-4}	7.2×10^{-5}
Workspace	3.6×10^{-3}	5.2×10^{-5}	3.9×10^{-4}	3.1×10^{-5}

spaces. The highest variance demonstrates the relative contribution of individual variables to the total variation. Based on these values, the use of workspace searches appears to lead to the retrieval of parts with the lowest overall coverage of both spaces. Since the dimensions of the appearance and functional embedding spaces differ, variances should be compared within (by search type) and not across (appearance vs. functional) the respective embedding spaces. Total variation is expected to be lower in the functional embedding space, since there are 64 parameters, compared to 128 in the appearance embedding space. At a high level, these results suggest that the search modality used impacted the breadth and diversity of inspirational stimuli discovered.

Discussion

In the present work, a multi-modal search platform was developed and used to study how designers search for inspirational stimuli. A cognitive study was conducted to investigate the impact of searching with different modalities to retrieve inspirational stimuli in the form of 3D-model parts. Findings related to the design of the search platform and results of the cognitive study are further discussed in this section with added insight from qualitative results.

Discussion of multi-modal search platform development and behavior

The design, development, and behavior of a multi-modal search platform are presented in this work. Deep-neural networks were trained to model relationships between 3D-model parts from the PartNet dataset. By selecting a large dataset of 3D-model parts as inspirational stimuli, data-driven, deep-learning-based methods could be leveraged. 3D-model parts specifically contain rich information and allowed semantic, visual, and function-based similarities to be derived between stimuli. These computational methods were then also effectively used to develop a platform to retrieve examples based on these features. Therefore, beyond deriving multiple types of similarity, this work presents a platform that additionally provides the flexibility to search based on these characteristics of design stimuli. Various similarity definitions were considered to help us understand the retrieval behavior of the neural networks using a ranked-based measure, as introduced in the section "Quantitative retrieval behavior of neural networks". The lowest accuracy measures were observed when relevance was defined in terms of self-similarity (i.e., the model retrieves the same part as the input). However, in the context of this platform's use, self-similarity is a highly conservative estimation given the difficulty of this task, as noted in the section "Definition and results of self-similarity". Alternative metrics are therefore explored, including a concept-based (i.e., semantic) definition of similarity and similarity of physical form. At the concept level, the platform's appearance network reached a top-10 test set accuracy of ~68% for identifying text labels of the corresponding 3D model. By comparison, Zhang and Jin's deep-learning approach achieved up to ~82% accuracy when labeling clusters of 2D sketches with one of five categories (e.g., "canoe" vs. "car") (2020). Different from the present work, 2D images, and not 3D models, were used in this study. Limited instances of 3D-part-based retrieval, as it has been implemented in this work, exist in prior research to compare retrieval behavior in the context of physical-form-based similarity. In the application of the platform in the cognitive study, the task was designed such that the specific stimuli retrieved was less relevant than how search intent was expressed. Moreover, we would like to highlight that these definitions of similarity, while simplistic and intuitive, only provide very limited perspectives on the ability of the models in supporting design ideation. However, future work can explore further direct validation metrics and an evaluation of the accuracy of the retrieved examples from the user's perspective when performing our targeted task.

Discussion of cognitive study results

This platform was used to complete a search task during a cognitive study, which was administered such that participants used the three available search modalities during three distinct subtasks. Participants were instructed to search for parts using keyword, part, and workspace searches, in Tasks A, B, and C, respectively. The overall goal of the task was to save parts that served as inspirational for designing a multi-compartment disposal unit. A limitation of this study design is that the effects of learning with each task and the stage of the design process during each task may have influenced how search modalities were used. However, for the aims of this work, understanding how each search modality was used and interacted with was prioritized over capturing how designers may have naturally used them to achieve specific design outcomes. Each search modality was associated with distinct search behaviors and interactions with retrieved parts.

