Université de Montréal

Amélioration de la perception spatiale et émotive d'un environnement virtuel interactif par la gestion cinématographique des prises de vue

par

Marie-Elise Cordeau

Département d'informatique et de recherche opérationnelle

Faculté des arts et des sciences

Mémoire présenté à la faculté des études supérieures

en vue de l'obtention du grade de

Maître ès sciences (M.Sc.)

en informatique

Décembre 2007

© Marie-Elise Cordeau , 2007

Université de Montréal

Faculté des études supérieures

Ce mémoire de maîtrise intitulé

Amélioration de la perception spatiale et émotive d'un environnement virtuel interactif par la gestion cinématographique des prises de vue

présenté par

Marie-Elise Cordeau

a été évalué par un jury composé des personnes suivantes :

El Mostapha Aboulhamid
président-rapporteur

Victor Ostromoukhov
directeur de recherche

Pierre Poulin
codirecteur de recherche

Yann-Gaël Guéhéneuc
membre du jury

Sommaire

Les jeux vidéo utilisent l'image cinétique pour communiquer avec leur utilisateur. L'introduction de techniques cinématographiques reconnues dans la composition de cette image peut donc ajouter une dimension émotionnelle à la communication, de façon à améliorer l'expérience de jeu de l'utilisateur. La majorité des jeux vidéo optent pour une caméra subjective, aussi appelée *point of view* ou *first person*, pour immerger l'utilisateur dans leur monde virtuel. Bien que ce type de caméra facilite certaines actions, il est limité pour l'expression d'émotions dans l'image. Ce mémoire de maîtrise couvre donc la production automatique de séquences de plans non subjectifs représentant le personnage, dans lesquels trois émotions sont transmises : la perception spatiale de l'environnement, l'identification psychologique avec le personnage, et un niveau de tension prédéterminé. La Caméra *SIT* développée permet ainsi d'appliquer des procédés cinématographiques documentés à l'image du jeu, en associant des pondérations émotionnelles aux différentes caractéristiques du plan, et de sélectionner pour chaque plan le cadrage qui représentera le mieux le contexte émotionnel complétant la séquence en cours. La visibilité du personnage est aussi considérée dans la solution de la Caméra *SIT*, de façon à éliminer les prises de vues montrant un personnage occulté. Les séquences produites incluses dans ce mémoire montrent les choix effectués par la Caméra *SIT* pour résoudre l'équilibre entre la perception spatiale et l'identification avec le personnage, et pour exprimer une tension donnée.

Mots clefs :
Infographie, Cinématographie, Contrôle de caméra, Environnement virtuel interactif.

Abstract

Video games communicate with their users through moving images. Compositing these images with cinematographic techniques can contribute an emotional dimension to the communication, thus improving the overall user's gaming experience. Most video games use the subjective camera, also called point-of-view camera or first-person camera, to immerse the user in their virtual world. Although this type of camera facilitates some game actions, it is limited for expressing emotions in images. This research addresses the automatic production of sequences of nonsubjective shots showing the character in action. Three emotions are targeted in these shots : space perception of the environment, psychological identification with the character, and tension level. The developed *SIT* Camera applies documented cinematographic techniques to game images by associating emotions to weights on various shot properties. It determines for each shot the framing best fit for the emotional context of the sequence in progress. Estimated visibility is considered to eliminate shots of an occluded character. The resulting sequences show the choices carried out by *SIT* Camera in order to keep the balance between the perception of space and the identification with the character, while expressing a desired tension.

Keywords :

Computer graphics, Cinématography, Camera positioning, Virtual interactive worlds.

Table des matières

Remerciements **xi**

1 Introduction **1**

1.1 Contexte d'application 1

1.2 Portée du projet 3

1.3 Spécifications de la Caméra *SIT* 5

2 Éléments de cinématographie **8**

2.1 Références 8

2.2 Le plan 8

2.3 Échelle de plan 9

2.4 Position de caméra 11

2.5 Angle de cadrage 12

2.6 Composition de l'image 13

2.7 Mouvement de caméra 14

2.8 Le montage par continuité 15

3 Travaux reliés en infographie **18**

3.1 Approche idiomatique 19

3.2 Approche spatiale 21

3.3 Approche émotionnelle 23

3.4 Discussion 26

4 Implantation de la Caméra *SIT* **28**

4.1 Fonctionnalités de la Caméra *SIT* 28

4.2 Intégration dans *Crystal Space* 29

4.3 Fonctionnement de la Caméra *SIT* 32
4.4 Implantation des paramètres de la caméra 36
4.5 Règles d'esthétique 38
4.6 Vérification de la visibilité 40
4.7 Cadrage du personnage 44
4.8 Variables émotionnelles *SIT* 46
4.8.1 Variables émotionnelles I et T 47
4.8.2 Variable émotionnelle S 51
4.9 Sélection du prochain plan 58
4.10 Gestion de l'axe de jeu 61
4.11 Transitions 64

5 Présentation des résultats 67
5.1 Méthode utilisée 67
5.2 Analyse d'une séquence communiquant une tension nulle 70
5.3 Analyse d'une séquence communiquant une tension élevée 78
5.4 Temps de calcul 81

6 Conclusion 91
6.1 Résumé de la méthode implantée 91
6.2 Évaluation de la solution 92
6.3 Travaux ultérieurs 94

Bibliographie 100

Table des figures

2.1 (a) Plan d'ensemble, (b) Plan moyen, (c) Plan américain [Int84]. 9

2.2 (a) Plan rappoché, (b) Gros plan, (c) Très gros plan [Int84]. 10

2.3 Positions de caméra relatives à l'axe de jeu défini entre les personnages X et Y [Sta07]. 11

3.1 Champ-contrechamp dans le cadre d'un idiome de conversation à trois. Images tirées de [HCS96]. 19

4.1 Environnement de jeu, château médiéval. 30

4.2 (a) Échelles de plan, (b) Angles de vue appliqués à une échelle de plan. 32

4.3 (a) Angles d'élévation, (b) Angles de vue. 33

4.4 (a) Plan d'ensemble, (b) Plan moyen, (c) Plan rapproché, (d) Gros plan. 33

4.5 (a) Angle 0 °, (b) Angle 45 °, (c) Angle 90 °, (d) Angle 135 °, (e) Angle 180 °. 34

4.6 (a) Élévation neutre, (b) Élévation positive, (c) Élévation négative. . . . 35

4.7 Visibilité du personnage. 40

4.8 Calcul de la distance entre la position initiale et la position centrée du personnage. 44

4.9 (a) Fonction de Verhulst, (b) Vecteur de spatiosité. 50

4.10 Points de spatiosité. 51

4.11 (a) Valeurs de SI en fonction de l'élévation, (b) Valeurs de S en fonction de l'élévation, (c) Valeurs de T en fonction de l'élévation. 60

4.12 Côté de la caméra suite à un changement de direction d'un personnage. 62

5.1 Secteurs de l'environnement de jeu. 68

5.2 Séquence *SIT*, tension=0. L'abscisse des graphes représente la séquence de plans. 71
5.3 Séquence semi-aléatoire. 73
5.4 Échelles de plan, séquence "The Birds". 73
5.5 Statistiques de connaissance spatiale. 74
5.6 Séquence *SIT*, tension=0.4. 77
5.7 Séquence *SIT*, tension=0.2. 78
5.8 Séquence *SIT*, tension variable. 78
5.9 Séquence *SIT*, tension=0. 82
5.10 Séquence *SIT*, tension=0. 83
5.11 Séquence *SIT*, tension=0. 84
5.12 Séquence *SIT*, tension=0. 85
5.13 Séquence *SIT*, tension=0. 86
5.14 "Psycho" (Alfred Hitchcock, 1960). 87
5.15 "Taxi driver"(Martin Scorsese, 1976). 88
5.16 Séquence *SIT*, tension=0.4. 89
5.17 Séquence *SIT*, tension=0.2. 90

Liste des tableaux

4.1 Configuration des échelles de plan. 37
4.2 Pondérations de S, I et T par échelle de plan. 50
4.3 Pondérations de I et T par angle de vue. 50

Remerciements

Bien qu'un mémoire de maîtrise soit généralement associé au travail d'un seul individu, sa réalisation comprend l'assistance des diverses personnes l'entourant. Je désire ici prendre le temps de remercier chaleureusement tous ces collaborateurs qui m'ont grandement aidé à atteindre mes objectifs.

