

# Understanding Modal Synergy for Exploration

Niklas Rönnerberg<sup>\*</sup>  
Linköping University  
Sweden  
niklas.ronnerberg@liu.se

Jonas Löwgren  
Linköping University  
Sweden  
jonas.lowgren@liu.se

## CCS Concepts

•Human-centered computing → HCI theory, concepts and models;

## Author Keywords

Musical sonification, Modal synergy, Images, Interaction

## 1. INTRODUCTION:

### DEVELOPING PHOTONE

When working on an empirical user study exploring how sonification can support visual perception of brightness, we used the green color channel for the visual stimuli [3]. Various sound elements, such as sound level, pitch, and harmony, were mapped to the brightness level of the green color and explored in the study.

This work inspired a more speculative thought: What if all three color channels were sonified independently and simultaneously? Could we then listen to an image, or could the image be used to explore musical sounds? To start investigating this idea, we created an interactive installation called Photone [2, 4] that combines color images with musical sonification to explore sonification of color values of individual pixels by moving a cursor over the image. The sonification is driven by pixel values of hue and brightness without any notion of what the image actually portrays. In other words, we sonify syntactic properties of an image (e.g., hue and brightness, edges, shapes, spatial frequencies) rather than semantic ones.

We believe that this combination of color images, sonification algorithms, and interaction represents the potential for a different way of seeing – similar to how someone learning to draw portraits needs to unlearn ideas of noses, ears and facial expressions in order to see lines, shapes and hues. What is more, we find that image, sound and interaction together form something more than a mere addition of supplementary modalities. The resulting experience of Photone cannot be understood as a simple sum of visual and auditory stimuli; our tentative analysis is that the fine-grained details of interaction form the glue between the visual and auditory modalities and the experience emerges as

a synergistic whole. We use the label of modal synergy to characterize this experiential quality.

## 2. QUESTION:

### WHAT HAPPENS IN THOSE LONG SESSIONS?

When we first exhibited Photone [2], we observed that most users in a public exhibition space only toyed with it briefly, while others spent a considerable amount of time exploring the images and musical sounds. Based on these observations and on feedback provided by some of the users, we reasoned that for some users, maybe for those with musical aptitude and reflective mind, Photone was more interesting and engaging.

To get a better understanding of the use of Photone, we continued the general-audience work [4] with a slightly updated version and by registering interaction patterns. In total interaction patterns for 233 users were registered; the analysis at the time focused on the longest 10% ( $n = 23$ ) of these (see 1 and 3).

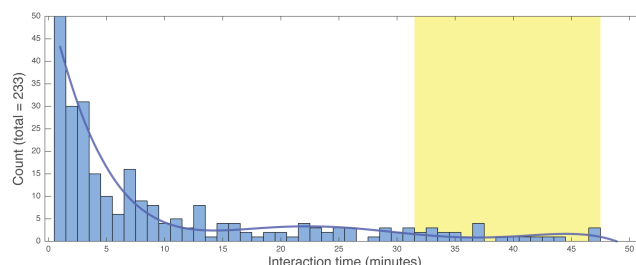


Figure 1: The interaction times (in minutes) for all users. The analysis of the interaction patterns focused on the longest 10% of these (with yellow background).

Briefly, what we found suggests that longer interaction engagements seem to have more interaction in image areas which do not contain visually salient objects in the conventional, semantic sense but rather syntactic features that may yield engaging visual-auditory synergistic effects. The question is: What happens in these long sessions, and how is it experienced?

## 3. LOOKING CLOSER AT THE INTERACTION PATTERNS

To begin addressing these questions, we re-examined the interaction patterns from the second study. When exploring the registered data further we compared the 50% ( $n = 129$ )



Licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). Copyright remains with the author(s).

2nd Audiovisual Symposium 2024, 6 December, Falun, Sweden.

longest interaction times with the top 10% ( $n = 23$ ) longest, and with the top 5% ( $n = 12$ ). Two examples from this analysis can be seen in 2 and Figure 3.

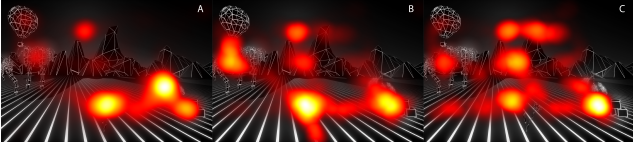


Figure 2: The leftmost image (A) shows the interaction pattern for the 50% longest sessions. The middle (B) shows the 10% longest, and the right (C) shows the top 5% longest sessions.



Figure 3: The leftmost image (A) shows the interaction pattern for the 50% longest sessions. The middle (B) shows the 10% longest, and the right (C) shows the top 5% longest sessions.

This analysis indicates that mainly clearly defined visual objects are targets for user exploration in general. However, for the longer sessions this becomes less obvious and a broader variety of areas are explored. This finding suggests that there is a different behavior and interest in exploring Photone for users that spend longer time compared to the ones that only spend a few minutes exploring it. The new look at the old logs highlights this difference, but still cannot provide any insights into the experiences and motivations of the most persistent users. The rest of this paper presents our plans for a first step towards satisfying that curiosity.

#### 4. CONNOISSEURIAL EXPLORATION

Human-computer interaction (HCI) and UX has a solid heritage of focusing on usability, of minimizing user error and making interaction efficient and easy. Arguably, this heritage has motivated spending much research and development efforts on walk-up-and-use situations and on users on the lower end of the expertise spectrum. Relatively little attention has been devoted to highly proficient users and uses.