Affected outcomes include search frequency and how search inputs were specified. Most searches occurred in Task B, by part search. Prior work has shown that, when presented with random examples, high click frequency on examples occurred to examine them until something desirable was found (Lee *et al.*, 2010). Increased searches made with new part selections or similarity slider positions may indicate this exploration for desirable stimuli. When conducting workspace searches in Task C, more searches made were new and introduced a new part to the workspace input, than modified with adjusted similarity sliders from the previous search. The same result was not observed with part searches. One explanation for this finding is that the ability to make incremental modifications to the main search input by adding parts to the workspace may encourage more new searches. An analogous incremental manipulation to visual features of the search input in part searches is absent. Observed differences in these inputs suggest that users value the ability to conduct searches that vary individual parameters one at a time.

When interacting with the retrieved parts, most parts viewed in context were those retrieved from keyword searches. One

https://doi.org/10.1017/S0890060422000130 Published online by Cambridge University Press

participant explicitly described their use of this function when commenting on their keyword search strategy: "I was inspired by some of the parts in the 'view in context' like the 'lid'". While participants could make part and workspace searches using a previously retrieved part, text-labeled images of parts from the view in context function may inspire subsequent keyword search inputs. Stimuli combining semantic elements and images were also found by Han *et al.* to help designers generate creative ideas (2018). However, the provided stimuli may not directly inspire new ideas, but help divert designers onto a new train of thought to enable new ideas (Howard *et al.*, 2011). A similar process involving indirect stimulation was also observed by Chen *et al.* during the use of a mind-mapping tool, where retrieved results prompted further querying (2022).

The final outcome of the cognitive study relates to what participants searched for and discovered. The lowest coverage of the search space occurred when searching by workspace, as assessed using metrics of variance within the appearance and functional embedding spaces. Increased breadth of coverage may occur when defining new keyword and part searches through inspiration by external concepts. For instance, parts discovered when utilizing view in context may inspire a new keyword search based on a shown text-labeled part or a part search for functionally similar parts. He *et al.* observed that concept-space exploration using external information was common during interaction with a concept-space visualization tool (2019). Future work may investigate these cognitive processes and motivations for conducting through think aloud protocols or in-depth post-task interviews.

Implications for understanding and supporting how designers search

The insights gained from the cognitive study aim to advance the understanding of how designers search for inspirational stimuli, and how search modalities can differently support these cognitive processes. Distinct interactions within the platform, as discussed above, may reflect the different cognitive processes underlying search.

Active and passive search strategies

As introduced in the section "Cognitive processes underlying search for inspiration", search behavior can be broadly divided into active versus passive strategies, which support situations in which a specific goal exists versus where random encounters with inspirational stimuli occur. In general, participants can be assumed to be engaged in active search processes when defining a search query (i.e., there is intention underlying search). Other interactions within the platform can also afford the ability to engage in search processes to passively inspire their next search. As mentioned, passive search can be supported by information gained by viewing parts in context. Previous work suggests that participants want to be struck by inspiration and to search more randomly (Herring *et al.*, 2009; Goncalves *et al.*, 2016). Increased engagement with parts may therefore be a strategy to randomly encounter inspiration and inspire subsequent searches. Parts retrieved from keyword searches were the most engaged with, specifically by being viewed in context, which may indicate that sources of inspiration not explicitly searched for may be especially helpful when searching with a directly articulated input, such as by text. Introducing additional means for passive search through random discovery of inspirational stimuli may formally achieve what participants found useful about viewing parts in context.

Exploratory search strategies

The platform enables part and workspace searches to specify the levels of appearance and function-based similarity of results to the input. While adjustment of sliders provides a method to specify desired search results, qualitative responses link the use of sliders to, counterintuitively, more exploratory behavior. When asked to "describe any strategies [used] when conducting part searches", one participant noted the use of sliders as supporting search when a distinct goal was missing: "I would try both combinations of functionality and appearance because I didn't really know what I was looking for and I wanted to see all my options". The use of similarity sliders is also mentioned as a way to explore limits of the design stimuli space, in one participant's part search strategy: "I mainly used this as a way to look at possible new ideas I had not considered before by moving the functionality slider to max and the appearance slider to the lowest setting" and another participant's use of workspace searches: "I was trying several factors that could play with changing the appearance and functionality levels while adjusting it from the opposite to all being very similar". Previous work on searching with inputs specifying desired similarity and variety of results has also shown that these parameters are helpful for finding relevant examples (Lee *et al.*, 2010). These responses support the use of providing mechanisms to conduct searches by adjusting parameters that assist with wider exploration. Search can be specified based on desired diversity or variety of stimuli, for example.

