

MULTI-MODAL SEARCH FOR INSPIRATIONAL EXAMPLES IN DESIGN

Elisa Kwon

Dept. of Mechanical Engineering
University of California, Berkeley
Berkeley, CA, USA
Email: elisa.kwon@berkeley.edu

Forrest Huang

Dept. of Electrical Engineering and
Computer Science
University of California, Berkeley
Berkeley, CA, USA
Email: forrest_huang@berkeley.edu

Kosa Goucher-Lambert

Dept. of Mechanical Engineering
University of California, Berkeley
Berkeley, CA, USA
Email: kosa@berkeley.edu

ABSTRACT

Inspirational stimuli are known to be effective in supporting ideation during the design process. However, minimal prior work has allowed individuals to search using multiple modes of input simultaneously, which is more representative of real design behavior. In the current work, we developed a multi-modal search platform that retrieves 3D model parts based on text, appearance, and function-based search inputs. This work presents the results of an experimental study ($n=21$) in which the search platform was used to find parts identified as potentially useful for inspiring solutions to a design challenge. Participants were asked to engage with three different search modalities: search by keywords, by curated 3D parts, and by user-assembled 3D parts in their workspace. When searching by parts that are curated or in their workspace, additional control over the similarity of appearance and function of results in reference to the input was available to participants. The results of this study demonstrate that the modality used affects search behavior, such as in the frequency of searches, how participants engage with retrieved search results, 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 multi-modal search should enable intentional search for desired goals through direct search inputs (e.g., by keyword) and incremental adjustments to features of visually represented search inputs. Moreover, enabling discovery of inexplicitly searched for examples through related information or more randomly encountered examples may assist exploratory search behavior.

Keywords: Multi-modal search, Computational tools, Inspirational stimuli

1. INTRODUCTION

Designers benefit from external inspirational stimuli during ideation to help achieve desirable design outcomes such as greater novelty, feasibility, or innovativeness [1-3]. The usefulness of potentially inspirational stimuli during idea generation for a design problem depends on many features of the given stimulus. One relevant feature is the representation of the stimulus. Visual stimuli compared to physical examples or when combined with textual examples have been shown to increase idea novelty [4, 5]. Analogical distance between examples from the design problem is also important, where both far-field and near-field examples have been found to contribute to desirable outcomes such as novelty or fluency, respectively [1, 3, 6]. Furthermore, inspirational stimuli may be more useful depending on when they are accessed during the design process [2, 7].

In this paper, we focus on the search processes involved in the retrieval of potentially inspirational stimuli by multiple modalities. To support this work, characteristics of inspiring stimuli, such as how they are derived and the level of design-related information they contain, are considered. Also described are the relevant cognitive processes designers may employ to search for inspiring examples and tools that have been developed to support search.

1.1 Inspirational stimuli in design

In prior studies that investigate the impact of inspirational examples on design, participants are often provided with fixed sets of stimuli while working on a given design problem. These stimuli may be derived from a variety of sources. 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 [1, 3, 8]. Design solutions and ratings collected

from crowd workers can also be provided to designers as sources of inspiration [9, 10]. Inspirational stimuli additionally 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 [11], other images [12], or in contrast to interactions with physical products [4]. The level of abstraction of inspirational examples also impacts their influence on the design process. Design stimuli at the concept level may, for example, provide more rapid inspiration, but miss the richer design details available in more comprehensive documents like patents [13]. Examples can differ further by being provided with descriptions that are more general vs. domain-specific [5] or constitute concrete design examples vs. abstract system properties [14].

1.2 Search for inspirational stimuli

The cognitive processes used by designers to search for inspirational stimuli described above are also important to understand. Early work on the role of search processes in design identified incidental experience and intentional learning as relevant sources of knowledge [15]. 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 [16]. In this process, active approaches to find specific stimuli more intentionally or passive approaches to randomly encounter relevant stimuli may take place [16, 17]. Active search refers to the deliberate search for a particular stimulus with a specific goal in mind [18]. 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 [17]. In information retrieval theory, search behavior has classically been categorized as exploratory vs. specific (or lookup) [19]. Lookup search activities involve precise search goals whereas exploratory search is related to knowledge acquisition and evolving needs [20]. In open-ended exploratory search tasks, users have been found to examine more results than during lookup tasks [21]. For computational tools to successfully support search for inspiration, user studies suggest that they should provide control and flexibility over the level of abstraction vs. literalness of search terms [22]. 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.

1.3 Search tools and modalities to retrieve inspirational stimuli

Computational methods that provide designers with, or allow them to search for, inspirational stimuli often rely on semantic relationships between search inputs and outputs. Tools include those leveraging functional relationships to facilitate search for analogies in biologically inspired design [23, 24],

function-based content from patent databases [25, 26], or visual stimuli using semantic information, instead of low-level visual features [27]. Text-based search using semantic relationships may, however, limit discovery of inspirational stimuli to concepts that are well enough defined to express using words. As described above, search processes using more passive or exploratory processes are sometimes preferred and may not be as well supported by tools requiring such direct input [16]. Introducing new modes of expressing search may be one approach to aid different search strategies when needed during the design process.

In contrast to verbal representation, sketching is a modality by which designers can more flexibly express their current thought process. Sketching can be used to search for ideas directly, during which the act of sketching can assist in idea formation [28]. The use of visual representations, including sketch as inputs, may be beneficial to the search process. While search for examples, particularly with non-text-based modalities, has been less explored, emerging methods to support visual analogy in design have leveraged image-based search for image retrieval. Recent work has demonstrated how visual analogy can be supported by sketch-based retrieval of visually similar examples [29] and how visual information can be used in the search for relevant examples in a patent database [30, 31].

