D 5.1

SPECIFICATION OF CINEMATIC IDIOMS

Project Number	FP7-ICT-231824
Project Title	Integrating Research in Interactive Storytelling (NoE)
Deliverable Number	D5.1
Title of Deliverable	Specification of Cinematic Idioms (Revised Version)
Workpackage No. and Title	WP5 - Cinematography
Workpackage Leader	UNEW
Deliverable Nature	Report
Dissemination Level	Public
Status	Final
Contractual Delivery Date	30th June 2009 (original submission)
Actual Delivery Date	15th July 2010 (revised submission)
Author(s) / Contributor(s)	Guy Schofield (UNEW), Patrick Olivier (UNEW) Marc Christie (INRIA)
Number of Pages	65

DOCUMENT HISTORY

Version	Date	Comment	Author(s)
1.1	August 30, 2009	Draft version for approval	UNEW: Guy Schofield and Patrick Oliver; INRIA: Marc Christie
1.2	July 15 2010	Revised version for approval	UNEW: Guy Schofield and Patrick Oliver; INRIA: Marc Christie

LIST OF ABBREVIATIONS

ICT	Information & Communication Technology
IS	Interactive Storytelling
FOV	Field of view
POV	Point of view
LOI	Line of interest

TABLE OF CONTENTS

Executive Summary

In order to fulfil the objectives of WP5, a thorough understanding of cinematography is required. An example scenario: a scene from the film 1984 (Radford, 1984) is used to demonstrate how certain cinematographic techniques are used by film directors to support narrative.
Beginning with standard idioms for placing and moving cameras and continuity editing, we examine a range of approaches both independently and with reference to specific genres. Each is discussed both in context and with reference to is possible implementation in IS.

Soundtrack is also discussed both in terms of film composition conventions and how these might be applied to IS contexts. A number of standard techniques from game design for implementing interactive musical scores are examined.

It is important to note the impact of changes in context brought about by transplanting approaches and techniques from cinema to IS. The possible effects of this are discussed with particular reference to how cinematic techniques may lose or alter their meaning in IS contexts.

1.0 Cinematography and narrative

In narrative film, each sequence of shots must not only be legible in terms of spatial continuity but must support and expose the greater narrative, allowing the film to be read as a whole. In the case of Interactive Storytelling this is complicated by the possibility of changes occurring within the narrative as it happens, meaning considerations of pace, continuity and the legibility of settings, characters and plots become ever more important.

Existing texts such as Arijon's *Grammar of the Film Language* document cinematographic idioms and allow for the construction of sequences of shots adequately describing spatial relationships between characters in a scene. Arijon also documents how effects such as the prominence of individual characters in the narrative might be emphasised through framing and camera placement. However film-makers also use camerawork and editing to establish more complex narrative relations between characters and the worlds they inhabit.

The specifications discussed in this document assume a certain relationship between action and the camera that differs from the approach of many directors. Specifically, it is assumed that the dialogue and staging is under authorial control and is therefore independent of the camera and editing. The main cinematographic devices we review are specified in terms of canonical camera properties describing the content of the screen, the camera parameters and relations to targets. This level of specification relies on a formalized cinematographic language that enables the transformation of these properties into constraint and optimisation-based systems.

The solutions described here focus mainly on a continuity approach to editing, prioritising spatial and temporal legibility over the aesthetic or psychological effects afforded by alternative approaches such as montage.

In this report we take the "Canteen Scene" from Michael Radford's film 1984, and extract and re-state the principle cinematic idioms in terms that emphasise the contribution of the cinematography (principally camerawork and editing) to the narrative goals of the scene itself.
The presentation is an attempt to restate Arijon's mechanistic presentation of shot collections into idioms (many of which cannot actually be achieved with standard camera set-ups) in terms of the more subtle means by which narrative progression is achieved in Radford's film. For this scene traditionally strong notions of "action" that are the focus of most research in interactive storytelling are mostly incidental. The subtleties of the characters' interactions with each other and the environment are the principal means by which the relational and emotional information in the scene is conveyed and consequently the cinematography (choice of shots and edits) is the director's vehicle for achieving this.

A further section included is a brief survey of contemporary directorial style, covering a number of significant innovations in camera movement, editing and post-production techniques and their possible application to IS with special reference to their meaning in narrative contexts.
We also include a discussion of the relationship of soundtrack to the moving image and examine a number of techniques which might be used to support narrative in IS systems, included a number of possible implementations.

2.0 Basic cinematographic devices

The techniques specified in this section may be considered a set of components that can be assembled to bring a story to the screen in a way that allows the viewer to easily understand the spatial relations/context of each scene, and the sequence of actions occurring within it. Included are idioms for dealing with camera placement and movement in a variety of settings involving two or three main characters in dialogue – our goal is to characterise these sufficiently to allow us to articulate these in the form of constraints and objectives to be solved by the automated camera control framework. For this purpose, we first propose a description of the canonical cinematographic properties we employ to specify the idioms.

2.1 Identification of canonical properties

In the following section we propose a set of cinematographic properties to precisely specify a shot in terms of composition (layout of elements on the screen), relations between the camera and the targets (distance, high and low angles) and camera motions. This set has been carefully designed to encompass and express most cinematographic properties proposed in the domain (see (N Halper, 2000), (P Burelli, 2008)and (W Bares, 2000)).

Regardless of the tasks and the application, the control of a virtual camera relies on the maintenance of properties over the screen, on the camera and on the camera path. Properties are indexed by time expressed either as a fixed value (t) or a range (t0-t1). In the fixed case, the property should hold at the specified time value t. In case of a range of values, the property should hold for all instants between t0 and t1. The symbol '|' in the parameters stands for a disjunctive OR operator, meaning either value can be employed. A number of properties are specified over a target entity (denoted by **ent** in the following table). This entity stands either for an object, a group of objects, a character or a group of characters.

Property	***Description***
`Orientation(ent, view, time)`	constrains the camera to view given entity **ent** from the specified angle **view** (eg. front view). The angle is relative to the intrinsic orientation of the target entity, and **view** is either a single value, or a range of angles, or a range with a preferred value angle. Aliases have been set for common values, such as FRONTVIEW, REARVIEW, LEFTVIEW, TOPVIEW, THREE_QUARTER_FRONT_LEFTVIEW,…
`Size(ent, size, time)`	constrains the projected area of a given entity **ent** on the screen where **size** is percentage of the screen area, or a range of percentages, or a range with a preferred value. The property considers the entity size regardless of it being (partially) occluded by other entities, and considers only the area projected onto the screen (in the case the frame cuts the entity).
`Width(ent, width, time)`	constrains the projected width of a given entity **ent** on the screen where **width** is a percentage of the

	screen width, or a range of percentages, or a range with a preferred value. The property considers the entity width regardless of it being (partially) occluded by other entities, and considers only the width projected onto the screen (in the case the frame cuts the entity).
Height(ent, height, time)	constrains the projected height of a given entity **ent** where **height** is a percentage of the screen height, or a range of percentages, or a range with a preferred value. The property considers the entity width regardless of it being (partially) occluded by other entities, and considers only the width projected onto the screen (in the case the frame cuts the entity).
Framing (ent, in | out | percentageIn, frame, time)	constrains the projected entity **ent** to be **in**side, **out**side, or for a given percentage **in**side a rectangular frame **frame** defined in screen coordinates (defined between 0 and 1 on both axis, starting at top left corner). The frame can partly lie outside the screen. The property considers the entity position regardless of it being (partially) occluded by other entities. For convenience, two default values for parameter **frame** are provided: RIGHTSCREEN and LEFTSCREEN that respectively represent the whole rightside of the screen (0.5 to 1.0) and the whole leftside (0.0 to 0.5)
RelativeSpatialLocation (ent_1, ent_2, right| left| above| below| infrontof| behind, time)	constrains the projection of entity **ent_1** to appear **right** (or left, or above, ...) the projection of entity **ent_2**. For example, **left** relation is true whenever the leftmost geometric extent of **ent_1** is left of the leftmost geometric extent of **ent_2**, and the rightmost geometric extent of **ent_1** is left of the rightmost geometric extent of **ent_2**.
ScreenSeparation (ent_1, ent_2, sep, time)	constrains the projected entities **ent_1** and **ent_2** to be separated by a distance, or a range of distances, or range plus a preferred value. The distance is measured by the minimum distance between the geometric extents of both entities.
DistanceToCamera(ent, dist, time)	constrains the camera to be placed at distance **dist** from entity **ent**, where **dist** is a value, or a range of values, or a range with a preferred value.
Occlusion(ent, yes | no | percentage, time)	constrains the entity **ent** to be occluded (**yes**) | not occluded (**no**) | occluded for the given percentage of its surface (**percentage**).
InView(ent, yes | no | percentage, time)	constrains the the entity **ent** to be fully projected onto the screen (**yes**), out of the screen (**no**), or for a given percentage of area onto the screen (**percentage**).

`FOVAngle(value, time)`	constrains the camera to the specified aperture **value**, range of values or range with a preferred value.
`Travelling (horizontal|vertical|any, time)`	constrains the camera motion to a linear translation movement either **horizontal**, **vertical** or along any axis (**any**).
`Panoramic (horizontal|vertical|any, time)`	constrains the camera motion to a rotational movement, either around its yaw axis (**horizontal**), its tilt axis (**vertical**), or around any angle (free rotation).
`Arcing(horizontal|vertical|any, ent, time)`	constrains the camera to an arcing motion (rotation of the camera around the center of entity **ent**). Rotation is either **horizontal**, **vertical** or around **any** axis.
`Zoom(In|Out, time)`	constrains the camera to a zoom motion (a change in the camera aperture), by either increasing the aperture angle (**Out**) or decreasing it (**In**).
`Roll(Left|Right, time)`	Constrains the camera to a **Left** or **Right** rotation around its roll axis.

2.2 Framing

As described by Arijon and Katz, particular framings of characters can be used to cover specific situations. Close-ups allow a character's facial expression and eyes to be emphasised whereas medium and full shots allow physical actions or attitudes involving the whole body to be read. To avoid aesthetically awkward compositions, the body is usually cut below the armpits, chest, crotch, waist or knees. In specifying the shots (see Fig. 2), we introduce some terms to identify the parts of the body we need to constrain: **head**, **headtorso** (from head to torso), **headknee** (from head to knee). The term **body** describes the entire character.

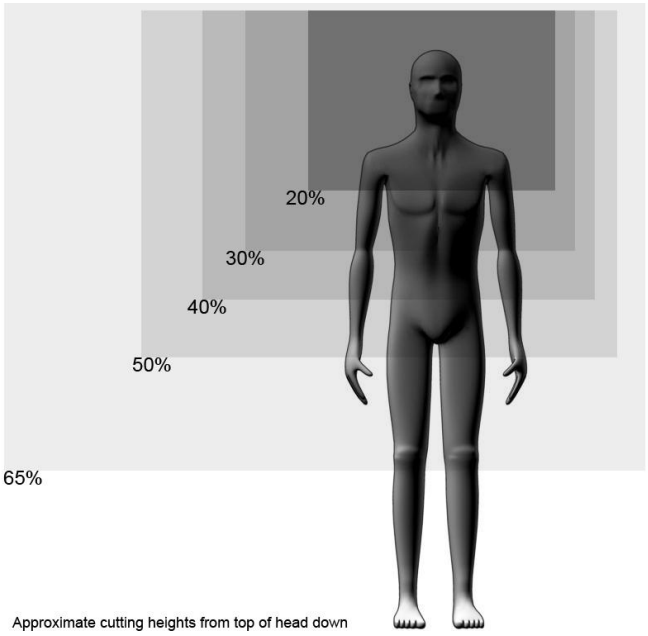
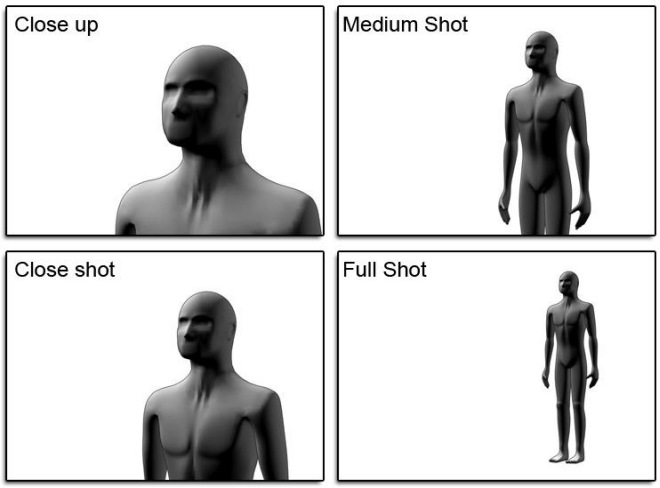


Figure 1 : Conventional cutting heights

Beyond fulfilling requirements such as legibility of physical attitude, facial expression etc. exact frame selection is largely a matter of style, with different directors favouring certain framings. In 1984, Radford uses close-ups for most dialogue shots, cutting to medium or close shots for sequences involving movement. Other intermediate or extreme shots are proposed by the literature (Medium Close Up, Extreme Close Up, Extreme Long Shot).



Close up	`Height(head, [40-60]%, t);`
	`InView(head, yes, t);`
	`Occlusion(head, no, t);`

Close Shot	`Height(headtorso,[60-80]%, t);`
	`InView(headtorso, yes, t);`
	`Occlusion(headtorso,no,t);`

Medium Shot	`Height(headknee,[80-90]%,t);`
	`InView(headknee, yes, t);`
	`Occlusion(headknee,no,t);`

Full Shot	`Height(body, [80-90]%, t);`
	`InView(body, yes, t);`
	`Occlusion(body, no, t);`

Figure 2: Conventional framings for human characters

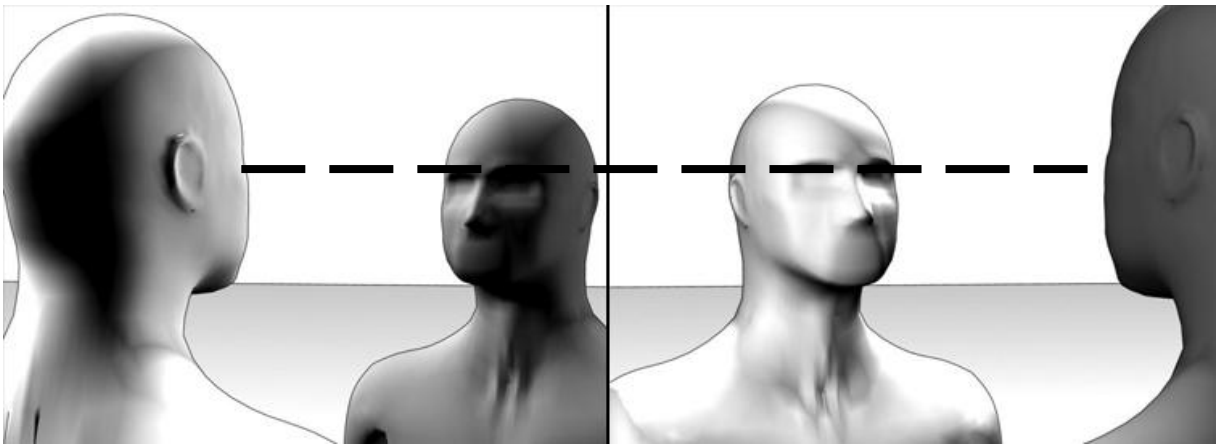
Figure 3: Typical close-up framing in 1984

2.3 Composition and balance

Two characters engaged in dialogue with each other are typically represented in paired shots. A character's speech is accompanied by their image, followed by their opposite number's reaction. In considering the composition of possible shots, continuity and continued legibility of the players' relative positions in space is vitally important. Mirroring or the selection of shots of roughly opposite composition is a common device used here. Set-ups where the character occupies an equal proportion of the frame in each shot establishes each character in space without making either seem dominant.

The exact composition of dialogue shots is a often matter of individual directorial style rather than any aesthetic formalism, however as a guideline, in Radford's 1984, each character in the canteen scene dialogue occupies roughly 40% of the frame.

Figure 4: Smith and Syme occupy opposite portions of the frame mirroring each other.



Left shot	`Height(head1, [30-50]%, t);`
	`Occlusion(head1, no, t);`
	`Height(eye1, [0.3-0.4], t);`
	`Orientation(head1, FRONTVIEW, t);`
	`Framing(head1, RIGHTSCREEN, t);`
	
	`Occlusion(head2, no, t);`
	`Height(eye2, [0.3-0.4], t);`
	`Orientation(head2, THREE_QUARTER_RIGHT_REARVIEW, t);`
	`InView(head2, [90-100]%, t);`
	`RelativeSpatialLocation (head1, head2, right, time);`

Right Shot	`Height(head2, [30-50]%, t);`
	`Occlusion(head2, no, t);`
	`Height(eye2, [0.3-0.4], t);`
	`Orientation(head2, FRONTVIEW, t);`
	`Framing(head2, LEFTSCREEN, t);`
	
	`Occlusion(head1, no, t);`
	`Height(eye1, [0.3-0.4], t);`
	`Orientation(head1, THREE_QUARTER_LEFT_REARVIEW, t);`
	`InView(head2, [30-50]%, t);`
	`RelativeSpatialLocation (head1, head2, right, time);`

Figure 5: Eyelines consistent across reverse shots. In the specification head1 refers to the dark-shaded character, head2 to the light-shaded one

In a dialogue where two or three characters are at roughly the same elevation, maintaining a consistent eye-level across shots helps maintain spatial legibility. In the canteen scene in 1984, eyelines between characters in dialogue in successive shots are within 10% of each other against the vertical axis of the frame.

2.4 Line of interest

It is considered an almost hard-and-fast rule of cinematography, that the line of interest is inviolate and that shot selection must take place along one side or another but never both. Typically, the line of interest connects the gazes of two characters engaged in dialogue. However, in certain circumstances characters engaged in dialogue may not be looking towards each other. The Line of Interest still joins their heads, regardless of gaze.