These contributions of our work encourage the further development of multi-modal search systems, as well as research on cognitive processes relevant to the search for inspirational examples to support design. Improved understanding is needed regarding when different approaches to search are more useful (e.g., direct and active vs. exploratory and passive), and how to identify and promote these processes through interactions with features of search interfaces.

Conclusion

The work presented in this paper provides insight into how search modality affects the processes designers use to search for and retrieve inspirational stimuli to support design ideation. We describe the development of a new multi-modal search platform and the results of a cognitive study investigating the role of modality in search. The first main outcome of this work is the design, development, and illustration of behavior of a multi-modal search platform. A deep-learning approach was leveraged to construct deep-neural networks based on semantic, visual, and functional relationships between design stimuli from a large dataset of 3D-model parts. The platform affords inputs based on text, 3D-model parts, and assemblies of 3D-model parts to search for additional parts. A variety of similarity metrics were used to quantitatively understand the platform's retrieval behavior using rank-based accuracy measures. Secondly, the results of the cognitive study conducted using the search platform were presented. When engaging with the platform to search for parts to inspire a solution to a given design challenge, differences between the three modalities were observed in terms of frequency of search, how search inputs were defined, interactions with retrieved results, and the resulting coverage of the search space. Behaviors such as increased search frequency and modified adjustments to search inputs are proposed to indicate random exploratory behavior, which can be enhanced in future creativity-support tools. Other interactions leading to random external stimuli discovery that

https://doi.org/10.1017/S0890060422000130 Published online by Cambridge University Press

inspired new search inputs can be more formally implemented to assist designers during different stages of the search process. Overall, the results of this study contribute to recent work on new search modalities to retrieve inspirational stimuli to enhance design ideation. This study supports the need for further research on both the search process itself, as well as on how modality affects and aids how designers search.

Data availability. The data that support the findings of this study are available from PartNet (https://partnet.cs.stanford.edu). Restrictions apply to the availability of these data, which were used under licence for this study.

Financial support. This work is supported in part by the National Science Foundation under grant 2145432 – CAREER. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Conflict of interest. The authors declare none.

References

Arora R, Darolia I, Namboodiri VP, Singh K and Bousseau A (2017) Sketchsoup: exploratory ideation using design sketches. *Computer Graphics Forum* **36**, 302–312.

Athukorala K, Głowacka D, Jacucci G, Oulasvirta A and Vreeken J (2016) Exploratory search: from finding to understanding. *Journal of the Association for Information Science and Technology* **67**, 2645–2651.

Borgianni Y, Rotini F and Tomassini M (2017) Fostering ideation in the very early design phases: how textual, pictorial and combined stimuli affect creativity. *Proceedings of the 21st International Conference on Engineering Design, ICED17*. The Design Society.

Botella M, Zenasni F and Lubart T (2018) What are the stages of the creative process? What visual art students are saying. *Frontiers in Psychology* **9**, 2266.

Cer D, Yang Y, Kong S-y, Hua N, Limtiaco N, St. John R, Constant N, GuajardoCespedes M, Yuan S, Tar C, Strope B and Kurzweil R (2018) Universal sentence encoder for English. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*. Brussels, Belgium: Association for Computational Linguistics.

Chakrabarti A, Sarkar P, Leelavathamma L and Nataraju B (2005) A functional representation for aiding biomimetic and artificial inspiration of new ideas. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* **19**, 113–132.