Creativity support tools can assist with inspirational stimuli retrieval in computer-aided design (CAD) using 3D sketching, to support emergence and reinterpretation processes. Shape emergence, where designers perceive emergent patterns not initially intended [32], is important in design exploration to trigger new mental images and thus new ideas for design [33]. Reinterpretation of visual representations is needed to enable alternate interpretations and restructuring of design problems. Since CAD drawings record design ideas held in mind with high precision, they do not typically support reinterpretation of the drawing once made [34]. Unlike in 2D sketching, 3D sketching is based on the unambiguous selection and placement of different elements to build a model, which leaves less room for new ideas to emerge or old ideas to be reinterpreted. Retrieval of inspiring examples based on developing 3D sketches can importantly allow for emergence and reinterpretation in the design process. While CAD modeling enhances visualization and communication of ideas by giving a form to early design ideas, it may also cause premature design fixation and limited ideation [35]. Support is needed to enable the reinterpretation of visual representations of early ideas using CAD without limiting creativity.

Addressing these gaps, we developed a search platform that retrieves 3D model parts from both text and visual search queries. An experimental study was conducted in which this platform was used to search for relevant inspirational stimuli with the different modalities. The impact of search modality on the processes designers use to search for, engage with, and select inspirational stimuli could thus be explored, leading to insight into the use and limitations of multi-modal search.

2. MATERIALS AND METHODS

In this section, the development of the platform we built that enables multi-modal search for 3D modeling parts is described. To understand the processes and behaviors associated with searching for and exploring design examples, a participant 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.

2.1 Multi-modal search platform

Our multi-modal search platform uses deep neural networks to model similarities between various 3D-model parts and natural-language-model keywords in a large dataset. This deep-learning approach directly consumes 2D snapshots of 3D modeling parts and utilizes knowledge from large text corpora, which subsequently enables the efficient retrieval of relevant examples in large datasets. Deep neural networks are a suitable candidate 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.

2.1.1 Neural network development

The PartNet dataset was used to train the deep-neural-network and provide users with relevant 3D examples. The PartNet dataset contains 26671 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). These object categories are everyday objects at various scales (e.g., microwave, scissors, tables). Our system's ability to provide inspirational stimuli from these objects may be limited, since they represent a small subset of possible objects that mechanical engineers might design. We instead provide part-based data within these objects, which are commonly seen in object categories beyond those in the dataset (e.g. legs, cover, lid). The networks were trained on 70% of the data and performance was evaluated with a held-out dataset (10% of the data). The entirety of the dataset was then exposed to users to query 3D model parts from in the user study.

The platform relies on three neural networks to embed raw 3D model data to high-level concepts and modeling parameters. First, the *text network* in the platform relies on the Universal Sentence Encoder [36] pre-trained on web text to find parts with names similar to the keyword queries provided by the users. The Universal Sentence Encoder is trained on non-technical 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).

Second, an *appearance network* was trained by embedding knowledge from 2D snapshots of 3D model parts. Eight snapshots from different angles of the same model were rendered

using blender to train the network with a contrastive loss function (a common metric-learning paradigm in the deep learning community), which considers these snapshots as similar in the latent space. The model was also trained to consider snapshots randomly sampled from other data samples as dissimilar to the first model. On a high level, this appearance network primarily considers the overall appearance and geometric presence of the 3D model parts.

To quantitatively evaluate the network's performance, snapshots of 2D renderings from random angles of the 3D model were used as search queries. The percentage of models that had snapshots of the same model ranked as top-1/top-10 nearest neighbors was then measured. Top-k accuracy excludes the input query itself as a "neighbor" since the input query's embedding is identical to its own embedding in the dataset. For training/validation/test sets of this model, the top-1 accuracy is 1.16%/3.58%/2.28% and the top-10 accuracy is 5.87%/17.2%/11.4%. These metrics are a highly conservative estimation of the model's performance as there are far more relevant screenshots of models than those generated from the same model as the input query (i.e., there are many types of chair legs in the dataset).

Finally, based on this appearance network, the *functional network* was trained by taking embeddings of the appearance network and transforming them into function-aware embeddings. **Figure 1** displays how the functional embeddings are derived from appearance embeddings using the described networks.

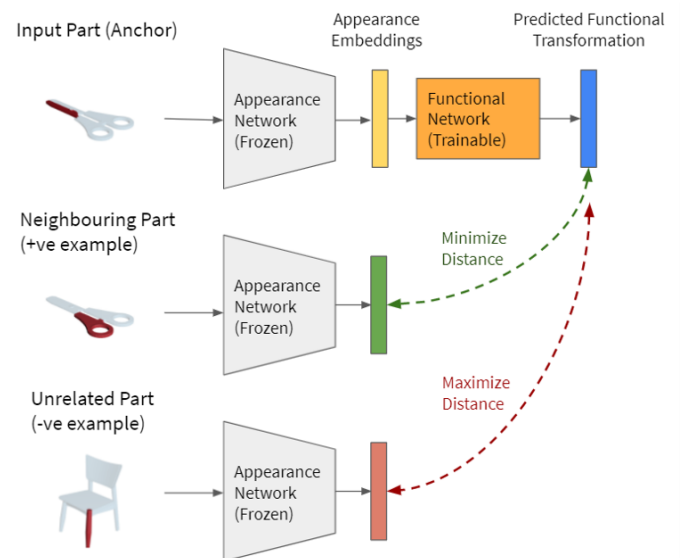


FIGURE 1: Overview of process of transformation of embeddings from appearance network to functional embeddings. Appearance embeddings of **input part** (anchor) were used to generate a predicted functional transformation using the functional network. Functional network was then trained by considering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

Models in the dataset are considered functionally similar if they have similar neighborhoods, meaning 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 functional network was trained using a contrastive loss function, similar to the approach used to train the appearance network. In this case, similarities were instead mapped across different 3D model parts that are close neighbors within the assemblies. The network takes as input the appearance embeddings of two 3D model parts that are either neighbors or non-neighbors of each other. They are then passed through a few trainable fully connected layers in the functional network to compute new functional embeddings. Loss measures ensure these new embeddings are near/far from each other.