J'aimerais tout d'abord souligner le support de mes directeurs Victor Ostromoukhov et Pierre Poulin qui ont accepté de me diriger et sans lesquels ce mémoire n'aurait jamais vu le jour. Remerciement tout particulier à Pierre Poulin qui, malgré son emploi du temps chargé, a su me faire profiter de son perfectionnisme, et a ainsi grandement contribué par ses conseils à améliorer la qualité de ce mémoire.

Je désire aussi remercier Bernard Perron et Caroline San Martin qui m'ont permis d'accroître et de formaliser mes connaissances cinématographiques et de développer mon esprit critique en la matière.

Remerciements à Ubisoft pour le soutien financier et à Rogers qui a motivé et en partie financé ce retour aux études et cette réorientation partielle de carrière.

Remerciements chaleureux aux membres du LIGUM et ses visiteurs : Simon Bouvier-Zappa, Yann Rousseau, Di Jiang, François Duranleau, Luc Leblanc, Philippe Beaudoin, Fabrice Roussel, Cynthia Papon, Nils (que j'ai très peu connu), David Vanderhaeghe, Myriam Lasserre, Noam, Romain Pacanowsky, Jean-Fraçois Dufort, Neil Stewart, Michelle Laprade, Eric Bourque, Vincent Nivoliers, Mathieu Gauthier, Dimitrios Touloumis, Frédérik Rozon, Wu Zhe et Eric Chang Jianghao. Vous m'avez gratifié de votre aide et de votre amitié et avez grandement contribué à égayer mes journées tout au long de cette trop courte étape. Merci pour les belles discussions, les parties de munchkins, Fantasia, les bières et les repas partagés, le café piston, Xkcd.

De façon générale, je voudrais remercier tous ceux qui m'ont offert un support technique ou moral pendant la réalisation de cette maîtrise : mon conjoint André Décary,

ma belle soeur Monique Beaudin et mes parents et amis.

Chapitre 1

Introduction

1.1 Contexte d'application

Les jeux vidéo représentent la catégorie la plus connue d'environnements virtuels interactifs. Ils immergent leur utilisateur dans un environnement virtuel complet en lui permettant d'explorer cet univers et d'interagir avec lui via un avatar, personnage appartenant au monde virtuel mais contrôlé par l'utilisateur.

Depuis trois décennies, les jeux vidéo se sont appropriés une place prépondérante dans le domaine du divertissement. Dès les débuts de leur développement, l'image cinétique a constitué l'interface principale entre les jeux vidéo et leurs utilisateurs. Cette image, initialement 2D pour les premiers jeux, a évolué vers une image 3D réaliste, qui s'assimile généralement à une fenêtre donnant accès au monde virtuel du jeu. Cette image permet de communiquer les différents éléments diégétiques (ensemble des composantes du monde fictif) du jeu, dont les lieux impliqués et l'interaction entre l'avatar, son environnement et les autres personnages évoluant dans le jeu.

L'utilisation de cette image cinétique révèle une parenté évidente entre une catégorie de jeux vidéo contemporains et le cinéma. Tout comme le film, le jeu se compose d'une narration (bien que souvent sommaire), impliquant un certain nombre de personnages, révélée au travers de l'image cinétique. L'avatar, personnifiant le joueur dans le monde virtuel, s'assimile au protagoniste du film. Une différence fondamentale entre le jeu et le film consiste en l'attitude spectatorielle. Celle-ci, de passive dans le cas du film où le spectateur reçoit la narration sous une forme entièrement prédéterminée, devient active dans le jeu, le joueur pouvant, grâce aux actions de son avatar, influencer les événements

narratifs et contrôler les informations qu'il désire obtenir du monde diégétique.

L'intégration de techniques cinématographiques à l'image du jeu pourrait ainsi contribuer à améliorer l'efficacité de sa communication. Différentes techniques développées depuis plus d'un siècle par les cinéastes sont disponibles à cette fin. Celles-ci touchent principalement l'éclairage, la composition de l'image, la mobilité de son cadre, l'utilisation des couleurs et le montage. L'ensemble de ces techniques constitue un langage cinématographique dont le but est de présenter directement ou indirectement les éléments narratifs au spectateur. Le langage cinématographique permet de manipuler la perception de l'espace et du temps, de contrôler l'exposition du flot informatif, et de communiquer les émotions et sous-textes relatifs à la narration.

L'efficacité des figures de style couramment utilisées en cinéma est souvent attribuable au développement de schémas cognitifs dont la facilité de reconnaissance provient de leur répétition dans l'ensemble des films visionnés. Le style classique Hollywoodien (*Hollywood style*) [BTS85] a largement contribué à la diffusion de ces stéréotypes cinématographiques. Ce style regroupe la majeure partie des films produits aux États-Unis entre 1917 et 1960. La période couverte par le *Hollywood style* est délimitée par l'établissement des grands studios et la fin de la période du *studio system*, organisation de la production liée au *Hollywood style*. Le *studio system*, conçu pour la production de masse des films et leur distribution, appliquait le principe manufacturier de systématisation à la réalisation des films. La mise sous contrat de l'ensemble des artisans par les studios permettait aussi une uniformisation dans le style des films appartenant à cette période.

Du point de vue stylistique, le *Hollywood style* se traduit principalement par une narration simple, et une expression de la spatiosité et de la temporalité subordonnée à la compréhension narrative et basée sur l'orientation du spectateur et la continuité. Le système narratif repose sur la causalité et les motivations finales et intermédiaires des personnages impliqués. Les systèmes spatial et temporel se distinguent par la standardisation des techniques cinématographiques et une redondance dans l'application de celles-ci. Le montage par continuité (voir section 2.8), aussi appelé montage transparent, est un procédé caractéristique du *Hollywood style*.

1.2 Portée du projet

Parmi les différentes techniques cinématographiques disponibles, le contrôle de la position de la caméra a été retenu comme ayant le plus grand potentiel d'implantation. Alors que l'éclairage se résume en quelques techniques classiques, et que l'utilisation de la couleur relève souvent de la métaphore, laquelle est difficilement automatisable, la gestion de la caméra couvre à la fois la composition de l'image et la mobilité de son cadre. Tel que développé dans le chapitre 2, la position de caméra utilisée pour un plan est déterminée par l'échelle du plan, l'angle horizontal avec le sujet et l'angle de cadrage, chacune de ces composantes comportant ses qualités expressives propres.

Une majorité de jeux vidéo commerciaux optent pour une caméra subjective, souvent appelée *point of view* ou *first person*, où l'image emprunte la perception visuelle de l'avatar. Ce type de caméra est parfaitement adapté aux scènes dans lesquelles le personnage doit interagir directement avec son environnement, comme les scènes de combat, d'où la nomenclature *first person shooter* identifiant une catégorie de jeux de combat.