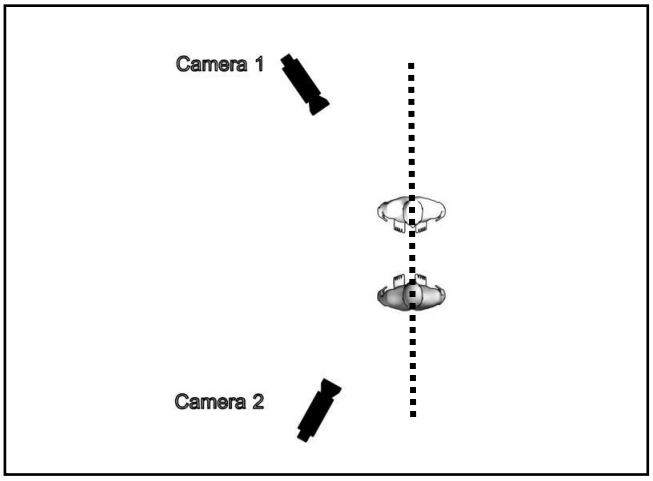


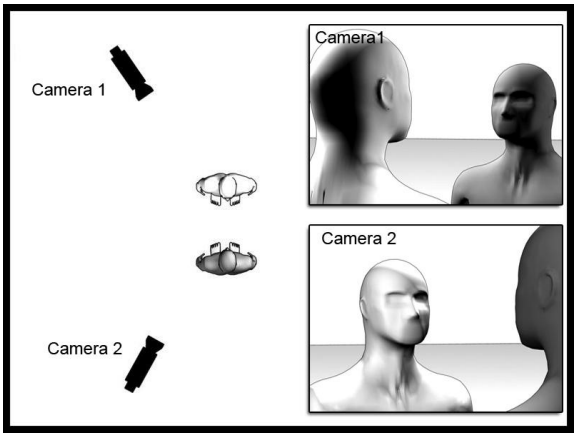
Figure 6: Plan view of a conventional 2 camera setup showing line of interest

2.5 Camera placement for two characters

Commonly used setups for shooting two characters in dialogue are shown below. First and foremost, each shot of a particular idiom reinforces for the viewer the relative positions of the characters.

2.5.1 External reverse shots

The advantage of using external camera positions is that spatial description of the characters takes place in a single shot: the viewer is presented with both characters. However, it is important that in each shot, characters are presented in the same portion of the frame.



Camera1	Height(head1, [30-50]%, t); Occlusion(head1, no, t); Height(eye1, [0.3-0.4], t); Orientation(head1, FRONTVIEW, t); Framing(head1, RIGHTSCREEN, t); Occlusion(head2, no, t); Height(eye2, [0.3-0.4], t); Orientation(head2, THREE_QUARTER_RIGHT_REARVIEW, t); InView(head2, [90-100]%, t); RelativeSpatialLocation(head1, head2, right, time);
Camera2	Height(head2, [30-50]%, t); Occlusion(head2, no, t); Height(eye2, [0.3-0.4], t); Orientation(head2, FRONTVIEW, t); Framing(head2, LEFTSCREEN, t); Occlusion(head1, no, t); Height(eye1, [0.3-0.4], t); Orientation(head1, THREE_QUATER_LEFT_REARVIEW, t); InView(head1, [30-50]%, t); RelativeSpatialLocation(head1, head2, right, time);

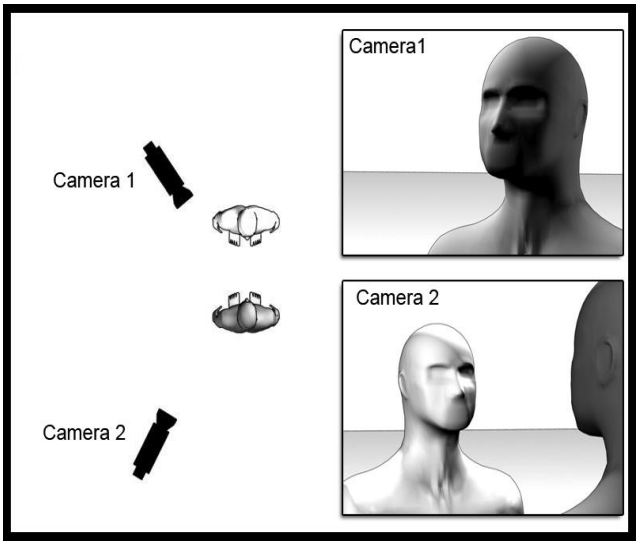
Figure 7: External reverse shots showing framings with their formalization

2.5.2 Internal reverse shots

Internal reverse shots allow characters to occupy a larger proportion of the frame. As the opposite character is not seen, maintaining the eyeline and keeping the camera on one side of the line of interest becomes even more important.

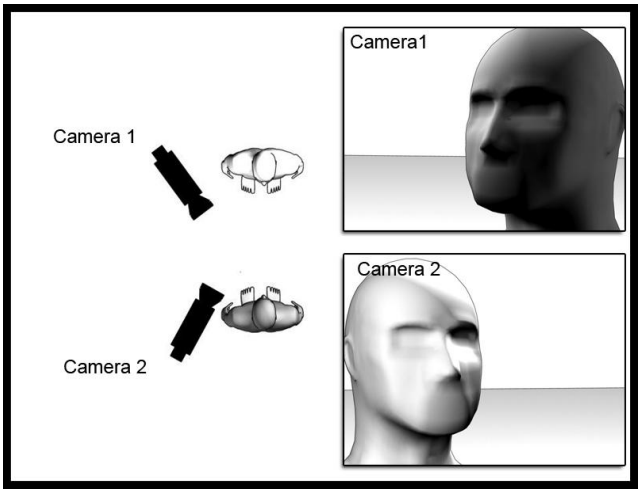
2.5.3 Mixed external/internal shots

Internal and external reverse shots can be used either consistently or alternately. For example, an external reverse shot can be inserted into a sequence of internal reverse shots to re-state the characters' relative positions.



Camera1	Height(head1, [60-80]%, t); Occlusion(head1, no, t); Height(eye1, [0.3-0.4], t); Orientation(head1, THREE_QUARTER_LEFT_FRONTVIEW, t); Framing(head1, RIGHTSCREEN, t); InView(head2, no, t);
Camera2	Height(head2, [30-50]%, t); Occlusion(head2, no, t); Height(eye2, [0.3-0.4], t); Orientation(head2, FRONTVIEW, t); Framing(head2, LEFTSCREEN, t); Occlusion(head1, no, t); Height(eye1, [0.3-0.4], t); Orientation(head1, THREE_QUATER_LEFT_REARVIEW, t); InView(head1, [30-50]%, t); RelativeSpatialLocation (head1, head2, right, time);

Figure 8: Mixed internal and external reverse shots with their formalization

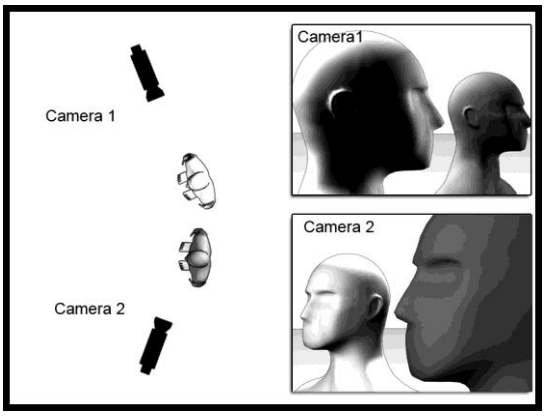


Camera1	Height(head1, [80-90]%, t); Occlusion(head1, no, t); Height(eye1, [0.3-0.4], t); Orientation(head1, THREE_QUATER_FRONT_LEFTVIEW, t); Framing(head1, RIGHTSCREEN, t); InView(head1, [90-100]%, t); InView(head2, no, t);
Camera2	Height(head2, [80-90]%, t); Occlusion(head2, no, t); Height(eye2, [0.3-0.4], t); Orientation(head2, THREE_QUATER_FRONT_RIGHTVIEW, t); Framing(head2, LEFTSCREEN, t); InView(head2, [90-100]%, t); InView(head1, no, t);

Figure 9: Internal reverse shots (and formalization)

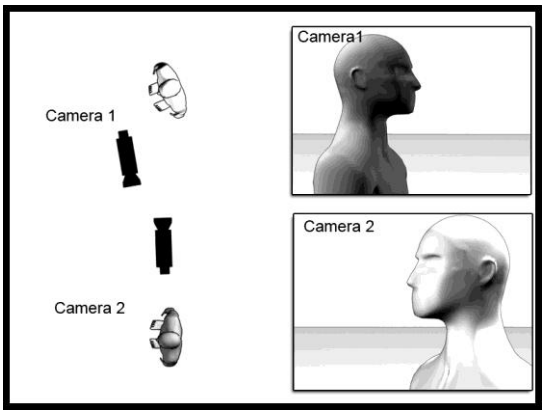
2.5.4 Variations for different configurations

Maintaining the line of interest and characters positions in the frame is vital regardless of the characters exact location and orientation to each other. Side-by-side and L-shaped configurations still require the same methods to maintain continuity.



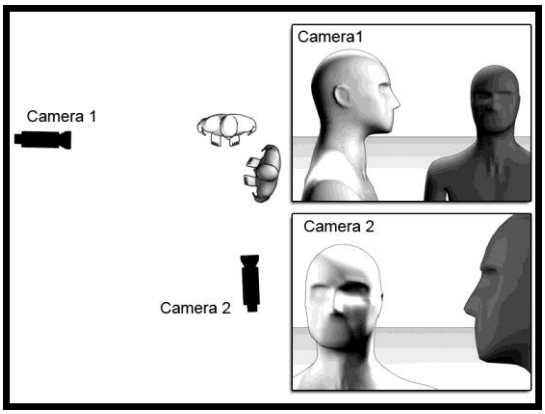
Camera1	`Height(head1, [40-60]%, t);`
	`Occlusion(head1, no, t);`
	`Height(eye1, [0.4-0.5], t);`
	`Orientation(head1, RIGHTVIEW, t);`
	`Framing(head1, RIGHTSCREEN, t);`
	`InView(head1, yes, t);`
	
	`Height(head2, [80-90]%, t);`
	`Occlusion(head2, no, t);`
	`Height(eye1, [0.4-0.5], t);`
	`Orientation(head2, RIGHTVIEW, t);`
	`Framing(head2, LEFTSCREEN, t);`
	`InView(head2, [90-100]%, t);`

Camera2	Symmetric to Camera1 modulo `InView()`



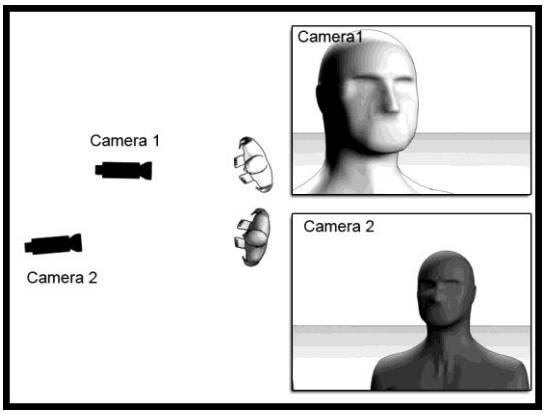
Camera1	`Height(head1, [30-50]%, t);`
	`Occlusion(head1, no, t);`
	`Height(eye1, [0.3-0.4], t);`
	`Orientation(head1, RIGHTVIEW, t);`
	`Framing(head1, LEFTSCREEN, t);`
	`InView(head1, yes, t);`
	`InView(head2, no, t);`

Camera2	Symmetric to Camera1 modulo `InView()`



Camera1	`Height(head1, [30-50]%, t);`
	`Occlusion(head1, no, t);`
	`Height(eye1, [0.3-0.4], t);`
	`Orientation(head1, FRONTVIEW, t);`
	`Framing(head1, RIGHTSCREEN, t);`
	`InView(head1, yes, t);`
	
	`Height(head2, [30-50]%, t);`
	`Occlusion(head2, no, t);`
	`Height(eye1, [0.3-0.4], t);`
	`Orientation(head2, RIGHTVIEW, t);`
	`Framing(head1, LEFTSCREEN, t);`
	`InView(head1, yes, t);`

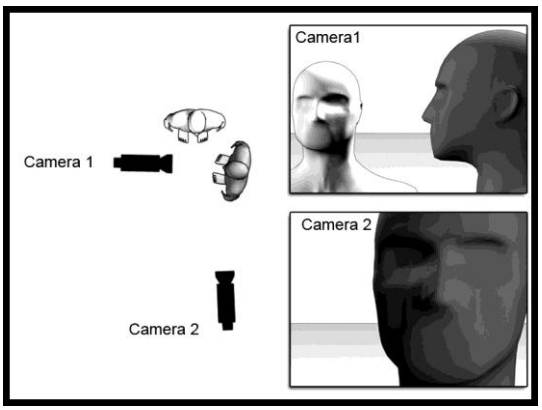
Camera2	Symmetric to Camera1 modulo `InView()` and `Orientation()`



Camera1	`Height(head2, [60-80]%, t);`
	`Occlusion(head2, no, t);`
	`Height(eye2, [0.3-0.4], t);`
	`Orientation(head2, FRONTVIEW, t);`
	`Framing(head2, LEFTSCREEN, t);`
	`InView(head2, yes, t);`
	`InView(head1, no, t);`

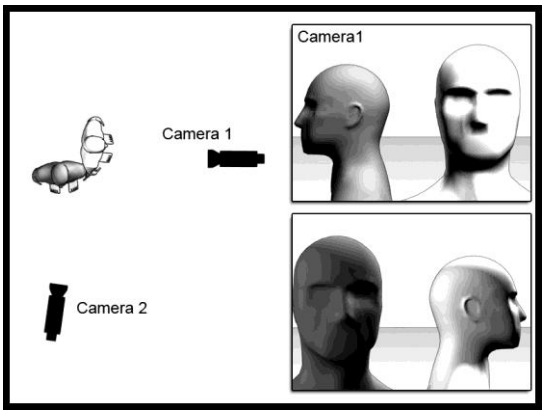
Camera2	Symmetric to Camera1 modulo `Height()`

Figure 10: Covering figures side-by-side using external and internal reverse shots



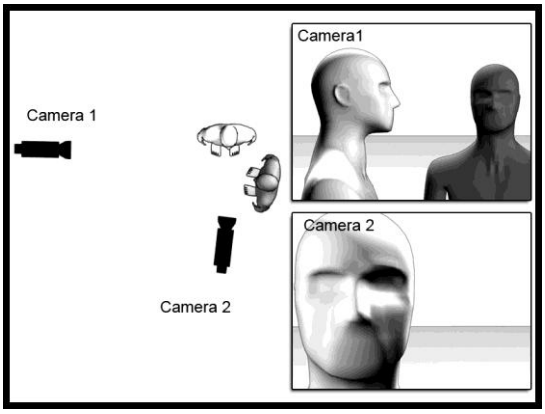
Camera1	`Height(head1, [60-80]%, t);`
	`Occlusion(head1, no, t);`
	`Height(eye1, [0.3-0.4], t);`
	`Orientation(head1, LEFTVIEW, t);`
	`InView(head1, [70-80]%, t);`
	`Height(head2, [40-60]%, t);`
	`Occlusion(head2, no, t);`
	`Height(eye2, [0.3-0.4], t);`
	`Orientation(head2, FRONTVIEW, t);`
	`Framing(head2, LEFTSCREEN, t);`
	`InView(head2, yes, t);`
	`RelativeSpatialLocation (head1, head2, right, time);`

Camera2	`Height(head1, 100%, t);`
	`Occlusion(head1, no, t);`
	`Height(eye1, [0.3-0.4], t);`
	`Orientation(head1, FRONTVIEW, t);`
	`InView(head1, [70-80]%, t);`



Camera1	`Height(head1, [40-60]%, t);`
	`Occlusion(head1, no, t);`
	`Height(eye1, [0.3-0.4], t);`
	`Orientation(head1, LEFTVIEW, t);`
	`Framing(head1, LEFTSCREEN, t);`
	`InView(head1, yes, t);`
	
	`Height(head2, [40-60]%, t);`
	`Occlusion(head2, no, t);`
	`Height(eye2, [0.3-0.4], t);`
	`Orientation(head2, FRONTVIEW, t);`
	`Framing(head2, RIGHTSCREEN, t);`
	`InView(head2, yes, t);`

Camera2	Symmetric to Camera1…



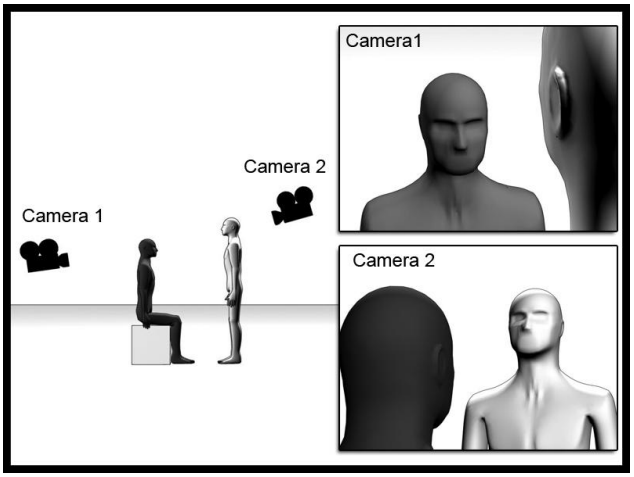
Camera1	`Height(head1, [30-40]%, t);`
	`Occlusion(head1, no, t);`
	`Height(eye1, [0.3-0.4], t);`
	`Orientation(head1, FRONTVIEW, t);`
	`Framing(head2, RIGHTSCREEN, t);`
	`InView(head1, yes, t);`
	
	`Height(head2, [30-40]%, t);`
	`Occlusion(head2, no, t);`
	`Height(eye2, [0.3-0.4], t);`
	`Orientation(head2, LEFTVIEW, t);`
	`Framing(head2, LEFTSCREEN, t);`
	`InView(head2, yes, t);`

Camera2	See top picture (Camera2)

Figure 11: Covering L-shaped configurations

2.5.5 Variations in height

Characters who vary in height can still be covered by the same sets of paired shots, although the camera must be tilted to compensate for the difference in eyelines.



