Catch and Release: Ludological Dynamics in *Fatal Frame II: Crimson Butterfly*

William Huber
University of California
San Diego, CA

Abstract

Videogames are software-based artifacts, and their nature as such affords researchers the opportunity to study a range of stable semiotic interface elements. Players engaged with a game become increasingly skilled at decoding these games over a range of semiotic registers to navigate through the game, interpret elements in the visual field, and respond effectively to game challenges. Using techniques drawn from scientific analysis of visual images and video data, the author has examined the videogame *Fatal Frame II: Crimson Butterfly*. After identifying modes of play corresponding to clusters of operations and representations in recordings of game navigation, the author identifies patterns of repetition and suspense which support the aesthetics experience appropriate for a horror game; however, this effect does not operate in isolation.

The article also criticizes approaches toward the analysis of videogames which rely on stark distinctions between operational and representational modes of reception, arguing instead for a simultaneous decoding of elements in the visual field. As signifiers are decoded within a given register, whether operational or diegetic, those signifying elements remain latently available for decoding in others. These latent, deferred decodings – evoking the resemblances and eerie identities described in the traditional discourse on the uncanny – are considered another element in the production of uncanny experience in *Fatal Frame II: Crimson Butterfly*.

Author Keywords

Ludology; culture; emotion; meaning.

One persistent problem in the study of videogames has been the difficulty in resolving them both as playable, dynamic software-based media and as screen-based systems of representation, especially as narratives. While few researchers still consider it viable to relegate either of these elements into the background and restrict their attention to the other, the tensions and omissions which result from methodologies ill-prepared to handle the hybrid nature of videogames still vex efforts to create critical readings of videogames as representational and gameable systems.

Two years ago, collaborating with a colleague in the art-history department in which my doctoral studies are taking place, I played and wrote about the Japanese horror game *Fatal Frame II: Crimson Butterfly* (Tecmo, Limited 2003.) My colleague focused on a Freudian reading of the references to twins, doubles, and dolls in the game, while I sought to see a reflection of those

themes in the relationship between the player and the avatar(s) and in the navigational mechanics of the game.

While my colleague's work was well-executed, I remained somewhat unsatisfied with our methodology in relocating the traditional reading of uncanny aesthetics into the fabric of a videogame. I began to sense that the analysis was based around an awkward gear-shifting, from discussion of the representational elements as if they constituted a unitary object of analysis, to that of the operational modes as if they too were unitary and ubiquitous. On reflection, this was inconsistent with the way that I - and others to whom I speak - actually play and experience games, particularly horror games. The narrative aspects of the game had been selectively reconstructed from recollection of the traversal of the game in a way that dismissed as incidental the concrete details of the interactive reception of the work. The language used to describe the relationship between the ludic and representational elements of videogame experience too often relies on a metaphor of a representational, semantic skin wrapped around an interactive, procedural and syntactic skeleton. In the spirit of generosity that has arisen in the wake of the conflict between so-called ludologists and narratologists (e.g., see Frasca, 2003), most critics recognize the necessity of interactive as well as representational elements in the analysis of games. In practice, however, criticisms privilege one element and pay only passing respect to the other. Of the over 130 papers presented at the 2009 conference of the Digital Game Research Association, about 30 could be said to be primarily based on the reading or analysis of the game as a cultural artifact (rather than being ethnographic studies, psychological and cognitive theories of play, design and engineering research, etc.1) Of these, about half privileged representational elements and relegated mechanical and game-dynamical elements to the background, while the other half did the opposite. Only Ian Bogost's keynote address at the 2009 *Digital Games Research Association conference*, "Videogames are a mess," suggested an ontology which is resonant with the temporal experience of play.

