

## The Style of Tetris is...Possibly Tetris?: Creative Professionals' Description of Video Game Visual Styles

Stephen Keating  
University of Washington  
[keatings@uw.edu](mailto:keatings@uw.edu)

Wan-Chen Lee  
University of Washington  
[leew23@uw.edu](mailto:leew23@uw.edu)

Travis Windleharth  
University of Washington  
[travisw@uw.edu](mailto:travisw@uw.edu)

Jin Ha Lee  
University of Washington  
[jinhalee@uw.edu](mailto:jinhalee@uw.edu)

### Abstract

*Despite the increasing importance of video games in both cultural and commercial aspects, typically they can only be accessed and browsed through limited metadata such as platform or genre. We explore visual styles of games as a complementary approach for providing access to games. In particular, we aimed to test and evaluate the existing visual style taxonomy developed in prior research with video game professionals and creatives. User data were collected from video game art and design students at the DigiPen Institute of Technology to gain insight into the relevance of the existing taxonomy to a professional audience. Using a think-aloud card sort method, we observed their thought process for describing and categorizing visual styles of video games, and also collected candidate terms for revising the taxonomy. The results of this research will inform ongoing metadata work in the field to develop a standard for cataloging video games and interactive media, and will be useful to information systems that sort and classify games for users and cultural preservation.*

### 1. Introduction

Video game retrieval remains a challenging task despite many existing systems of organizing information. In the library and information science domain, there have been continuous efforts devoted to the establishment of classification schemes, such as the Dewey Decimal Classification, and cataloging rules like the Resource Description and Access. There are also metadata schemas like Dublin Core that expand their application to non-book materials. However, these existing standards were developed with a focus on describing textual materials and thus, are limited in their applicability to describe visual information appropriately [10]. With the advent of the Internet, the digital world created a glut of visual information that needs to be organized and accessed. Video games, as a prominent type of visual information, prosper in the digital age, and there are increasing needs to retrieve video games and game-related information. However, as noted in prior research, current search and discovery

tools cannot fully meet the needs of retrieving visual information and digital materials [2] [21]. In addition, current access points for retrieving video games are limited, with browsing options often restricted to platform or genre [9] [10] [14]. In order to improve the retrieval performance of video games in current information systems, standards for describing and organizing video games are necessary. Among the many types of metadata related to video games, visual style is an important but overlooked piece of information [5]. The lack of standard taxonomy for describing visual information, with regard to video games, is a primary reason for this problem. This study focuses on improving the organization and retrieval of video games through visual style information.

Our goal is to conduct specialized user-based research for modeling visual styles as they relate to video games. In particular, we explore how individuals close to the creation process of video games perceive and describe visual styles of games. The participants of this study were recruited from the DigiPen Institute of Technology, a university offering strong undergraduate programs focused on video games and digital art creation. This group was selected specifically for their close involvement in the video game creation process.

This study applies think-aloud and card sorting methods to collect user-generated descriptors of video game visual styles. We compared the collected terms with the existing taxonomy established by the GAMER (GAME Research) group at University of Washington Information School [6] [18]. This video game visual style taxonomy was created as part of a larger research agenda establishing a metadata schema for video games, which targets general users and information professionals who describe, organize, and preserve video games. The Video Game Metadata Schema (VGMS) was established as a result of this project. The schema was created based on multiple approaches involving domain analysis and a series of user studies, and was tested and modified several times [3] [10] [14]. The existing visual style taxonomy was also created and tested in a prior user study [3], but primarily with general game players. With the data contributed by the particular group of participants in

our study, we can examine the level of conformity between the user-generated terms and the visual style taxonomy established by the GAMER group. Through this study, we explore the following research questions in particular:

RQ1. How do individuals close to the creation process of video games describe and organize the visual styles of video games?

RQ2. Does the existing visual style taxonomy accurately reflect the cognitive descriptors applied by creatives to visual styles?

RQ3. Are there terms specific to video games and visual styles that should be added to the existing taxonomy?

This work will contribute to 1) improving our general understanding of how specialized users describe and organize visual styles of games, and 2) evaluating and understanding what might be improved upon in the current metadata schema and encoding scheme with relation to video games and visual styles. This work also highlights the importance of adapting metadata to specific users in order to facilitate information retrieval, given the highly specialized organizational terms resulting from the user research.

## 2. Relevant Work

Examining video game visual styles requires an understanding of both the context of conceptual image retrieval as well as the role of visual styles in video game retrieval. The following sub-sections will focus on reviewing literature related to each of these topics. Understanding the context of conceptual image retrieval helps elucidate the need for concept-based search in image retrieval, its inherent biases and limitations, and informs, through transference, how video game image retrieval is handled [5]. In addition, the second sub-section will focus on discussion of video game retrieval, mainly how metadata contributes to video game discovery, and how users respond to searching and browsing for games by visual style.

### 2.1 Conceptual Image Retrieval

In prior image-based retrieval research, there are two kinds of retrieval methods that are prominently discussed: content-based retrieval and concept-based retrieval [4]. Content-based image retrieval focuses on examining the image material, or what is contained within the image itself. Concept-based image retrieval, on the other hand, is focused on examining the visual relationships the work establishes via description. Chu identifies literature elucidating the schism between

content and concept image-based retrieval [4]. The history and focus of computerization is heavily influenced by content, rather than concepts, largely because content-based retrieval is an easier process to automate [5]. This resulted in a research field dominated by content-based work in spite of the rich but different research of concept-based work [4].