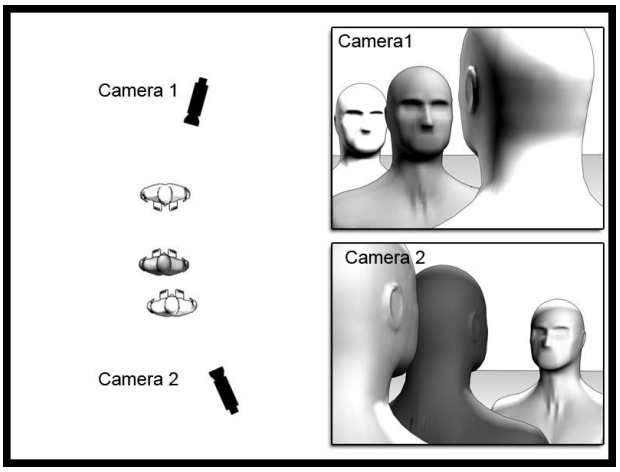
Camera1	`Height(head1, [40-60]%, t);`
	`Occlusion(head1, no, t);`
	`Height(eye1, [0.45-0.55], t);`
	`Orientation(head1, FRONTVIEW, t);`
	`Framing(head1, LEFTSCREEN, t);`
	`InView(head1, yes, t);`
	
	`Height(head2, [80-90]%, t);`
	`Occlusion(head2, no, t);`
	`Height(eye2, [0.2-0.3], t);`
	`Orientation(head2, REARVIEW, t);`
	`Framing(head2, RIGHTSCREEN, t);`
	`InView(head2, [20-30]%, t);`

Camera2	`Height(head1, [60-80]%, t);`
	`Occlusion(head1, no, t);`
	`Height(eye1, [0.45-0.55], t);`
	`Orientation(head1, REARVIEW, t);`
	`Framing(head1, LEFTSCREEN, t);`
	`InView(head1, yes, t);`
	
	`Height(head2, [20-30]%, t);`
	`Occlusion(head2, no, t);`
	`Height(eye2, [0.2-0.3], t);`
	`Orientation(head2, FRONTVIEW, t);`
	`Framing(head2, RIGHTSCREEN, t);`
	`InView(head2, yes, t);`

Figure 12: Covering characters whose heads are at different heights

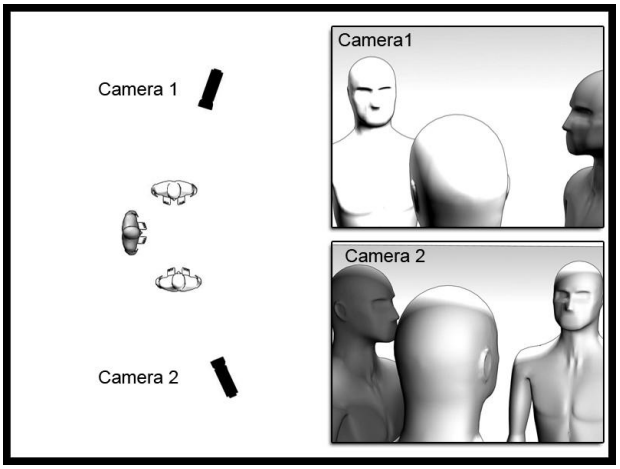
2.6 Camera placement for three characters

Shooting more than two characters naturally requires greater attention to establishing the space in which the dialogue occurs. Below are idioms allowing coverage of conversation between 3 characters.



Camera1	`Height(head1, [40-50]%, t);`
	`Occlusion(head1, no, t);`
	`Orientation(head1, FRONTVIEW, t);`
	`Framing(head1, LEFTSCREEN, t);`
	`InView(head1, yes, t);`
	
	`Height(head2, [40-60]%, t);`
	`Occlusion(head2, no, t);`
	`Orientation(head2, FRONTVIEW, t);`
	`Framing(head2, LEFTSCREEN, t);`
	`InView(head2, yes, t);`
	
	`Height(head3, [60-80]%, t);`
	`Occlusion(head3, no, t);`
	`Orientation(head3, REARVIEW, t);`
	`Framing(head3, RIGHTSCREEN, t);`
	`InView(head3, [80-100]%, t);`

Camera2	Symmetric to Camera1



Camera1	`Height(head1, [30-40]%, t);`
	`Occlusion(head1, no, t);`
	`Orientation(head1, FRONTVIEW, t);`
	`Framing(head1, LEFTSCREEN, t);`
	`InView(head1, yes, t);`
	
	`Height(head2, [40-60]%, t);`
	`Occlusion(head2, no, t);`
	`Orientation(head2, REARVIEW, t);`
	`Framing(head2, CENTERSCREEN, t);`
	`InView(head2, [80-100]%, t);`
	
	`Height(head3, [30-40]%, t);`
	`Occlusion(head3, no, t);`
	`Orientation(head3, RIGHTVIEW, t);`
	`Framing(head3, RIGHTSCREEN, t);`
	`InView(head3, [40-100]%, t);`

Camera2	Symmetric to Camera1

Figure 13: Possible Camera positions for covering a three-way dialogue.

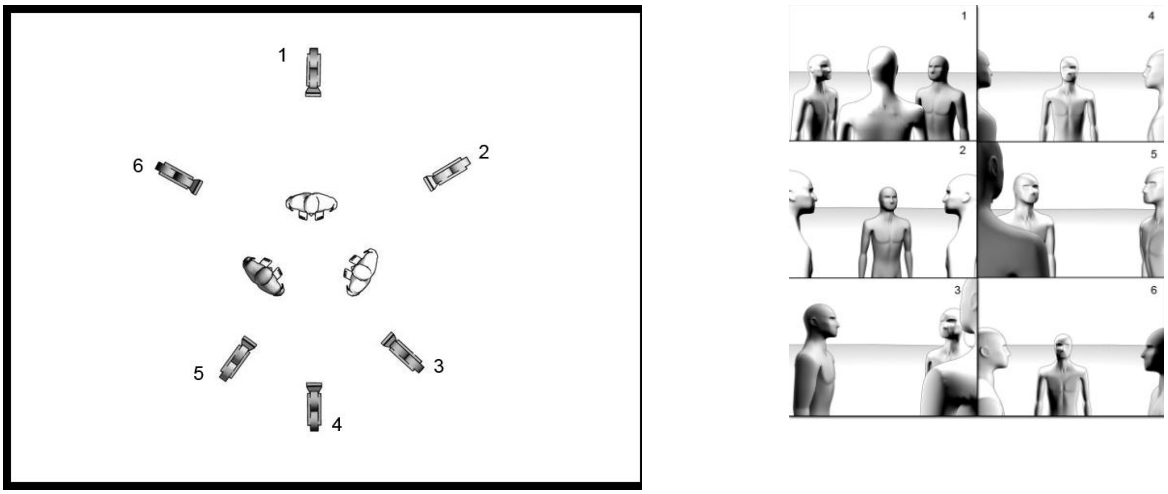


Figure 14: Camera positions for characters in a triangular configuration.

2.7 Character movement

Shooting characters moving from place to place involves careful planning if the viewer is to maintain his/her understanding of the spatial makeup of the scene. Covering a change of orientation in a character by cutting to a closer camera reinforces the character's new direction. Actions covered in more than one shot must follow a consistent direction across the frame.

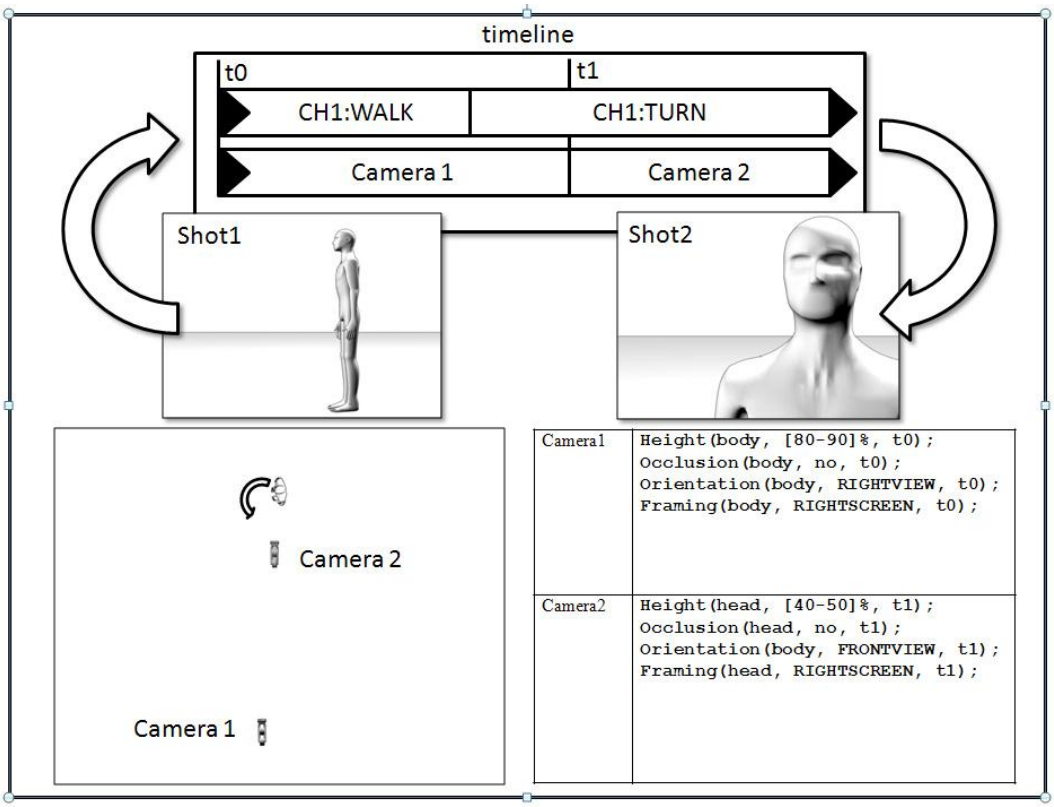


Figure 15: Covering a turning character by advancing the camera

Here, a character changes orientation during a dialogue. External Reverse angles maintain the character's position in the frame.

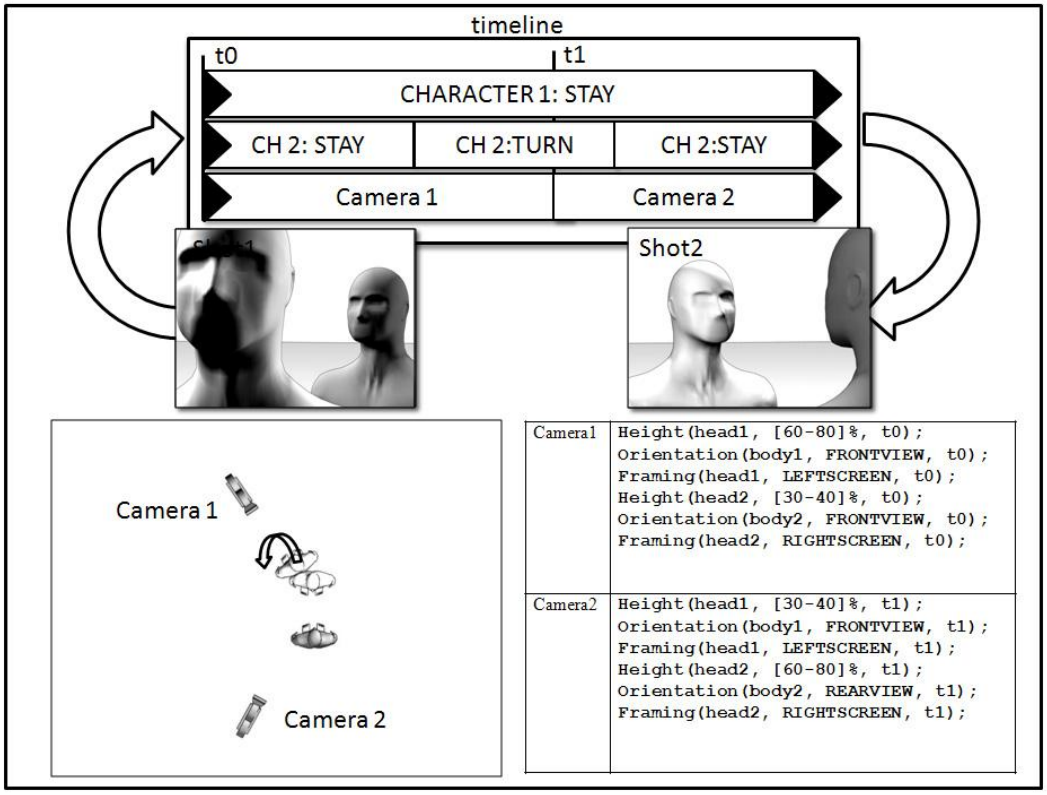


Figure 16: Covering a character through a 180degree turn

Characters may change direction to face away or towards the camera. The Line of interest however is still considered to be between them. Here the situation is covered by cutting from external to internal reverse shots.

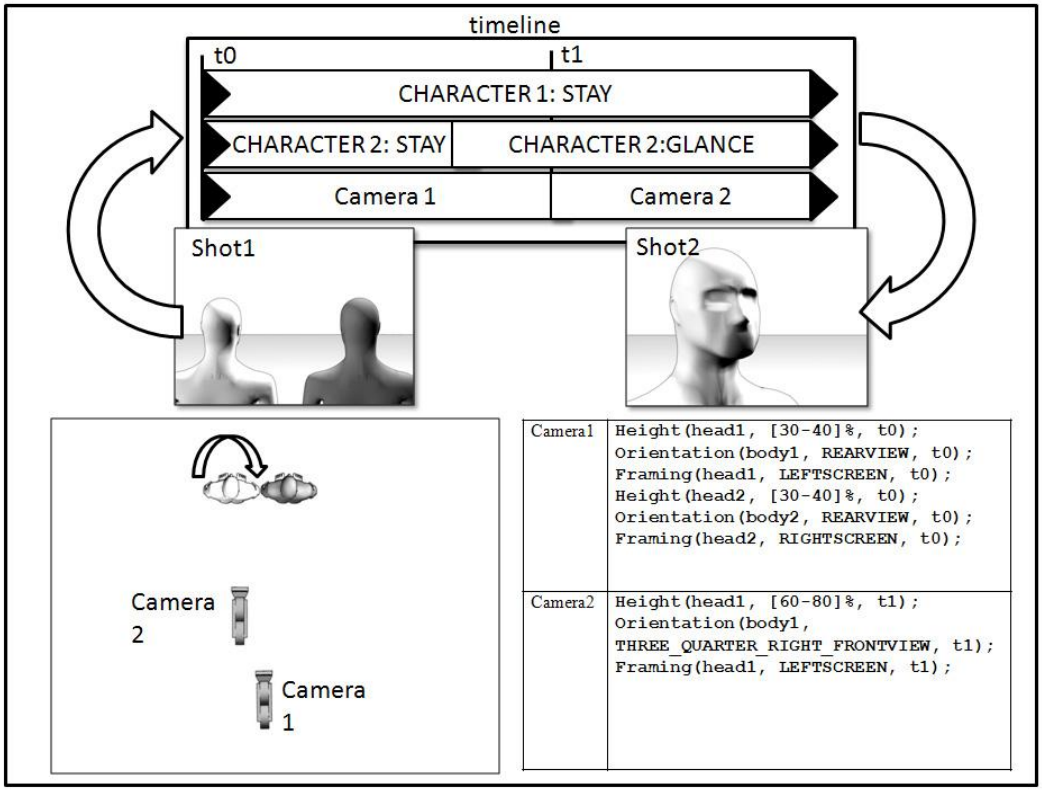


Figure 17: Covering a character glancing backwards

Character actions such as rising or sitting often change the composition of a scene dramatically. Cutting to a wider or tighter shot is an effective way of adjusting for the character's new position in the frame. As previously stated, the cut should occur 1/3 of the way through the character's motion.

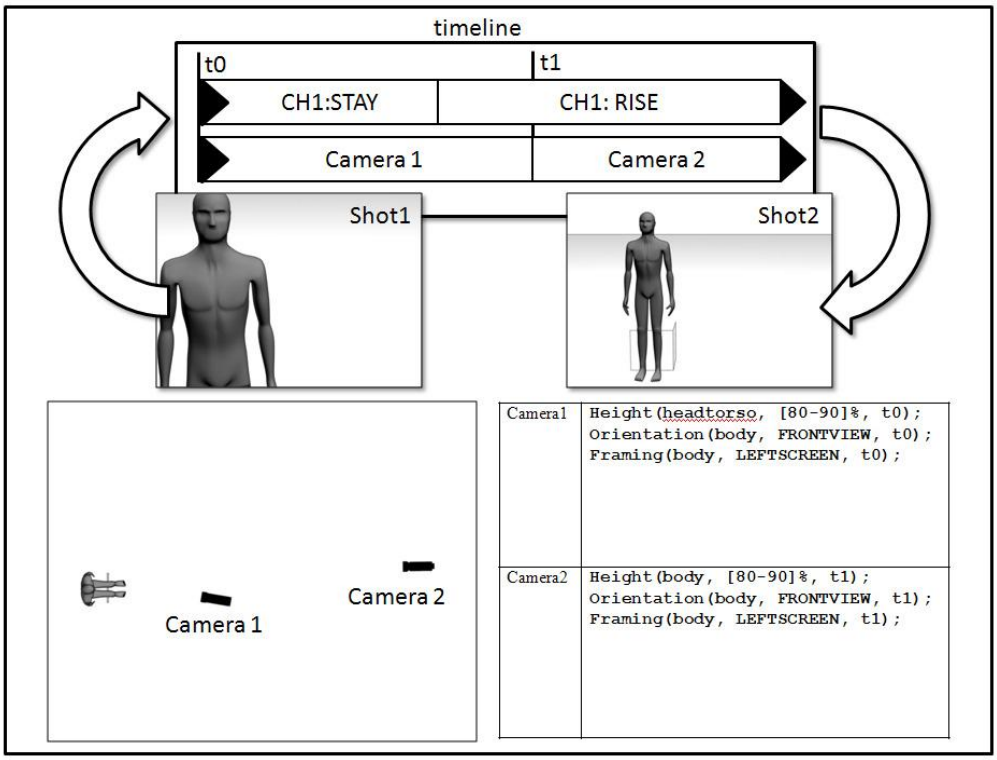


Figure 18: Covering a character rising from a sitting position

Covering a character crossing a large space should be treated in the same way, with the action divided 1/3 vs 2/3.