Visualizing ludological dynamics

At the Software Studies Initiative at the California Institute for Telecommunications and Information Technology (Calit2), in UC San Diego Jeremy Douglass, Lev Manovich and myself have assembled a suite of techniques for the analysis of videogame recordings. These tools and methods produce linear representations of the gameplay experience. While these temporally flattened visualizations of videogame play can be visually provocative, more importantly, they suggest a re-conceptualization of the signifying functions of a videogame. Videogames are not simple portmanteaus of interactive software and systems of representation and diegesis. Instead, these aspects of the game are phenomenologically bound up into the rhythms and tempos of gameplay, their signs getting discharged in unison across a range of semiotic registers.

How can we approach the videogame as an object of analysis itself? Does the gameplay session (or at least a representation of it) provide a better object for critical reading? We might begin to use this intermediary approach to address questions about the nature of videogame aesthetics and experience – in this case, the question of the Unheimlich (uncanny) and other elements of the aesthetics of fear in a videogame. This approach both troubles and reconciles the separation between the object as autonomous and reception as the context for the production of meaning.

The game/play analysis methods make use of software written to analyze image data from the biological sciences, particularly ImageJ, an application currently being developed by the National Institutes for Health. Recordings of gameplay footage are saved as digital files. These files are then converted to still frames at a sample rate which varies according to the scope and scale of the research question at hand: in the case of this project, most of the footage was sampled at the rate of 2 frames per second.

Subsets of these images are then sorted into bins of distinct (and non-exclusive) modes reflecting the player-researcher's reading of the game itself. In this sense, the process is really a quantification of the subjective experience of gameplay, insofar as these modes are derived from reflection on the experience of play. These collections of manually-sorted frames within a mode are then composited to create a single image: the stable visual elements within that image are taken as mode-identifiers. It is then a simple matter to match those elements to other frames throughout the footage of the entire traversal of the game, no matter how many hours of gameplay has been recorded.

Elements of franchise

The game *Fatal Frame II: Crimson Butterfly* is part of a franchise of games on various platforms, which began with the 2001 release of *Project Zero* on the PlayStation 2. The game's concept was the work of Keisuki Kikuchi; it was both developed and published by Tecmo. Kikuchi was the design lead for the first three games, and (was) the co-lead for the fourth, a title created for the Nintendo Wii. The game has been so-well received that it has become a staple in academic videogame criticism (along with the *Silent Hill* and *Resident Evil* series: see Taylor, 2006 and Perron, 2004) even if in some ways it is idiosyncratic as a member of the survival horror genre.

Videogame franchises are rhizomatic, complex entities. Some, like *Pokémon*, were conceived from early in their production as transmedia properties and merchandising opportunities. Others have members which share a handful of tropes, iconic elements, or gameplay elements to create a relationship between its individual titles on various media and platforms, without sharing even a fictional world. Yet others implicitly claim to be portraying a narrative trajectory unfolding over a long period of time. The *Fatal Frame* series maintains its coherence as a franchise both in its diegesis and in its game mechanics. The titles are a series of loosely inter-related returns-of-the-repressed, atavistic hauntings, possessions and assorted traumatic memories and re-experiences involving a constellation of characters with various relationships to each other. The core game mechanical component has remained stable throughout the series: the use of the camera obscura, a device strongly reminiscent of the paranormal technologies of the Victorian era. There are others in the game: the spirit radio, for example. The auditory field in the game does not, however, have the same rich ludic-diegetic function as the visual field. When the auditory does have a ludic function, it is generally extra-diegetic.

The camera obscura replaces traditional forms of weaponry used in most other games, and is at the center of the most important player activity: the simultaneous documentation and exorcism of

spirits. What importance in this case means is not simply a function of time, nor of centrality to the narrative, but rather it is foregrounded at the activity of peak excitation, that in which the player is in greatest peril of failure, the activity which makes the most demands on the immediate attention of the player throughout the course of the game. This action is distinguished from the practice of navigating the space of the game in a trade-off that is made between freedom of motion and freedom of view: when the player is in a navigational mode, the camera is completely controlled by the CPU; when the player is in camera obscura mode, the player has full freedom of rotation for a 360º view, but is unable to move forward or backward. The visual reception of the game even in its most diegetic moments and modes is subject to the transition between these modes.