In the 1980s, Shatford [24] discussed the relative absence of studies for what users are searching for in images as well as how they are searching for it. By the early 2000s, image searching and organization emerged as its own field due to the explosion of visual content generated by the Internet. As indicated by the numerous studies Jørgensen references in 2003, the amount of research for querying digital images goes from practically non-existent in the 1980s to necessitating multiple studies by multiple organizations in less than a decade [11]. The importance of understanding user-querying of images and their needs is becoming increasingly important today. As users become more specialized, a need arises for the ability to search and organize information catered to interests that are not just general, but also highly specific. As early as the 1990s, Keister [12] noted that searching becomes specific to user types based on queries structured by specialization. A better understanding of how specialized users query their semantic world thus becomes vital for being able to better serve user needs, ideally at the level of the individual.

Though digital specialization results in the creation of a huge amount of image content, it also results in much of that information lacking context. Conceptual metadata creation is a time-consuming process and frequently motivated by personal interest or knowledge rather than empirical data from users. As a result, much of the research for image organization is still based on content analyses performed by individuals or machines rather than multi-user image organization analyses [7]. The emphasis on image content creates a semantic gap as defined in the image retrieval community [19], which prevents users from being able to discover concepts contained in the image. In addition, it also misses how the audience is using the image as well as what they perceive as important. For purposes of discovery and acquisition, how an audience finds images serves as the primary challenge, and what is often missing tend to be terms not described purely by content alone.

Visual styles are important methods for referring to concepts in an image and also provide a linguistic and descriptive organizational frame for understanding conceptual content. To assist retrieval of conceptual content, there must be some criteria that determine what is important within an image. Beaudoin's [2] research on image users across professions and

specializations serves as a guidepost for our research framework of open coding as well as identifying the organizational and search practices of individual users as they use and refer to images.

Lee et al. [14] identified evaluation criteria for organizational frameworks in metadata and their associated quality in two important information retrieval concepts, namely those of expectation conformance and consistency. If these two important factors are missing from the organizational process, it is highly unlikely that users will have the ability to find what they are looking for. This article focuses on expectation conformance and consistency as measures for understanding how specialized users organize and describe information about visual styles. Cho et al. [6] also explain the descriptive characteristics of visual styles, particularly for video games, as being difficult to pin down due to the difficult nature of defining and organizing art. Rather than trying to establish a descriptive or guiding definition, this research focuses instead on letting users describe their own understanding of visual styles as it relates to the subject of video games. By using an open coding method, the users are simultaneously free to organize the information according to their understanding, as well as use language that is unique to them. Later, this information is collated to establish a sort of communal language among participants for representing the visual style of video games.

## 2.2 Video Game Metadata

How video game metadata for visual styles and its descriptive standards play into the retrieval of video games is an important but simultaneously nascent field. None of the prominent conceptual models, metadata or cataloging standards, such as Functional Requirements for Bibliographic Records, Cataloging Cultural Objects, or Resource Description and Access can properly describe video game visual styles because they lack elements to comprehensively and accurately represent video games [1] [8] [20]. The intersection of a nascent video game metadata and under-researched visual styles in conceptual image retrieval adds an extra layer of complication. Visual styles are concept-based and thus not frequently researched in the image retrieval field, while the metadata for video games are also just starting to be established and defined.

Subsequently, while the need for search and discovery of video games have increased drastically due to the popularity of games, search/browse by visual style is not typically available in existing systems. On popular platforms for selling games such as Steam, Amazon, and Gamestop, no options exist for searching and browsing games by visual style. In a

recent user study involving 1,257 survey responses by Lee et al. [14], however, visual style was identified as useful metadata for seeking new games and for games the users currently play (53.4%). Furthermore, 55.3% of 237 participants with game-related professions responded positively to the need of visual style metadata as a search/browse method [17]. The visual styles of games were also mentioned as important factors when people decide which games they obtain and play, and also one of the strong appeals of games to particular types of users [15]. Because users have frequently responded positively to visual style as metadata for information retrieval, it is even more important to be able to discover video games by searching for their visual style explicitly. However, as a concept-based method of description, retrieving video games by their visual style is challenging due to a lack of standard on how to describe them.

## 3. Research Methods

As our primary research method, we employed a think-aloud card sorting activity. This was accompanied by some questions about the participants' organization process at the end of the activity. We chose observation of a card sort activity as our method because it is a useful technique for developing taxonomies with a specific user group. It can be used to "generate options for structuring information, as it can identify different schemas for organizing user navigation, menus, and taxonomy" [22]. Specifically, "the method can identify items that may be difficult to categorize or perhaps aren't as important as others" [22]. The goal of the card sort was to generate a list of terms that were prevalently used among participants, which we can compare against the existing visual style taxonomy in VGMS to identify areas for improvement.

