

Puzzling

Tomás Laurenzo and Ernesto Rodríguez

Instituto de Computación, Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay

Abstract

In this paper we present an augmented-reality installation that allows its users to interact with a deconstructed cinematic space. The installation presents some scenes of pre-selected and well-known movies, mapping different portions or aspects of each frame to tangible blocks of wood tagged with fiducial markers. Users are presented with an augmented cinematographic space that subverts their role within the cinematic experience and introduces new possibilities, playful and narrative, while –at the same time– providing them with the likes of a non-linear edition tool, whose products are ephemeral showing the creative aspect of playing. In its construction we applied direct-manipulation –one of the fundamental HCI concepts– aiming to facilitate users’ engagement.

Categories and Subject Descriptors (according to ACM CCS): H.5.1 [Multimedia Information Systems]: Artificial, augmented, and virtual realities.

1. Introduction

Augmented Reality (AR) –also sometimes called Mixed Reality– stands for the juxtaposition of virtual, or computational, images and *real* ones –that is, acquired by some mechanism whose transduction we have learned to ignore–, and providing some interaction scheme for its users.

AR techniques rely strongly on this assumption of reality – what Vilem Flusser called “the technical image” [Flu85]– and also on the self-describing computational interfaces and interaction styles that the system proposes with the addition of virtual images.

Rendering synthetic images over real-time acquired images can be also seen as type of multichannel cinema (picture-in-picture), which adds in the notion of spatial montage to the traditional temporal montage, bringing to cinema the aesthetics of other realms such as video games or military graphical interfaces [Rho05].

Videogames, for instance, do not use film-like cuts as a narrative resource (except, perhaps, in the non-playable cinematic sequences that sometimes are included for storytelling and use the traditional language of film [Bar05]), but rather use multiple simultaneous channels.

These parallel channels, like the borrowed-from-the-military first person shooters’ head-up displays, superimpose game information (e.g. the character’s energy left) improving data bandwidth, playability and providing new ways of interactivity; although sometimes also hurting the players’ suspension of disbelief and their immersion into game play [Wil06, Bar05].

It is worth noting that similar resources have been used – very succinctly– in cinema, being an easy example the common split-screen telephone conversations [Rho05].

There are also examples of AR projects that focus on creating or re-creating traditional film narrative with a twist. Examples of this are MacIntyre, Moreno *et al*’s AR setups aiming to immerse spectators *into* the narrative by changing their point of view and providing an interactive approach to the film [McI02] [Mor01].

Puzzling, our installation proposes the exact opposite to these experiments: in it, the traditional cinematic material provides the computational imaginery that is rendered onto the “real” images.

There exists at least one previous experiment on the same vein: G.A. Rhodes’ *52 Card Psycho* [Rho08], where the 52 individual shots of the shower scene of Hitchcock’s *Psycho* (1960) are superimposed onto a deck of 52 cards imprinted with a unique fiducial identifier.

Rhodes’ setup, while inspiring, does not question the basic cinematic unity –the frame–, and provides only one interaction scheme with the art piece, relying strongly in the incongruence provided by the mundane gestures of card playing. This association (between cards and shots) does allow “a material interaction with the ‘cinema screen’”, where users can create “spreads of time”, recomposing the cinematic flow.

We find particularly interesting that 52 Cards Psycho lays the 52 scenes in unison, allowing users to create spatial compositions that show the chosen scene simultaneously. Yet, this maintains the frame as basic unit, while losing the narrative’s intrinsic dimension of time.

Our work, on the other hand, attempts for a deconstruction of the film’s frame, allowing for its deconstruction and montage (both temporal and spatial), also positioning itself both as a direct manipulation tool of filmic material *and* a video game of sorts, questioning spectators’ role, offering them the possibility to choose *how* to relate with the installation and the filmic material.

In the rest of this paper we will present our installation. We will start with some concepts that drove our work, then we

will present its setup and we will conclude with some of the technical aspects of its implementation.

2. Direct manipulation

Direct manipulation is an interaction style, which stands for interactive systems with continuous representation of the domain of interest, with rapid, reversible, incremental actions and continuous feedback. Direct manipulation allows users to *feel* that they are operating directly with the objects presented to them with a direct representation of the domain of interest. [Nor88]

A direct manipulation system proposes a world metaphor [Tre94], where the task domain is explicitly represented and the interaction language is implicitly coded in this metaphor. The orders users give to the system and their responses are implicitly coded in this metaphorical world.