2.1.2 Search modalities: keyword, part, and workspace-based

Using the neural networks described above, participants could search for parts in the dataset by three types of input. The first search type is *keyword-based*, where text input by the participant is embedded using the aforementioned text network. 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 Fig. 2A.

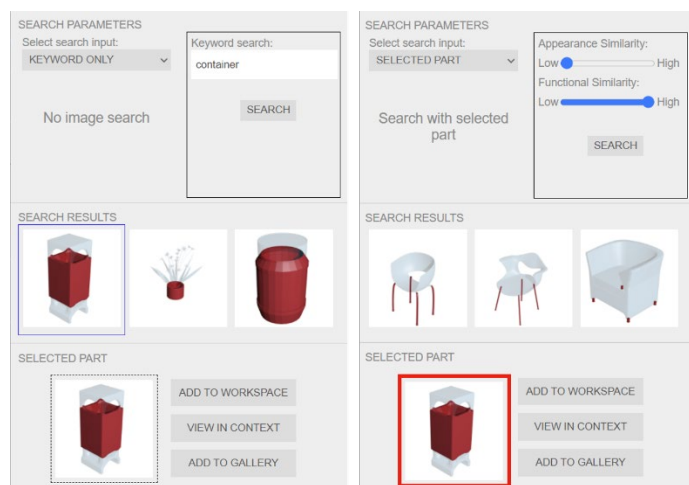


FIGURE 2: A (left) - search results for a keyword search of the term “container”; B (right) – search results for a part search of a result from keyword search for “container”

The second and third search types are *part-based* and *workspace-based*, where new parts are retrieved using snapshots taken of a selected 3D-model part or the participant’s current workspace (composed of 3D-model parts), respectively. 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. 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 **Fig. 2B**, parts are searched for with low similarity in appearance but high similarity in function to the selected container. Based on the results retrieved, participants are then able to modify these inputs to continue to search for new results.

2.1.3 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 **Fig. 3A** shows how parts can also be added 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. 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.

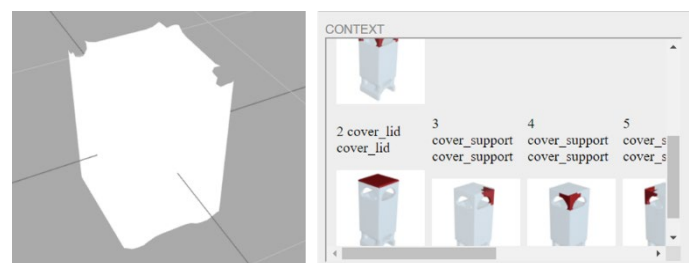


FIGURE 3: Interactions with selected part in Fig. 1 - A (left): added to workspace; B (right): viewed in context

2.2 Experimental design

A 30-minute study was administered to understand how participants engage with the three search types available in the platform. Participants searched for parts using each search modality in three separate subtasks and worked towards collecting inspirational stimuli for a given design challenge.

2.2.1 Participants

Participants were recruited for the study 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 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.

2.2.2 Study objective

The study objective presented to participants was to use the platform provided 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 was chosen such that the parts available in the dataset were relevant, without being overly constrained to a particular domain or set of solutions.

2.2.3 Study overview

The study was divided into three subtasks, where each task involved the use of a different search type (keyword, part, or workspace), but worked towards 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 instructions and directions, 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. An overview of the experiment is summarized in **Figure 4**.

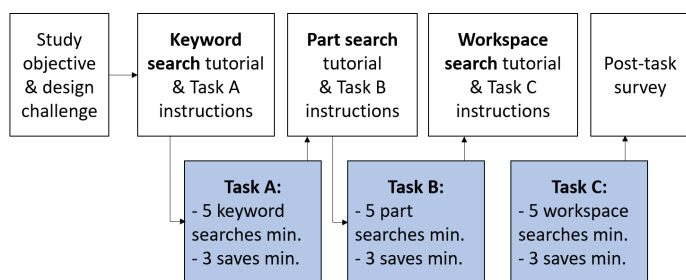


FIGURE 4: Overview of tasks and search types used during the study

Task A: In Task A, all participants were instructed to first search by keyword beginning with the term “container” (**Fig. 1A**). They were instructed to make 4 additional keyword searches (min.) and to save min. 3 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. 4 additional part searches and save min. 3 more parts. Participants were also instructed to not make any additional keyword searches.

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 4 additional workspace searches were made and a min. of 3 parts were saved, without making any new keyword or part searches.

The motivation for the ordering and division of tasks was to easily teach participants how to engage with the search platform. Pilot testing revealed that learning about each search type at study onset overloaded participants with too much information. Tasks were ordered to first use the most intuitive search mode (keyword) and to last introduce the least familiar and most difficult mode (workspace). Future work aims to study use of these search modalities in a more freeform manner.

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 breadth of their exploration of the part repository and of their final gallery of saved parts on five-point Likert scales.