Malgré le fait que la caméra subjective place directement le spectateur ou joueur à la place de l'avatar, ce procédé est insuffisant pour induire un sentiment d'identification psychologique entre le joueur et son avatar. Cette identification serait, au contraire, renforcée par l'utilisation de plans représentant le protagoniste, préférablement de face. De plus, il existe un nombre limité de techniques cinématographiques, communiquant les émotions reliées au contexte diégétique, disponibles pour le procédé subjectif.

Conséquemment, le choix d'implantation fait pour l'engin cinématographique développé pour ce mémoire porte sur la génération de séquences de plans non subjectifs (ou à la troisième personne), pour lesquels la position de la caméra tente de communiquer certaines émotions prédéfinies. Tel que mentionné précédemment, les scènes d'interaction directe entre l'avatar et son environnement se représentent avantageusement par la vue subjective. Des scènes moins impliquantes sont donc visées par ces séquences non subjectives. L'application porte, plus précisément, sur la représentation d'un personnage en déplacement à l'intérieur des lieux virtuels. Pour des raisons de simplification, un personnage seul est actuellement utilisé par le système.

Certains jeux commerciaux utilisent une caméra à la troisième personne. Les algorithmes actuels de gestion de caméra n'adaptent que rarement l'image produite au contexte émotionnel de l'avatar. Dans les jeux plus anciens comme "*Resident Evil 1*" (*Capcom*, 1996), la caméra est fixe et liée à la région où se situe l'avatar. Cette approche, appliquée à un jeu d'horreur, a l'avantage de dissimuler, par le cadre de l'image, une portion de l'environnement de l'avatar. Cette technique, familière des films d'horreur, augmente l'anticipation du joueur en limitant sa connaissance des événements narratifs immédiats.

Les jeux plus récents utilisent plutôt une caméra mobile qui effectue des travellings et des panoramiques afin de conserver l'avatar centré dans l'image et à une certaine distance de la caméra. Cette caméra peut quand même être rattachée à une région définie, de façon à changer directement de position lorsque le personnage sort de la région couverte pour entrer dans une autre. Ce fonctionnement peut être observé dans le jeu "*Indigo Prophecy*" (*Quantic Dream*, 2005). La caméra peut aussi constamment adapter son mouvement aux déplacements du personnage pendant une scène complète, tel qu'implantée dans le jeu "*The Lord of the Rings, The Two Towers*" (*Electronic Arts*, 2003).

Le domaine d'application, le jeu vidéo, possède ses contraintes propres qui doivent être considérées pour les spécifications de l'engin cinématographique. Contrairement aux personnages d'une production cinématographique dont les positions et déplacements sont connus et planifiés, l'avatar du jeu est manipulé interactivement par l'utilisateur. Les plans produits ne peuvent donc pas être conçus à partir d'une position désirée ou en fonction d'un déplacement élaboré. L'engin cinématographique doit ainsi calculer la position de caméra correspondant à la position courante du personnage, et s'adapter à ses déplacements arbitraires. La géométrie de l'environnement entourant le personnage doit aussi être considérée afin d'obtenir une vue non occultée du personnage. En outre, l'engin cinématographique doit générer chaque nouveau plan en temps réel pour ne pas interférer avec l'interaction entre l'utilisateur et son avatar et avec la fluidité de l'animation.

1.3 Spécifications de la Caméra *SIT*

L'engin cinématographique développé dans le cadre de ce projet de maîtrise a été nommé Caméra *SIT*, en référence aux trois émotions véhiculées par les séquences de plans qu'il génère. Ces séquences représentent un personnage seul en déplacement dans un environnement spatial arbitraire. Le personnage est entièrement contrôlé interactivement par l'utilisateur. Les techniques cinématographiques reliées à la composition de l'image sont appliquées au contexte cognitif et émotionnel courant de l'animation pour déterminer la position de la caméra à chaque plan. Ce contexte concerne la perception spatiale (S), l'identification avec l'avatar (I) et la tension ressentie dans la scène (T). La position de caméra doit être calculée en temps réel.

D'un point de vue créatif, la caméra virtuelle possède un avantage indéniable sur une caméra réelle. Sa masse et son volume nuls lui permettent d'utiliser toutes les positions dans l'espace, créant ainsi tous les types de cadrages désirés. Elle peut aussi réaliser des mouvements qui sont actuellement proscrits à une caméra réelle et utiliser des lentilles idéales qui ne restreignent plus la profondeur du champ visuel dans l'image. L'utilisation de stéréotypes pour les positions de caméra contraignant les capacités expressives inhérentes à ce médium, une prémisse de base à la conception de la Caméra *SIT* était de considérer valide pour ces positions l'intégralité de l'espace entourant le personnage. L'application de techniques cinématographiques dans le but d'obtenir l'expression d'une émotion donnée ajoute cependant une composante classique aux plans. L'utilisation d'une théorie cinématographique simple, ne permettant pas de s'extraire complètement du style classique Hollywoodien, est préférable pour l'automatisation des choix de cadrage, l'analyse stylistique d'un langage cinématographique plus avant-gardiste étant moins accessible à cette fin.

Les trois axes émotionnels perception spatiale (S), identification avec l'avatar (I) et tension ressentie dans la scène (T) constituent la composante émotionnelle de la Caméra *SIT*.

Appartenant davantage au domaine du cognitif que de l'émotif, la perception spatiale mesure le niveau de connaissance de l'environnement spatial fictif dans lequel évolue l'avatar. Cette connaissance est essentielle pour l'interaction avec le jeu. Elle permet d'une part une navigation adéquate de l'avatar, et est d'autre part un support narratif

important.

Le sentiment d'identification psychologique avec l'avatar est lié à l'empathie du joueur pour le protagoniste. L'émotion d'empathie accroît la volonté de protéger et de favoriser un destin avantageux pour l'avatar [Per04]. En outre, l'empathie permet au joueur de mieux intégrer les émotions fictives relatives au personnage [Tan96], ce qui améliore l'expérience globale du jeu.

La tension, émotion liée aux moments de suspense d'un film ou d'un jeu, est souvent recherchée par les joueurs, et ciblée par toute une catégorie de jeux. Elle prépare le joueur à réagir lors d'une confrontation anticipée en dirigeant son attention vers les détails de la scène [Per04]. La différence élevée entre l'anticipation de la connaissance narrative et la connaissance narrative réelle, caractéristique des périodes de suspense, entraîne l'intérêt du joueur. L'intérêt, considéré comme une émotion, tend à augmenter l'attention portée à la narration et à poursuivre le jeu [Tan96]. L'émotion de tension est aussi plus facilement exprimée par des techniques cinématographiques documentées, puisque toute une esthétique du suspense a été développée.

À chaque nouveau plan, la Caméra *SIT* utilise les valeurs contextuelles reliées aux variables émotionnelles S (perception spatiale), I (identification avec le protagoniste) et T (tension ressentie dans la scène) pour établir les paramètres de la caméra. Ces paramètres, définissant la position spatiale de la caméra, sont la distance entre la caméra et le personnage, qui détermine l'échelle de plan, la position horizontale de la caméra relativement au personnage et son élévation. La gestion de l'occultation du personnage est aussi intégrée à la sélection de la position de caméra.

