

A handful of pixels of blood

A Historical and Technological Perspective on Understanding Video Game Graphics

Good afternoon, colleagues. Today, I'd like to share with you parts of my research on video game programming practices of the 1980s and 1990s, with a particular focus on graphics. This work is an integral part of my dissertation, where I'm exploring the technological foundations of video games as a popular medium.

The advent of homebrew game development in the 1980s and 1990s fascinates me because it was such a formative period for video game development. During this time, countless individuals and small teams established techniques that would become foundational to the medium. I believe that video games, as a form of ludic software, have been instrumental in shaping our access to and understanding of the digital world. That's why I think it's crucial to comprehend how the technological underpinnings of this medium came into being.

In my research, I've faced a significant challenge: we rarely have access to both the programmers and the source code from this period. However, I've found that we can examine the products themselves—the games—to draw meaningful conclusions. I see this approach as a form of critical code studies with historical material, but with code viewed through the lens of video game graphics. ## Datasets, Framework, and Method

To conduct this research, I'm working with two primary datasets. The first is called the Ludens Images Dataset. It contains a bit more than 3800 images from 35 games relevant to the Confoederatio Ludens project, all of which have strong connections to Switzerland. We would have around 220 video games in our corpus, but due to the obscurity of many of these games, it is difficult to get a hold on screenshots and I've primarily had to rely on video stills from Let's Play videos to populate this dataset.

The second dataset is much larger and more comprehensive, the Video Game History Screenshot Dataset, or VHS-Dataset for short. It contains approximately 113,000 images from around 4,300 games, encompassing nearly all pre-1990s games available on the video games information platform MobyGames. I created this dataset to provide a map or a distant view of video game graphics history and its development.

Further, I'm using the Framework for the Analysis of Video game Representation (FAVR), which was introduced in a seminal 2015 paper. I find this framework particularly useful because it emerged from a fundamental critique of existing methods for analyzing video game images. The authors argued, and I agree, that vocabulary borrowed from art history and film studies was insufficient for capturing the unique aspects of video games.

FAVR is built around five core elements: modes, composition, construction, spaces, and planes. Modes refer to the overall structure of what's visible on screen and its specific function, like title screens or gameplay screens. Composition and construction represent different levels and aspects

of visual analysis. Composition focuses on the overall arrangement and organization of spaces on the screen surface, providing a broad view of how visual elements are laid out, and foundational understanding of the screen's visual organization.

In contrast construction delves deeper into the layered nature of video games images, recognizing that the visual representation is composed of multiple conceptual planes. These can be agents, in-game environment, off-game environment, and interfaces. They take formal aspects into account, such as graphic materials and projection methods. These attributes help reading into technical aspects of what we see on screen, a bit like art historians that know what kind of pigments were used by looking at a painting's colors in conjunction with when and where it was painted. Planes offer a more granular analysis of the technical and conceptual elements that make up the game's images.

As part of my research, I translated the framework into an ontology, based on CIDOC CRM, and created Tropy templates to be able to properly annotate images. I've found it also necessary to expand FAVR by another element. I've incorporated the concept of ludemes as another type of plane. A ludeme is the minimal unit of play or the smallest element that can be grasped by a player, in terms of game rules or game mechanics. Think of the +4 card in Uno. They're essential for gameplay and can take many forms, such as these switch-orbs-and-barrier-puzzles in The Legend of Zelda: A Link to the Past. Ludemes combines both the structural aspect of game design and the experiential aspect of player engagement, and can be composed of graphics, sounds, and game mechanics. What I'm particularly interested in is their service as an interface between playing and creating games. They reflect the dual materiality of video games: playable forms and underlying code. I see ludemes as the meeting point between player agency and the game's visual representation of its internal states.

Findings

So far, FAVR has proven invaluable in two specific cases in my research. First, in an exploratory case study, I investigated the ratio of static to dynamic content to infer computational resource requirements. This helped me understand how the technical limitations of the time influenced game design and vice versa. I mentioned before how computing resources were limited. So, by simply looking at how complex and dynamic parts of a video game image are, one is able to deduct certain technological aspects. Was the game made for a certain system, the dynamics of video game graphics can conclude if a game was written in BASIC or Assembly.