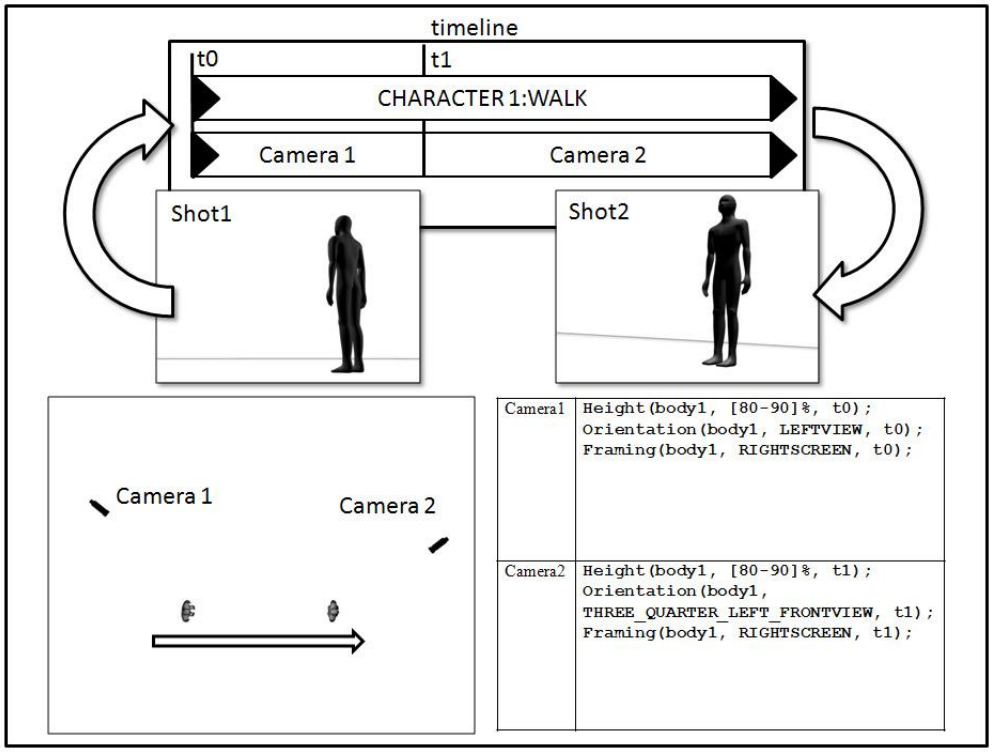


Figure 19: Framing a character crossing a space

2.8 Camera movement

Camera movement in traditional cinema is usually used only for special purposes. Motions in the literature encompass travelling, tracking, dollying, panning and tilting techniques. More complex devices such as camera cranes enable elaborate motions. Camera motions have a range of narrative motivations:

- For establishing contexts. In order to ensure the viewer has a proper understanding of the spatial relations in a scene, or in order to present a large openspace (eg. a battlefield), camera motions are intensively employed. Crane devices enable the fly-over of large scenes, as well as the transition between groups of characters or actions. In smaller scenes, panning or travelling motions provide a description of the context. When preceded by a reaction shot on a character, the viewer reads the motion as the character surveying the scene.

- For establishing relations. Making a panning motion between two key subjects is a well-known device to establish a relationship. Whether the camera represents a subjective or neutral view, the viewer uses the shot to complete his understanding.

- For tracking targets. Tracking a character or a group of characters is easily performed using camera motions such as to-and-fro pannings (complex motions) or travellings (for long linear motions). In such contexts, a high dynamicity of the camera (speed at which the camera moves) can help to drive a sense of tension in the scene. All these camera motions can be used together. Tracking a target and establishing relations are often interweaved: the camera tracks a character, pans to a second one and tracks it, before panning to the first one again.

- For transitions between focus and context or context and focus. Another specific motivation for the use of camera motions is to establish the relation of a key subject in its context. The camera focuses on a character and then moves back (generally to an establishing shot) to display the character in his context. A specific application of this is found in closing a scene or a film (see Figure 20): the moving camera expresses the increasing distance between the viewer and the story.

Figure 20: Use of a camera crane in the "Into the Wild" movie (Focus to Context)

Modern cinematographic movies explore the use of new camera devices which impact strongly the resulting shots. Steadycams and hand-held cameras, for example, have strongly changed the way cinematography is performed (Section 5.0 provides a detailed study). In the canteen scene in 1984, the camera constantly makes small panning and tilting movements, adding to the uncomfortable tension in the scene. However, these movements are small enough not to change the framing of the scene.

2.9 Editing

Multiple camera idioms are used in cinematography because of their ability to arrange the narrative in space in a way that can be easily read by the viewer. Cutting between cameras is also a highly formalised process. In line with traditional Continuity Editing techniques (Thompson, 2006) shot selection in dialogue scenes is usually formed around the script and is temporally continuous. During each utterance the speaker is pictured, and cuts occur when another character reacts in some way to the utterance or replies with another. More complex editing techniques involving montage, temporal ellipsis and others are discussed in section 3. If other actions occur within the scene, it may be necessary to cut from one camera to another to follow the action. A rule-of-thumb used in this situation is to include on third of the action at the end of the first shot and two thirds in the beginning of the second.

2.10 Translation of basic cinematographic devices to Interactive Storytelling

From the description of basic cinematographic devices (focusing on camera control), we draw insights on their practical application within the context of interactive storytelling by exploring three dimensions: characteristic viewpoints, composition and idioms.

2.10.1 Characteristic Viewpoints

A first key element is the ability to identify and compute characteristic shots of key-subjects. Given that the IRIS approach to IS mainly focuses on 3D environments, it is necessary to characterize regions around the key-subjects: each region gathers a set of viewpoints which provide a similar shot (in considering the size and the relative orientation of the character). Partitions between regions represent changes in size (eg. from a close shot to a medium close shot) or changes in orientation (e.g. from facing the character, to a ¾ shot). In case of multiple characters, such spatial partitions can be computed similarly by considering the size of the key-subjects, and their relative locations on the screen. Viewpoints around subjects could therefore be characterized by high-level semantic information (external shot, apex shot, medium closeup shot). Due to the interactive nature of IS systems and the unpredictability of events, it is necessary for these regions to be recomputed in real-time.
Additionally, visibility of key-subjects is an implicit issue that needs to be handled in 3D environments. While in real cinema, elements of the decor may be removed, or actors change places between successive shots, a real-time IS system needs to compute and handle the visibility of its key-elements.

2.10.2 Screen composition

Since composition of shots is often a matter of directorial style, an IS system should offer the means to control the way key-elements are portrayed. Furthermore, narrative dimensions such as dominance, affinity or isolation between characters are partly conveyed through composition. While a general response to the problem of screen composition is way out of scope of the IRIS project (and of Computer Science in general), there are some basic elements that can be enforced:

- layout of characters on strong lines in the frame (rule of the third)
- consideration of eye levels in the composition
- consideration of gaze

2.10.2 Cinematic idioms

Idioms represent idiosyncratic ways for conveying actions. Whether to transpose and formalize such idiom is a key question. Most previous approaches to virtual cinematography have considered idioms as the only way to formalize shots and transitions [He, 96][Elson 2007]. While their straightforward application provide legible results, the underlying models (Finite State Machine, or Hierarchical representations) are implicitly unable to handle variations in style. In trying to propose a new approach to the problem of camera control in interactive storytelling, it is necessary to perform a clear shift from the straightforward encoding of cinematic idioms in order to handle large variations in directorial style, a key element to more expressive interactive movies.

3.0 Advanced narrative devices

The idioms and constraints discussed so far serve only to make the actions in Interactive Stories legible in terms of space and continuity; however, these concerns form only part of the techniques needed to properly elucidate the goals of any narrative. Camerawork and editing not only serve to document the action but may also be called upon to expose narrative information not visible from every possible viewpoint.

In the case of Michael Radford's 1984, relying only on these basic rules to shoot the dialogue within the film would leave the viewer entirely unaware of the characters' motivations and consequent effects on the development of the story. In the canteen scene in particular, the dialogue has little bearing on the narrative: the conversation mostly revolves around day-to-day and seemingly mundane themes such as food and work. In terms of imparting information, more is achieved by Radford's employment of cinematographic devices to expose and amplify subtle actions that would otherwise pass unnoticed.

In linear cinema, narrative information is imparted to the viewer over the course of the film in line with temporal conventions with which the viewer is already familiar. In IS, this might not always be the case. However, by considering each scene as a unit, with its own narrative goals within the greater structure of the story, visual information to be shot can be usefully divided into two types:

3.1 Information pertaining to the characters and plot development.

Clearly, many actions such as dialogue make sense and contribute to the story only when presented in a particular sequence. Actions such as the introduction of important characters or particular objects, actions of speech which are crucial to the development of the plot or the viewer's understanding of the narrative depend on being presented at particular junctures within the story. In 1984, Smith enters the main space of the canteen and notices Julia is watching him. The timing of this encounter early on in the scene is crucial as through this single pair of shots the viewer is:

1. *introduced* to Julia and prepared for her participation in later scenes;
2. *reminded* that surveillance and watching one's peers is a central theme of the narrative;
3. *alerted* to the fact that Smith is more interested in the people around him than his friend Syme which affects the viewers understanding of all subsequent dialogue.

This information is imparted through framing and editing rather than by any single action on the part of either character. Both characters are almost motionless during the sequence.

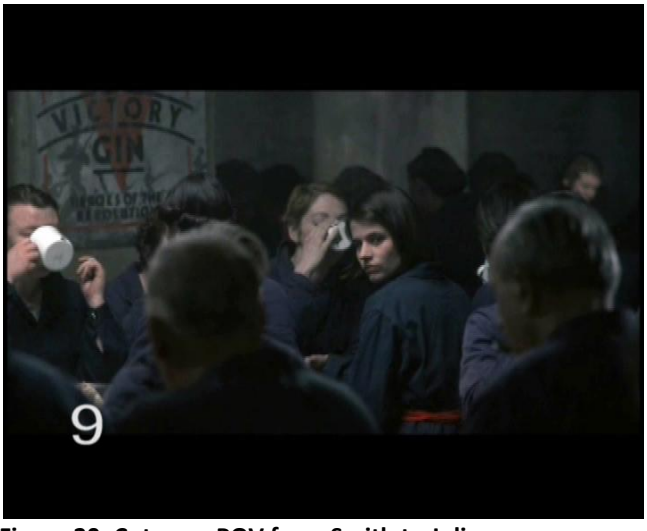


Figure 20: Cutaway POV from Smith to Julia.

3.2 Information about the world

Other equally important information serves to support and enrich the narrative as a whole rather than form integral parts of its structure and in the case of 1984, often depends little on the sequence in which it is presented. This information could be considered a series of discrete facts that are drip-fed into the action over the course of the film. This supporting narrative information can be exposed through the actions of characters other than the main protagonists or through the inclusion of particular objects and spaces.

An example of this in 1984 is the establishing shot in which an anonymous extra watches the telescreen for several seconds. The viewer learns from this that the omnipresent telescreen are significant to the society in which the film is set.

Figure 20:Establishing Shot with extra

Another piece of peripheral narrative information concerns the food served in the canteen which is presented in 3 shots. In the first, the food is slopped into bowls, in the second Smith hesitates for several seconds before putting it in his mouth, in the third Parsons expresses pleasure despite stating that there is no meat in the stew. From this the viewer is reminded that conditions in Oceania are poor for the characters and party members are expected to be enthusiastic nevertheless.

Figure 21:Following the food theme: significant objects are often shown 3 times

3.3 Composition and proportion

Composition, apart from establishing space, can also be used to state the relationships between characters. Used simply, framing a character so that he/she appears tiny in relation to the whole frame can emphasise his or her vulnerability or insignificance. In dialogue shots, a character framed to appear larger will seem dominant and through the consequent enlargement of their features,

facial expressions can be emphasised. By gradually increasing the amount of screen-space a character inhabits in subsequent shots by either narrowing the field of view or moving the camera towards them, the viewer's attention can be drawn especially to the character.

In shots 12, 14, 16, the camera tightens slowly on Syme during his speech about the Newspeak Dictionary, increasing the amount of screen space he inhabits. This has the effect of making him appear more enthused about his subject and gives his utterances more importance. However, as he still occupies slightly less screen space than Smith, the viewer still identifies Smith as the central character in the narrative.

Figure22: Tightening slowly on Syme

3.2 Depth of field

A shallow depth of field can be used to ascribe prominence to characters and objects, especially within visually busy or cluttered shots. This technique can counter the distracting effects of objects or characters moving between the camera and subject.

Figure 23: A zoom lens with its characteristic shallow depth of field distinguishes subjects from their background

Equally a deep field can highlight a character's association with others in the scene or emphasise dramatic actions, such as a character attempting to hide within a crowd.

3.3 Occlusion and isolation

By pairing a shot in which one character is alone with a shot in which multiple characters are present, even in the background, a sense of the first character's isolation can be emphasised.

Figure 24: Contrasting Smith's isolation with Syme's companions in the party

Obscuring the face of a character, whether through framing or through use of an occluding object, can communicate to the viewer that they are secondary to the plot. This is used in 1984 to characterise the Proles as a faceless mass and not to be considered as individual characters. To counterbalance the distracting effect of characters or objects in the foreground, a shallow depth of field must be used (see previous section).

Figure 26: Obscuring the faces of the proles

3.4 Objects

Composition can be used to introduce or revisit narrative threads. Including significant objects within the composition of a shot can keep particular themes in mind, even if they are only seen

peripherally. Throughout the canteen scene in 1984 the presence of the telescreen over Smith's shoulder reinforces the idea that Smith is being watched and increases the impression of tension within the scene.

Figure 27: The Telescreen clearly visible over Smith's shoulder

3.5 Direction of gaze

The line of interest is of paramount importance to conventional linear film making. This compositional line linking the heads of characters in dialogue is crucial because as human beings the eye is our most important directional indicator.

Characters in dialogue generally look off camera, as looking directly out of the frame tends to place the viewer in the action. The exception to this is during a POV shot, where the direct gaze of characters allow the viewer to understand that their own gaze has become contingent with that of a character.

Conventions such as following a character's gaze with a POV shot capitalise on this. In figure 28 Smith's attention is clearly caught and the camera cuts to follow his gaze. As Julia is staring directly at the camera, the viewer understands that we in this shot we are looking through Smith's eyes.

Figure 28: A cutaway to Julia following Smith's gaze. N.B Julia is looking directly into the camera, placing the viewer in a subjective position

Through the act of looking at an object or other person, a character automatically ascribes significance to it. Conversely, by following a character looking away from a dialogue, for instance to a part of the frame where the viewer knows no character is located, the camera can indicate lack of interest on the part of a character.

3.6 Translation of narrative devices to Interactive Storytelling

While the basic cinematographic devices makes the actions legible in terms of space and continuity, elaborate narrative devices are necessary to support the narrative goals of the movie and to introduce information not explicitly displayed on the screen or provided by the soundtracks. The use of such devices in the context of IS requires that the system shares a detailed knowledge of the underlying narrative structure together with the narrative goals of each scene. A minimum level of information required by a cinematography tool encompasses:

- a structured representation of the narrative into scene units.
- within the scene units, a set of primitive *narrative elements*.
- the causal relations between the set of narrative elements.

Each *narrative element* has a narrative purpose, should it be related to the explicit action of a character (eg *Smith takes his spoon*), to a non explicit action (eg. *Establish the relation between Syme and Julia*), or to the understanding of the context (eg. *Establish the context of the Canteen*). Narrative elements have a duration and may overlap in time. The aim of the cinematography is to analyse an convey such a narrative structure in a way that properly provides an understanding of the story while supporting cinematographic conventions.

Given that there are many means to convey the same narrative element, we associate all the possible devices with the event. For example, to establish the relation between Syme and Julia, one can employ:

- an establishing shot displaying both characters (characters need to be isolated, either spatially, or by using depth of field)
- a front shot of Julia after displaying a from shot of Syme
- an over-the-shoulder shot of Syme and Julia with a focus varying from one to another
- a panning shot from Syme to Julia
- display both characters in the background of a third (unknown) character.

The selection of the appropriate device is then a matter of directorial style, together with continuity and coherency in edits.

4.0 Linking narrative, editing and shot selection

4.1 Further editing issues

In Continuity Editing, dialogue is edited using match-on-action pairs of shots, where a character is pictured speaking and his/her opposite number is seen reacting in a shot that is roughly a mirror image of the first. Figure 29 shows the length of each shot compared to each utterance in the canteen scene in 1984. Overlaps where cuts occur demonstrate the pattern of shot and reaction-shot used to cover both the speaker and respondent.

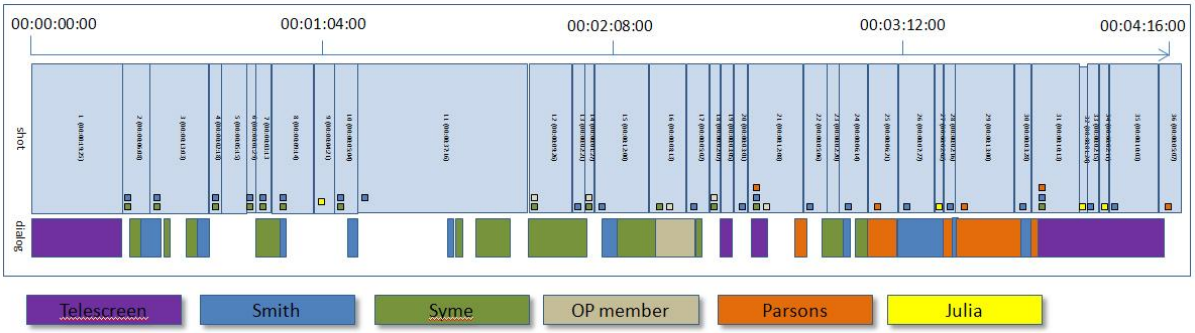


Figure 29: Shots in the Canteen sequence scaled by time. Colour bars indicate speaking character. Colour squares indicate character onscreen

This pattern can adequately convey the sense of a conversation with the viewer following much as he/she would in real life but can also be subverted in order to achieve specific effects. By failing to cut to a character speaking and instead remaining on another character, the attention of the viewer can be diverted away from a dialogue. This can be used to indicate disinterest on the part of the character pictured or their sudden interest in something else within the scene. This is clearly visible in the diagram above, where, in Shots 11 and 15, the camera remains on Smith even though Syme is speaking. This is often accompanied by a drop in the volume of the dialogue and can be followed by a cutaway to the new source of interest.