Les variables contextuelles *SIT* expriment la quantification de la perception émotionnelle évaluée à un moment donné de l'animation. Elles sont calculées à partir de l'évaluation du niveau de ces émotions, exprimé par un certain nombre de plans antérieurement présentés. Les niveaux de perception spatiale et d'identification avec le protagoniste étant généralement opposés dans un plan, les choix de positions de caméra visent à obtenir, sur une séquence de plans, un équilibre entre les niveaux de ces deux variables émotionnelles. Le niveau de tension est, quant à lui, arbitrairement déterminé par l'utilisateur à l'aide d'une glissière (*slider*).

La manipulation d'une séquence de plans implique l'application de principes de montage pour effectuer les raccords entre les plans. Les principes du montage par continuité (section 2.8) sont utilisés de façon à fournir au joueur une perception spatiale et une orientation cohérentes. Dans sa version actuelle, la Caméra *SIT* développée n'utilise pas de mouvements de caméra, raccordant tous ses plans avec des coupes franches.

La position de la caméra pour le prochain plan est toujours calculée au moment du changement de plan, à partir de la position courante du personnage. Aucune anticipation basée sur le vecteur de déplacement du personnage n'est utilisée pour tenter de prévoir les prochains plans. La Caméra *SIT* n'effectue pas non plus de modifications optiques sur l'image. Nous discuterons de ces limitations à la conception de la Caméra *SIT* avec les extensions possibles au chapitre 6.

Le reste de ce mémoire s'organise comme suit. Le chapitre 2 fournit les définitions et la théorie cinématographiques sur lesquelles s'appuie l'analyse des fonctionnalités de la Caméra *SIT*. Le chapitre 3 récapitule les travaux antérieurs en infographie couvrant la problématique de positionnement cinématographique de caméra. Une description plus approfondie de l'implantation de la Caméra *SIT* dans un engin de jeu, introduite dans la présente section, est fournie au chapitre 4. Le chapitre 5 présente les résultats obtenus par l'utilisation de la Caméra *SIT* développée. Le chapitre 6 propose des pistes de recherches ultérieures visant à améliorer la Caméra *SIT*.

Chapitre 2

Éléments de cinématographie

Ce chapitre introduit les diverses notions de cinématographie utilisées pour l'élaboration des règles de sélection des plans. Ces notions sont strictement présentées sous l'angle de la cinématographie. Leur application dans l'implantation de la caméra virtuelle sera détaillée dans le chapitre 4.

2.1 Références

La littérature contient de nombreux ouvrages traitant d'art cinématographique et d'analyse filmique. Bien que les concepts décrits dans cette section soient communs à l'ensemble de ces sources, leur terminologie peut cependant différer. Deux livres ont principalement été utilisés pour la rédaction de cette section. "L'art du film : une introduction" de Bordwell et Thompson [BT00] est un ouvrage de référence traitant de l'ensemble des composantes contribuant à l'esthétique du film. "Grammar of the film language" d'Arijon [Ari76] est davantage un manuel de réalisation décrivant les règles de base de composition des plans et les schémas d'enchaînement de plans.

2.2 Le plan

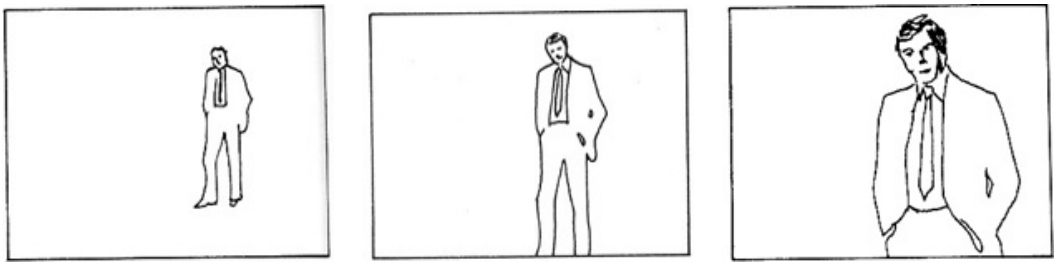
Le plan représente l'élément de base d'un film. Il est composé d'une prise de vue montrée en continu, étant ainsi caractérisé par une composante spatiale et une composante temporelle. La durée d'un plan est variable, pouvant n'être que de quelques secondes à l'intérieur d'une séquence montée ou de plusieurs minutes dans la cas d'un plan séquence.

Le plan est visuellement défini par le contenu du cadre de son image. La position de la caméra par rapport à la scène filmée détermine l'échelle du plan ainsi que l'angle avec lequel les éléments de la scène sont montrés. Les mouvements de caméra et la distance focale de l'objectif contribuent aussi au style du plan. La distance focale modifie la perspective de l'image, permettant d'aplanir ou de dilater l'espace. Elle en influence aussi la profondeur de champ (épaisseur de la zone de foyer).

2.3 Échelle de plan

Déterminée par la distance comprise entre la caméra et le personnage ou objet filmé, l'échelle du plan définit la proportion de l'image occupée par ce dernier. Chaque échelle possède des qualités descriptives spécifiques, en montrant le lieu de l'action dans son ensemble ou en s'attardant sur certains détails.

Plan général Correspondant à une caméra éloignée, le plan général est utilisé pour montrer les lieux de l'action dans leur globalité et non pour y situer les personnages. Elle correspond à une maison montrée dans son ensemble ou un paysage extérieur filmé de loin.



(a) (b) (c)

FIG. 2.1 – (a) Plan d'ensemble, (b) Plan moyen, (c) Plan américain [Int84].

Plan d'ensemble Le plan d'ensemble montre les personnages en relation avec leur environnement. Contrairement au plan général, les personnages y sont identifiables, mais le décor reste dominant. Ce type de plan est utilisé pour permettre au spectateur d'établir une ébauche mentale de la disposition des lieux. Il peut aussi montrer plusieurs personnages avec leurs positions relatives. Un plan d'ensemble est souvent utilisé lors de l'introduction d'un nouveau lieu, suivi de plans plus rapprochés détaillant

les diverses portions de l'espace filmique. Lorsqu'utilisé pour cette fonction, le plan d'ensemble est appelé plan de situation. Ce schéma, très répandu dans la production cinématographique, reproduit l'approche cognitive de la perception visuelle. Le plan d'ensemble peut aussi amener un détachement psychologique par rapport à l'action en cours.

Plan moyen Le plan moyen montre le personnage en pied. Il permet de suivre l'ensemble de ses mouvements et de montrer l'espace immédiat l'entourant. Ce type de plan est souvent utilisé lors de scènes d'action.

Plan américain Cette échelle de plan peut cadrer le personnage aux genoux ou à mi-cuisse. Identifiée aux westerns puisqu'elle y montrait les personnages avec leurs revolvers, elle est tout indiquée pour souligner les mouvements du haut du corps et des bras. Les plans américain et moyen peuvent montrer quelques personnages dans un même plan.

(a) (b) (c)

FIG. 2.2 – (a) Plan rappoché, (b) Gros plan, (c) Très gros plan [Int84].

Plan rapproché Le plan rapproché cadre le personnage au niveau de la taille ou de la poitrine. Il correspond à la perception qu'une personne a de son interlocuteur lors d'une conversation. Ce type de plan induit un sentiment d'intimité avec le personnage et situe le spectateur au milieu de l'action. Généralement, un seul personnage est présent dans un plan rapproché. Les plans plus serrés peuvent contribuer au sentiment de tension en augmentant la portion du hors champ (espace filmique non visible dans le cadre de l'image), restreignant ainsi la vision du spectateur de même que l'information narrative lui étant fournie.