Second, FAVR has been crucial in my conceptualization of ludemes. It's helped me bridge the gap I've observed in my field between image studies and technical analyses. Either they are concentrating on a images content and semantic discourses or they focus on technical aspects and describe how these were the base for certain formal outcomes. But they rarely bridge that gap. If we want to have a better understanding of video game programming practices we need to be able to do interdisciplinary analysis. Programmers don't just write code, they also have to pre-conceptualize the results, and work in several domains at the same time. That was certainly the case in 1980s video game development. I believe that ludemes are crucial in building such a bridge, since they are at the intersection of code and visibility in game design.

And now for a brief detour. Working on the VHS-Dataset has provided valuable insights. So let me briefly talk about that, since it is shining another light on the same issue. I used the self-supervised transformer model DINOv2 and UMAP for dimensionality reduction, followed by k-means clustering. Interestingly, I found that clusters primarily formed around either formal or semantic aspects, or in some cases bringing these two together to build genre clusters.

An example of a cluster around formal aspects is the isometric projection cluster. This cluster picked up on the specific formal element of isometric or oblique projection, which is a way of constructing a 3-dimensional representation without perspective. Regarding semantics we can pull up this nsfw-ish one. The ability to include images, graphics and illustrations, of course, introduced visual material to video games, that is deemed not safe for work. This includes pornographic images, display of violence and gore, and material that shows ideologically problematic material, like war and fascist symbolic. Lastly and is an example of a clusters around a video game genre. The *Boxing* cluster is an example of a clustering, in which formal and semantic aspects mingle and evolve in the direction of a genre. Despite being of fundamental different game play and mechanics, and aesthetics, the cluster is held together by a few semantic-formal elements.

I'm at the moment in the process of compiling the research on the VHS-dataset into a paper, so I'm not going into any more details for now. The important takeaway for this presentation is that DINOv2 can identify modes. Many of the clusters that were produced, for example only contain start screens or in-game screens with levels with an isometric projection. But, the model struggles with ludemes. Those typically manifest as repeated graphical elements within a game, but were not recognized as an important visual or semantic unit within the clusterings. Ludemes are most often intuitively comprehended by players, if the game developers and designers did their job. But for now they pose a semantic category, that is beyond what the computer vision model can see. Beyond epistemological questions, this also raises the need to continue the work on the FAVR ontology, in order to teach the computer to see these elements.

Implications

How does this research align with the theme of the conference: “Historical Research, Digital Literacy, and Algorithmic Criticism.”? I'm examining how digital matter can take various forms—source code, algorithms, executable software, on-screen image—and exploring ways to formalize aspects that traverse these forms, like ludemes. I believe this to be an important take away for this conference. Algorithms don't manifest only in social media feeds, online marketplaces or stock trading software. They are also imbued in everyday things such as modern car radios or video games. Often, studying these digital objects means looking at them from a variety of perspectives, in a multiplicity of forms, and of course their context.

By combining historical analysis, technological understanding, and visual examination, I believe we can gain deeper insights into the evolution of video game graphics and programming practices. This multifaceted approach not only enriches our understanding of video game history but also provides valuable perspectives on the interplay between technology, design, and player experience in the broader context of digital media. Moreover, I see it contributing to digital

literacy and algorithmic criticism in a broader sense, helping us better understand the digital artifacts that increasingly shape our world.

Of course, this research comes with its challenges and limitations. The scarcity of source material from the period I'm studying is a significant hurdle. Additionally, the subjective nature of visual analysis means that some of my interpretations may be open to debate. However, I believe that by being transparent about these limitations and by combining multiple analytical approaches, we can still draw meaningful and valuable conclusions.

As of now, the two datasets, as well as the formalized FAVR ontology stand and are published on Zenodo. I wish I could have already done concrete case studies and ask questions to the datasets, through the application of FAVR. But for now the creation and maintaining of the datasets as well the construction of the ontology used far more time than I anticipated. After finishing the VHS-dataset paper I will continue with this work here and start to annotate specific examples from the datasets. If you have some free time on your hand, I'd be more than welcoming if you want to join this endeavour. And if not, I would be eager to hear your thoughts and be open to any questions or discussions about this work.

Thank you for your attention.