Chan J, Fu K, Schunn C, Cagan J, Wood K and Kotovsky K (2011) On the benefits and pitfalls of analogies for innovative design: ideation performance based on analogical distance, commonness, and modality of examples. *Journal of Mechanical Design* **133**, 081004.

Chan J, Dow SP and Schunn C (2015) Do the best design ideas (really) come from conceptually distant sources of inspiration? *Design Studies* **36**, 31–58.

Chaudhuri S and Koltun V (2010) Data-driven suggestions for creativity support in 3D modeling. *ACM Transactions on Graphics* **29**, 1–10.

Chen T, Mohanty RR and Krishnamurthy VR (2022) Queries and cues: textual stimuli for reflective thinking in digital mind-mapping. *Journal of Mechanical Design* **144**, 021402.

Cheong H, Chiu I, Shu LH, Stone RB and McAdams DA (2011) Biologically meaningful keywords for functional terms of the functional basis. *Journal of Mechanical Design* **133**, 021007.

Devlin J, Chang M-W, Lee K and Toutanova K (2019) BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: human Language Technologies (NAACL-HLT)*. Minneapolis, MN, June 2–7, pp. 4171–4186.

Do EY-L (2005) Design sketches and sketch design tools. *Knowledge-Based Systems* **18**, 383–405.

Eckert C and Stacey M (2003) Sources of inspiration in industrial practice: the case of knitwear design. *Journal of Design Research* **3**, 16–44.

Fu K, Cagan J, Kotovsky K and Wood K (2013a) Discovering structure in design databases through functional and surface based mapping. *Journal of Mechanical Design* **135**, 031006.

Fu K, Chan J, Cagan J, Kotovsky K, Schunn C and Wood K (2013b) The meaning of "near" and "far": the impact of structuring design databases and the effect of distance of analogy on design output. *Journal of Mechanical Design* **135**, 021007.

Goel AK, Rugaber S and Vattam S (2009) Structure, behavior, and function of complex systems: the structure, behavior, and function modeling language. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* **23**, 23–35.

Goel AK, Vattam S, Wiltgen B and Helms M (2012) Cognitive, collaborative, conceptual and creative — four characteristics of the next generation of knowledge-based cad systems: a study in biologically inspired design. *Computer-Aided Design* **44**, 879–900.

Goncalves M, Cardoso C and Badke-Schaub P (2016) Inspiration choices that matter: the selection of external stimuli during ideation. *Design Science* **2**, 1–31.

Goucher-Lambert K and Cagan J (2019) Crowdsourcing inspiration: using crowd generated inspirational stimuli to support designer ideation. *Design Studies* **61**, 1–29.

Goucher-Lambert K, Moss J and Cagan J (2019) A neuroimaging investigation of design ideation with and without inspirational stimuli—understanding the meaning of near and far stimuli. *Design Studies* **60**, 1–38.

Goucher-Lambert K, Gyory JT, Kotovsky K and Cagan J (2020) Adaptive inspirational design stimuli: using design output to computationally search for stimuli that impact concept generation. *Journal of Mechanical Design* **142**, 091401.

Gross MD (2001) Emergence in a recognition based drawing interface. In Gero JS, Tversky B and Purcell T (eds), *Visual and Spatial Reason in Design II*. Sydney, Australia: Key Centre of Design Computing and Cognition, pp. 51–65.

Han J, Shi F, Chen L and Childs PR (2018) The combinator–a computer-based tool for creative idea generation based on a simulation approach. *Design Science* **4**, e11.

Han J, Sarica S, Shi F and Luo J (2022) Semantic networks for engineering design: state of the art and future directions. *Journal of Mechanical Design* **144**, 020802.

He Y, Camburn B, Liu H, Luo J, Yang M and Wood K (2019) Mining and representing the concept space of existing ideas for directed ideation. *Journal of Mechanical Design* **141**, 121101.

Herring SR, Chang CC, Krantzler J and Bailey BP (2009) Getting inspired! Understanding how and why examples are used in creative design practice. *Proc. of SIGCHI Conference on Human Factors in Computing Systems, CHI'09*. New York, NY, USA: Association for Computing Machinery.