A total of eight participants were involved in this research. All participants at the site were chosen based on their prior educational and professional work on video games. They were undergraduate students in at least their third year, or master's students, and all were active members of project teams developing video games. All participants studied video game art, design, or programming. Additionally, none of the participants had read or were aware of the specific taxonomy for visual style on which the cards in the activity were based. Participants were also not informed of previous or in-development taxonomies on video games or their associated visual styles. This kept the method free of bias which was important for discovering the points of divergence from the existing taxonomy.

Participants in the study were given a set of 48 cards containing video games screenshots, which

included multiple examples representing a broad array of video games with different visual styles. The screenshots were still images taken from video games as they are being played. The examples were carefully selected to ensure a balanced representation of all the terms from the visual style taxonomy developed in a prior study by Donovan et al. [6]. We included at least two screenshots of games representing each of the visual styles identified in the taxonomy.

Their main task was to sort the images into groups based on their similarities in visual style and then label the visual style of each group. Participants were read a set of instructions, indicating they were to sort the images into categories based solely on visual style, and that the sorting was intended to reflect participants' perceptions of video game visual styles. Participants were instructed to sort the cards into as many or as few stacks as they felt necessary, with as many or as few cards in each stack as they felt appropriate. They were informed that the exercise was not a test and that there were no right or wrong answers. Participants were also asked to think aloud while sorting, so additional data about their decision making process could be collected via video recording. They had as much time as they needed to complete the sort, which took users an average of 28 minutes to complete.

At the end of the sorting activity, participants were given a set of blank labels and were asked to write and place a label on or near each stack of cards they created. The researchers requested that these labels describe the stacks by visual style, using the participants' own language as descriptors. The information was recorded using photographs and video. At the end of the activity, researchers also asked questions about the sorting and labeling process, if they have noticed something unusual or interesting. This information was also recorded in video.

We completed the data analysis using the NVIVO qualitative data analysis software. All video recordings of the card sort activity were loaded into NVIVO, and comments made during the activity were openly coded for visual style [25]. The resulting code categories represent all the terms gathered during the sort activity describing the visual style.

## 4. Results

A total of 121 terms were generated from the recordings of the think-aloud process, and the labels participants assigned to the sorted piles of cards. Among the terms generated by the participants, we merged derivations of the same term (e.g. *Cartoon* and *Cartoony*, *Realism* and *Realistic*). Four prominent groups of visual styles emerged: *Abstract*, *Realistic*, *Stylized*, and *Text*. The card sort process showed that

participants did consistently use a few terms from the four prominent groups of visual styles, as they were commonly assigned to similar groups of cards. For example, among the five participants assigning the term *Abstract*, all of them included the identical two cards in their piles. Besides the four groups of visual styles, the participants also mentioned terms about different aspects of visual styles (e.g., techniques used to display particular visual styles, moods evoked from the visual representation) or terms representing information which is preserved in other metadata elements in the VGMS [18] (e.g., 2D in the dimension element, and real-time strategy games in the gameplay genre element). Since the instruction for the card sorting activity emphasized visual style, our expectation is that the terms collected are representing some aspects related to visual styles, at least in our participants' minds. The occurrences of terms representing information other than what we may typically consider as visual styles suggest close relationships between visual style and other metadata elements participants mentioned (e.g., mood, genre, point of view, and dimension). Overall, there was also a long tail of unique terms (56 out of 121) that were only used once by a single participant.

### 4.1. Prominent Visual Styles

Figure 1 presents the taxonomic structure we propose, combining the terms that emerged from our user data with unique terms in the current taxonomy (in italics). The shaded boxes mark our proposed changes to the current taxonomy. The four prominent types of visual styles that emerged from the user data are *Abstract*, *Realistic*, *Stylized*, and *Text*. The term *Stylized* functions as an umbrella term that covers descriptors of various visual styles showing simplified versions of real objects and/or environment. *Stylized* is defined in the existing taxonomy as "an illustrative style capturing or distorting distinctive features of a character or object in order to capture a visual likeness. Simplified characters or objects can be considered stylized." (p.10) [18]. The participants' assignment of this term, and other narrower terms such as *Anime*, *Handicraft*, and *Lego* were consistent with this definition. The earlier version of the visual style taxonomy (Version 1.0) [6] included hierarchies between the term *Stylized* and other relevant terms. The subcategories of *Stylized*, identified in our user data, included terms like *Anime*, *Cartoon*, *Handicraft*, and *Lego*. Some participants assigned even more specific terms to describe styles like *Handicraft*; for example, assigning terms like *Clay/Claymation*, *Knitted/Stitched*, or *Yarn* by indicating the textures shown in the screenshots.