3. RESULTS

In order to study how participants search for examples to inspire idea generation, the interactions made within the search platform are explored. To reiterate, participants performed keyword searches in Task A, part searches in Task B, and workspace searches in Task C. Throughout the study, they could take the following actions on any part retrieved from search: adding it to their workspace, viewing it in context, or saving it to their gallery. This work considers how different modalities of expressing search revealed by these interactions affect and support the search process. Specific research objectives are to identify differences in search modality by how participants (1) search for new parts and (2) engage with parts retrieved from search, as well as the (3) breadth of their search space coverage.

3.1 Searching for new parts

To characterize the process of searching for new parts, the frequencies of each search modality used and slider movements made are examined. New searches are defined in Task A by unique keywords, in Task B by unique parts used as search input, and in Task C by workspaces with newly added parts. In Tasks B and C, participants may also search using the same part or workspace from a previous search, with adjustments made with sliders to appearance and functional similarity. In these tasks, it is also of interest to determine how parts with varying similarity from the input may be differently sought by search type. Counts of new searches and searches with modified similarity parameters are thus separately considered.

3.1.1 Highest frequency of searches are part-based

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 significantly differs ($\chi^2(2, N=677) = 9.8, p < 0.01$), where the number of part searches (264) compared to keyword (207) and workspace (206) searches is the highest. For part and workspace searches, these frequencies include both searches that are new and are modified with different similarity parameters from a previous search with the same input.

3.1.2 More part searches and modified than new

Total occurrences across participants of part searches using new parts and those with the same part from a previous search with changes to the appearance and/or functional similarity are summarized in **Table 1**. Also shown are the number of modified part searches by the similarity parameter adjusted, and whether the adjustment increased (+) or decreased (-) similarity of results from the input.

TABLE 1: Frequencies of new and modified part searches (+: 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	131

As shown in **Table 1**, more total number of searches are conducted with the same part (104) than a different part (131) 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 ($F(1,20) = 0.55, p=0.5$), as shown in **Fig. 5**.

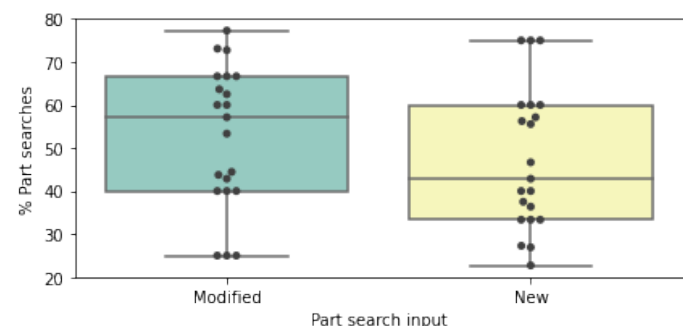


FIGURE 5: % Part searches modified by sliders or made with a new part as input (n = 21)

Table 1 shows that 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) ($\chi^2(2, N=131) = 9.3, p < 0.01$). Most modified searches only adjust the appearance similarity of

results. However, when comparing the proportion of searches modified by appearance, function, or both within each participant, no significant difference was revealed ($F(2,40) = 0.03, p=0.97$). 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 ≥ 5 appearance-modified searches.

3.1.3 More workspace searches are new than modified

The same analysis was performed to identify the processes used to conduct workspace searches, as summarized in **Table 2**.

TABLE 2: Frequencies of new and modified workspace searches (+: increasing similarity, -: decreasing similarity)

Search input		Search counts
New part added to workspace		105
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		76

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 demonstrated by a repeated measures ANOVA test ($F(1,20) = 7.43, p < 0.05$). The increased proportion of new over modified workspace searches is shown in **Fig. 6**.

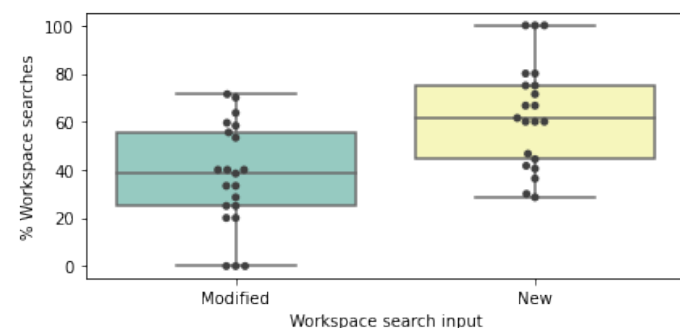


FIGURE 6: % Workspace searches modified by sliders or made with an added part to the workspace (n = 21)

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. This difference in the two search modes may possibly explain the higher frequency of part searches that modify consecutive searches using the similarity sliders.

3.2 Engagement with parts retrieved from search

Once a search is made, three results are shown, which participants may decide to engage with further or not. Participants may decide not to interact with search results beyond seeing them once retrieved. Alternatively, a participant can engage with a part by viewing it in context (gaining contextual information), adding it to the workspace (which may then be seen and manipulated in its 3D representation), or saving it to their gallery. The number of times each interaction is made is counted to determine how results from each search type are engaged with differently.

3.2.1 Lowest engagement observed with part search results

Frequencies of interactions with search results are compared across tasks to assess differences in how participants engage with results retrieved from different search modalities. There is a significant difference between tasks in both the total number of search results that users engaged with ($\chi^2(2, N=106) = 18.6, p < 0.001$) and did not engage with ($\chi^2(2, N=581) = 23.0, p < 0.001$), as shown in **Fig. 7**.