Gros plan Le gros plan cadre la tête du personnage. Celui-ci occupe alors la majeure proportion de l'image et le décor devient à peine perceptible. Le gros plan a un aspect psychologique très fort et augmente la tension, l'aspect dramatique de la scène ainsi que l'identification avec le personnage. Le gros plan d'un objet est aussi utilisé pour en souligner l'importance narrative.

Très gros plan Le très gros plan agrandit un détail d'un visage ou d'un objet, comme un oeil ou une bouche. Cette échelle de plan extrême n'est utilisée que pour les scènes très fortes psychologiquement ou pour augmenter la portée dramatique d'une situation.

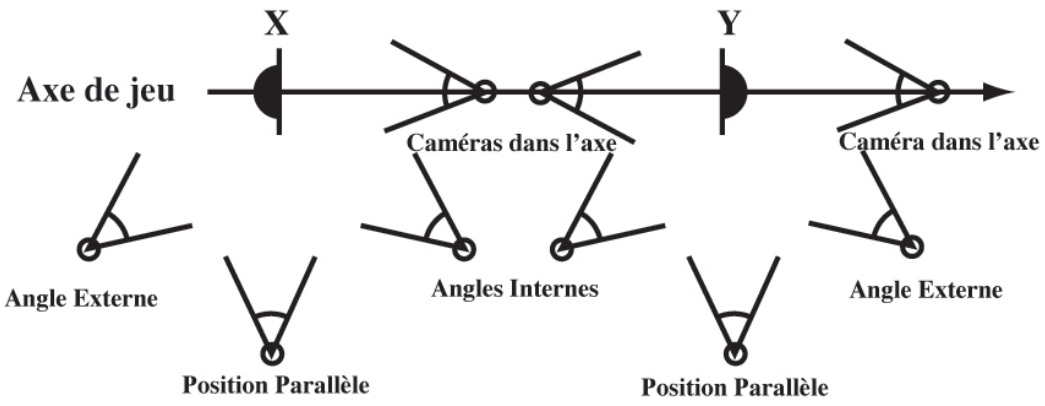


FIG. 2.3 – Positions de caméra relatives à l'axe de jeu défini entre les personnages X et Y [Sta07].

2.4 Position de caméra

La position de la caméra détermine l'angle dans le plan horizontal avec lequel sont montrés les personnages. Les positions usuelles sont l'angle interne, l'angle externe, la position parallèle et la caméra dans l'axe. Ces positions se définissent par rapport à l'axe de jeu. Cette ligne imaginaire, qui relie les têtes de deux personnages impliqués dans la scène ou s'assimile à la direction du déplacement d'un personnage, divise l'aire de jeu de la scène. Cette division logique est utilisée pour déterminer les positions de caméra pour les plans d'une même séquence (voir la règle des 180 °, à la section 2.8). Ainsi, les

positions décrites dans cette section, appartenant au principe du triangle, peuvent être situées d'un côté ou de l'autre de l'axe de jeu.

Angle interne La caméra placée à angle interne montre le personnage de trois-quarts face. Utilisé dans un champ-contrechamp, figure usuelle pour la représentation des dialogues, montrant chaque personnage alternativement, l'angle interne ne montre qu'un seul des protagonistes à la fois.

Angle externe La caméra placée à angle externe montre le personnage de trois-quarts de dos. Dans un champ-contrechamp, son interlocuteur serait vu de trois-quarts face, par-dessus son épaule.

Position parallèle La caméra est placée perpendiculairement à l'axe de jeu et montre le personnage de profil. L'orientation de la caméra reste la même pour filmer un second personnage situé face au premier ou pour un plan d'ensemble montrant les deux personnages.

Caméra dans l'axe Dans cette composition, la caméra est placée directement sur l'axe de jeu et montre le personnage complètement de dos ou de face. Un personnage montré de face, regardant directement la caméra, interpelle le spectateur, le rendant ainsi conscient du processus narratif [Bor85]. Un plan de dos peut, de son côté, s'assimiler à un plan subjectif (*point of view*) puisqu'il fournit au spectateur la même information visuelle que le personnage, induisant un sentiment d'anticipation des événements.

2.5 Angle de cadrage

L'angle de cadrage vertical définit l'angle formé entre la direction de l'objectif et la verticale au sol. Trois types d'angle vertical sont utilisés pour la composition d'image : l'angle normal où la caméra est orientée perpendiculairement à la verticale, la plongée où une caméra généralement haute est orientée vers le bas et la contre-plongée où une caméra généralement basse est orientée vers le haut.

Alors que l'angle normal exprime une vision neutre de la réalité, les plans pris en plongée ou en contre-plongée ont une fonction plus expressive. Une généralité concernant l'angle vertical stipule qu'il exprime l'état de domination du personnage, une plongée

écrasant le personnage alors que la contre-plongée en souligne la puissance. Ce type de métaphore est parfois utilisé, mais on note que l'angle vertical représente le plus souvent un point de vue posé sur la scène, la caméra empruntant le regard d'un observateur, appartenant ou non à l'univers fictionnel du film. Par exemple, les plans montrant alternativement les protagonistes impliqués dans un dialogue utiliseront l'angle vertical correspondant à la position de l'interlocuteur.

Expressivement, une contre-plongée tend à augmenter l'effet dramatique d'une scène, et un plan en plongée à en augmenter l'effet de tension. Des positions extrêmes de caméra, à vol d'oiseau ou au ras du sol sont souvent utilisées dans des scènes très angoissantes ou pour en souligner la violence. L'utilisation massive de l'angle normal, correspondant au regard d'un individu debout, par le cinéma classique hollywoodien expliquerait l'inconfort provoqué par les angles verticaux accentués, les spectateurs n'ayant développé que peu de schémas cognitifs pour l'interprétation de ce type de prise de vue [BTS85].

D'un point de vue strictement perceptuel, la plongée amplifie l'indice de profondeur spatial relié à la hauteur relative des objets dans le champ visuel et apporte une meilleure connaissance topographique des lieux [Cut05]. Elle gomme cependant certaines informations visuelles usuelles, nuisant par exemple à la reconnaissance des personnages.

Plus rarement utilisé, l'angle horizontal définit l'angle formé entre les lignes horizontales du cadre de l'image et l'horizon. Un cadre incliné peut augmenter l'impression de déséquilibre par rapport au personnage ou à l'action.

2.6 Composition de l'image

La composition de l'image détermine la portion de l'image occupée par chaque personnage ou élément de la scène. Elle est utilisée pour gérer la prédominance visuelle dans l'image et ainsi l'attention du spectateur. Les types de compositions les plus couramment utilisés sont la composition par tiers où l'image est divisée verticalement en trois zones de surfaces équivalentes, et la composition par moitiés où l'image est divisée verticalement en deux zones de surfaces équivalentes. Chaque portion de l'image peut contenir un personnage ou une partie du décor.

Le cinéma hollywoodien classique utilise généralement la centralité du personnage dans l'image, alliée à sa frontalité pour souligner son importance narrative. Des mou-

vements de caméra sont utilisés pour le recadrer lorsqu'il se déplace. La gestion de l'attention par la centralité serait mieux adaptée à notre fonctionnement cognitif que l'utilisation de l'illumination sélective, qui consiste à augmenter l'intensité relative de l'éclairage dans une partie de l'image. Les objets qui retiennent notre attention se retrouvent naturellement au milieu de notre champ de vision et, psychologiquement, notre intérêt se centre sur eux. La majeure partie de l'évolution de la perception de la lumière ayant été soumise à la seule lumière naturelle, la compréhension innée d'une illumination sélective serait limitée [Gro05].