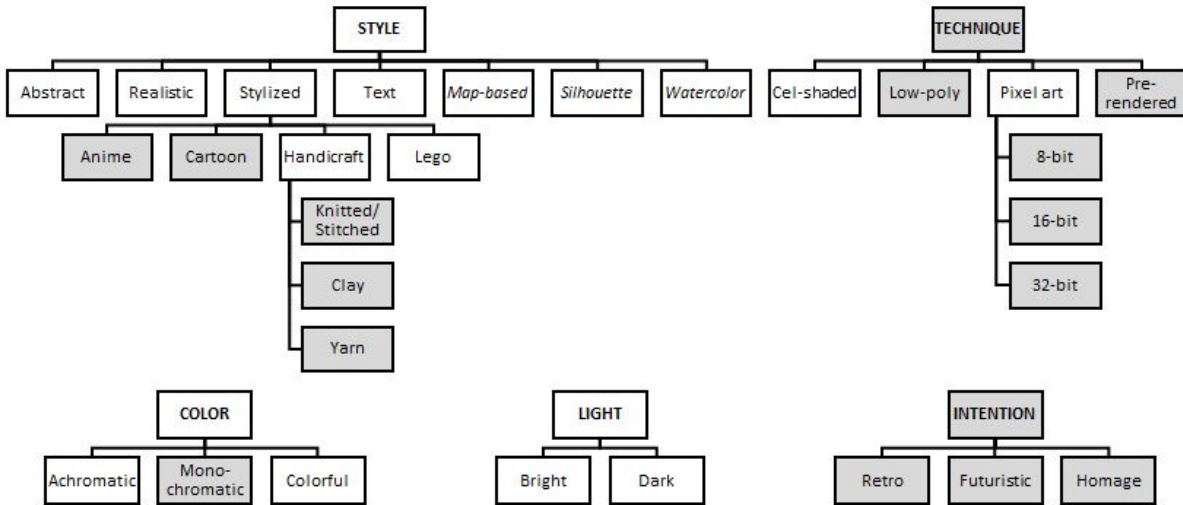


Figure 1. Proposed revision to taxonomy

*Abstract* was assigned by five of the eight participants, and *Realistic* and *Text* were assigned by six participants. The card sort process and result show that the participants not only commonly use the three terms, but also seem to have pretty consistent ideas about what they mean. Different participants came up with similar piles that they later described using one of these terms. Participants assigned *Realistic* to cards portraying photo-like details with the aim of recreating real-world visual experiences. *Abstract* was frequently assigned to screenshots without identifiable characters or objects. *Text* was used to describe cards presenting substantial amount of text. Some variants of the label *Text* include *Text-based* and *ASCII*.

The Technique group includes descriptors of the various techniques applied to create certain visual styles, namely *Cel-shaded*, *Low poly*, *Pixel art*, and *Pre-rendered*. *Pixel art* is further divided into *8-bit art*, *16-bit art*, and *32-bit art*. The earlier version of the taxonomy [6] also had a Technique facet, which included *Pixel art*, *Cel-shaded*, and *Low poly*, and some additional terms. The current version (Version 1.5) [18], however, does not have the Technique facet; *Cel-shaded* and *Pixel art* are present, but are shown within a simple list of different visual styles without a hierarchical structure. Taxonomy for professional users such as game artists and programmers may need to be structured with different hierarchical levels to accommodate more specific and technical terms.

#### 4.2. Metadata Elements Related to Visual Styles

When the participants were searching for their own words to describe visual styles, they often mentioned terms that are closely related to, but not technically

about visual styles (e.g., *Colorful*, *Bright*, *3D*, *Real-time strategy game*, *Depressing*, *First person perspective*). The terms like *Colorful* and *Bright* can be represented as terms organized under different facets (i.e., *Light* and *Color*) in the existing taxonomy. Several other terms describe some other metadata elements in VGMS [18]. The language and sorting used specifically by the participants is discussed at length in Section 5.3.

We matched the user terms with identified metadata elements in VGMS, which are Dimension, Gameplay genre, Mood, Point of view, and Subject (i.e., the main object depicted in a card). Similar findings were described in Windleharth et al. [29] where the authors collected and analyzed user-generated tags from Steam, a popular game distribution platform [28]. They observed five different aspects of visual styles mentioned in user tags: 1) visual looks or appearances, 2) visual mood (e.g., cute, dark), 3) visual representation of gameplay mechanics (e.g., grid-based movement), 4) visual techniques (e.g., hand-drawn), and 5) visual motifs (e.g., blood, gore). The first and fourth aspects are what we are treating as visual style terms in this work. However, the number of other relevant aspects mentioned in both this study and Windleharth et al. [29] (e.g., mood, mechanics) demonstrates the complexity of describing visual styles and also affirms the importance of including elements in the current schema that encompass these terms.

Another type of metadata that was sometimes mentioned by participants when describing the visual style was specific game titles or series (e.g., *Tetris*, *Mirror's Edge*). In some cases, specific games were mentioned because participants believed that those were the games from which the images were taken. In other cases, they were mentioned as a proxy term

representing a particular visual style that is highly associated with the games. For instance, they used the term *Street Fighter* style to refer to colorful, 2D representation of stylized characters in a fighting mode, or the term *Mirror's Edge* to refer to a style using extremely bright and saturated colors with reflective objects in the environment.

## 5. Discussion

### 5.1. Purposes and Structure of Metadata Schema

The result of this study shows that the participants, as expert users who create video game art, tend to assign more terms about techniques that are used to create particular visual styles. Unlike the prior user testing of the existing visual style taxonomy by Cho et al. [3], where the authors found that general game users did not typically use technique-related terms when they were asked to describe the visual style of sample game screenshots [3], our participants did self-generate and offer several technique-related terms. This is especially true for the six artists (as opposed to the two programmers) who tended to provide more descriptors on different techniques. The emergence of technique-related terms in our data demonstrates the higher importance and relevance of techniques and its relationship with visual style from the perspective of expert users. As mentioned above, the first version of the visual style taxonomy recognized the importance of techniques by including a Technique facet. This facet was removed in later versions of the taxonomy to make it more usable for catalogers [6] [18]. The taxonomy also placed an emphasis on the overall schema serving general users' needs [6]. The shift of inclusion to exclusion of the Technique facet corresponds to the shift of the schema's purpose from catering to multiple user groups including game players, scholars, game designers/developers, curators, parents, and teachers, to serving general game players' needs of searching and discovering video games.