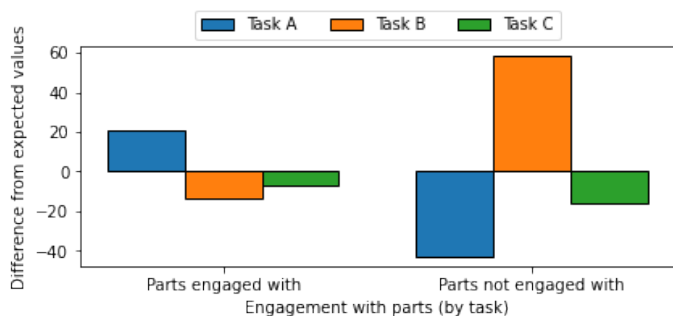


FIGURE 7: Differences between observed and expected values of parts engaged with and not engaged with, shown by task

The differences in frequency between the expected and observed values for each set of results are plotted in **Fig. 7**. 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 in Task A by keyword search, while results not engaged with were mostly those retrieved in Task B by part-based search.

3.2.2 Most contextual information sought about keyword 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 ($\chi^2(2,$

$N=104) = 13.3, p < 0.01$). Displayed in **Fig. 8**, more results from keyword search are viewed in context than expected and fewer results from workspace search are viewed than expected.

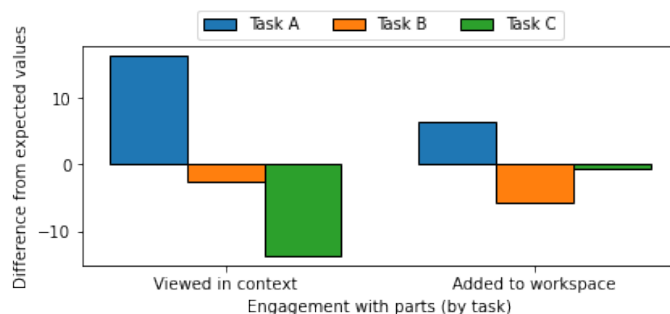


FIGURE 8: Differences between observed and expected values of parts viewed in context and added to the workspace, shown by task

As in **Fig. 7**, expected values in **Fig. 8** refer to the total numbers of parts viewed in context (104) or added to the gallery (101), divided evenly between tasks. Numbers of parts added to the workspace do not differ significantly between tasks. Combined, these results suggest that part-based search in Task B does not encourage increased engagement with individual results, while keyword search does. Motivations for participants to more frequently view results from keyword searches in context are explored next.

3.2.3 Actions following add and view actions vary by search type

In the previous results, part search results were found to be the least explored, and workspace searches the least frequently viewed in context. The direct effects of adding parts to the workspace and viewing parts in context can be known by examining actions immediately following these actions. **Figure 9** displays mosaic plots of the occurrences of actions (adding to the workspace, saving to the gallery, searching, or viewing in context) after parts were added to the workspace (left) or viewed in context (right) during each task.

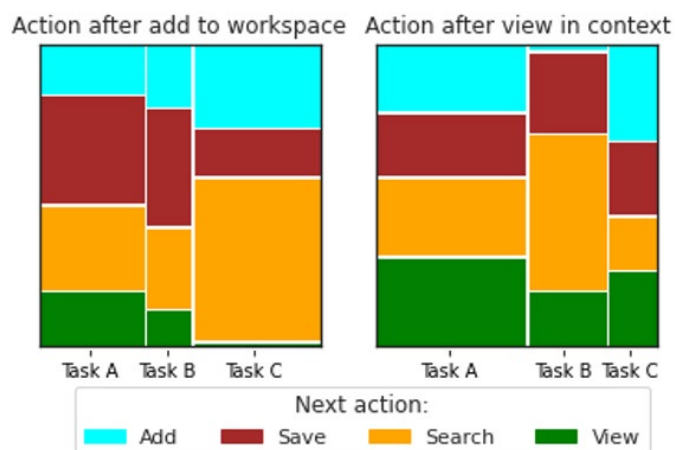


FIGURE 9: Mosaic plots of frequencies of each action type after adding a part to the workspace and viewing a part in context

The frequency of each action type taken after adding parts to the workspace significantly vary between tasks ($\chi^2(6, N=245) = 32.9, p < 0.001$). In tasks A and B, participants saved a higher proportion of parts immediately after adding them to the workspace than in Task C, where participants mostly continued to make searches. Adding parts may inform participants' decisions to subsequently save them, in Tasks A and B. Adding parts before searching in Task C makes sense, given explicit instructions to add parts to the workspace. Consecutive addition of parts to the workspace may occur when multiple results from a single search were added to the workspace.

Tasks also differed in the frequency of each action type taken after parts were viewed in context ($\chi^2(6, N=159) = 20.0, p < 0.01$), where a large contribution is from the high proportion of searches following views in Task B. Viewing of other parts related by context may directly encourage further part-based search since participants can then search for more or less similar parts based on what is seen. More frequently in Tasks A and B, parts are added to the workspace after being viewed in context.

3.3 Coverage and selection of retrieved parts

For each task, participants were instructed to make a minimum of 5 searches and save a minimum of 3 parts to their galleries that were identified as potentially inspirational. To measure how broadly the dataset of parts was explored in each task, the dataset coverage of retrieved results by each participant is computed. 3D models are defined as covered by the user when they are within a Euclidean distance of 0.0375 in the appearance model embedding space from a model the user has interacted with. The coverage measure computed is the percentage of covered models relative to the dataset. This calculation is made assuming each 3D model has on average ~ 102 similar neighbors in the embedding space. Similarly, to analyze the collection of parts saved by each participant, the coverage of their saved parts relative to the dataset is computed.