2.7 Mouvement de caméra

Dans un plan fixe, le cadre est immobile et les personnages évoluent dans la portion du décor ainsi délimitée. Les mouvements de caméra permettent de modifier les limites du cadre durant le plan et ainsi les portions de l'espace filmique contenues dans le champ et le hors champ. Les deux principaux types de mouvements de caméra utilisés sont le panoramique où la caméra pivote sur place et le travelling où la caméra effectue un déplacement.

Panoramique Le panoramique s'assimile à une rotation de la tête qui permet de détailler un espace. Sa fonction est donc principalement descriptive. Il peut attirer l'attention sur un objet, ou servir de conjonction entre deux éléments mis en relation. Le panoramique a aussi une fonction de recadrage dans le cas où le personnage effectue un léger déplacement.

Travelling On distingue le travelling avant, arrière ou latéral, selon la direction de la caméra. Le travelling latéral peut avoir la même fonction descriptive qu'un panoramique, mais il est perçu moins passif. Il fournit aussi un indice supplémentaire de la perception de profondeur dans l'image par rapport au panoramique, le *motion perspective*. Causé par l'effet de parallaxe, le *motion perspective* provient de la vitesse relative de déplacement des objets qui est inversement proportionnelle à leur distance dans l'image.

Le travelling avant permet de souligner l'importance d'un objet ou d'un personnage et implique une augmentation de l'intensité dramatique de la scène et de la proximité psychologique avec le personnage. Inversement, un travelling arrière amène une distan-

ciation par rapport à l'action en cours et une atténuation de l'effet d'identification avec le personnage.

Le travelling arrière, en augmentant la portion de l'espace filmique contenue dans le champ, révèle peu à peu l'environnement du personnage. Ce type de plan, dévoilant progressivement le contenu de la scène, est appelé structure par exposition lente. Il propose une dynamique descriptive opposée à l'utilisation classique du plan de situation qui débute une séquence en situant par un plan large les différents éléments composant la scène. L'exposition lente est moins passive et provoque l'anticipation chez le spectateur.

Les *zoom-in* et *zoom-out* ainsi que le *cut-in* remplissent la même fonction que les travellings avant et arrière. Le *zoom*, aussi appelé travelling optique, se distingue du travelling par la modification de la distance focale. Le *cut-in*, composé d'une série de plans fixes pris avec le même angle et se rapprochant du sujet, augmente l'impact de l'image.

Un travelling peut aussi servir à accompagner un personnage en déplacement. Le travelling d'accompagnement possède une connotation affective supérieure au panoramique de recadrage et augmente l'identification avec le personnage. Le travelling avant, où la caméra suit le personnage, provoque un sentiment d'anticipation quant à sa destination. Il permet au spectateur de découvrir les lieux visités simultanément au personnage.

La vitesse des mouvements de caméra possède des qualités expressives spécifiques. Un mouvement lent appelle à la description neutre, alors qu'un même mouvement effectué rapidement attire davantage l'attention du spectateur sur le sujet montré. Un travelling d'accompagnement latéral conservant le personnage centré souligne son importance, mais lui confère une impression d'immobilisme. Par opposition, le même travelling effectué avec une vitesse variable ou de façon saccadée (style caméra à l'épaule) permet de dynamiser la scène.

2.8 Le montage par continuité

Le montage définit la manière de combiner les plans formant une séquence de façon à ce qu'ils dégagent un sens donné. Le montage expressif, associé au cinéaste soviétique Sergei Eisenstein (1898-1948), associe des plans originalement disjoints afin de générer une nouvelle interprétation, différente de celles directement liées aux plans eux-mêmes. Le montage par continuité, aussi appelé montage invisible ou découpage, par opposi-

tion au montage expressif, se subordonne quant à lui à la compréhension de la narration. Systématisé par le réalisateur américain D.W. Griffith (1875-1948), le montage par continuité a été largement utilisé dans les années 1920, au début de l'époque des films parlants, alors tournés en studio. Représentatif du style classique hollywoodien, sa pratique reste dominante dans l'ensemble de la production cinématographique.

Les règles développées pour le montage par continuité favorisent la perception cohérente et continue de l'espace filmique et l'utilisation de raccords fluides entre les plans. La principale règle observée, la règle des 180 °, stipule que l'ensemble des plans formant une séquence doivent être pris à partir d'un même côté de l'axe de jeu (voir section 2.4). Cette règle permet de conserver d'un plan à l'autre les positions relatives des personnages et éléments du décor. Ainsi, un personnage se trouvera du même côté par rapport à son interlocuteur dans tous les plans. La règle des 180 ° garantit aussi une constance dans les directions dans l'image, tant au niveau des mouvements que des orientations des regards des personnages. Un personnage effectuant un déplacement de gauche à droite sortira d'un cadre fixe par la droite et entrera par la gauche dans le prochain plan. Cette figure, appelée *frame cut*, est très fréquente dans la cinématographie classique, car elle souligne l'importance de la centralité, les bords du cadre servant à atteindre le centre de l'image. La représentation d'une portion de l'espace commune aux plans facilite aussi la compréhension de la disposition des lieux, les éléments répétés d'un plan à l'autre agissant comme repères spatiaux. Bien que la règle des 180 ° préconise l'utilisation d'un seul côté de l'axe de jeu pour la composition des plans, l'utilisation de certaines positions de caméra, dites positions neutres, permet d'effectuer un changement de côté d'axe. Les positions de caméra dans l'axe, où la caméra est située de face ou de dos par rapport à l'axe de jeu permettent d'effectuer ce changement. L'axe de jeu peut aussi se déplacer dans le cours d'une séquence, selon les changements de position des personnages.

La règle des 30 ° est une autre règle caractéristique du montage par continuité. Cette règle demande que la différence entre les positions de caméra utilisées pour deux plans consécutifs d'un sujet fixe soit d'au moins 30 °. Deux plans consécutifs pris d'une même position de caméra et impliquant un changement au niveau de l'échelle de plan résulte en un changement subit de volume dans l'image, générant ainsi une discontinuité visuelle. Un changement dans l'angle de la caméra permet de détourner l'attention du

spectateur de cette discontinuité et redonne la fluidité au raccord. Le même effet est obtenu par un mouvement dans l'image, impliquant qu'un raccord effectué pendant le mouvement d'un personnage peut se libérer de la règle des 30 °. La gestion de la variation scalaire dans les plans obéit à la même règle, une variation de plus d'une échelle de plan diminuant la fluidité du raccord. L'utilisation de raccords plus frappants dirige l'attention du spectateur sur un objet ou un personnage et augmente l'aspect émotionnel et dramatique d'une scène.

La durée relative des plans détermine le rythme de la séquence montée, manipulant ainsi la perception de la durée d'une action, soit en la compressant ou en la dilatant. Une séquence composée de plans de plus en plus courts participe à l'accélération de l'action ou augmente l'anticipation d'un événement important de la narration. À l'inverse, des plans de longueur égale donnent une impression de lenteur et d'attente. Outre les relations rythmiques, la durée d'un plan est généralement proportionnelle à son échelle. Les plans larges, comportant un plus grand nombre de composantes, requièrent davantage de temps d'analyse visuelle que les plans serrés ne montrant qu'un nombre restreint d'éléments.