Howard TJ, Culley S and Dekoninck EA (2011) Reuse of ideas and concepts for creative stimuli in engineering design. *Journal of Engineering Design* **22**, 565–581.

Hu R, Yan Z, Zhan J, van Kaick O, Shamir A, Zhang H and Huang H (2018) Predictive and generative neural networks for object functionality. *ACM Transactions on Graphics (Proc. SIGGRAPH)* **37**, 151:1–151:14.

Hua M, Han J, Ma X and Childs P (2019) Exploring the effect of combinational pictorial stimuli on creative design performance. *Proc. of the 22nd International Conference on Engineering Design, ICED19*. Cambridge University Press.

Jiang S, Luo J, Ruiz-Pava G, Hu J and Magee CL (2020) A convolutional neural network-based patent image retrieval method for design ideation. *Proc. ASME 2020 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Paper No. DETC2020-22048, Virtual, Online, August 17–19.

Jiang S, Luo J, Ruiz-Pava G, Hu J and Magee CL (2021) Deriving design feature vectors for patent images using convolutional neural networks. *Journal of Mechanical Design* **143**, 061405.

Jiang S, Hu J, Wood KL and Luo J (2022) Data-driven design-by-analogy: state-of-the-art and future directions. *Journal of Mechanical Design* **144**, 020801.

Kazi RH, Grossman T, Cheong H, Hashemi A and Fitzmaurice G (2017) Dreamsketch: early stage 3D design explorations with sketching and

https://doi.org/10.1017/S0890060422000130 Published online by Cambridge University Press

generative design. *Proc. of the 30th Annual ACM Symposium on User Interface Software and Technology, UIST '17*. New York, NY, USA: Association for Computing Machinery.

Kittur A, Yu L, Hope T, Chan J, Lifshitz-Assaf H, Gilon K, Ng F, Kraut RE and Shahaf D (2019) Scaling up analogical innovation with crowds and AI. *Proceedings of the National Academy of Sciences* **116**, 1870–1877.

Kwon E, Pehlken A, Thoben KD, Bazylak A and Shu LH (2019) Visual similarity to aid alternative-use concept generation for retired wind-turbine blades. *Journal of Mechanical Design* **141**, 031106.

Kwon E, Huang F and Goucher-Lambert K (2021) Multi-modal search for inspirational examples in design. *Proc. ASME 2021 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Paper No. DETC2021-71825, Virtual, Online, August 17–19.

Lee B, Srivastava S, Kumar R, Brafman R and Klemmer SR (2010) Designing with interactive example galleries. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '10*. New York, NY, USA: Association for Computing Machinery.

Linsey J, Wood K and Markman A (2008) Modality and representation in analogy. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* **22**, 85–100.

Luo J, Sarica S and Wood KL (2021) Guiding data-driven design ideation by knowledge distance. *Knowledge-Based Systems* **218**, 106873.

Marchionini G (2006) Exploratory search: from finding to understanding. *Communications of the ACM* **49**, 41–46.

Menezes A and Lawson BR (2006) How designers perceive sketches. *Design Studies* **27**, 571–585.

Mo K, Zhu S, Chang AX, Yi L, Tripathi S, Guibas LJ and Su H (2018) Partnet: a large-scale benchmark for fine-grained and hierarchical part-level 3D object understanding. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 909–918.

Mougenot C, Bouchard C, Aoussat A and Westerman S (2008) Inspiration, images and design: an investigation of designers' information gathering strategies. *Journal of Design Research* **7**, 331–351.

Murphy J, Fu K, Otto K, Yang M, Jensen D and Wood K (2014) Function based design-by-analogy: a functional vector approach to analogical search. *Journal of Mechanical Design* **136**, 101102.

Nagel JK and Stone RB (2012) A computational approach to biologically inspired design. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* **26**, 161–176.

Purcell AT and Gero JS (1992) Effects of examples on the results of a design activity. *Knowledge-Based Systems* **5**, 82–91.