Modes of play

Clarification of the term "mode" for the intermediary elements which comprise the basis of the following analysis is appropriate. Conventionally, the term "mode" is used to designate variations in the style and structure of the game system as a whole: e.g., "mulitplayer mode", "capture the flag mode", "cooperative mode", etc. These variations are designed features of the game, and frame the player's engagement with the piece. However, I use the term to describe the (mostly) visual cues and guides that order the attention and activity of the player over the course of play of the game: the semiotic regimes through which fictive, gamic, and meta-game representations are organized. These modes are phenomenological and experiential, yet can be inferred from the elements on the screen at any given instant at play. These elements include the representations which indicate the momentary affordances of the mode (to heal, to shoot, to navigate, to detect the presence of foes, etc.). The nature of videogames as software-based artifacts makes the semiotics of the visual field relatively stable (vis-a-vis film or video), making it easy to analyze the digitized footage of gameplay with image analysis tools as described earlier. These modes can be thought of as a relationship between semiotic registers, in which one may dominate the others: for example, while in a navigational mode, during which the player can move through the 3-dimensional space of the game, the dominant registers are those discharged in spatial cognition: surfaces, orientations, the apparent ability of a door to open or a surface to be climbed, etc. The textures of these spaces, denoting aged wood, or moss-covered rocks, or hidden nooks, are also available to the player, signifying diegetic information relevant to the fictions of the game. Reducing the navigational to a simple, straightforward operational register would be a mistake.

I have identified the following modes in the game *Fatal Frame II* (and an illustrative montage of representative frames accompanies each description).

The first is the cutscene (Figure 1): a mode in which the player's interactive agency is removed and their attention is called to various narratological, aesthetic, or game-diegetic elements (e.g., clues for the solution of puzzles). In many games, these are indicated by letterboxing and/or FMV. In *Fatal Frame II* on the PlayStation 2, there is no letterboxing; however, they can generally be recognized by the absence of other display elements.

Figure 1. The cinematic mode: cutscenes and intertitles

Another is the navigational (Figure 2) mode, in which the player's interaction with the controller moves the avatar through the rendered three-dimensional space of the game; this mode constitutes the majority of most players' experience of the game. The depth and nature of player attention demanded in this mode can vary. It may not be far off to describe this mode as the canvas for the others; the default spatio-temporal posture in the Lost Village. In *Fatal Frame II*, this is a timeless mode; one generally invites danger only if one moves.

Figure 2. Navigational mode: avatar moving through 3-dimensional space; the player's camera is orthogonal to the avatar's point-of-view.

I have already described the operation of the camera obscura as a franchise-signature element. The camera itself is discovered in the first chapter of the game, in the first of a series of buildings

which the player explores. Camera mode (Figure 3) is accessed from navigational mode: the player presses the square button to shift between the two modes. Shifting into the camera mode arrests mobility. The stable visual elements that identifiy this mode are the partially transparent focus ring which is superimposed over the visual plane and the icons on the top row of the image, indicating the type of film currrently loaded in the camera, the number of frames of film left, and other indicators of the operational state of the camera in the context of the scene.

Figure 3. Camera mode: the avatar's point-of-view is synchronized with the player's point of view, as the fictive camera becomes the player's camera-eye. Progressive motion is arrested; the player has, instead, complete f360-degree freedom of motion in the round.

Another mode of play is combat mode (Figure 4), distinguished by the inclusion of specific visual indicators to the camera mode described above. The element which indicates combat is the filament at the top center of the frame in camera mode; it is presaged by a glow in a nearly identical filament, placed in the lower-right corner of the screen, in navigational mode. It should be thus noted that these modes are not mutually exclusive: they are denoted by the semiotics of the visual field (or rather, the significations of the visual field indicate the systems of meaning and reference which constitute a mode) but they are not identical to them. Thus, combat mode can be characterized as the union of the sets of frames in either navigational or combat mode in which the glowing filament appears.