3.3.1 Coverage of retrieved search results

To compare how broadly participants search using each modality, the coverage of the dataset by retrieved search results is explored. The coverage of all results is compared by task in Fig. 10, which shows the distribution across participants.

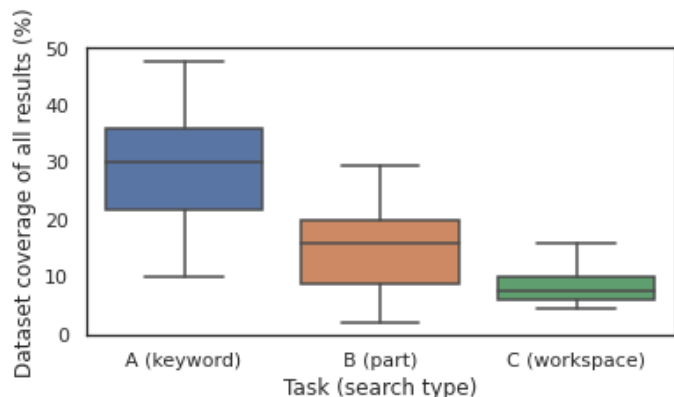


FIGURE 10: Comparison of dataset coverage of search results by task, with highest coverage observed in Task A (keyword search)

A significant difference was revealed between search types in the dataset coverage of all results using a repeated measures ANOVA ($F(2,40) = 27.0, p < 0.001$). The highest coverage of the dataset is observed in Task A. While previous results show that high frequencies of searches are observed in Task B, the coverage of part search results is significantly lower than the coverage of keyword search results, within participants ($t(1,20) = 12.5, p < 0.001$). The coverage of results from workspace searches is also significantly lower than from part search results ($t(1,20) = 12.0, p < 0.01$).

3.3.3 Coverage of results saved by participants

The coverage of a participants' saved results relative to the dataset is next examined to determine how broadly participants select parts. The same significant difference between tasks is seen in the coverage of saved parts as before ($F(2,40) = 20.4, p < 0.001$), where the highest coverage is observed in keyword search and lowest in workspace search. This relationship may suggest that participants retrieving a broader set of results also select and save a broader set of parts. Pearson correlations reveal that this is true in Tasks A and C, as shown in Fig. 11, where the coverage of saved parts vs. all results in each task are plotted for each participant.

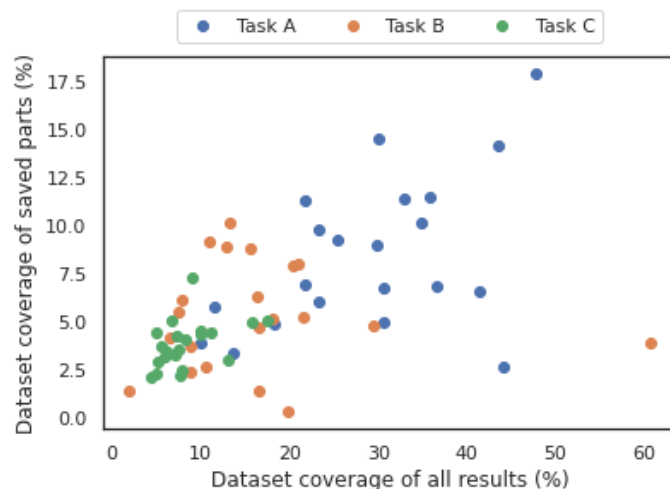


FIGURE 11: Comparison of dataset coverage (%) of saved parts and all parts retrieved by participants during each task

There is a moderate correlation between the dataset coverage of saved parts and all parts retrieved in Task A ($r(19)=0.49, p < 0.05$) and Task C ($r(19)=0.43, p = 0.05$) that is not observed in Task B. Therefore, when using part search, participants retrieving a broader set of parts did not necessarily also select and save a broader set of parts. Use of part-based search may therefore encourage a narrow exploration and selection of parts overall. However further testing is needed to account for the number of searches, retrieved results, and saved parts to fully clarify differences in coverage observed between tasks.

4. DISCUSSION

The platform developed in this work provides three options for users to express what they are searching for with search modes rarely studied in design. The first contribution of this work is the development of this platform and its effective use during a participant study. Specifically, we built a multi-modal search platform that enables search for 3D model parts by text, visual, and/or functional components of other 3D model parts and assemblies. A participant study was conducted to observe and compare search behavior when engaging with these various search modalities to collect parts to inspire solutions to a design challenge. Results of this study offer a preliminary investigation into how individuals search for, interact with, and select relevant inspirational stimuli using various inputs.

The second contribution of this paper lies in the insights gained regarding the association of search modalities with different interactions with the platform interface. Parts retrieved from keyword searches were the most engaged with, specifically by being viewed in context. Results retrieved and saved from keyword searches were also associated with the broadest coverage of search space. Increased search frequency occurred during part search, where more searches were modified using similarity sliders than made with new parts as search inputs. Additionally, during part search, the lowest engagement with search results was observed. When conducting workspace searches, participants made more new searches with changes to the workspace than with changes to similarity sliders only. Parts found and saved from workspace searches resulted in the lowest overall coverage of the search space.