The alternative is to offer a conversation metaphor, where the interaction language becomes explicit and the task domain is mediated by descriptions in the conversation's language (e.g. a command line interface). Thanks to its verbal abstraction of the users' domain of interest, a conversation metaphor can provide a more powerful language of expression to its users, at the cost of a steeper learning curve.

When designing a direct manipulation based interaction, the representation of the task domain (the world) is a most important aspect of the systems appeal and usability. When users are not familiar (perhaps because it is a novel system) with the metaphorical representation, the design of this virtual world is key to the system's success.

Because of the difficulties of providing a well-designed world (both in terms of the objects presented and their interactive behaviors), some domains –live video editing– have not been benefited by the multiple advantages of direct manipulation.[Dra08].

that when the matrix F is symmetric, it is possible to reformulate the iterative formulation of the radiosity equation in terms of the eigensystem of the form factor matrix. Finally, Ashdown generalizes the results of Di Laura and Franck to any F .

3. The installation

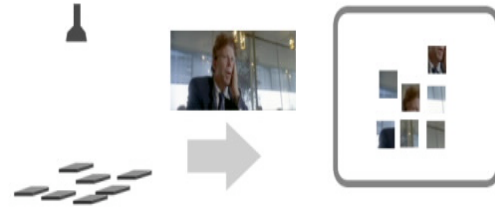
Puzzling is a small and simple Augmented Reality art installation, live cinema performance tool and video game.

Its base idea is to deconstruct the frame –in augmented reality space– to equally sized portions, with each bit superimposed onto tangible pieces, each one tagged with a unique identifier, (see Figure 1), creating a jigsaw puzzle of sorts.

Users are presented with the tagged pieces on a table (its number can vary, but we usually use nine, as we found it

provides the best results), which is viewed by a mounted zenithal camera.

Accompanying the table is a real-time video projection, in front of the users, showing the *screen space*, that is, the images acquired by the camera with the synthetic images



superimposed (see Figure 2).

Figure 1: The camera acquires the pieces' positions and the system then renders the augmented image with the decomposed frame.

In the screen space the Augmented Reality is shown: there coexist the table, the portions of the frame (of course, still playing), and the hands of the users, a hybrid space appears where the direct manipulation of the cinematic material is possible.

3.1 Interaction

Users can interact with the cinematic material in different ways. First of all, and probably most important, is the spatial reconfiguration of the frame: by just moving the wooden pieces the frame disposition changes allowing for new configurations. It is possible, also, to map more than one frame (or portions of different frames) into the pieces enabling users to create new compositions, distorting the existing or creating new narratives.

But *Puzzling* also allows for the control of different parameters of the reproduction. For example, it is possible to alter the speed of the video reproduction rotating the pieces. The frame then becomes unsynchronized, with portions speeding –and soon showing images of the future– or ralenting, showing images of the past of the film.

We found interesting one of the setups, where the image cannot be spatially coherent (in the sense of recreating the original disposition) and synchronized at the same time. If the geometry is recreated, then the frame is temporally disonant and viceversa. This creates a tension for the users to resolve: what is to be favored, temporal or spatial coherency?

3.2 Interaction modes

Puzzling, with its very simple underlying idea, offers a very rich interaction scheme, with many possible gestures

and effectively creates the idea of direct manipulation of the filmic material.

This approach to using *Puzzling* is very similar to live cinema's one: the real-time editing and composing of pre-



We have found that depending on the users' intention when using the installation, it becomes a game *or* a performance tool (and, stretching the definition it can be seen as a film analysis or edition tool).

The setup can suggest the mode of interaction. For example, if *Puzzling* is to be used as a game where the users' goal is to recompose the frame (again, like an animated jigsaw puzzle), *Puzzling* can add a visual clue when two or more pieces are together in their correct position (see Figure 3. In the right, one of the three fragments is tinted showing that its position is not correct, the other two are rendered with their original colors).

When *Puzzling* is used as a game, users focus almost exclusively on the detection of the internal coherence of the frame, easily forgetting its narrative possibilities as the actual content only matters to them as information to solve the task.