However, in other fields such as art and design criticism, acknowledging higher levels of expertise is not only customary but also recognized as valuable. A “connoisseur” in these contexts is not an ironic slander against elitism but rather a valuable source of knowledge and insight – the productive work of an erudite art or design critic, the “expert reading”, relies entirely on connoisseurial judgments, insights and perspectives [1].

A simple analogy to illustrate the difference might be to consider a musical instrument. Most traditional musical instruments are fundamental failures in terms of instant usability, and an obvious UX fix for a violin if the task is to produce a pleasant tune would be to propose a handheld music player with tune selection, start and stop buttons. However, that would kind of miss the point of a violin. The expressive possibilities that a skilled violinist enjoys are of an entirely different nature; they rather represent connoisseurial experiential qualities.

We are obviously not claiming that Photone would be comparable to a violin. It might perhaps be productively viewed as a simple musical instrument – in fact, one of us has “played” Photone in several live performances for fair-sized audiences – or maybe it is merely an interactive composition, providing readymade components but leaving the detailed ordering and exposition of phrases and passages to the user. Be that as it may, our observations so far suggest that it can offer worthwhile experiences to some users who engage for longer times in more explorative ways. Our provisional position is that these might be examples of connoisseurial exploration of the modal synergy sensations that Photone can provide. The next and final section describes our plans for starting to look into such situations and articulate their experiential characteristics.

#### 5. PLANS FOR EXPLORATIVE CONNOISSEUR INTERVIEWS

The aim of the continuation of the work with Photone is to gain a deeper understanding of modal synergy. We intend to do this by a qualitative approach of observations and interviews focusing on connoisseurial exploration of Photone. At least in the field of sonification this is an important yet understudied subclass of use situations, as most often students or lay people are recruited for generic sonification experiments.

The implementation and design of Photone will be similar to the second iteration of it [4]. The composition and sound design will similarly reflect different levels of energy in the image material. An image that we consider to have more energy in terms of, e.g., brighter colors and more contrasts, will have a musical sonification or composition with brighter timbre, more dissonance, and more complex harmonies. However, in this version of Photone, the rhythmic aspect will be omitted as we found that it directed the user more towards visual areas with more contrasts. In this work we want the user to be free to explore Photone in any way they see fit.

In the first installation of Photone we used photographic images that were rich in syntactic information (e.g., varying in colors, contrasts, brightness levels, shapes and visual objects). The second installation used computer generated images that consisted of a background and a clear visual object. We felt that this type of artificial images was less interesting to explore in Photone compared to photographic images. Therefore, in the coming version of Photone we aim to use image materials rich in colors, contrasts and different brightness levels, and where all areas of the image are more or less equally important. Specifically, we plan to use Paul Cezanne’s painting “Mont Sainte-Victoire”, Sonia Delaunay’s painting “Portuguese market”, and Paul Klee’s painting “City with the three domes”.

For the explorative connoisseur interviews we will invite individuals with a profound experience and interest in music, art and design, and/or interactive art. We believe that such individuals will be able to both appreciate the experience of Photone, and have the ability to articulate this experience. Photone is a way for us to explore the phenomenon of modal synergy, and this knowledge and these interests are (probably) equally important when users want to explore Photone for a long time.

We want to find components of a language that describes the phenomenon of modal synergy and the kind of experience Photone enables. Our tentative method is to let the user choose an image and then freely explore this for as long as they want, then they continue with the next image

and finally the third and last. Between the exploration of each image there will be a break to clear the mind (and the ears). The interaction (mouse position and movement speed) is recorded in the background while the user explores Photone.

When the user is done with exploring all three images, a session of conversation and prompted recall starts. The conversation starts with the image that the user found most interesting. The user's interaction is replayed (visually together with the musical sonification) to catalyze a semi-structured interview/conversation that includes the experience in general and passages of particular interest in terms of music, sound, color, image, and interaction. After this the interaction for the least interesting image is replayed. Also for this image the more interesting passages are discussed and reflected upon. This conversation will also highlight the experienced differences between the first and the second image and the musical expression. Finally the interaction for the last image is reviewed and discussed.

What once started with a question about what different pixels in an image sound like, how the auditory modality could be used to explore the visual, turned out to be an exploration of something larger. We do not see modal synergy as a translation between sensory modalities, but rather a blend of modalities that creates something larger. The visual and auditory modalities contribute to an experience of a synergistic whole, catalyzed by interaction which involves

the corporeal modality of proprioception. With the methodological approach described in this paper, we hope to better understand the experience of Photone, and provide a better understanding of the effect of modal synergy.

## 6. REFERENCES

- [1] J. Bardzell. Interaction criticism: An introduction to the practice. *Interacting with computers*, 23(6):604–621, 2011.
- [2] R. Niklas and J. Löwgren. Photone: Exploring modal synergy in photographic images and music. In *Proceedings of the 24th International Conference on Auditory Display (ICAD 2018), June 10-15, 2018, Michigan Technological University, Houghton, MI, USA.*, pages 73–79. Georgia Institute of Technology, 2018.
- [3] N. Rönnerberg. Musical sonification supports visual discrimination of color intensity. *Behaviour & Information Technology*, 38(10):1028–1037, 2019.
- [4] N. Rönnerberg and J. Löwgren. Traces of modal synergy: studying interactive musical sonification of images in general-audience use. In *Proceedings of the 25th International Conference on Auditory Display (ICAD 2019), June 23-27, 2019, Northumbria University, Newcastle-upon-Tyne, UK*, pages 199–206. Georgia Institute of Technology, 2019.