4.1.1 Ellipsis

Several easily legible techniques can be used to distinguish between actions occurring along 'in real time' and actions which are to be read as temporally discontinuous. Use of transition effects such as dissolves or fades as opposed to straight cuts are usually read as signalling either ellipsis or the introduction of a new temporal or spatial context.

4.1.2 Pace and meter

By varying the pace at which shots are selected, various changes in pace can be achieved. Simple measures such as increasing the average speed of cutting can enhance the sense of tension, speed or chaotic action within a scene. Choosing a moving camera over a stationary one also has this effect: a slightly moving frame often being read through its association with the documentary as placing the viewer more completely in the action.

These effects can be highly effective when contrasted against their opposites. In 1984, the camera constantly shifts slightly, enhancing the viewer's impression of the characters' unease. However when the telescreen speaks, the camera halts completely, reflecting the stillness and tension of the characters. When Smith sits and gathers his thoughts at the table in the canteen scene, the camera lingers on him for 34 seconds, emphasising his isolation and unease.

Figure 30: Shot 11 as Smith gathers his thoughts, lasts significantly longer than any other in the sequence.

Equally, choosing a fast-moving camera in a succession of quick cuts may be used in contrast to previous and following scenes, having a direct effect on the overall structure of a narrative.

4.1.3 Establishing shots

Establishing shots are often used at the beginning of a scene to introduce characters, spaces or objects to the viewer. They often function as an overview of the space in which the action takes place, locating the characters and their spatial relationship to each other. Movements such as slow tracking or panning of the camera are often used to encompass as much of the scene in detail as possible.

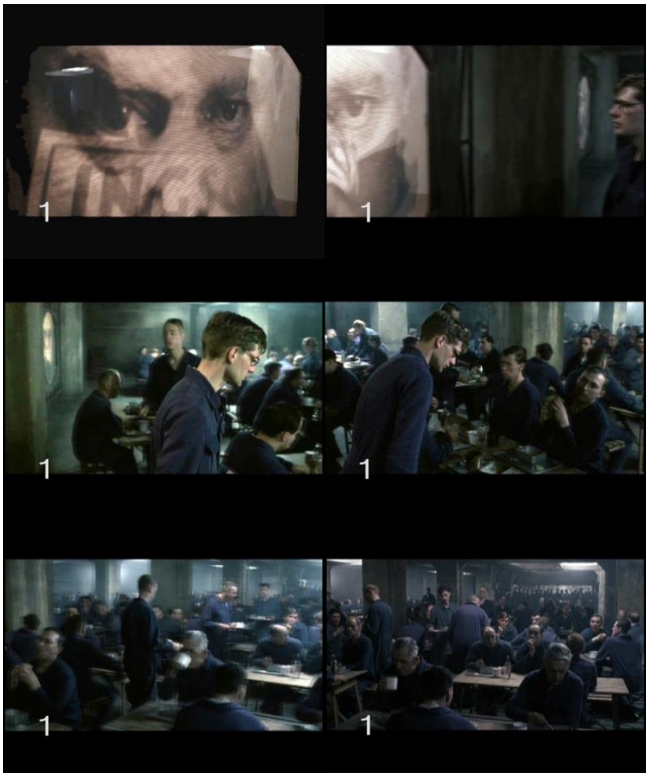


Figure 31: Establishing shot panning around the room

Once again, in addition to functions of spatial description, establishing shots perform a vital narrative role through being an opportunity for the viewer to establish the visual components of the scene before dialogue occurs.

In introducing particularly complex or unfamiliar scenes, sequences of establishing shots can be used. These need not be part of a cohesive idiom of spatially related shots as the viewer does not necessarily need to know the exact layout of all parts of the scene.

4.1.4 POV shots and the subjective camera

Generally, cameras move only to track the action in a scene. An exception to this rule is the POV shot, which has several unique properties. By moving the camera forward at eye level (especially if some secondary movement such as camera-shake is discernible) the shot can be read as looking out through the eyes of a character.

4.1.5 Montage

Montage techniques can be usefully employed to impart narrative information to the viewer in contexts where the psychological or intellectual content of the narrative is more important than the spatial legibility of the sequence. The juxtaposition in subsequent shots of 2 sequences can be used to emphasise meaning or affect.

4.2 Cinematography and narrative

Within a film, narrative information is imparted to the viewer over time. This information may be presented as audio information (through the dialogue, soundtrack or diagetic sound) or through displaying actions. Clearly, actions that are not included in the dialogue must be presented entirely through cinematography and call for careful shot selection. With respect to IS, certain factors may be altered, as a result of which the direction of the scene must be recomputed in order to preserve its meaning. These factors are primarily the scene length, the number of cuts, the spatial configuration of characters and the actions of characters including dialogue.

The cinematic style of the film seeks to support the narrative, and this can be viewed as a particular parameterisation of idioms described in the previous section which are applied before composing individual shots in order to attempt to retain the outcomes of the scene. The following rules are specific to the direction of 1984 and attempt to retain the directorial style of the original.

- Main character must appear in 70% of shots.
- Narrative information must be included in required proportions (see above).
- Shots during dialogue must follow the speaker and then cut to reaction.
- Dialogue must be covered using internal reverse shots where possible.
- If an action occurs involving a character's gaze, a POV shot should cover it.
- Character movement should be covered using one of the techniques below.
- Beginnings and endings of sequences should be covered using a fade.
- Camera should be static: changes of view should be covered by a cut.
- Shots less than 30 degrees apart should be consolidated.
- Shots should be between 1 and 10 seconds.

By considering each scene to have start and end states, a body of narrative information can be established which allows multiple routes through the scene (in terms of both staging and cinematography). The development or introduction of characters and themes may be considered as properties for judging the success of a particular edit, in the same way that certain compositional properties may be applied to individual shots.

In the context of IS, it is useful to categorize these actions scene by scene in order to best present them. For instance, certain actions may represent themes developed earlier in the story while others (such as the introduction of new characters) may represent entirely new narrative threads. There may also be a need for the subtle presentation of actions, objects or characters that may play a greater part in the narrative later on.

Looking at the Canteen scene in 1984, it is possible to break down the huge number of actions (including cinematographic events such as placing a character or object onscreen) into a number of facts. Some of these facts are presented in a single shot and indeed in the case of minor details (such as the prevalence of Victory Gin advertisements in 1984) may not even be the subject of the shot.

Figure 32. Gin is dispensed from a hatch

Figure 33. Julia is watching Smith

More important factors, such as Smith's unease around Syme may be presented a number of times in multiple shots and both through dialogue and cinematography. For example in separate actions, Smith lies to Syme about razor blades, Smith is uncomfortable around Syme on multiple occasions as Syme evangelises about the regime. From previous scenes, the viewer knows that Smith is isolated and silently opposes the regime: therefore Syme's support of the regime contributes to the developing sense that Smith fears and dislikes Syme. For the Michael Radford's "Canteen Scene" is 1984 we can enumerate what narrative information is introduced, emphasised and whether this is achieved conveyed through the cinematography, the dialogue or both (see table 1).

Narrative information	*New*	*Number of shots*	*Cinematography*	*Dialogue*
Telescreens are everywhere	NO	14	YES	NO
People stop to watch telescreens	YES	1	YES	NO
Canteen is where Syme and Smith eat	YES	35	YES	NO
Syme and Smith are friends	YES	17	YES	YES
Proles serve the OP	NO	7	YES	NO
Smith lies to his comrades	NO	1	NO	YES
Hangings are entertainment	NO	2	NO	YES
OP drink gin	NO	9	YES	YES
Victory gin is the only brand	NO	2	YES	NO
Smith is cautious around Syme	YES	17	YES	YES
Gin is dispensed from a hatch	YES	2	YES	YES
Julia is watching Smith	YES	3	YES	NO
Smith is aware of Julia	YES	5	YES	NO
Syme works on the dictionary	YES	5	NO	YES
The language is shrinking	YES	1	NO	YES
Newspeak involves cleaning the language	YES	2	NO	YES
Syme evangelises about the regime.	YES	5	NO	YES
The meat in the canteen is awful	YES	4	YES	
OP members talk freely to each other about matters of doctrine	NO	2	NO	YES
OP members spontaneously make ovations	NO	2	YES	NO
The telescreens rewrite the facts	YES	1	NO	YES
OP members believe the telescreens	YES	1	YES	NO
Parsons is unquestioningly enthusiastic	NO	5	YES	YES

Table 1. Narrative information and the role of cinematography for the "Canteen Scene"

In order to construct an edit that has the same start and end states as the original film, all these facts need to be presented to the viewer over the course of the scene.

To incorporate this information in the editing process it is useful to consider the proportion of the scene through which the information is presented, so for instance, a simple fact such as 'Smith and Syme are eating in the canteen' needs to be restated in the vast majority of shots in order to maintain the sequence's sense of place. The omnipresent telescreen also need to be included in a large number of shots, regardless of the length of the sequence. By changing these proportions the meaning of the narrative may be modified. For example, if Smith's nervous reactions to Syme's statements were omitted in favour of the camera staying on Syme, their relationship might appear quite different.

From these considerations it is now possible to apply certain properties to the construction of multiple edits of the same scene, regardless of its length or spatial configuration. If the dialogue between Syme and Smith took place in a different location (for instance, they remained in the queue) it may still be possible to create a scene with the same meaning). So long as the proportions of information remain the same (for instance around 13% of all shots include dialogue or coverage of actions which reinforce the idea that Parsons is unquestioningly supportive of the regime) the meaning of the scene should remain intact.

This approach also allows for the modification of the outcome of a scene. If these proportions are changed, the contribution of the scene to the overall narrative progression may change. If Smith's looking at Julia is only presented once, the viewer may well consider it to be of far less importance to the plot than in the original edit.

5.0 Contemporary Cinematographic Techniques

Modern screen-based media including television, film and games rely on constantly evolving visual and narrative conventions built up over hundreds of years. The extreme commercial pressure of the media industries and their constant search for novelty has led to a number of important technical innovations in the last decade or so which have profoundly influenced the way that audiences and film makers think about film. Innovations such as digital tape, electronically stabilised camera mounts and high-quality portable cameras have extended the range of techniques open to film makers, while CG animation and post-production technologies have led to an increasing plasticity of the filmed image. On the other hand, many film-makers, liberated by the availability of low-cost camera and editing equipment have conceived of new ways to approach cinematography. Likewise, the constant tension between play and narrative in modern video games has led to a number of novel approaches in interactive media.

Modern cinema, as an art-form more than a century old, constantly refers to itself, relying on the high level of cinematic literacy of its audiences. Depending on a common vocabulary negotiated over many decades, film-makers commonly use references to other styles of films as a short-cut to a set of associations: for example using Western-style crash zooms to add an epic, heroic atmosphere to a conflict in a more mundane setting. Modern films continually reference other films, perhaps most notably in the work of directors such as Quentin Tarantino, who borrows characters, costumes and entire sequences directly from older movies. Film-makers also borrow from other screen-based media, often taking techniques from television or recently from video games to achieve particular dramatic effects (see below). In recent years this has led to a number of films presented (in the manner of the epistolary novels popular in the 19th Century), as genuine documents. In certain cases, such as *The Blair Witch Project* and *Cloverfield* this extends far beyond the level of individual sequences, to the entire narrative structure of the piece.

This section offers a brief survey of a number of technical and conceptual innovations in various types of screen-based media and their possible application to IS, including conceptual consequences of this application.

6.0 Notable Genres

6.1 Television Drama, Soap Opera and Situation Comedy

One of the most significant ways in which TV dramas and other related genres differ from theatrical film is in terms of time. While few cinema productions exceed 2 hours in length, many soap operas and serial dramas have storylines which play out over many hours of programming, spread out over months of broadcasts. In many respects this allows far greater potential for character and plot development, however the fragmentary nature of this type of storytelling, coupled with the logistical and financial constraints common to TV production (such as limited production time and the necessity to maintain audiences over long periods of time) often mean that directors of soap operas and sitcoms have little scope for experimentation.

Certain television dramas rely heavily on camera movement to support the narrative. In the TV series *ER* (Crichton, 1995-2007), camera movement is used extensively to set the tone and pace of the drama. The camera rarely remains still, alternating between handheld and steadicam tracking during emergency and action scenes. The camera is often moved around the set during scenes of intense action (for example the scenes of surgery upon which much of the drama hinges) with the camera operator stepping between and around the actors. In this way, even inert objects such as medical apparatus acquire a sense of movement and consequently significance and dramatic tension. Fast tracking or 'whip-pans' from one subject to another are used in the place of cuts to heighten dynamicity.

Figure 34. A whip pan in one of ER's trademark surgery scenes.

In quieter scenes, different types of camera movement are used. Crane shots are occasionally used to demonstrate the direction of a scene's attention. Slow dollying towards a character is often used to signify dramatic intensity or tension.

Figure 35. A quieter moment in ER: the camera dollies towards the character over 10 secs.

In *Eastenders* (BBC, 1985-2010), the overlapping lives of the characters are often connected using exterior shots. Making the most of the fact that the action is set in a single square of houses, tracking shots are sometimes used to connect different plotlines. Often characters from a previous scene are included in the background of a subsequent shot, reinforcing the temporal structure of the plot.

In many Sitcoms and Soap Operas, a small number of exterior establishing shots are used to establish the context of the drama. In many cases these function as simple descriptions of the location of the action: for example four or five exterior static shots are used to bookend scenes in comedy series *Friends*((Marta Kauffman, 1994)) and in soap opera *Eastenders*, tracking exterior shots are often used to open an episode. Sometimes however, these shots serve to present the social or dramatic context. In *The OC* (Schwartz, 2003) for example, montages of shots of the opulent area of California in which the series is set are used to divide scenes, reinforcing the aspirational aspects of the series.

Figure 36 An establishing shot from the OC, used in nearly every episode

6.1.2 Application to IS

Performing a Whip Pan

The expression of a Whip Pan in an IS system requires that ability to extract the characters locations and orientations, as well as to guarantee their visibility during the motion. As in the ER series, whip pans are usually employed in a sequence to cover multiple characters. The ability to compute paths between viewpoints is therefore critical. By building upon these requisites, Whip Pans can be expressed as fast-paced panoramic shots between characters The duration of the whips pans may be controlled;in ER, these pans typically take between 0.2 and 0.8 seconds. However the appropriate

controller for the duration of Whip Pans is more related to the desired pacing and dynamic nature of the scene than to a expected duration.

Performing a Slow Dolly
In line with the Whip Pans, slow dollies require the ability for the IS system to compute paths between locations of characteristic shots (eg, from a medium shot to a medium close shot). Since motivations for slow dollies is strongly related to the character in the story (eg. by providing an increasing sense of proximity between the viewer and the character), their use is triggered by a narrative element.

Use of Establishing Montages
In IS, non-interactive montages of establishing shots of this type might not only be used to perform narrative tasks as described above but, played as pre-rendered video in between real-time 3D sections could be used to ameliorate loading times.

6.2 Action Films

Hollywood action films, relying heavily on spectacle and technical novelty often borrow techniques from other styles of film-making but also - as a staple of mainstream cinema - establish standard formal frameworks which are used and subverted by other genres. For example, *The Matrix* (Larry Wachowski, 1999) appropriated many stunt and cinematographic motifs from Hong Kong cinema, bringing them to mainstream audiences and has, due to its combining of these motifs with innovative post-production techniques, influenced countless films, games and television dramas since.

The huge budgets of mainstream action films ensure that their production teams are often the first to showcase new developments in cinematographic technology such as novel camera and post-production techniques. In particular 'bullet time' and other techniques for manipulating time are highly prevalent in the action genre (see below) and are often referenced in particular in video games.

With the advent of digital film-making technologies, a number of directors, most notably Zack Snyder (*300*) and Frank Miller (*The Spirit, Sin City*) have pioneered an approach to film making which dispenses almost altogether with the need for locations and conventional sets. In these films colour-keying technology and CGI is used to realise the world in which the drama takes place, giving the film maker unprecedented control over all aspects of the image.

The consumption of spectacular cinema has been described as alternating between immersion (in the drama made more engaging by the spectacular elements) and admiration (for the ingenuity of the film-maker, the sophistication of the effects or the athleticism of the performers) (Ndaliansis, 2000). This model of consumption relies heavily on the audience's understanding of the mechanics of the real world and its relationship to film. For example, the central premise of *The Matrix* relies on the audience's understanding that the characters' actions are beyond the capabilities of a person in the real world. They also understand (even if not cognisant of the technical details involved) that, as the illusion is convincing and engaging, a high level of ingenuity and technical sophistication was involved in its creation.

For IS, this situation is clearly problematic, as techniques designed to engage the audience in this way rely on their relationship to the physical world, a link which is not present in computer graphic

environments. For this reason, certain techniques commonly used by film makers may lose much of their meaning when used in computer graphic environments as discussed later in section (111)

6.3 'Found Footage' and Pseudo-documentary style.

In most mainstream films, the mechanics of cinema are invisible. The elaborate lighting, machinery and people needed to capture the actors' performance on film is carefully hidden outside the frame and the camera itself is an invisible entity. Carefully placed to avoid self-shadowing, it moves seamlessly through walls and the audience is meant to consider it an invisible all-seeing eye. Many film-makers have attempted to subvert this, making the camera an agent in the drama or at least according it some semblance of presence and weight. Modern film-makers continue to manipulate the camera in this way for a variety of reasons, sometimes in order to emphasise the physicality of a sequence (as in the case of handheld cameras used in action sequences) and sometimes to simulate the structure and formal aspects of documentary, lending a semblance of truth to the drama.