Based on the result of this study and the earlier taxonomy [6], we suggest reinstating Technique as an optional facet for specific user groups. Organizations that typically cater to general game users can choose to opt out of applying this facet, while organizations that serve expert users can provide richer metadata related to visual styles to meet their needs. To develop a more comprehensive list of technique-related vocabularies, future study can explore 1) a wider variety of technique-related terms collected from a thorough literature review and user study involving a larger group of people, 2) the extent of agreement among expert users about the definitions and usages of these

terms, and 3) how these terms influence the organization and retrieval of video games in a game browser or retrieval system such as the one presented in [17].

### 5.2. Hierarchical Relationships between Descriptors

The participants of this study, who are artists and programmers contributing to the design and creation of video games, often assigned terms with different levels of specificity in written labels, and sometimes suggested hierarchical relationships between terms. For instance, participant one assigned the label "8-bit→16-bit (pixelated)" to a pile of cards, and explained while thinking aloud "I would say this is like 8-bit/16-bit art. That is what I meant by pixelated." This suggests a hierarchical relationship between the broader term "pixelated" and narrower terms: 8-bit and 16-bit, because when users say pixelated, they may be referring to 8-bit, 16-bit, or 32-bit pixel arts. Cases like this helped us identify and establish a hierarchical structure among the terms. While the current taxonomy does not employ a hierarchical structure for the goal of simplification and serving the needs of general users, the results from this study well demonstrate how the purpose and the needs of targeted user group can influence the structure of the taxonomy. More broadly, the difference between the current taxonomy and the result of this study also highlights the importance of flexibility in developing metadata schemas and encoding schemes. Instead of viewing the difference between the structure of current taxonomy and what we propose as a conflict, one possible approach to accommodate this issue may be to incorporate a hierarchical structure, but offer different recommendations to different types of organization as to how deep they should go into the structure when applying the visual style terms. Based on the needs of the users these organizations serve, and the resource and expertise catalogers have, they can decide to apply specific terms from a lower level in the hierarchy, or choose to only use the terms from the top level.

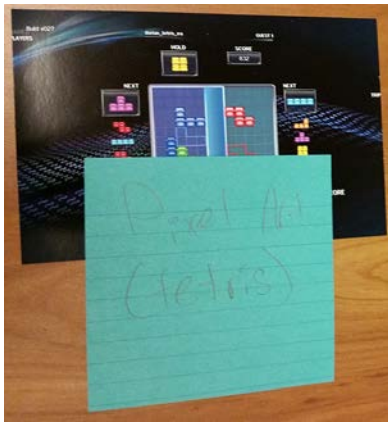
### 5.3. Card Sort Reactions

A number of interesting behaviors were observed during the participants' card sort activity. Here, we discuss a few notable themes in more detail.

**Lack of user vocabulary and uncertainty:** One of the major themes emerged from user observation was their challenge in coming up with vocabularies for describing visual styles in general. While most of the users had no problem grouping and labeling certain

images representing visual styles like *Abstract* or *Text*, they were much more uncertain about other images.

There were some games that were particularly more challenging to users. The screenshot of *Tetris* was one such example (Figure 2). Participant two stated, “I don’t know what to label *Tetris*, other than possibly *Tetris*”. While the overall style of *Tetris* can be described as *Abstract*, it is such a well-known game with a prominent visual style that the participants had a hard time coming up with terms other than *Tetris* to describe the visual style. The users’ uncertainty sometimes resulted in them re-sorting the piles of images. For instance, participant nine re-sorted the piles after being reminded that he had to label the piles, because he felt he lacked a language to describe some of the piles. This highlights the ambiguous nature of the concept of visual style, and how labeling affects the organization. For this user, the initial grouping strictly based on visual similarity of the images ended up not being preserved.



**Figure 2. Visual Style Labeling Difficulties - Tetris**

**Different aspects of visual styles:** When examining individual card sorts, we observed that participants tend to focus on different aspects of visual styles. For instance, participant eight assigned the term *Group battle* to screenshots of games depicting scenes where a group battle is occurring in the game. This ties back to visual mechanics, one of the five aspects mentioned in Steam tags, as discussed in [29]. Participant eight, on the other hand, frequently focused on subjects (i.e. prominent parts of characters) and labeled a card as *Body parts* to indicate a focus on visual styles that accentuated forms as the central focus over defining the image in its entirety. This participant also tended to assign terms like *Cute* or *Friendly*, terms conflated with moods, to images with bright or pastel use of color. The use of mood terms was also observed in participant four, six, and seven. Participant nine had a strong focus on colors and light, assigning terms like

*Black and white*, *Colorful*, and *Bright*. There were other instances of a long tail of visual style terms that emerged, and each of them tended to focus on some visual aspect that rang true to the individual participant as contributing to a specific visual style, though many of these would often fall under the taxonomy of *Technique* or *Mood*.