Finally, the third contribution we make is to suggest reasons as to why the uses of each search modality resulted in distinct interactions within the platform and how they may reflect different cognitive processes underlying search. As introduced earlier, search behavior can be broadly divided into active vs. passive strategies, which support situations in which a specific goal exists vs. where random encounters with inspirational stimuli occur. Of the three modalities, searching by keyword offers the most direct means for employing active search. However, the platform also affords the ability to engage in passive search processes during keyword search. When a well-defined search term is unknown by the user, they may take other actions to help passively inspire their next search. One participant explicitly described this relationship when commenting on their keyword search strategy: *"I was inspired by some of the parts in the 'view in context' like the 'lid'".* This inspiration-seeking behavior is in line with observations from previous work that participants want to be struck by inspiration and to search more randomly [16, 17]. Increased engagement with parts may therefore be a strategy to randomly encounter inspiration. In other prior work, users presented with random examples frequently clicked on examples to examine them until something desirable was found [37]. Viewing parts in context and adding them to the workspace, where they can be perceived in a new representation, may be strategies to inspire subsequent searches. Additional methods to modify search parameters in keyword search, as currently afforded by similarity sliders in part

and workspace search, may help further fulfill this need to encounter more random examples when a new search term cannot be articulated.

During part and workspace searches, mechanisms to guide active search processes include the selection of the part-based search input and configuration of 3D parts in the workspace. The higher proportion of new than modified workspace searches, compared to more modified than new part searches, may demonstrate the effectiveness of enabling incremental modifications to the search input itself. Observed differences in these inputs suggest that users value the ability to conduct searches that vary individual parameters one at a time. 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 [37]. Multi-modal search should therefore support increased flexibility in defining the search input.

While adjustment of similarity sliders also provides a method for users to specify desired search results, qualitative responses reveal other reasons for changing slider positions during search. When asked to "describe any strategies [used] when conducting part searches", one participant noted: *"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 in 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."* These responses link the use of sliders to a more exploratory search strategy that can counterintuitively be useful for supporting exploratory behavior.

Some limitations of the current study exist, including that the experimental design of this study, where individual tasks employed a new search type, may have contributed to the confounding of results with other aspects of the study, such as ordering. The ordering of search types, from most to least intuitive, may have introduced other behaviors to the results independent from the modality used. The parts users want to find may also change over the duration of the study, such that more directed searches are made as their potential design solutions become more specific. Extensions to the current work can focus on participants using the search platform in a more freeform manner to understand which modalities are preferred and when.

These contributions of our work encourage and will hopefully inspire 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., a more direct and active vs. exploratory and passive strategy), and how to identify and promote these processes through interactions with features of search interfaces.

5. 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 participant study investigating the role of modality in search. There are many opportunities for future work to further the results of the current study. In the development of multi-modal search platforms, future work can introduce more interactions that allow users to flexibly search for inspirational stimuli. These interactions may be to alter the search input itself (e.g., towards freeform sketch-based inputs), or the desired properties of the results retrieved from search. Visual and functional relationships between search inputs and results were the main focus of this study. However, parameters such as the level of diversity within a set of results may be useful to specify whether random vs. specific inspirational examples are desired. Search platforms supporting these and other interactions can better 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.

ACKNOWLEDGEMENTS

The authors would like to thank Ahan Sabharwal for his assistance developing the search platform, and the participants who completed the user study.

REFERENCES

- [1] 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," *J. Mech. Des.*, 133(8), p. 081004.
- [2] Goucher-Lambert, K., Gyory, J. T., Kotovsky, K., Cagan, J., 2020, "Adaptive Inspirational Design Stimuli - Using Design Output to Computationally Search for Stimuli that Impact Concept Generation," *J. Mech. Des.*, 142(9), p. 091401.
- [3] Fu, K., Chan, J., Cagan, J., Kotovsky, K., Schunn, C., and Wood, K., 2013, "The Meaning of 'Near' and 'Far': The Impact of Structuring Design Databases and the Effect of Distance of Analogy on Design Output," *J. Mech. Des.*, 135(2), p. 021007.
- [4] Toh, C. A., Miller, S. R., 2014, "The impact of example modality and physical interactions on design creativity," *J. Mech. Des.*, 136(9).
- [5] Linsey, J. S., Wood, K. L., Markman, A. B., 2008, "Modality and Representation in Analogy," *AI EDAM*, 22(2), pp. 85–100.
- [6] Goucher-Lambert, K., Moss, J., 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, pp. 1-38.
- [7] Siangliulue P., Chan, J., Gajos, K. Z., Dow, S. P., 2015, "Providing timely examples improves the quantity and quality of generated ideas," *Proc. of 2015 ACM SIGCHI Conference on Creativity and Cognition (C&C '15)*, Glasgow, United Kingdom, June 22-25, pp. 83-92.
- [8] Cheong, H., Chiu, I., Shu, L. H., Stone, R. B., McAdams, D. A., 2011, "Biologically meaningful keywords for functional terms of the functional basis," *J. Mech. Des.*, 133(2), p. 021007.
- [9] Goucher-Lambert, K., Cagan, J., 2019, "Crowdsourcing inspiration: Using crowd generated inspirational stimuli to support design ideation," *Design Studies*, 61, pp. 1-29.
- [10] Kittur, A., Yu, L., Hope, T., Chan, J., Lifshitz-Assaf, H., Gilon, K., Ng, F., Kraut, R. E., Shahaf, D., 2019, "Scaling up Analogical Innovation with Crowds and AI," *PNAS*, 116(6), pp. 1870-1877.
- [11] Borgianni, Y., Rotini, F., Tomassini, M., 2017, "Fostering ideation in the very early design phases: how textual, pictorial and combined stimuli affect creativity," *Proc. of 21st International Conference on Engineering Design (ICED17)*, Vancouver, BC, Canada, Aug. 21–25, pp. 139–148.
- [12] Hua, M., Han, J., Ma, X., Childs, P., 2019, "Exploring the effect of combinational pictorial stimuli on creative design performance," *Proc. of 22nd International Conference on Engineering Design (ICED19)*, Delft, The Netherlands, Aug. 5–8, pp. 1763–1772.
- [13] Luo, J., Sarica, S., Wood, K., 2021, "Guiding data-driven design ideation by knowledge distance," *Knowledge-Based Systems*, 218, p. 106873.
- [14] Vasconcelos, L. A., Cardoso, C. C., Sääksjärvi, M., Chen, C. C., Crilly, N., 2016, "Inspiration and fixation: The influences of example designs and system properties in idea generation," *J. Mech. Des.*, 139(3), p. 031101.
- [15] Purcell, T., Gero, J., 1992, "Effects of examples on the results of a design activity," *Knowledge-Based Systems*, 5(1), pp. 82-91.
- [16] Goncalves, M., Cardoso, C., Badke-Schaub, P., 2016, "Inspiration choices that matter: the selection of external stimuli during ideation," *Design Science*, 2(10).
- [17] Herring, S. R., Chang, C. C., Krantzler, J., Bailey, B. P., 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)*, Boston, MA, USA, April 4-9, pp. 87-96.
- [18] Eckert, C., Stacey, M., 2003, "Sources of inspiration in industrial practice: the case of knitwear design," *J. Design Research*, 3(1).
- [19] Sutcliffe, A., Ennis, M., 1998, "Towards a cognitive theory of information retrieval," *Interacting with Computers*, 10, pp. 321-351.
- [20] Marchionini, G., 2006, "Exploratory search: From finding to understanding," *Com. ACM*, 49(4), pp. 41–46.
- [21] Athukorala, K., Glowacka, D., Jacucci, G., Oulasvirta, A., Vreeken, J., 2016, "Is exploratory search different? A comparison of information search behavior for exploratory