Several recent films have been presented, at least in part as documentaries, once again borrowing the structure and formal aspects of a non-synthetic medium to lend weight and immediacy to fantasy subjects. In *District 9* (Blomkamp, 2006) (Reeves, 2008), the greater part of the action is presented using conventional dramatic structures, however, in certain sequences characters appear to be taking part in a documentary about the events unfolding within the drama. These sequences are seamlessly integrated into other scenes, with only the presence of certain key cues to announce to the audience that they are watching a film within a film. These cues include:

- Characters speaking directly to the camera
- Handheld character movements
- Characters talking to characters outside the frame without cutting to that character.
- Graphical cues: News Idents, Recording Info etc including corresponding changes in picture quality

Figure 37 A 'news' item presented in District 9

Certain film makers have extended this simulation of the documentary to affect the entire structure and narrative of films. By substituting conventional 35mm film cameras for non-professional technology such as home video cameras, film-makers can distance themselves from the familiar production styles of Hollywood and achieve a greater sense of immediacy and 'rawness'. *The Blair Witch Project* (Eduardo Sanchez, 1999), is an example of a genre sometimes referred to as 'found

footage'. Shot on Hi-8 video and 16mm film, the movie purports to be assembled from rushes shot by a documentary film crew who disappeared in the backwoods of New England. The cameras are supposedly operated by three film-school students, who in between presenting pieces directly to camera, routinely speak to the cameraman, pass the camera between them etc.

Cloverfield (Reeves, 2008) takes this conceit even further. Apparently shot on low-quality digital tape, the film begins as a home-video of a house party, which is overtaken by the arrival of a 'godzilla' style monster which destroys the city. The action follows the characters from the party shot from the POV of one of the characters. Few cuts are used and the camera is never locked down: the audience are constantly reminded that the camera operator is also a character.

6.4 Application to IS

Characters speaking Directly to Camera
Usually, a character looking directly into the camera in a dialog situation, is read by viewers as the opposite character's POV. For such documentary styles, cameras can be easily computed in interactive contexts by using basic cinematographic devices. Typically one can select the front shot of the character (subjective shot) and perform a neutral screen composition. For more complex documentary shots such as those in which a cameraman is following characters (eg in a street fight), IS camera control needs to rely on path planning techniques enforcing specific constraints such as human walking/running speed and camera jumpiness.

2D Graphical Elements
News Idents and other 2D graphical features can be loaded as overlays in IS systems. In cinema, these techniques are often reinforced by a change in picture quality, for example, simulating the scan lines of television in a news item.

7.0 Handheld and Moving Cameras

Many camera techniques used by contemporary directors rely not only on the particular technical affordances of a camera mount but on the range of associations its use engenders in the minds of the audience. As discussed previously, the handheld camera can be used to readily suggest a sense of immediacy and dynamicity through its association with the documentary. Distinctive handheld camera movements may include:

- Small random and rapid changes of rotation and location of the camera.

- Uneven tracking of target objects including leading and overshooting moving targets.
- Periodic downwards tilts when moving the camera forward (as the camera operator checks his footing).
- Slight tilt around Y axis (suggesting improvised camera placement or shoulder mounting)
- Rapid changes in focus or target object.
- Partial cropping and out-of-focus subjects (suggesting errors on the part of the camera operator)

Figure 38. The Blair Witch Project (left): camera operator checks his footing, Cloverfield(centre): characters often appear partially cropped, tilted and out of focus. Likewise in District 9 (right)

Stabilised camera mounts such as those made by Steadicam and Libra allow camera operators to place cameras in previously inaccessible situations, allowing precise tracking over uneven ground, steady coverage of fast moving subjects and extremely smooth movement, however for dramatic reasons, many film makers choose to retain a certain amount of instability in camera movements either to suggest pace or to reference documentary film.
Movements particular to stabilised cameras include:

- Damping of sudden movements: due to the sophisticated suspension used in Steadicam and other stabilised mounts, jolts and sudden movements are often almost completely absorbed: however this often leads (especially in the case of steadicam shots) to a distinctive undulating motion.
- Slow acceleration and deceleration in camera movements due to the weight and damping effect of the suspension

7.1 Application to IS

While these motions in real cinema are easier to setup and provide considerable flexibility to the filmmaker, their simulation into interactive storytelling system actually raises a number of issues:

- motions of hand-held cameras are constrained by the complex physics of human displacements.
- natural tracking of events and targets, anticipation of cameraman, erratic motions, motions due to checking one cameraman's footings are very difficult phenomena to reproduce.
- navigation of a virtual cameraman in a virtual environment remains a complex problem.

However a number of basic effects can easily simulate hand-held and steadycam motions. We provide in the following some cues the camera and viewpoint motions:

- small and rapid changes in the camera location/orientation (30 cm in location, 1 to 3 degrees in orientation).
- overtracking of all mobile objects on the screen.
- perform rapid changes of focus given the distance of the focus point.

8.0 Editing and Time Manipulation

8.1 'Bullet Time' and Extreme Slow Motion

In a technique developed by Manex Visual Effects for *The Matrix*, hundreds of digital still cameras are used to capture multiple viewpoints of a single action. When re-sequenced in post-production this allows the appearance of freezing the action in a scene while still retaining the camera's ability to move around the scene. In other sequences, extremely high-speed cameras are used to render frenetic moments of action into extremely slow motion.
Both these technique are used to expose small details of the action and render them more significant by extending their onscreen time. In *The Matrix* they are used to cover elements such as piles of shell cases pouring from the breeches of machine guns and walls being destroyed by gunfire, emphasising the destruction.

Figure 39. An extreme slow motion sequence from the Matrix, emphasizing the violence of the scene

8.2 Time Ramping

Time-ramping is commonly used in contemporary action films during action sequences. In *300* (Snyder, 2006), a single violent action such as a sword stroke is often covered in 3 segments, the first at full speed, the moment of impact in slow motion and the follow-through at full speed. These techniques are applied firstly to render confusing and fast-moving actions temporally and spatially legible, to draw attention to particular details of the action (for instance the huge amount of ammunition expended in *The Matrix*, the extravagant blood-letting in *300*) and to render particular actions significant by extending the amount of time they are onscreen.

Figure 40. In 300, Leonidas approaches an enemy at normal speed, the sword stroke is covered in slow motion and he returns to normal speed afterwards.

8.3 Application to IS

In IS both these features can be achieved by varying the playback speed of either the whole scene or animations within it.

Bullet Time and Extreme Slow Motion
A bullet time effect or extreme slow motion can be achieved by lowering the playback speed of the whole scene while maintaining the pace of the camera, including its speed of movement and cutting scheme.

Time Ramping
To achieve a simple Time Ramping effect, markup of the significant action can be used to cue a 50% drop in playback speed of the animation for the duration of the action, however this relies on a highly detailed description of the scene actions and their timings.

9.0 Transitions

Post-production techniques are used increasingly to link shots together. In *The Fast and the Furious* (Cohen, 2001), computer graphic sequences are used to transition seamlessly between camera positions. At the beginning of a motor race, the camera tracks quickly from the character Dom's face to his hand changing gear and a computer graphic sequence follows the internal mechanism of his car's gearbox through the engine and out of the exhaust where it is linked to a live shot of the car accelerating away. This technique, also used in *The Matrix*, allows extra narrative information about the character's world to be included between the two shots, reinforcing the atmosphere of the film. In the case of *The Fast and The Furious*, the audience learns from the transition that the simple gearshift involves both intricate and exciting technology and violent forces (as we follow flaming gas through the car's exhaust system).

Figure 40. A transition between shots in The Fast and The Furious

In *The Fast and the Furious*, remote control Libra camera mounts are used to stabilise the camera during complex motions. Combined with post-production match-moving techniques this allows the director to morph seamlessly from one shot to another, with the camera apparently moving between two points at high speed: a shot that would have been impossible to capture using a single camera. This technique, similar in effect to the whip pans used by *ER*'s camera operators, adds an extra level of dynamicity to the sequence, apparently changing not only the rotation but the location of the camera.

9.1 Application to IS

The effect of linking shots by transition from camera to camera can be achieved simply in the by interpolating quickly between camera positions. The complex camera paths used in The Fast and the Furious could be achieved by tweening between multiple positions. In the Fast and The Furious, although the movement is continuous, the camera apparently transitions between a number of distinct shots, rather than simply following a particular direction as if on a rail.

10.0 Colour and Image Quality

Carefully controlling the colour balance in a scene, whether achieved through lighting, filtering in-camera or in post-production is often used not only to emphasise the physical aspects of an environment (such as using reddish or orange filters in desert environments) but can also be used to symbolic effect.
In the plot of *The Matrix,* the real world is revealed to be a simulacrum created by machines and much of the plot hinges around the protagonist Neo's instinct that something is wrong with the world. To emphasise this, the exterior sequences of the film, although shot conventionally in a modern city, are heavily treated with a green filter (a colour associated with computers in other parts of the film) while the post-apocalyptic world which the characters actually inhabit is, although stylised in terms of production design more conventionally colour-balanced. In a film which features confusing jumps back and forth from the Matrix to the real world, this colour-coding helps to keep the audience oriented.

Figure 41. Differences in colour balance between the world of the matrix and the real world in The Matrix

10.1 Application to IS

Colour symbolism can be easily simulated in most game engines by either using 2D filters over the 3D scene (applied according to cues from the narrative) or by applying an adjustment to the lighting and ambient colour in the scene.
For example, to simulate the effect used in the Matrix in the model of Radford's 1984 scene, the following adjustments could be made.

Applied to each Light: R=-10 G=+10 B=-10
Applied to Ambient light: R=-10 G=+10 B=-10

11.0 Cinematography in Games

Developers of interactive 3D applications, such as computer games, are expending increasing levels of effort on the challenge of creating more narrative experiences in virtual worlds. As a result, there is a pressing requirement to automate the cinematography and develop camera control techniques that can be utilized within the context of interactive environments in which actions are not known in advance.

In contrast with films which are realised actions (the viewer has no influence on the unfolding of the story), games are framings of events and sequences (Walther 2004). Films will concentrate on the conveyance of the spatial, temporal and causal understanding of its content, together with a generally identifiable directorial and strong aesthetics flavour. Games focus more on supporting the user interaction; that is the appropriate selection of viewpoints that trigger the user decisions and the next step in the unfolding story. Interactivity in games imposes a number of constraints on the underlying camera control system, and while games have clearly drawn from cinematographic knowledge, the specificities of this interactive media make it a genre to be studied in IS.

We briefly expose the range of camera control techniques used in 3D games and we detail how cinematography has strongly influenced games.

Games will generally cover two cinematographic logics (Hancock 2002): a filmic representation that uses the potentials of interaction and a cinematic mode which covers any non-interactive section or scene element in the games (such as cut-scenes). Historically, the complexity in the simultaneous control of all camera parameters has strongly constrained the means by which 3D computer games have been influenced by cinematography. Visibility of targets and necessity to maintain playability, for example, dictated the use of specific cinematographic devices. Walther (Walther 2004) considers two general techniques adopted in games:

- Focalisation. Focalisation considers the point in space from which the scene is viewed, and its straight correlation with the viewer. Most first person shooter games (FPS) are in inner focalisation mode (the player is the viewer, and the viewer is the player). Outer focalisation modes are employed in games, where the viewer has one or more representations of himself from an outside view (Prince of Persia, Tomb Raider, and many sport games). Focalisation has been extensively used in games due to the simplicity in its implementation; inner focalisation is trivial since the character and the viewer are collocated, and in 3rd person shooter games, the camera merely orbits around the target using a spherical coordinate system. Fully automated systems presented a number of flaws in specific situations (in Tomb Raider's first version, as the character backed up on a wall the camera swivelled to display the front view of the Lara Croft and failed to focus on the opponents). In a number of cases, the user controls the camera together with the character to provide an active viewing of the environment (eg. Assassin's Creed). The mapping between the user inputs and the camera parameters remain intuitive (two-dimension orbiting operations).

- Montage. Montage in games relate to the disruption of scenes and sequences. In a major difference with films, editing in games need to be in line with the interaction possibilities of the user. Multiple montage forms are considered:
    a) *cinematographic montage* appears in games with strong narrative structures and visual expression. "*Alone in the Dark*" initiated this move by appropriately selecting

shots within a room, depending on the user location, interaction and on the context of the scene. The strength of this game came from the careful control of the pacing and tension in the unfolding, in particular through the use of many wide angle establishing shots, using dead angles and subjective views from unknown characters (see Fig. 42). Many titles were inspired from this novel genre (Resident Evil, Silent Hill).

Figure 42. Two shots from the *Alone in the Dark* game (1992). Use of wide angle shots, slow pacing and subjective views from unknown creatures creates tension and fear.

a) *long-take montage (or anti-montage)* consist in building a single continuous shot all along the game. This is a characteristic and common montage appearing in first and third person shooting games. It remains the best suited editing style for games where the user has a full interactive control of the character (displacements, shooting, etc).

b) *cut-scenes montage and automatic replays* are positioned between interactive games sequences. These rely on the full range of cinematographic techniques and serve the purpose of unfolding the global structure of the narrative, while improving the sense of the players presence in the game by displaying the consequences of the players acts (see Figure 43). Cut-scenes encompass conversation scenes, mood scenes, introduction and conclusion to plots and foreshadowing of events. Specific cut-scenes (such as instant replays) can be inserted after remarkable actions (eg. Need for Speed game). The appropriate balance between interactive sequences and cut-scenes strongly influence the playability and life-time of games (players refer to cut-scenes as "interaction killers").

c) *multiple viewpoints in one scene* is more a game-specific editing where world-in-miniature metaphors assist the players/viewers in the understanding of the spatial context (see Figure 44). Overview maps, located in the screen corners, are representative examples of multiple viewpoints. Some rare games (Fahrenheit, Heavy rain) rely on simultaneous multiple viewpoints, mimicking split-screen movies such as Time Code or the "24" TV series.

Figure 43. Extremely realistic cut-scenes from the game Call of Duty 4: Modern Warfare (Activision) that accompany and motivate the unfolding of the narrative.

A major problem of appropriating cinematic styles for games is that using the standard idioms for covering characters, the viewer is presented with a succession of very narrow views of the scene. As most games rely on far more detailed knowledge of the spatial configuration of the player's environment than in cinema, this 'blinkering' effect can make navigating the game space impossible. By using the typically wide-angle camera views common in video games however, the player loses the extra information gained through filmic composition.

To conquer this obstacle, Heavy Rain (Quantic Dream, 2010) uses a split-screen approach, allowing the designers to harness the allusive and descriptive qualities of cinematic style (in the form of composition and camera movement) while covering the scene more thoroughly in order to facilitate interaction with the characters and maintain both playability and presence (Figure 44).

Figure 44. A shot from the game "Heavy Rain" (Quantic Dream Studios). Multiple viewpoints are displayed simultaneously to show the subtle interactions between the characters, maintain the playability and the feeling of presence.

The major challenge in transposing cinematography techniques to games therefore lays in achieving the appropriate balance between aesthetics and functionality (ability to maintain the interactivity and flow of the game). Furthermore, in the context of Interactive Storytelling, the challenge is more to achieve a shift from the "*winning*" attitude of players (what is the sequence of actions I should perform to end the game) to the "*experiencing*" attitude (similar to watching movies). This in turn requires a fundamental change in the way people view and play.

12.0 Soundtrack and Interactive Media

In contemporary visual media such as films, games and television, music often forms an integral part of the presentation. Brown describes non-diagetic film music as functioning on at least 3 levels: as a 'wallpaper soporific to allay fears of darkness and silence (in cinemas)' (Brown, 1994) as an aesthetic counterbalance to the iconic/representational nature of the cinematic signs, which although they do not require music to validate the language they create as artistic, get that help anyway and as a co-generator of narrative affect that skews the viewer/listener towards a culturally determined reading of the characters and situations.

The formats and conventions of cinema have evolved slowly over a century of innovation, allowing modern film makers and audiences to draw upon a huge common vocabulary of signs and techniques when creating or reading films. Although, film music has not been subjected to anything approaching the level of theoretical attention directed toward cinematography (Kalinak, 1992), a number of composers and theorists have helped to establish and identify a number of techniques and devices which are commonly used to support or subvert meaning in the film. Modern audiences are highly educated (often without being aware of it) in the languages of film music: being able to read the narrative structure of the score almost as clearly as the plot of the film (Kalinak, 1992). This document examines some of the most commonly used techniques in film composition and examines their potential application to IS.

Interactive media, including games also draw on techniques established by the film industry, although differences in approach are necessary to take into account the different modes of consumption at work during watching films, playing games or using IS systems. Included in this document is a survey of techniques which may be used in the creation of soundtrack for interactive applications.

13.0 Classical Hollywood Approaches

A large proportion of the musical conventions applied to visual media today derive from techniques first described by film composers in the 1930s and 1940s (Kalinak, 1992). The outbreak of World War 2 caused the emigration of large numbers of classically-trained composers from Europe to the USA, where they found work in composing new music for the rapidly expanding film industry. (Bruce & Herrmann, 1985) The particular schooling of these composers along with the highly structured Hollywood Studio system with its standardising effect on film production, lead to the dominance in film of 19th Century Symphonic/Romantic styles of composition which still form the basis for most film scores today.

Initially, composers applied musicological theories to the newly emerging need for musical accompaniment. The most dominant of these was Wagner's idea of the *gesamtkunstwerk* or total art work, in which he demanded that each aspect of a production (staging, music, dialogue) should contribute to an integrated whole (Brown, 1994). In film composition, this meant that the principle task of the soundtrack should be to support the image track. In contemporary cinema, orchestral scores still dominate the soundtrack of most Hollywood films, operating within a range of commonly used conventions discussed below.

14.0 Alternative Approaches

As the Hollywood studio system imposed its own rules and conventions on mainstream film-making, many directors took different approaches, in many cases reacting against American domination of the industry on ideological grounds. The *gesamtkunstwerk* was viewed with suspicion by Adorno, Eisler and Eisenstein who saw it as an attempt to seduce the audience with an immersive and anaesthetic form of cinema (Adorno & Eisler, 1947). In direct opposition, they established a system based on counterpoint where the score would assert an opposite effect to the image track. Calm or static images would be set against stirring martial music, battles would be juxtaposed with slow orchestral scores. The intention was to cause a dichotomy in the viewer, forcing them to question the content of both image and soundtrack and take a far more critical standpoint than the type of passive consumption engendered by Hollywood films.