Figure 4. Combat mode: usually a subset of camera mode, although a player will move to the navigational mode as needed to re-orient themselves, alternating between rotational and linear freedom of movement. Signifying elements which were latent in the previous modes indicate the need for combat, and new iconic elements become foregrounded.

Other modes of player-game activity include inventory screens, camera-upgrade menus, photo-album modes, various diegetic inscriptive systems (discovered notebooks, discovered newspaper clippings, the spirit radio and the discovered ethnographic film footage). There are interesting aesthetic and narrative effects created by these elements; however, my observation is less about the specifics of *Fatal Frame II* and more about the relationship between the overall experience of the game as a horror video game and the aesthetics of the uncanny and the eerie. For the sake of clarity, we have isolated two of these modes as significant: a generalized "menu mode" in which various auxiliary operations occur, and the save-screen, where a player saves or loads a game.

The dynamics of Fatal Frame II: Crimson Butterfly

The modes described above were produced by capturing the videostreams of the full traversals of the game. We describe as a traversal the collection of game-sessions from the beginning of a new game (the conventional "new game" button that reserves an area of long-term storage on a memory device to record the state of the game between play sessions) through the final game sequence and the closing credits. A traversal is thus an artifact of an extended instance of reception, a transcription of the "putting-into-play" of the game software's assets and algorithms by a player.

The videogame recording is then processed to produce a linear series of still frames, sampled at various rates as appropriate for the analysis at hand. For purposes of determining the modal

transitions of game traversals, we generally use two-frame-per-second (fps) sampling, with fifteen fps sampling when we wish to "zoom" into the dynamics of shorter play-segments.

We were able to compare four different players' traversals of the game: three were fairly typical (in terms of genre and platform literacy and enskilment) videogame players (one undergraduate and two graduate students); the fourth was a speed-run performance uploaded to YouTube by a player under the pseudonym Persona2 (Figure 5). In general, the patterns of alternation between the three researchers were stable, despite the range of ability and familiarity with the game and its genre between them. The speed-run traversal (Figure 8) was distinctive: the speed-runner was, of course, interested in minimizing the duration of any mode as much as possible. Cinematic sequences are, however, of fixed duration; thus, they constitute a greater percentage of the speed-runner's full traversal when compared to the other players. The speed-runner has essentially removed a semiotic register entirely: any given sign or icon is treated only as a cue to a well-rehearsed, somatized activity.
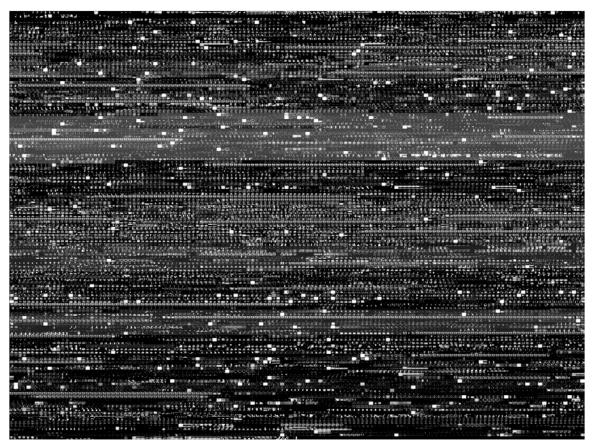

Figure 5. The complete traversal of Fatal Frame II as a 2-hours speed-run, visualized as a montage (sampled at one frame per second.) The player has learned optimal strategies to eliminate unnecessary expenditure of time. White frames indicate that an enemy has been defeated: the defeat of an enemy is accompanied by a bright, white flash which fills the screen completely. This element can also be used as an indicator for a (successful) battle.