- and lookup tasks,” *J. Assoc. Inf. Sci. Technol.*, 67(11), pp. 2635–2651.
- [22] Mougenot, C., Bouchard, C., Aoussat, A., Westerman, S., 2008, “Inspiration, images and design : an investigation of designers’ information gathering strategies,” *J. Design Research*, 7(4), pp. 331–351.
- [23] Goel, A. K., Vattam S., Wiltgen, B., Helms, M. E., 2011, “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(10).
- [24] Chakrabarti, A. Sarkar, P., Leelavathamma, B., Nataraju, B. S., 2005, “A functional representation for biomimetic and artificial inspiration of new ideas,” *AI EDAM*, 19, pp. 113–132.
- [25] Murphy, J., Fu, K., Otto, K., Yang, M., Jensen, D., Wood, K., 2014, “Function based design-by-analogy: a functional vector approach to analogical search,” *J. Mech. Des.*, 136(10), p. 101102.
- [26] Fu, K., Cagan, J., Kotovsky, K., Wood, K., 2013, “Discovering structure in design databases through functional and surface based mapping,” *J. Mech. Des.*, 135(3), p. 031006.
- [27] Setchi, R., Bouchard, C., 2010, “In search of design inspiration: A semantic-based approach,” *J. Comput. Inf. Sci. Eng.*, 10(3), p. 031006.
- [28] Botella, M., Zenasni, F., Lubart, T., 2018, “What are the stages of the creative process? What visual art students are saying,” *Frontiers Psych.*, 9, p. 2266.
- [29] Zhang, Z., Jin, Y., 2020, “An unsupervised deep learning model to discover visual similarity between sketches for visual analogy support,” *Proc. of ASME 2020 IDETC/CIE, Virtual, Online*, Aug. 17–19, p. DETC2020-22394.
- [30] Jiang, S., Luo, J., Ruiz-Pava, G., Hu, J., and Magee, C. L., 2021, “Deriving design feature vectors for patent images using convolutional neural networks,” *J. Mech. Des.*, 143(6), p. 061405.
- [31] Jiang, S., Luo, J., Ruiz-Pava, G., Hu, J., Magee, C.L., 2020, “A convolutional neural network-based patent image retrieval method for design ideation,” *Proc. of ASME 2020 IDETC/CIE, Virtual, Online*, Aug. 17–19, p. DETC2020-22048.
- [32] Soufi, B., Edmonds, E., 1996, “The cognitive basis of emergence: implications for design support”, *Design Studies*, 17(4), pp. 451–463.
- [33] Menezes, A., Lawson, B., 2006, “How designers perceive sketches”, *Design Studies*, 27(5), pp. 571–585.
- [34] Gross, M. D., 2001, “Emergence in a recognition based drawing interface,” in “Visual and Spatial Reason in Design”, II Key Centre of Design Computing and Cognition (eds Gero, Tversky, Purcell).
- [35] Robertson, B. F., Walther, J., Radcliffe, D. F., 2007, “Creativity and the use of CAD tools: lessons for engineering design education from industry,” *J. Mech. Des.*, 129(7), pp. 753–760.
- [36] Cer, D., Yang, Y., Kong, S., Hua, N., Limtiaco, N., St. John, R., Constant, N., Guajardo-Cespedes, M., Yuan, S., Tar, C., Strope, B., Kurzweil, R., 2018, “Universal sentence encoder for English,” *EMNLP*, pp.169–174.
- [37] Lee, B., Srivastava, S., Kumar, R., Brafman, R., Klemmer, S. R., 2010, “Designing with interactive example galleries,” *Proc. of SIGCHI Conference on Human Factors in Computing Systems (CHI ‘10)*, Atlanta, GA, USA, April 10–15, pp. 2257–2266.