**Language and semantic complexity:** During the card sort process, participant nine asked if he could assign labels in his native language (Chinese), because he had a hard time coming up with the ideal term in English. One of our group members shares the same native language with the participant, and thus was able to translate his terms into English. The linguistic struggle adds another layer of complexity to the generation of descriptors to visual style, which is already an ambiguous concept. An example is the label *Black* assigned by the participant. Through the think-aloud information, we came to understand that what he really meant was *Dark*. However, since black and dark can share the same term (黑) in the participant’s native language, there can be semantic vagueness in the process of encoding descriptors.

**Mix of multiple visual styles in games:** Our participants sometimes found it difficult to process images from certain games where they employ a combination of multiple visual styles. For instance, participant nine placed a card with a Lego figure with a realistic background in the *Realistic* pile, along with cards with realistic textures (e.g., yarn), and cards with overall realistic-ness (i.e., photorealistic screenshot), although he also had a separate *Lego* pile. This is an example of grouping by different interpretations of the *Realistic* style. It highlights the multi-dimensional nature of terms, and surfaces the challenges of defining the descriptors.

**Varying textual expressions of the same concept:** Comparing the labels assigned by different participants, we observed a wide variety of different textual expressions of the same concept. For instance, participant eight used *Visual trick architecture* to describe the same card participant one described as *Escher-style game art* and participant two described as *Escher-esque*. It demonstrates the lack of consensus of vocabularies people use to describe visual styles, and the importance of developing a taxonomy, clearly grouping the near synonyms, and specifying the preferred and lead-in terms to improve video game retrieval. This also indicates how a specific knowledge of a visual style that is particular to an artist (M.C. Escher) can be helpful for representing similar styles.

**Different level of specificity:** There was a variance in how specific each participant tried to describe certain visual styles. For example, while most participants assigned some terms to represent a pile of realistic images, participant six assigned terms with different degrees of realism, such as *Exaggerated*, *PS2 realism*, *Realistic*, establishing that there is a spectrum of realism across which visual styles cover, rather than it being a singular conceptual structure.

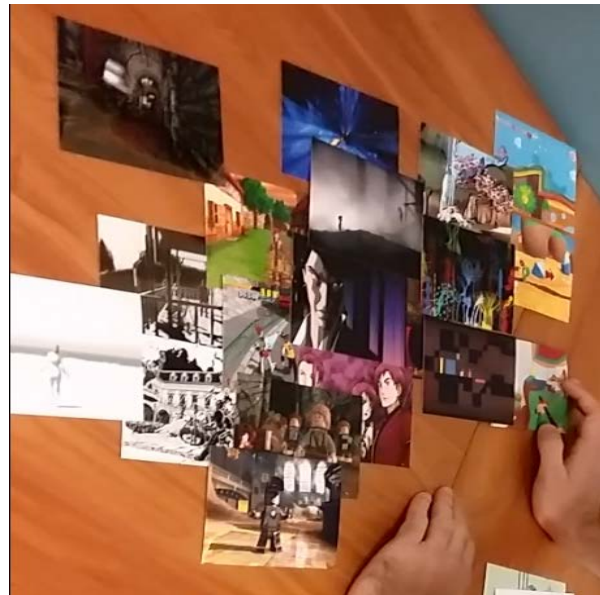
**Dealing with challenging images:** When participants came across challenging images, they often put the card aside at first, hoping to see other cards with a similar visual style so they can put this card in the same pile. While they were instructed that the pile can contain any number of games (not necessarily multiple games), they still sought to find a “group” of similar images.

However, some participants got to a point that they determined they were not able to figure out a good category for certain cards. Participant one created a pile labeled *General (hard to quantify style)*. Participant eight also used *Didn't know* as a label.

**Conceptualizing the relationships in the physical space:** During the process of sorting and arranging cards, some interesting behaviors emerged in the manner in which users arranged card stacks during the exercise, across the physical space of the table. Several participants (participant one, two, four and nine) not only created and labeled categories, but also attempted to portray some manner of conceptual proximity of one category to another by thoughtfully arranging the stacks on the table (Figure 3). For instance, participant one initially created a spectrum of cards across the table to indicate varying levels of abstraction. While pointing at an image on the left side of an array of cards he stated, “even though this is blocky, it is also stylized, and not minimalist either...probably over here”. He then moved two cards to mark where he felt the relative position of that image was in terms of abstraction, relative to the others.

Similarly, participant four arranged the card piles in a specific manner on the table during the sort, in this case, after the sorting was complete. The participant used his finger to indicate a tight cluster of seven piles on the table, and stated, “I guess you could say that these are stylized genres. It's like, they're not fully abstract, but they are stylized”. The participant followed up by indicating where that selection of stylized categories would exist in the spectrum of abstraction. He stated, “so it would be if this pile was a pop-out of a pile right here”, and then indicated a place on the table between the “abstract” cards and the “realistic” cards. He then stated, “So this is on the far

edge of like, less strict, and this would be higher poly”. By “less strict”, he indicated the more abstract side of the table, and by “higher poly”, he indicated the side of the table with more realistic and stylized images.