It is worth noting that even when the frame is decomposed in a few pieces (our setups usually use nine), the task of rearranging the frame can be very difficult.

Other similar games can be easily imagined and added to *Puzzling*, we have experimented with tasks that comprise using several frames, arranging them spatially and chronologically, or tasks where users need to find the similarities between different scenes.

In the first case we used footage from *2001, A space odyssey*, and on the second one we used material from *Short Cuts* (see next section for a discussion on the filmic material).

But the installation can be an expression tool (or a toy, i.e. a game with no intrinsic goal, apart from enjoying using it), where the spatial and temporal arrangements of the cinematic material become a new way of expression, a narrative orthogonal to the film's intrinsic one.

recorded footage becomes a visual art performance.

4. Cinematic material

While –as it is inferred from the previous section–, *Puzzling* –as a tool– can be used to play with any cinematic material, our installation uses two referential movies: Stanley Kubrick's *2001: A Space Odyssey* and Robert Altman's *Short Cuts*.

2001: A Space Odyssey.

Stanley Kubrick's 1968 iconic film, considered by many as one of the best films ever made [BFI02], contains several scenes instantly recognizable by a great amount of people.

Partially responsible for this recognizability are the film's extremely rich visual language, and distinctive pace and music. In effect, its visual richness is so compelling that only few scenes are actually driven by dialogues.

This distinct visual density, when deconstructed in *Puzzling* give its users strong cues on how to reconstruct the original frame, while at the same time allowing for new configurations (both spatial and temporal) that effectively turn *Puzzling* not only into a fun game but also into the aforementioned expression instrument.

While some of the scenes are extremely easy to reconstruct, some others –like those depicting weightless sequences– lack, for example, directional cues (there is no up or down), allowing for different semantically sensible reconstructions.

Our main criterion for choosing scenes was to focus on those displaying geometrical patterns (e.g. the early men's scene on the first act), specifically those that allow for multiple different, yet sensible, visual configurations.

Short Cuts

Where *2001...* provided *Puzzling* with slow paced narrative that somewhat encouraged its users to search for new in-frame configurations, Robert Altman's *Short Cuts* (1993) narrative speed and parallel storytelling, when deconstructed with *Puzzling*, urge its users to re-combine and create new lectures of the stories or even new stories altogether.

Short Cuts "consciously refuses to single out any particular character" [Azc05], showing them all –twenty-two arranged in nine groups–, and their (independent but sometimes intertwined) stories with equal importance.

These stories are presented in an already deconstructed style: hundreds of segments ending with a cut ("never a dissolve, never a fade to black, never a wipe. Cuts only" [Bal96]), linking the themes from the stories, which –most of the time– are concurrent.

But what makes *Short Cuts* extremely apt for *Puzzling*, are the techniques used by Altman to interlink the stories. In the words of David Balcom [Bal96] "The film offers associations between characters, and between mini-sequences, that an active viewer can construct along with the stories they see on the screen. There are subcurrents flowing in the text that link thematic material throughout the film. In fact, most themes are linked in more than one direction (from material that preceded it, and toward material that later recalls it), and to more than one theme." Several scenes reappear in different forms providing formal links between the stories, with different characters in similar positions or situations, providing "a tremendously rich body of associations at work in the film that simple narrative analysis cannot account for".

But perhaps *Puzzling* users can.

4. Implementation

The installation consists of the following elements: a table where users can manipulate the pieces while looking at the output video, a number of tagged wooden pieces, a camera that views the table and a projector or screen that shows the augmented video to the user and the audience.

Puzzling is, basically, a work of software. All the hardware used (apart from the pieces of wood), is standard, off-the-shelf, only requiring a webcam and a reasonably powerful computer.

The installation's software is platform independent, being able to run in any modern operating system (it has been deployed on both Windows and Mac OSX).

We coded it using the software framework OpenFrameworks [Ope08]: a collection of libraries that help enormously with many common tasks of multimedia applications.

The recognition of the pieces was implemented using the ARToolitPlus library, "a software library that can be used to calculate camera position and orientation relative to physical markers in real time" [Art08].

This physical markers, commonly called fiduciary markers or fiducials, are patterns that can be easily identified and, when attached to objects, allow for the objects recognition and tracking.