D'un point de vue stylistique, les séquences formées de plans montés s'opposent aux longs plans séquences. Les séquences montées contrôlent l'attention du spectateur par l'utilisation d'échelles de plan et de positions de caméra qui déterminent les différents points de vue posés sur la scène. Une profondeur de champ étroite, fixe ou variable (*rack focus*), procédé souvent associé à la séquence montée, participe à la même finalité. Le plan séquence, présentant une action dans sa continuité en respectant sa durée réelle, suscite la participation psychologique du spectateur. Les mouvements de caméra peuvent tenir lieu de découpage à l'intérieur d'un plan séquence. Une profondeur de champ exacerbée, alliée à l'utilisation de plans séquences est caractéristique du style cinématographique développé dans les années 1940. Ce type de construction visuelle utilise la hiérarchisation de l'image pour permettre au spectateur d'analyser lui-même la scène lui étant présentée.

Chapitre 3

Travaux reliés en infographie

La problématique visant à appliquer des procédés cinématographiques comme le contrôle de la caméra à une animation arbitraire relève tant de l'intelligence artificielle que de l'infographie. Ainsi, des travaux effectués dans ces deux disciplines s'y sont attaqués, y apportant des solutions de natures variées. De façon générale, les différentes solutions proposées contraignent le domaine des positions potentielles de caméra et résolvent le système ainsi contraint afin d'obtenir la position la mieux adaptée aux caractéristiques désirées.

Les travaux retenus pour ce chapitre se divisent en trois approches distinctes : l'approche idiomatique, l'approche spatiale et l'approche émotionnelle.

L'approche idiomatique se base sur le formalisme des idiomes cinématographiques pour générer les séquences de plans. Un idiome est un enchaînement de plans stéréotypé utilisé pour représenter une action définie pour un certain nombre de personnages. Le schéma classique de champ-contrechamp (figure 3.1), généralement utilisé pour représenter un dialogue entre deux personnages, est un exemple typique d'idiome. Plusieurs types d'actions sont définis sous forme d'idiomes, dont les conversations entre deux, trois ou plusieurs personnages, ou le déplacement d'un personnage.

L'approche spatiale utilise la géométrie de la scène et la position de chaque personnage pour déterminer la position de la caméra et le type de plan produit en tenant compte de l'occultation des personnages.

Finalement, l'approche émotionnelle, plus récente, détermine les caractéristiques des plans produits à partir de la quantification de certaines émotions à communiquer.

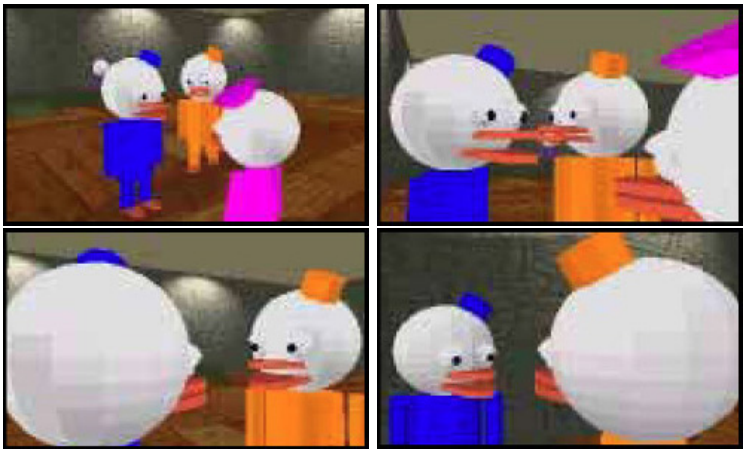


FIG. 3.1 – Champ-contrechamp dans le cadre d'un idiome de conversation à trois. Images tirées de [HCS96].

Dû à la nature temporelle de l'animation, et malgré la complexité de la problématique, la plupart des articles portant sur le domaine du positionnement de caméra, mis à part les thèses et mémoires cités, ne contiennent qu'une présentation assez sommaire de leurs résultats. Ce manque de résultats présentés, allié à la nature hautement subjective de l'évaluation de leur qualité, rend difficile la comparaison de leurs contributions respectives. L'évaluation effectuée sur les travaux mentionnés dans ce chapitre se base ainsi presqu'entièrement sur la description faite des techniques développées, dont les détails d'implantation sont souvent négligés.

Un effort de formalisme a été apporté à la présentation des résultats de ce mémoire, afin de permettre la comparaison des séquences de plans produites selon certains critères, comme la progression des échelles de plans et de l'élévation de la caméra dans la séquence. Cette analyse quantitative des séquences permet d'identifier les procédés cinématographiques priviligiés par la Caméra *SIT* pour exprimer les axes émotionnels.

3.1 Approche idiomatique

La *Camdroid* de Drucker et Zeltzer [DZ95] définit pour chaque type de plan un module de caméra composé d'un ensemble de contraintes reliées à la composition de l'image, comme la taille et la position de chaque personnage y apparaissant. La résolution de ces contraintes, appliquées à la localisation des personnages dans la scène, détermine la position de la caméra pour ce plan. Le système choisit, à partir du type d'activité

des personnages, le module de caméra à utiliser pour un plan donné. L'association entre le type d'activité détecté dans la scène et le module de caméra correspondant est définie par l'utilisateur. L'enchaînement des modules de caméra forme un idiome cinématographique.

Dans leur *Virtual Cinematographer*, He *et al.* [HCS96] utilisent les idiomes cinématographiques décrits dans le manuel de réalisation de Arijon [Ari76], encodés sous forme de machines à états finis. Chaque état de l'idiome correspond à un type de plan prédéterminé, défini par son échelle de plan et sa position de caméra. La navigation entre les différents états de l'idiome détermine l'enchaînement des plans composant la séquence produite. La sélection de l'idiome correspondant à un événement, décrit sous forme de description de personnages et d'actions, s'effectue par l'utilisation d'une hiérarchie d'idiomes dans laquelle les idiomes traitant les cas plus généraux sont éclatés en idiomes plus spécifiques. Les idiomes sont choisis en fonction des événements passés en temps réel au *Virtual Cinematographer*.

Christianson *et al.* [CAH$^+$96] reprennent l'utilisation des idiomes de Arijon [Ari76], en les décrivant avec leur *Declarative Camera Control Language* (*DCCL*). Leur *Camera Planning System* (*CPS*) se compose d'un planificateur de séquences, d'un compilateur de *DCCL* et d'un évaluateur heuristique. Le planificateur divise les séquences, reçues en temps réel sous forme de positions, vélocités et actions de personnages, en scènes, suivant les changements d'actions. Les idiomes, conservés dans une base de données, sont ensuite associés aux scènes précédemment extraites, selon l'action effectuée dans la scène. Tous les idiomes potentiels correspondant à une scène lui sont alors associés.

Le compilateur de *DCCL* génère, pour chaque idiome sélectionné par le planificateur de séquences, l'ensemble des plans (échelle et position de caméra) à partir de la description de l'idiome et des paramètres de la scène (mouvements des personnages).

L'évaluateur heuristique pondère chaque idiome d'une scène, selon des règles visant à maintenir des transitions harmonieuses entre les plans le formant, et choisit ainsi l'idiome le mieux adapté visuellement à la scène. L'enchaînement des plans formant les idiomes ainsi sélectionnés produit la séquence finale.

Funge *et al.* [FTT99] appliquent le formalisme de la modélisation cognitive afin de contrôler la durée des plans présentés. Ils définissent avec leur *Cognitive Modeling Language* les positions de caméra relatives aux personnages, ainsi que les actions et préconditions nécessaires à la détermination du moment de la transition vers un prochain plan. Les plans générés par leur modèle alternent entre des positions prédéfinies de caméra, comme les positions internes, externes et parallèles. Les changements de plans sont déterminés à partir de l'activité des personnages impliqués dans la scène, de façon à obtenir une séquence dynamique sans être trop rapide.