Robertson B, Walther J and Radcliffe DF (2007) Creativity and the use of CAD tools: lessons for engineering design education from industry. *Journal of Mechanical Design* **129**, 753–760.

Sarica S, Luo J and Wood KL (2020) Technet: technology semantic network based on patent data. *Expert Systems with Applications* **142**, 112995.

Sarica S, Song B, Luo J and Wood KL (2021) Idea generation with technology semantic network. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* **35**, 265–283.

Setchi R and Bouchard C (2010) In search of design inspiration: a semantic-based approach. *Journal of Computing and Information Science in Engineering* **10**, 031006.

Siangliulue P, Chan J, Gajos KZ and Dow SP (2015) Providing timely examples improves the quantity and quality of generated ideas. *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition, C&C '15*. New York: Association for Computing Machinery.

Song H and Fu K (2022) Design-by-analogy: effects of exploration-based approach on analogical retrievals and design outcomes. *Journal of Mechanical Design* **144**, 061401.

Soufi B and Edmonds E (1996) The cognitive basis of emergence: implications for design support. *Design Studies* **17**, 451–563.

Sutcliffe A and Ennis M (1998) Towards a cognitive theory of information retrieval. *Interacting with Computers* **10**, 321–351.

Toh CA and Miller SR (2014) The impact of example modality and physical interactions on design creativity. *Journal of Mechanical Design* **136**, 091004.

Tseng I, Moss J, Cagan J and Kotovsky K (2008) The role of timing and analogical similarity in the stimulation of idea generation in design. *Design Studies* **29**, 203–221.

Vasconcelos LA, Cardoso CC, Sääksjärvi M, Chen C-C and Crilly N (2017) Inspiration and fixation: the influences of example designs and system properties in idea generation. *Journal of Mechanical Design* **139**, 031101.

Vattam S, Wiltgen B, Helms ME, Goel AK and Yen J (2011) DANE: fostering creativity in and through biologically inspired design. In Taura T and Nagai Y (eds), *Design Creativity 2010*. London: Springer, pp. 115–122.

Zhang Z and Jin Y (2020) An unsupervised deep learning model to discover visual similarity between sketches for visual analogy support. *Proc. ASME 2020 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Paper No. DETC2020-22394, Virtual, Online, August 17–19.

Zhang Z and Jin Y (2021) Toward computer aided visual analogy support (CAVAS): augment designers through deep learning. *Proc. ASME 2021 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Paper No. DETC2021-70961, Virtual, Online, August 17–19.

Elisa Kwon is a Ph.D student at the University of California, Berkeley advised by Dr. Kosa Goucher-Lambert. She received her B.A.Sc. in Engineering Science (2017) and her M.A.Sc. in Mechanical Engineering (2019) at the University of Toronto. Her main research interest is in investigating human cognition during the engineering design process through human-subject studies and drawing on methods from cognitive neuroscience and psychology.

Forrest Huang is a Ph.D. candidate at the University of California, Berkeley advised by Prof. John F. Canny. He received a B.S. in Computer Science from the University of Illinois at Urbana-Champaign in 2017. His research focuses on developing deep-learning systems that support creative activities with sketch-based and natural-language-based user interaction. His research contributions also include large-scale novel UI design and sketch datasets, and interactive visualization and debugging tools for deep-learning workflows.

Kosa Goucher-Lambert is an Assistant Professor of Mechanical Engineering at the University of California, Berkeley, and Affiliate Faculty member in the Jacobs Institute of Design Innovation and the Berkeley Institute of Design. Dr. Goucher-Lambert is an expert in the field of engineering design theory, methods, and automation, and conducts research merging computational analyses of human-behavior in design with cognitive and neuroscience models of designer behavior. He is the recipient of an NSF CAREER Award and 2019 Excellence in Design Science Award. He has received several best paper awards from the American Society of Mechanical Engineers and the Design Society.

https://doi.org/10.1017/S0890060422000130 Published online by Cambridge University Press