Modal Rhythms

The oscillation of informational and operational modes is the "catch and release" of my title. The cinematic sequences during which the player is not called upon to act releases the compression created when the signs of the uncanny accumulated during operations are instead relased into another kind of reading. This release is never complete: even as the player is transfixed by the phantasmic spectacle on the screen, the sequence may contain clues and indications which the player will need to progress through the game. The result is a somaticization of the reading of the range of signs.

The perception of hazard is, in fact, greater than its reality, because the same signifier can come into play in various registers and modes. In this visualization of mode transitions during a typical play-session (Figure 6), what is striking is how little actual combat or near-combat activity there actually is: most of the activity occurs as cinematic material or in navigating the space of the game.3 The low ratio of actual jeopardy to the connotation of jeopardy is a simple mechanism for creating suspense. One detail which emerges after sorting through the transcripts of play sessions is the player's management of the codas and discontinuities of the play experience. Not only do players use save-points in the ways we might expect - to save progress to avoid having to repeat a passage, or in order to end a session - but menus are also called up to freeze play to take a break, as a pause mechanism. These patterns of saving and pausing are the player's way of managing the tempos of the game experience in the interstices between intense and semiotically dense sequences of play.
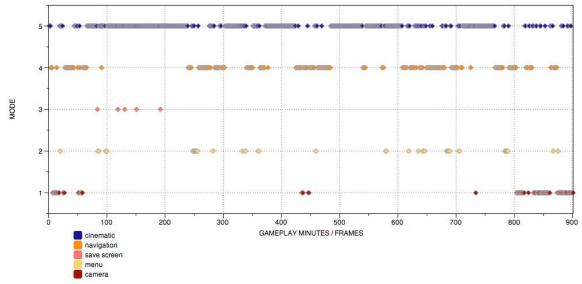

Figure 6. Typical modal transitions over 900 minutes of gameplay.

Modes encode the possibility of other modes, and the player's experience is marked by an indeterminacy: where are they? What is expected of them? We can see this ambiguity at work by looking at entirely typical screens and isolating dominant signifiers within them.

9

A typical frame from a player's navigational mode (Figure 7) features Mayu, the character/avatar (b) navigating a passageway. The filament (c) would begin to glow if there were spectral activity in this scene in the direction toward which the avatar (but not the player's camera) is facing. The region (a) is an area in which, at certain points in the game, a ghost could well appear, if the avatar turns in that direction. At the moment of this frame, however, these are mere possiblities. In fact, at this point, Mayu is exploring the deserted Kurosawa house. The filament remains at the periphery of attention, in the background, denoting that the player has the camera obscura equipped (at a later point in the game, this object is lost and the filament does not appear on the screen) and connoting the Edwardian aesthetics (i.e., the dawning years of the 20th century, when new media technologies such as the wireless radio were being first produced, and had the nature of rarities produced in workshop-like conditions), of the camera and the spiritual technologies which populate the game.
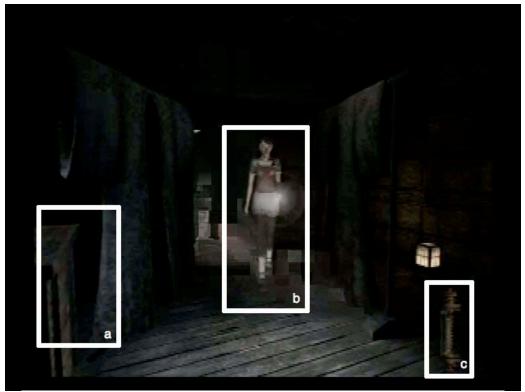

Figure 7. Semiotic elements from a sequence within the navigational mode.