**Figure 3. Spectrum of visual styles instead of groups**

These two examples illustrate the ambiguity involved in classifying images with varying levels of abstraction and stylization, and that images tended to exist on a continuum for some participants, leading from abstraction, to stylization, to realistic. This implies there will be boundary cases in categorizing video game images, particularly at the intersection of stylized and realistic, as well as abstraction and stylized. Some specification is advised in developing systems that sort and classify games by visual style, such as “abstract”, versus “stylized”, versus, “realistic”. This also implies that perhaps the games themselves in the information retrieval systems should be visually represented and organized in a digital space with some kind of spectrum, so that the navigation becomes more intuitive for users.

## 6. Challenges

### 6.1. Conflation of Descriptors

As noted in the previous section, we observed users often conflating the visual style with other types of metadata about the game or game images. This conflation resulted in users providing a mix of terms related to several different metadata elements. What became evident was that our participants, as creators of video game art, do not always agree on what counts as



visual style of video games. The result also indicates the inherent ambiguity of the concept of visual style. Unlike content-based image retrieval, conceptual organization takes on an aspect of abstraction that ultimately leads to disagreement about appropriateness or relevance, even if a user consensus is established. Without a clear definition users agree upon, descriptive complexity is unavoidable. The card sorts and open coding are bottom-up approaches which are ideal for generating user-centered opinions and terms. However, we also face the challenge of participants' inconsistent interpretations of key concepts, and particular to this study, the concept of visual style.

## 6.2. Accommodating Temporal Changes

The meanings of concepts and terms change over time. The tension between keeping up with current trends to reflect the evolving semantics and at the same time, remaining stable and consistent for practical application is a common challenge across many information organization systems, as discussed in [26] and [27]. In this study, the tension surfaced when we placed the *Pixel art* category under the Technique facet. Recently, pixel art has become a visual style a video game creator can choose to adopt. In the past, all video games used a pixel art style, not because of creators' choices, but because of technological limitations. What was once considered a realistic visual style may become retro as a result of technological or technique-oriented progress.

In order to accommodate the temporal changes of concepts, descriptors, and the relationship between a technique and a descriptor, we propose possibly adding an optional facet Intention to address the intention of video games creators. This must be an optional element because game creators do not always make explicit statements about their intentions, but sometimes they do (e.g., [23]), and the information can be helpful for video game retrieval. Possible values of this facet could be: *Retro*, *Futuristic*, and *Homage* (i.e., adopting a particular visual style from an existing game as a tribute). This would alleviate the problem of not being able to distinguish, for example, between "intended" pixelated look from recent games versus pixelated look in older games which might have been intended and perceived to be realistic. In addition, through this facet, we can describe the temporal aspect of a video game, as well as expressing potential relationships between games that are often not explicitly captured in catalogs and information retrieval systems [16]. For instance, we can link the inspiring game and the inspired games through the value *Homage*. *Pillars of Eternity* is one such game, where the author of the game makes explicit their intention to take inspiration from a

previous series, *Baldur's Gate*, which has a similar visual style. The use of such explicit visual style relationships is common in many media, including video games.

## 7. Conclusion and Future Work

This study takes a preliminary step toward exploring how professionals who create video games apply the concept of visual style in their works, and some main categories of visual style emerged. We observed several interesting behaviors as participants engaged in this card sort activity, which illustrates various issues and challenges that exist in describing visual styles of video games. A few terms prominently emerged from the user data, including *Abstract*, *Stylized* (and different sub-categories of Stylized), *Realistic*, and *Text*. Some terms overlap with the existing taxonomy [18], and are affirmed by the expert users' perspectives. The results from the card sort activity also highlighted other metadata that are closely related to visual style, including dimension, gameplay genre, point of view, subject, and mood in particular. We propose complementing the taxonomy by suggesting restoration of an optional Technique facet, possibly adding an Intention facet, and enriching vocabularies with higher specificity. Since there are Technique terms in the earliest version of the taxonomy [6] that were not mentioned by our participants, future works can expand the sample size to test if terms included in the earliest taxonomy recur, and to develop a more comprehensive list of terms.

## 8. References

- [1] Baca, M., P. Harpring, E. Lanzi, L. McRae, and A.B. Whiteside, *Cataloging cultural objects: A guide to describing cultural works and their images*, ALA, Chicago, 2006.
- [2] Beaudoin, J.E., "An investigation of image users across professions: a framework of their image needs, retrieval and use", Ph. D. dissertation, Drexel University, PA, 2009.
- [3] Cho, H., A. Donovan, J.H. Lee. "Art in an algorithm: A taxonomy for describing video game visual styles". Unpublished manuscript, Information School, University of Washington, Seattle, United States, 2016.
- [4] Chu, H. "Research in image indexing and retrieval as reflected in the literature", *Journal of American Society for Information Science.*, vol. 52, 2001, pp.1011-1018.
- [5] Colombo, P., C. Antoniotti, & M.R. Ciceri, "Pictures and Words: Analogies and Specificity in Knowledge Organization", In B. Apolloni, M. Marinaro, & R. Tagliaferri (Eds.), *Neural Nets: 14th Italian Workshop on Neural Nets*,

WIRN VIETRI 2003, *Revised Papers*, Springer, Berlin, Heidelberg, pp. 358-366, 2003.