Other movements in cinema dispense with music altogether. The Dogme 95 movement, listed diagetic music as forbidden in their 1995 ‘Vow of Chastity’ (Simons, 2005), again, on ideological grounds, mainly as an attempt to capture a more true and direct representation of the text, in opposition to the sense of illusion created by Hollywood cinema.
Other composers such as Philip Glass, Michael Nyman and more recently Clint Mansell’s scores for the films of Darren Aronofsky use classical ensembles but incorporate modern musical techniques, for example: *Reqiuem for a Dream*’s (Aronofsky, 2001) score performed by the Kronos Quartet uses constantly repetitive and evolving passages in-line with Aronofsky’s repeating high speed montages to highlight the cyclic processes the characters undergo over the course of gradual transformation during the narrative.

15.0 Techniques and Conventions: supporting the diagesis with music.

15.1 Style and Allusion

Although sharing a common musical vocabulary, classical approaches to film composition often use particular styles of music to support the diagesis. This may impact on instrumentation and orchestration, rhythmic, melodic and harmonic structures and motifs. Allusion and intertextuality is used extensively in the work of many composers either to reinforce the image track or to subvert it by offering an alternative reading (Kalinak, 1992). By working a recognisable phrase from another piece of music into the score, or even using an original recording, the composer can quickly reference a theme from outside the diagesis (such as a character's background) or reinforce a sense of period or location. This technique is often reinforced with appropriate instrumentation.

In *Ferris Bueller's Day Off* (Hughes, 1987), a slow motion shot of the priceless borrowed Ferrari flying through the air and over the camera is juxtaposed with John Williams' score for *Star Wars* (Lucas, 1977). The allusion to another familiar text reinforces the image track (*Star Wars* famously begins with a similar shot with a spaceship flying overhead) to comedic effect, also referencing a more abstract sense of excitement, adventure and fantasy.

Figure 5 The visual allusion to Star Wars (right) is reinforced by a passage from its Soundtrack in Ferris Bueller's Day Off

Trevor Jones soundtrack for Michael Mann's *The Last of the Mohicans* (Mann, 1992) although performed by a conventional orchestra is based on Dougie Maclean song, *The Gael*, written in a style reminiscent of Scottish folk music and contains folk passages. This allusion to 18th and 19th century folk traditions reinforces the film's period settings.

15.2 Leitmotifs

Leitmotifs are distinct musical phrases, recurring throughout the score of a film, the principle purpose of which is to reference a particular theme or character within the diagesis. A popular technique in the works of many 19th composers, the leitmotif is used extensively in Opera (Kalinak, 1992). Leitmotifs may range from a few notes to whole passages and may be repeated on different instruments, in different keys and in different meters throughout the film. Crucially, leitmotifs are not only used to reinforce a character's presence onscreen but can be used to refer to that character when he/she is absence. Certain composers apply this approach so consistently that the music becomes a constant commentary on the image track, mimicking the individual movements and expressions of characters, a technique known as 'Mickey Mousing' (Brown, 1994).

In John William's soundtrack to *Jaws* (Spielberg, 1975), a simple two note motif announces the presence of the shark at various points in the narrative. This is established during the opening titles

with an extended theme, dominated by the repeating sequence of notes, C, D.

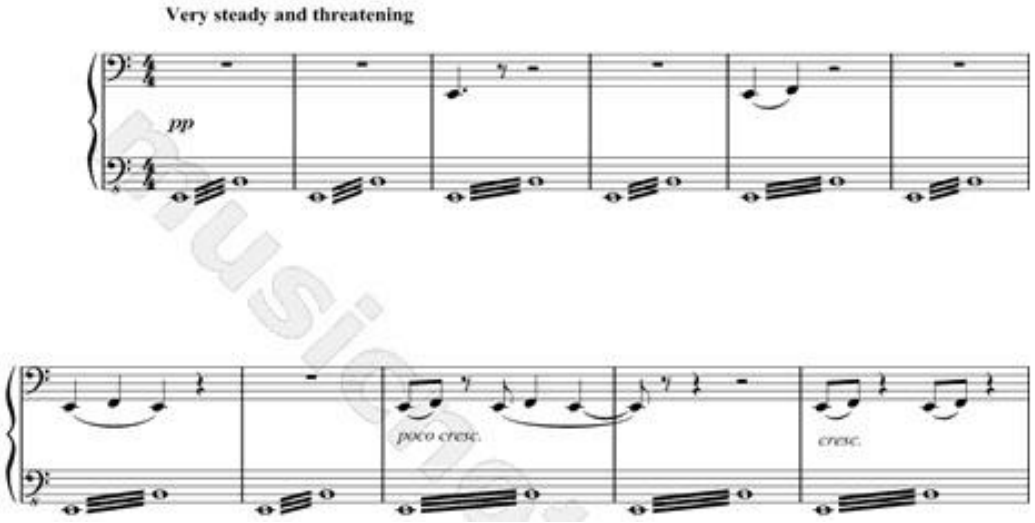


Figure 42 The Introduction to Jaws

15.3 Rhythm, Tempo and Timing

Changing the tempo of the score can radically affect the viewer's perception of the pace of the film and in conjunction with editing can convey a variety of effects. Rhythm and meter can often achieve similar results: for example doubling the speed of the percussion from ½ notes to ¼ notes. In *Jaws*, the main theme begins slowly played as ¼ notes is played as 1/8 notes as the shark approaches the swimmer. Although the tempo remains constant, the impression is one of increasing speed.

15.4 Dynamic Range

Rises and falls in the relative volume of parts of the soundtrack or the soundtrack's volume in comparison with dialogue and sound effects can be used to intensify or highlight particular themes and are extremely effective in generating tension. Complete silence in the score can be used in contrast with scored sections of the film to lend an immediacy and heightened sense of realism to a scene.

In the opening scene of *Jaws*, the swimmer's death in the water is accompanied by variations on the main theme at high volume. At the climax of the sequence, this is intercut with an almost silent shot of her boyfriend asleep on the beach, with no accompaniment. When the action cuts back to the shark attack, the sudden jump in volume serves to renew the sense of shock and violence central to the sequence.

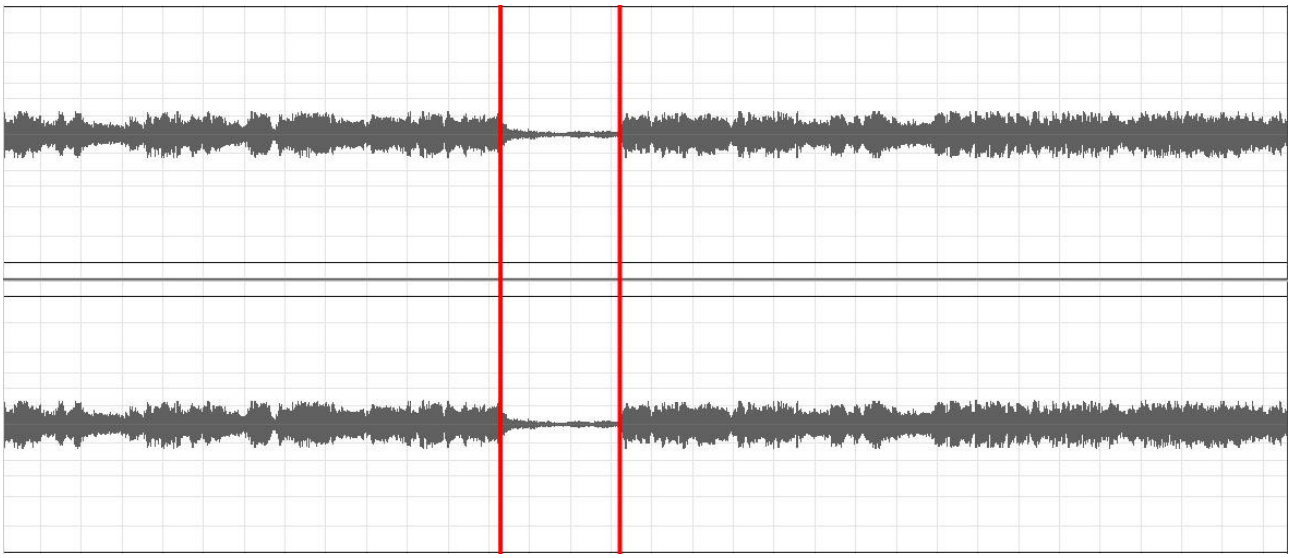


Figure 6 Excerpt from Jaws, the volume of the soundtrack drops abruptly as the swimmer's death scene is intercut with her boyfriend sleeping on the beach

15.5 Pitch and Timbre

Since the Common Practice Period (1600-1900), Western music has been dominated by the tonal system (Brown, 1994): a hierarchical approach to pitch, placing emphasis on each note's relationship to a tonic or central pitch. In the vast majority of western popular and classical music, major and minor scales are used to determine which notes are included within the structure of the piece (i.e. in the same key) and which should be considered outside it. Consonance or dissonance within this system is a powerful tool for suggesting harmony and discord, tension and release in a narrative. As the major and harmonic minor modes are so dominant, use of other scales, even by conventional ensembles can be powerful in terms of allusion. Relationships between pitches can be used to create specific effects or to refer directly to spatial elements of the drama, for example, slowly rising glissandos can generate tension while descending notes can mimic a character physically falling (Kalinak, 1992).

In the murder scene in Psycho (Hitchcock, 1960), a staccato series of dissonant chords underpins the stabbing in the shower, mimicking the action of the knife and heightening the uncomfortable quality of the image.

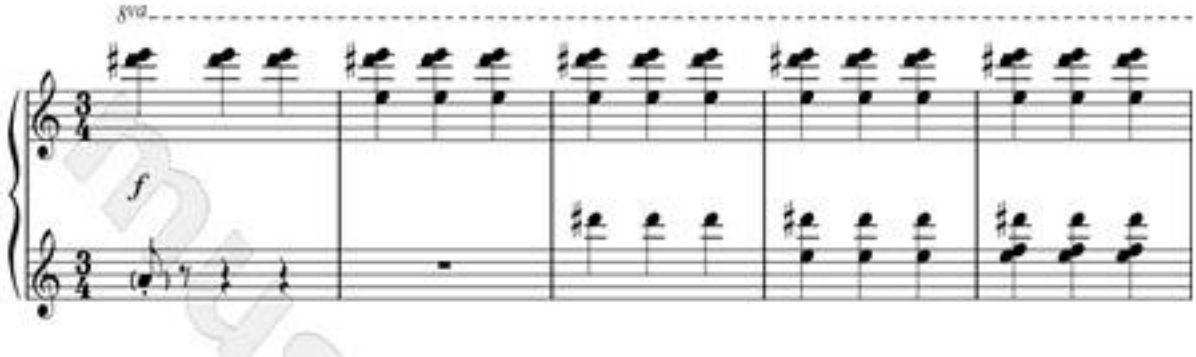


Figure 43. The theme from psycho, using dissonance to produce unsettling effects.

15.6 Instrumentation and Orchestration

Particular choices of instrumentation can be used to achieve several effects. Specific instruments, either used in isolation or with the backing of a more conventional ensemble can be used to make an explicit allusion to a particular environment, era or setting. A common technique used in film scores to set the scene is to include an instrument perceived by the audience as indigenous to the context of the film's settings or characters, for instance the use of instruments such as *Sao* bamboo flutes in Hollywood films to emphasise references to Vietnam.

The scale of orchestration can reinforce changes in the plot, especially with regards to tension and action. The replacement of a full orchestra by a single instrument can suggest a tightening of the frame of action from a crowd to an individual.

16.0 Soundtrack for Interactive Media

Many of the theoretical bases and practical approaches for dynamic soundtrack in interactive media come from the videogame industry. A large number of action, adventure and role-playing games borrow heavily from cinema in their approach to soundtrack, especially in terms of style. Since the advent of 32bit games consoles and modern high-quality sound cards, game designers have been able to commission orchestral soundtracks composed and recorded in the same commercial frameworks as film music (Collins, 2008). However, as discussed below, the differences between the way that films and games are consumed often necessitate a different approach to composition. When composing for films, the composer usually knows (as discussed previously) many of the dimensions of the final piece such as length, context, points of particular emotional content, pace, overall structure and moments of tension and release are already established, to a greater or lesser extent by the image track.

When composing for interactive media, many or all of these factors may be extremely flexible. While a conventional linear film plays over a fixed duration, in most games, the duration of the whole game and the episodes or levels of which it is comprised can vary from play to play (Collins, Game Sound, 2008). Characters often appear unpredictably and although moments of tension can be controlled to some extent at the highest level (by placing bottlenecks in the game environment through which the player must pass) these can rarely be predicted accurately enough to synchronise with a pre-recorded soundtrack.

16.1 Non-dynamic Approaches

Many video games still rely on a score composed of simple non-interactive audio tracks, which are triggered by high level events in the game. The overwhelming advantage of this simple approach is that music can be either commissioned without any precise temporal requirements or can be selected from the catalogues of record companies. The soundtrack of many video games, especially non-narrative genres such as racing games is often comprised of pop tracks cued up by a simple playlist (Collins, Game Sound, 2008).

Although this approach stops any detailed temporal link to the action, it can still be used to generate mood and authenticity in the game. The Grand Theft Auto series, in which the player spends large amounts of time exploring the city by car, relies on a semi-diagetic score realised through radio

stations which the player can tune into via the stereo systems of the vehicles he drives. These radio stations are composed for the most part of familiar hip hop and rock tracks but are interspersed with recorded monologues by the 'DJ' commenting on events in the game world, commercials and news flashes relating to the players progress.

16.2 Dynamic Approaches

Collins lists 10 approaches to variability in music for interactive media (Collins, Game Sound, 2008 p147)

- Variable tempo
- Variable pitch
- Variable rhythm /meter
- Variable volume/dynamics
- Variable DSP timbres
- Variable melodies
- Variable harmony
- Variable mixing
- Variable form (open form)
- Variable form (branching parameter based music)

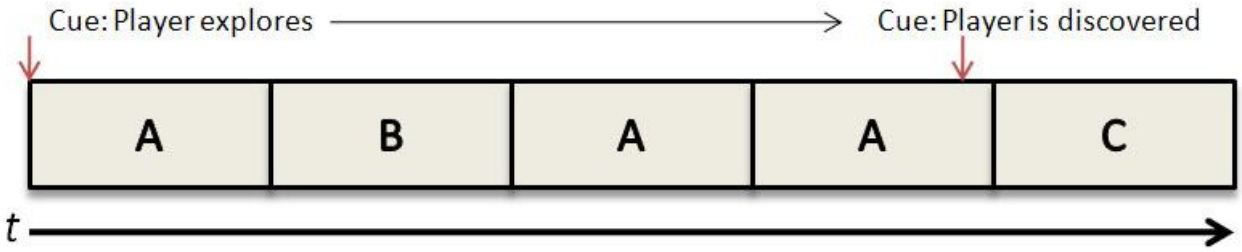
All or some of these may be used through a number of techniques in order to generate soundtracks which adapt to the unpredictable nature of interactive media including IS.

Each of the approaches below necessitate an audio engine able to accept inputs from the narrative in some way. Typical audio engines such as FMOD (FMOD, 2010) include the ability to play recorded tracks in time, may also have synthesiser capabilities (either using sample based synthesis or generative techniques such as FM) and, in the case of modern audio engines can often include real-time Digital Signal Processing.

16.2.1 Horizontal resequencing

With horizontal re-sequencing, a number of pre-recorded segments (often of a uniform length and time signature) are re-combined in different combinations (Collins, Game Sound, 2008) according to the dramatic requirements of each area or episode of the game, allowing variations in rhythm meter, and form. Typically, fast, energetic passages can be replaced with slow, reflective sections on the fly, with the audio engine reacting to preset cues in the games environment.
An obvious drawback of this approach is small number of possible combinations, meaning that either a large set of recordings must be used or the soundtrack will seem repetitive. Also, as each loop must be able to follow others in different orders, changes in key, tempo or instrumentation are difficult to effect without obvious transitions.

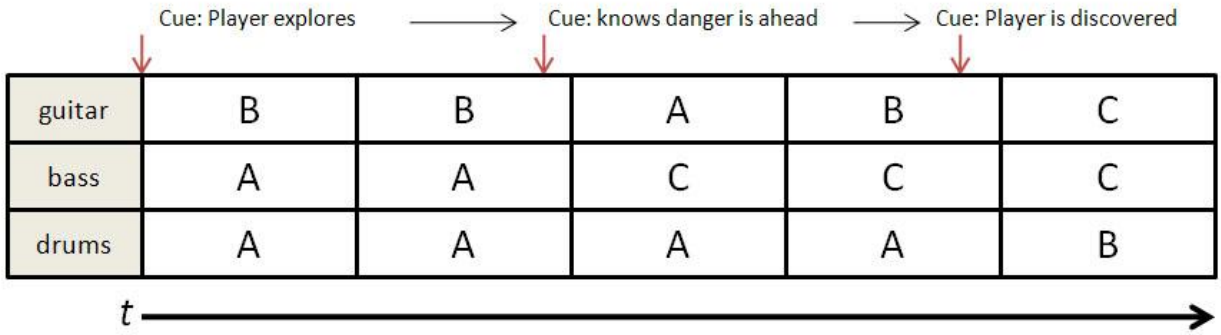
In a typical game sound scheme, a passage derived from a recorded ensemble piece is divided into 16 bar parts A and B (generic music suggesting exploration) and C (a more dramatic theme as the player is discovered). A and B can be combined randomly as the character explores the level. C is cued by triggering the alert state of NPCs. Importantly, C only starts to play at the end of the loop, avoiding a jarring break in tempo.



16.2.3 Multi-track resequencing

Many of the limitations of single-track re-sequencing can be overcome by layering segments on top of each other. By using separate tracks for individual instruments, a far greater number of combinations can be achieved, with more subtle variations between passages, moreover, branching structures can be used to ensure that the structure of the music is always appropriate for the changing context of the game or story.