A moment of greater player excitation (figure 8) occurs as the player-character (b) encounters a ghost which is not yet in the visual field (a), indicated by the red filament (c); that the filament is glowing red indicates that the ghost is hostile and must be dispatched by the player. The blue gauge (d) indicates that the player has already taken some damage, and must attend to their health-state or risk losing the game and being obliged to restart from a save point. We should take note of the transition between the above mode and the one illustrated here and observe that this is not a transition between a representational or fictive state and an operational one. The scene which did not indicate combat or jeopardy was also operational: however, that operation is navigation, rather than conflict. The subsequent scene contains a range of signifiers which must still be decoded in a fictive register: the identity of the ghost, the nature of the box in which it

appeared, etc. The transition between those modes is a change in excitory states which manage the tempo and tone of the experience of the game.

Figure 8. Signs presaging a modal transition.

The screen in mid-combat (Figure 9) reveals a rich range of icons and representations, which the player is compelled to decode appropriately in an operational register, while deferring their decoding in a fictive one. One gauge (a) indicates the quantity of a given type of film (c) that is left in the camera (different film types have different efficacy - the player learns over the course of a game to conserve the most effective film types for the most difficult battles.) The ghost (e) appears after examing a window (d) - after the combat is completed, some information about the nature of the ghost will become available to the player. At this moment, however, such considerations are deferred. Other icons indicate with what augmentations the camera has been equipped (f). The element on the bottom of the frame (g) is decorative: it also connotes the late-Victorian/early-Edwardian spiritual-technology motif. The player is probably not attending to it at the moment of this frame, although it shares visual features with elements (b), (c), and (h).

Figure 9. Combat mode: signifying elements across a range of ludological semiotic registers.

After dispatching a ghost such as that shown in figure 9, the player might either open their menu and restore their health, or return to a navigational mode, continuing the exploration of the architectural spaces in which the game is set. The exposition of the elements discharged by the encounter with the ghosts are read within the menu screens. Ironically, the "menu" is usually considered the least narratological and most operational aspect of a typical console game, yet it may contain the most explicit fictive exposition.

Observations

How does a modal analysis of this game contribute to a reflection on uncanny experience in the videogame? In Freud's 1919 essay, the uncanny as read through an analysis of E.T.A. Hoffman's story, *The Sand-Man*, is an effect of multiple decodings of the sign: a privileged one, (the 'Heimlich,' the homey, familiar reading of an event) and a hidden one, discomforting, menacing, alien. The transitivity between a reliable decoding of a familiar sign and its destabilization (when a human figure in motion is revealed to be a doll, or a corpse; when a shadow takes a human form; when speech is distorted back into noise) is more than a simple matter of suspense—it is the latency of interpretation that triggers the experience. Ernst Jentsch, in his original essay, plainly observes that "the effect of the uncanny can easily be achieved when one undertakes to reinterpret some kind of lifeless thing as part of an organic creature, especially in anthropomorphic terms, in a poetic or fantastic way" (Jentsch, 1997: 13-14). For both Freud and Jentsch, it is the undecidability of interpretation which maintains the sense of the uncanny.

If we accept this multivalence in decoding as characteristic of the uncanny, then the ludological uncanny can spread this multivalence across the range of semiotic registers of the game,

according to the ordering of those registers by the modes of gameplay. Thus, in navigation, a player will see the filament burn blue: this presages a shift to another mode, that of the camera (a player-directed mode switch); the space which was traversed when the filament began to glow becomes a space for a new, undetermined reading in the fictive register, which will be defered until the operation of manipulating the camera is completed. If a cutscene then begins, then the operational decoding is the one which is repressed, while the supernatural explanation for the event in the fictive register enters the foreground.

If the filament glows red, then the next, presaged mode is likely to be combat (flight is almost impossible in this game except as a temporary delaying tactic); conventional methods of foretelling may have already imbued the now-visible visage of the attacker with a reading in the fictive register, yet contemplation of this element must be deferred and repressed all the more until the threat is dispatched.