[6] Donovan, A., Cho, H., Magnifico, C., & Lee, J.H. "Pretty as a pixel: Issues and challenges in developing a controlled vocabulary for video game visual styles", In *Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries*, New York: ACM, 2013, pp. 413-414.

[7] Marsh, E.E., and M.D. White, "A taxonomy of relationships between images and text", *Journal of Documentation*, vol. 59, no. 6, 2003, pp. 647-672.

[8] Hagler, R., "Nonbook Materials: Chapters 7-11". In D. H. Clack (Ed.), *The Making of a Code: The Issues Underlying AACR2*, ALA: Chicago, 1980.

[9] Järvinen, A., "Gran stylissimo: The audiovisual elements and styles in computer and video games", In *Proc. CDGC*, Tampere University Press, 2002, pp. 113-128.

[10] Jett, J., S. Sacchi, J.H. Lee, and R.I. Clarke, "A conceptual model for video games and interactive media", *Journal of the Association for Information Science and Technology*, vol. 67, no.3, 2016, pp. 505-517.

[11] Jörgensen, C. & P. Jörgensen, "Image querying by image professionals", *Journal of the American Society for Information Science and Technology*, vol. 56, no. 12, 2005, pp. 346-1359.

[12] Keister, L. H., "User types and queries: Impact on image access systems", In R. Fidel, T.B. Hahn, E. M. Rasmussen, and P.J. Smith (Eds.), *Challenges in Indexing Electronic Text and Images*, Learned information, Medford NJ, 2012, pp. 7-22.

[13] Kyprianidis, J., J. Collomosse, T. Wang, and T. Isenberg, "State of the art: A taxonomy of artistic stylization techniques for images and video", In *IEEE Trans. Vis. Comput. Graphics*, IEEE, 2012, pp.1-20.

[14] Lee, J.H., R.I. Clarke, and A. Perti, "Empirical evaluation of metadata for video games and interactive media", *Journal of the Association for Information Science & Technology*, DOI: 10.1002/asi.23357.

[15] Lee, J.H., R.I. Clarke, and S. Rossi, "A qualitative investigation of users' discovery, access, and organization of video games as information objects", *Journal of Information Science*, 2015. <http://doi.org/10.1177/0165551515618594>

[16] Lee, J.H., R.I. Clarke, S. Sacchi, and J. Jett, "Relationships among video games: Existing standards and new definitions", In *Proceedings of the 77th ASIS&T Annual Meeting*, 2014.

[17] Lee, J.H., S. Hong, H. Cho, et al., "VIZMO game browser: accessing video games by visual style and mood", In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*, ACM, New York, NY, USA, 2015, pp. 149-152.

[18] Lee, J.H., A. Perti, H. Cho, et al. "UW/SIMM Video Game Metadata Schema: Controlled Vocabulary for Visual Style, Version 1.5., 2015. Retrieved from: [http://gamer.ischool.uw.edu/official\\_release/](http://gamer.ischool.uw.edu/official_release/)

[19] MacFarlane, A., "Knowledge organisation and its role in multimedia information retrieval", *Knowledge Organization*, vol. 43, no. 3, 2016, pp.180-183.

[20] McDonough, J., M. Kirschenbaum, D. Reside, N. Fraistat, and D. Jerz, "Twisty little passages almost all alike: Applying the FRBR model to a classic computer game", *Digital Humanities Quarterly*, vol. 4, no. 2, 2010.

[21] Midge, C., "Because you're you: factors influencing item selection in a digital sheet music collection", *The Electronic Library*, vol. 32, no. 6, 2014, pp. 884-897.

[22] Martin, B., and B.M. Hanington, *Universal methods of design: 100 ways to research complex problems, develop innovative ideas, and design effective solutions*, Rockport Publishers, Beverly, MA, 2012.

[23] Obsidian Entertainment, Project Eternity, 2012. Retrieved from: <https://www.kickstarter.com/projects/obsidian/project-eternity/description>

[24] Shatford, S., "Analyzing the subject of a picture: a theoretical approach", *Cataloging & Classification Quarterly*, vol. 6, no. 3, 1986, pp. 39-62.

[25] Strauss, A. L., and J.M. Corbin, *Basics of qualitative research: Grounded theory procedures and techniques*, Sage Publications, Newbury Park, California, 1990.

[26] Tennis, J.T., "Scheme versioning in the semantic web", *Cataloging & Classification Quarterly*, vol. 43, no. 3-4, 2007, pp. 85-104.

[27] Tennis, J.T., "The strange case of eugenics: A subject's ontogeny in a long-lived classification scheme and the question of collocative integrity", *Journal of the American Society for Information Science and Technology*, vol. 63, no. 7, 2012, pp.1350-1359.

[28] Welcome to Steam, n.d., Retrieved <http://store.steampowered.com/>

[29] Windleharth, T.W., J.H. Lee, M. Schmalz, and J. Jett, "Full steam ahead: A conceptual analysis of user-supplied tags on Steam", *Cataloging & Classification Quarterly*, vol. 54, no. 7, 2016, pp.418-441.