It is important to note that both the camera and the light sources need to be put in adequate positions and correctly calibrated or the Augmented Reality illusion will (mostly) go wrong.

Puzzling's software cuts the footage into a given number of equally sized pieces, and automatically maps each piece to a marker that appears on screen.

When a marker is identified, the system overlays the corresponding video, applying the same geometrical transformations recognized in the marker, completely substituting it in screen space.

4. Conclusion

Using *Puzzling* to manipulate the cinematic space shows two possible levels of representation of information, that is, two different and disjoint modes of user engagement; but in both cases, the direct manipulation of the footage questions the relationship between the viewer and the cinematic material.

In effect, the relationship between users and the narrative changes drastically, in one extreme, with users taking a playful role, the cinematic narrative is an accessory that provides coherence to the deconstructed bits of the film. The narrative itself is not important except for its cognitive density: it hints on the relationships between the different blocks.

Users are never part of the narrative action, the suspension of disbelief does not occur; the so-called fourth wall keeps all its strength.

But equally important is the approach of considering our installation as an ephemeral non-linear, spatially distributed tool.

Users' goals change. Not being playful, they can relate to the film's content and interact with it. They are no longer spectators but became part of the film: their role and the film's role change.

And the interaction's product is not more important than the interaction itself, its aesthetic: *using* the installation becomes an artistic performance, editing –in a way that we believe would satisfy Kuleshov's idea of edition's role– *becomes* the art fact (which is also consonant with Live Cinema's proposal).

But this art fact is an ephemeral one: each performance, each playing with our installation –our tool– could be unrepeatable and last only in the memory of the performers or their audience.

References

- [Flu85] FLUSSER V.: *Philosophy of Black box*, Sao Paulo Cultural Center, (1985).

- [Rho05] RHODES G.: Metonymy in the Moving Image: Multichannel Cinema, Master Thesis, State University of New York, Buffalo, (2005).
- [Bar05] BARWOOD, H.: Cutting to the Chase: Cinematic Construction for Gamers, Game Developer's Conference, (2000).
- [Wil06] WILSON G.: Off With Their HUDs!: Rethinking the Heads-Up Display in Console Game Design, Gamasutra, (2006).
- [McI02] MACINTYRE B., BOLTER J., VAUGHAN J., HANNIGAN B., MORENO E., HAAS M., GANDY M.: Three Angry Men: Dramatizing Point-of-View using Augmented Reality. In SIGGRAPH 2002 Technical Sketches, page 268, San Antonio, TX, (2002).
- [Mor01] MORENO, E., MACINTYRE, B., BOLTER J.: Alice's adventure's in new media. GVU Center, Georgia Institute of Technology, Atlanta, GA, (2001).
- [Rho08] RHODES G.: 52 cards psycho. International Symposium on Electronic Arts, Singapore (2008).
- [Nor88] NORMAN, D.: The psychology of everyday things, USA: Basic Books Norman, D. (1988) The psychology of everyday things, USA: Basic Books
- [Tre94] TREGLOWN, M.: Qualitative models of user interfaces. People and Computers: Proceedings of HCI '94, Glasgow, (1994).
- [Dra08] DRAGICEVIC P., RAMOS G., BIBLIOWICZ J., NOWROUZEZAHRAI D., BALAKRISHNAN R. , SINGH K. Proceedings of CHI 2008.
- [BFI02]
<http://www.bfi.org.uk/sightandsound/topten/poll/critics.html> "Sight and Sound: Top Ten Poll 2002". British Film Institute web site. Retrieved on 2006-12-15.
- [Azc05] "Making Sense of a Multi-Protagonist Film: Audience Response Research and Robert Altman's Short Cuts". *Miscelánea. A Journal of English and American Studies*, 32,2005, 11-22.
- [Bal96] "Short Cuts", Narrative Film and Hypertext".
http://www.mindspring.com/~dbalcom/short_cuts.html
 1996. Accessed: May 25, 1998
http://www.mindspring.com/~dbalcom/short_cuts.html
- [Ope08] <http://www.openframeworks.cc>
- [Art08]http://studierstube.icg.tugraz.ac.at/handheld_ar/artoolkitplus.php