Obstacles and limitations

This article concludes with some more thoughts on the methodology that is being developed here: the use of software to create intermediary objects that stand in-between the Ding an sich of the game (the thing-itself of an uncomputed string of binary information, whether algorithm or digitized asset, undischarged and not-yet-ordered by the act of play, existing only as a field of potentialities) and the player. The techniques being used to extract visual and computational information from play-sessions and the software itself are still being developed. We intentionally restrict ourselves to the range of information available to the player's senses: we are not mining the "back-end" of the data or parsing databases (except as they are also available to a player.) There are still elements of the player experience that we do not yet analyze, however: the auditory modalities being the most important of them. Nor are we yet transcribing the responses and reactions of the player through the controller. Capturing player activity is interesting insofar as the player conspires in the generation of meaning over time in the practice of play: obtaining information about the emotional, psychological or physiological state of the player, however, is conceptually outside of our goals.

A more essential limitation of our procedure is also, in some sense, a strength. It is not a fully automatable process. Currently, the process of identifying game-play modes is mostly manual, as the player-researcher identifies the modes of play, collects exemplary frames, and then begins tracking the stable and dynamic signifiers in regions of interest for changes. Supervised machine learning will improve this process and reduce error rates. We do not foresee the development of a set of completely game-agnostic mode-classification and mode-detection algorithms, nor would we want to; despite the use of methodologies of visual and computational analysis adapted from the natural sciences, our perspective is phenomenological, not positivistic.

References

Bogost, I. (September 3, 2009). *Videogames are a mess*. Retrieved 11/25/2009.
<http://www.bogost.com/writing/videogames_are_a_mess.shtml>.

Dolar, M. (1991, October). "I Shall be with You on Your Wedding-Night": Lacan and the
Uncanny. *Rendering the Real,* Vol. 58, 5-23.

Egenfeldt-Nielsen, S, Heide Smith,, J., & Pajares Tosca, S. (2008). *Understanding Video Games:
The Essential Introduction*. London: Routledge.

Frasca, G. (November, 2003). "Ludologists Love Stories, Too: Notes from a Debate that Never
Took Place." *Digital Games Research Conference* 2003. Utrecht, 4-6.

Freud, S. (1955). The "Uncanny. *The Standard Edition of the Complete Psychological Works of
Sigmund Freud*. James Strachey (Ed). London: Hogarth, 217-256.

Jentsch, E. (1997). On the Psychology of the Uncanny. *Angelaki: Journal of Theoretical
Humanities,* Vol. 2.1, 7-16.

National Institutes for Health, and Wayne Rasband. (2009). ImageJ: Image Processing and
Analysis in Java. Version 1.42o. <http://rsb.info.nih.gov/ij/>

Perron, B. (2004, May 13-16). Sign of a Threat: The Effects of Warning Systems in Survival
Horror Games. *COSIGN 2004: Computational Semiotics for Games and New Media*.
University of Split, Croatia.
<http://www.cosignconference.org/downloads/papers/perron_cosign_2004.pdf>.

Taylor, L. N. (2006). Cameras, Radios, and Butterflies: The Influence and Importance of Fan
Networks for Game Studies. *Fibreculture*, 8 . Retrieved 2/4/2009
<http://www.journal.fibreculture.org/issue8/issue8_taylor.html>.

---

[1] This should not be taken as a criticism of the program; DiGRA's charter is interdisciplinarity. See Egenfeldt-Nielsen et al 2008, pp 9-10 for a discussion of the various disciplinary positions in re: videogame studies.

[2] Attempts to contact the player were unsuccessful; according to his now-defunct web site, apparently he/she lives in Singapore.

[3] This segment, from a long game-session by Laura Hoeger, was fairly typical for the student players. The cinematic mode dominated the "speed-run" player's modal ratio for an obvious reason: it is the one mode which cannot be shortened by optimal play. This produces the superficially counter-intuitive result that the content of a speed-runner's game traversal transcription will be largely cinematic, rather than interactive.