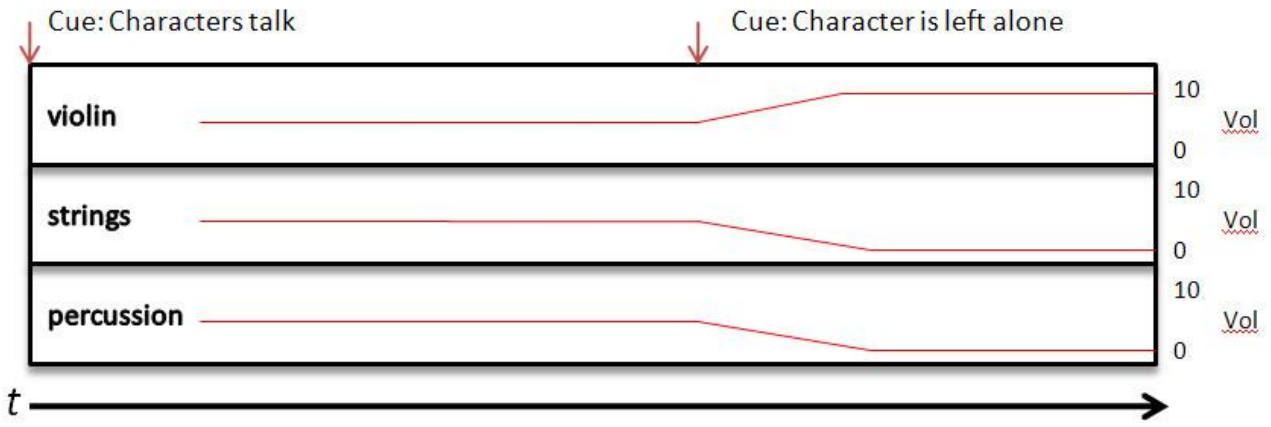
In the example below, parts A, B and C consist of recordings, for each instrument. During the exploration state, each instrument track can play A or B, allowing 8 possible 16 bar combinations. When the player draws within a certain radius of an NPC, one instrument at random must play C. When the player is discovered 2 instrument tracks must play C.



16.2.4 Vertical Reorchestration

Vertical re-orchestration involves a single linear recording divided into individual instrument tracks playing in sync. These individual tracks are faded in and out of a multi-channel mix, allowing lines to come to the front. This approach is particularly appropriate to large ensemble scores and has been used extensively in games with high-production values such as Bungie's Halo series.

Fig shows 3 instrument tracks representing strings, solo violin and percussion. The red lines represent the volume envelope. As the character is left alone, strings are faded down and percussion muted altogether, leaving only the violin audible. The other tracks continue to play and at a later point may be faded up.



16.2.5 Generative approaches

Common musical approaches used in games usually involve the restructuring of pre-recorded music. Generating musical scores on the fly represents a completely different approach from the combinatorial systems discussed previously as it involves to a greater or lesser extent replacing musician's performance with a set of rules and structures which are explored by the audio engine in real-time. This approach also necessitates real-time synthesis capabilities for the audio engine: either for generating its own tones (via fm synthesis for example) or by the manipulation of samples.

To produce melodies on the fly, generative approaches may include random or probabilistic techniques such as the use of Markov chains to generate musical phrases and melodies. By using chains in which each note can potentially be followed with another note based on a certain probability, generative melodies can be constructed, where changes in the ruleset can lead to variations in form, melody and harmony. (Collins, Game Sound, 2008)

16.2.6 Digital Signal Processing

To add further variety to their audio, many contemporary games use the digital sound processing capabilities built into most modern PC sound cards and next generation consoles to affect the final output of the soundtrack. Real-time audio effects and spatialisation of the soundtrack can be used to form a link from an otherwise non-diagetic soundtrack to the environment in which the diagesis takes place, for example adding a hall-like reverb effect to an orchestral score as the player explores an echoing cavern.

16.2.7 Mixing and dynamic range

The musical score of a piece of interactive media such as a game is typically experienced in an extremely cluttered musical environment, populated with spot effects, ambient sound and importantly dialogue. It is extremely important that the score takes this into account. Typically in game scores, individual samples are recorded at a far more uniform level than in films, where the location of peaks in dialogue or sound effects are already known, allowing entire parts of the score to be accurately controlled. This can lead to a 'cramming' of dynamic headroom, as the score is dominated by solid blocks of sound.

17.0 Interaction between Narrative and Soundtrack in IS

In practice, most game music systems use a combination of approaches where for instance, individual loops are combined in a multichannel scheme allowing huge combinatorial possibilities. In Fig, individual instrument tracks are comprised of multiple looping sequences, allowing subtle changes in instrumentation or completely new passages to begin (Collins, From Pac-Man to Pop Music, 2008).

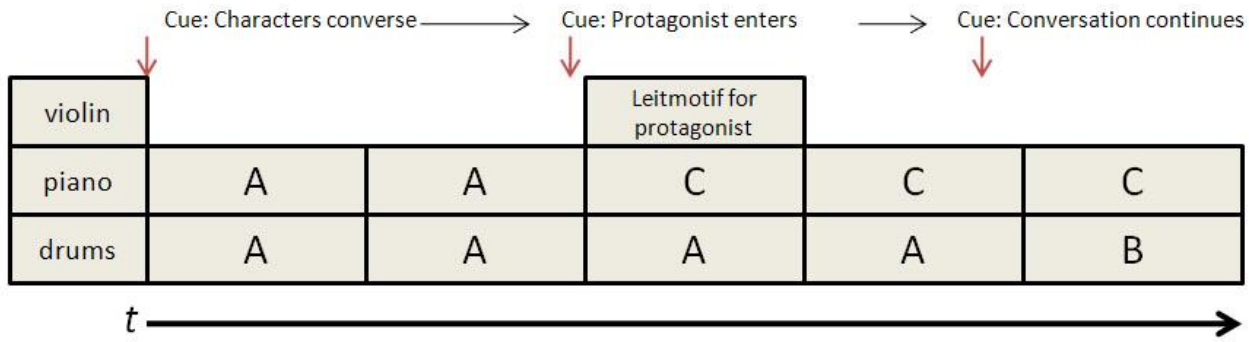
As with many elements of IS systems, constructing a coherent narrative in which the music supports the dramatic elements relies heavily on effective communication between the narrative engine (in whatever form) and the audio engine. The specification of cues by which the audio engine can be controlled by the drama are all important. Unlike in games, in some IS systems the length and structure of the narrative may be known, eliminating the need for total flexibility of form in the soundtrack. On the other hand, in other more character-driven systems, the emotional context of the drama may be far more subtle and unpredictable than even the most elaborate and flexible sandbox game.

In the production of most video games, it is usual for the sound designer and composer to 'spot' the game's design document, noting key aspects including objects, environments, characters and crucially emotional states(Collins p92). Simple approaches such as identifying points of tension and release and plotting graphs of possible rises and falls in the pace or intensity of the drama could form useful cues for the soundtrack.
All of the approaches discussed rely on communication between the narrative engine and the audio engine, however this can implemented at different parts of the system at runtime. In most applications, it is desirable that narrative events effect a state change or produce triggers to alter the behaviour of the audio engine, communicating in one direction only (fig).

17.1 Suggested approaches for soundtrack in IS

17.1.1 Approach 1

Audio behaviours may be set at a planning level, triggered by events in the narrative. In this case specific cues can be inserted into the planning system, or more usefully, the audio system can respond to specific classes of events. In the example below, certain events are cued based on characters entering the dialogue (leitmotif phrases are included), arguments beginning (rising tremolo strings volume)



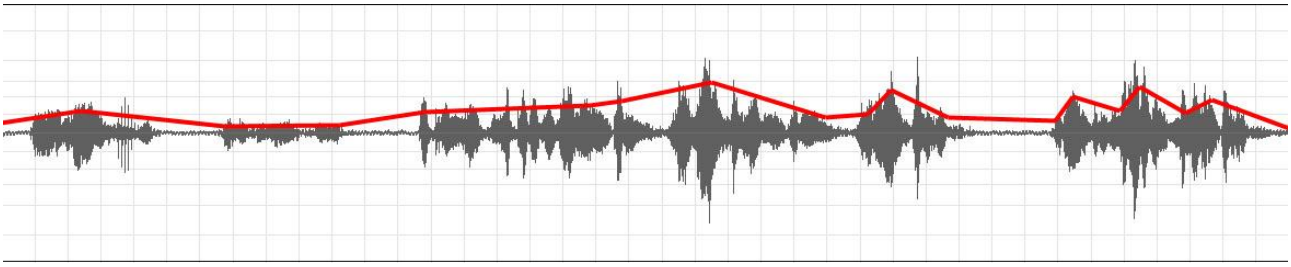
17.1.2 Approach 2

Alternatively the audio might be specified directly alongside cinematographic considerations.

In the example below (fig) using Semantic Volumes system, a system of vertical orchestration is combined with cinematographic constraints specifying growing tension between two characters. In the cinematic solver, this is realised by tightening the frame on the characters in successive shots. As each shot is selected, the audio system is cued to increase the volumes on track 3 (tremolo strings) and decrease the volume on track (piano) over a short period, ensuring that the soundtrack and image track are wholly integrated.

17.1.3 Approach 3

In IS systems where recorded dialogue is used, simple DSP analysis provides cues based on dynamic range of utterances, inferring tension or heightened emotional state from a combination of context and volume. Care must be taken to eliminate anomalous sounds such as loud percussives, however these may be discarded by a simple damping algorithm, averaging over a short period of time.



17.1.4 Approach 4: Realtime narrative responsive soundtrack

As discussed, in the vast majority of approaches to non-diagetic music in real-time scenarios, the audio engine is controlled in line with the image track. An alternative approach which may be appropriate for some interactive applications involves reversing this relationship.

A recent digital art piece, *OverWatch* (Schofield) presents a scenario where musical performance directs the projected visuals, determining content, action and editing. *OverWatch* is a narrative piece, in which the action is largely figurative and staged according to cinematographic conventions. However, the narrative is driven not by the actions of actors within the frame of the image but by the musical interaction of the performers outside it. This approach explores the boundary between diagetic and non-diagetic approaches to music in film and introduces a third possibility: music which does not exist within the context of the story or as an accompaniment but actually defines it, scripting the narrative in real-time. Techniques borrowed from film music such as style topics and leitmotifs are used in *OverWatch* to cue animations and camera movements.

OverWatch is a live performance piece designed primarily for a theatre setting. Two musicians occupy the stage area, beneath a large projection screen. As they play, live Audio and MIDI streams from synthesisers, an electric guitar, bass and microphones are transmitted via a USB audio interface into a laptop. This information is first processed in Ableton Live and then passed to MAX/MSP which analyses and translates it into a stream of OSC messages. This stream is then transmitted via a UDP network connection to a second laptop running a 3D graphics engine (Blender GE) which uses these messages to control not only much of the action in the environment but camera movements, framing and cutting.

Scenes and active elements within the visuals are introduced using *leitmotifs:* musical structures which are read as combinations of MIDI messages and are used to trigger loading or visibility of objects. The pace of the editing is set by counting MIDI clock messages and adjusting cuts

according to the resulting speed. Camera switching is affected by various instrumental effects depending on context: sometimes changes of chord, mode or key. Camera zoom and small camera movements are affected by the velocity of MIDI notes and amplitude of audio signals, connecting intensity in the score to psychological intensity in the narrative. Discord/harmony is calculated by comparing the notes played to the perceived key signature and is connected to the implementation of unsettling techniques: unbalanced compositions, slight camera roll and tilt etc.

18.0 Conclusions

Many of the camera techniques discussed here arise from physical features of real-world film cameras which must be deliberately simulated in computer graphic environments in order to create similar effects. In these cases the normal, often unwanted responses of real-world cameras to the physical world manifested as camera-shake and instability or loss of focus must be consciously built into the system.

Real-world cameras have weight, momentum and a physical presence which extends behind and around the lens. They must be supported and moved, protected from impacts and vibration, constantly fed with electricity, tape or film and kept clean if they are to function at all. To achieve smooth movement necessitates an experienced operator, while to elevate the camera more than a few metres above the ground or move it at more than a few kph requires expensive and complicated machinery. In contrast, cameras in 3D computer graphic environments have no weight, are invisible and can be moved smoothly in any direction, including through objects, consequently many of the techniques discussed here are far easier to achieve in 3D CG environments than in the real world.

Although often described as a synthetic medium, cinema differs from IS systems in that it usually holds a connection (however tenuous) to a real event or performance. The route from the actions performed on the film set to the cinema or television screen may be extremely tortuous, involving editing, effects, post-processing, colour grading and the changes effected by re-presenting the final result in a cinema but the connection to a real performance is often important.

Because of this distinction, many techniques used by contemporary directors must be carefully considered if they are not to lose their meaning when used in non-film contexts. The aspirations of the Dogme95 movement in particular, with their view of cinema as the recording of a performance take on a completely different meaning if applied to 3D CGI environments where there are no human actors and no performance outside the frame of the camera. Several of the effects discussed here are in fact allusions or simulations themselves, designed to make connections from narrative film to other styles.

Nevertheless, many of the same structures of reference and allusion used by film maker, if used carefully, can work well for authors of Interactive media including IS. In the case of documentary styles for instance, although the viewer understands perfectly well that in an Interactive Story there is no real camera or human subject, the sense of immediacy, dynamicity and the sense that the camera has some physical presence in the scene remains unchanged. Time manipulation techniques such as time ramping may lose some of the spectacular quality in the transition to CG but still retain their ability to render legible rapid movement. Colour symbolism works perfectly well in CG environments (as demonstrated by many video games). If used properly, many techniques used in modern cinema can be translated into powerful tools for makers of Interactive Stories.

In approaching soundtrack, makers of Interactive Stories are able to use the vast majority of techniques described here unaltered. As non-diagetic music by its very nature holds no connection to an original performance, the structures by which is applied to cinema should function perfectly well in interactive contexts, or to support a computer graphic image track, rather than a film of actors' performances.

Bibliography

Adorno, T., & Eisler, H. (1947). *Composing for the Films.* New York: Oxford University Press.
Arijon, D. (1976). *Grammar of the Film Language.* New York: Hastings House Publishers.
Aronofsky, D. (Director). (2001). *Requiem for A Dream* [Motion Picture].
BBC (Director). (1985-2010). *Eastenders* [Motion Picture].
Blomkamp, N. (Director). (2006). *District 9* [Motion Picture].
Brown, R. S. (1994). *Overtones and Undertones, Reading Film Music.* London: University of California Press.
Bruce, G., & Herrmann, B. (1985). *Film Music and Narrative.* Ann Arbor: University of Michigan Press.
Cohen, R. (Director). (2001). *The Fast and The Furious* [Motion Picture].
Collins, K. (2008). *From Pac-Man to Pop Music.* Aldershot: Ashgate.
Collins, K. (2008). *Game Sound.* Cambridge: MIT Press.
Core (Director). (1996). *Tomb Raider* [Motion Picture].
Crichton, M. (Director). (1995-2007). *ER* [Motion Picture].
Dream, Q. (Director). (2010). *Heavy Rain* [Motion Picture].
Eduardo Sanchez, D. M. (Director). (1999). *The Blair Witch Project* [Motion Picture].
Epic (Director). (2006). *Gears of War* [Motion Picture].
FMOD. (2010). *products*. Retrieved June 10, 2010, from www.fmod.org: http://www.fmod.org
Frank Miller, R. R. (Director). (2005). *Sin City* [Motion Picture].
Hitchcock, A. (Director). (1960). *Psycho* [Motion Picture].
Hughes, J. (Director). (1986). *Ferris Bueller's Day Off* [Motion Picture].
Hughes, J. (Director). (1987). *Ferris Bueller's Day Off* [Motion Picture].
Jones, D. (Director). (20009). *Moon* [Motion Picture].
Kalinak, K. (1992). *Settling the Score, Music and the Classical Hollywood Film.* Madison: University of Wisconsin Press.
Katz, S. (1991). *Film Directing Shot by Shot.* Michael Wiese Productions.
Larry Wachowski, A. W. (Director). (1999). *The Matrix* [Motion Picture].
Lucas, G. (Director). (1977). *Star Wars* [Motion Picture].
Mann, M. (Director). (1992). *The last of the Mohicans* [Motion Picture].
Marta Kauffman, D. C. (Director). (1994). *Friends* [Motion Picture].
Mascelli, J. (1965). *The Five C's of Cinematography: Motion Picture Filming Techniques.* Hollywood: Cine/Grafic Publications.
Miller, F. (Director). (2008). *The Spirit* [Motion Picture].
N Halper, P. O. (2000). A Camera Planning Agent. *Smart Graphics AAAI Spring Symposium*, (pp. 92-100).
Ndaliansis, A. (2000). Special Effects, Morphing Magic and the 1990s Cinema of Attractions. In V. Sobchack, *Meta-morphing: Visual Transformation and the Culture of Quick Change* (pp. 251-271). Minneapolis: University of Minnesota Press.
P Burelli, L. D. (2008). Virtual Camera Composition with Particle Swarm Optimization. *Proceedings of SG 2008: 8th International Symposium on Smart Graphics, Lecture Notes in Smart Graphics* (pp. 130-141). Springer Verlag.
Radford, M. (Director). (1984). *1984* [Motion Picture].
Reeves, M. (Director). (2008). *Cloverfield* [Motion Picture].
Schofield, G. OverWatch. *www.theturingtest.co.uk/guyschofield*. Newcastle University, Newcastle upon Tyne.
Schwartz, J. (Director). (2003). *The OC* [Motion Picture].
Simons, J. (2005). Playing The Waves. In M. H. Marijke de Walck, *Cinephilia: movies, love and memory* (pp. 181-196). Amsterdam: Amsterdam University Press.

Snyder, Z. (Director). (2006). *300* [Motion Picture].
Spielberg, S. (Director). (1975). *Jaws* [Motion Picture].
Thompson, D. B. (2006). *Film Art: An Introduction (8th Edition).* McGraw-Hill.
W Bares, S. M. (2000). Virtual Camera Composition from Frame Constraints. *Proceedings of the ACM International Conference on Multimedia (ACMMM 00)* (pp. 177-186). ACM Press.
Walther, B (2003) Cinematography and Ludology: In Search of a Lucidography, *Papers from spilforskning.dk Conference*, august 28.-29.