L'utilisation d'idiomes cinématographiques pour représenter une série de situations données génère des séquences de plans stéréotypées. Ces stéréotypes étant largement utilisés dans le cinéma classique, ils fournissent au spectateur des séquences efficaces puisque facilement identifiables. Par contre, ces séquences peuvent souffrir d'un manque de diversité dans les plans présentés. Une des prémices de conception de la Caméra *SIT* développée pour ce mémoire était de laisser la plus grande latitude possible dans le choix de la position de la caméra pour un plan, afin de conserver la souplesse inhérente à une caméra virtuelle.

3.2 Approche spatiale

À l'opposé de ces travaux, les prochaines approches présentées utilisent les informations spatiales courantes de la scène et des personnages pour déterminer les caractéristiques des plans générés.

La *ConstraintCam* de Bares *et al.* [BGL98, BL99] utilise la géométrie de la scène ainsi que des caractéristiques visuelles souhaitées pour chaque personnage pour construire ses plans. La *ConstraintCam* établit les prises de vue permettant de suivre un certain nombre de personnages évoluant dans une narration interactive, en utilisant un algorithme de résolution de contraintes. À chaque personnage impliqué dans la narration est associé un ensemble de contraintes. Celles-ci définissent l'inclusion du sujet dans le cadre, l'angle de vue désiré du sujet, l'intervalle des distances valides entre le sujet et la caméra permettant une bonne reconnaissance du sujet et la gestion de l'occultation du sujet par un autre personnage ou un élément du décor. Chaque type de contrainte

possède une pondération et un indicateur spécifiant si cette contrainte peut être relaxée ou non. Pour chaque contrainte de chaque personnage est calculée une région de l'espace englobant la position potentielle de la caméra. Les régions correspondant à l'ensemble des contraintes sont alors intersectées pour obtenir une région commune, valide pour la position de la caméra. Lorsque l'intersection des régions associées à une paire de contraintes est nulle, ces contraintes sont identifiées comme incompatibles. Le système cherche alors à résoudre ces incompatibilités en relaxant les contraintes par ordre de pondération. Si aucune solution ne peut être trouvée avec la relaxation des contraintes, le plan est divisé en deux ou plusieurs plans montrant les personnages séparément. Les incompatibilités de contraintes sont alors évaluées pour ces nouveaux plans. Si plusieurs plans doivent être utilisés pour présenter adéquatement l'ensemble des personnages, *ConstraintCam* permet de les montrer en utilisant des incrustations dans la fenêtre principale ou sous forme de séquences alternées.

ConstraintCam détermine la zone d'occultation d'un personnage en projetant les boîtes englobantes des objets de la scène sur une sphère entourant le personnage. La Caméra *SIT* de ce mémoire utilise cette même technique afin de restreindre les positions potentielles de la caméra.

Comme Bares *et al.* [BGL98], Halper *et al.* [HHS01] utilisent aussi la résolution de contraintes pour déterminer la position et l'orientation de la caméra, en y ajoutant une composante prédictive afin d'obtenir une plus grande cohérence temporelle entre les plans. La position future d'un objet est évaluée à partir de sa trajectoire et de son accélération, et les paramètres de la caméra sont calculés en fonction de cette position estimée. Ils sont ensuite interpolés pour les plans intermédiaires.

La gestion de l'occultation est effectuée par le calcul des *Potential Visibility Regions*. Cette technique se réduit au calcul des zones d'ombrage relatives à la géométrie de la scène par rapport aux points d'intérêt des plans. Les zones non ombragées sont considérées comme étant les plus favorisées pour l'application de la contrainte de visibilité.

Nam et Thalmann [NT99] utilisent la détection de mouvement pour déterminer la région d'intérêt à filmer lors d'une séance d'improvisation virtuelle où les participants sont associés à des avatars reproduisant leurs gestes. Selon la partie du corps d'un

participant impliquée dans le mouvement, le planificateur de caméra choisit le type de plan (échelle, transition et mouvement de caméra) correspondant. Une expression au niveau du visage entraîne un gros plan de l'avatar, un mouvement au niveau des bras ou du haut du corps un plan rapproché, alors qu'un déplacement est montré par un plan large. Le planificateur permet aussi d'inclure un objet dans le cadre lorsque celui-ci est pointé par l'avatar, générant ainsi un plan de conjonction. Une priorisation entre les types de plan permet de choisir l'avatar ciblé par le plan, selon le niveau de mouvement de chacun.

3.3 Approche émotionnelle

Les travaux cités jusqu'à présent n'exploitent aucunement l'aspect émotionnel relié à la cinématographie des plans. Les prochaines approches citées déterminent les caractéristiques des plans générés de façon à transmettre certaines émotions prédéfinies, correspondant aux événements décrits dans l'animation.

Dans son mémoire, Hornung [Hor03] décrit un module de caméra implanté dans un jeu vidéo qui, recevant en entrée une série d'événements narratifs décrivant l'action en cours, produit en temps réel les paramètres de caméra correspondant aux caractéristiques de plans sélectionnés. Les événements narratifs spécifient le type d'action à représenter, les personnages ou objets impliqués, incluant leur position et direction de déplacement respectives, ainsi que des variables quantifiant certaines composantes émotionnelles à communiquer dans le plan. Dans l'implantation décrite, les émotions représentées par ces variables couvrent la distance par rapport à l'action (exprimée en terme d'échelles de plans), l'expression de la dominance d'un objet ou d'un personnage, l'implication émotionnelle, le calme et la tension.

Les décisions des différents sous-systèmes du module de caméra sont effectuées à l'aide d'un réseau de neurones recevant en entrée une somme pondérée de valeurs de paramètres. La pondération correspondant à chaque paramètre est établie par apprentissage. Chaque neurone détermine sa valeur de sortie par l'application d'une fonction sigmoïdale.

Afin d'établir une continuité dans la séquence de plans produits, le module de caméra choisit le prochain événement à représenter parmi la série d'événements narratifs soumis,

selon sa priorité relative et le degré de cohérence entre cet événement et l'événement actuellement mis à l'écran. Le degré de cohérence entre deux événements narratifs est proportionnel au nombre d'objets impliqués communs aux deux événements. La durée du plan courant dépend aussi de la priorité du prochain événement sélectionné, ainsi que de la valeur des variables émotionnelles qui indiquent le niveau d'agitation à induire à la séquence.

Le module associe le type d'action et les variables émotionnelles de l'événement narratif sélectionné à un plan tiré de sa bibliothèque. La bibliothèque de plans contient un certain nombre de plans prédéfinis pour lesquels les paramètres de caméra (position relative, profondeur de champ, élévation et position des objets dans le cadre) sont entièrement déterminés.

La *CameraCreature* de Tomlinson *et al.* [TBN00] utilise un système combiné d'émotions et de motivation afin de déterminer les caractéristiques des plans produits. La valeur de la composante émotion provient d'une combinaison linéaire des valeurs attribuées aux six émotions de base : colère, dégoût, peur, joie, peine et surprise. *CameraCreature*, comme chaque personnage impliqué dans le monde virtuel, possède son propre ensemble d'émotions. À chaque plan, *CameraCreature* détermine ses valeurs d'émotions à partir des émotions des autres personnages, pondérées par leur importance à l'écran. L'émotion dominante de *CameraCreature* influence le style cinématographique donné aux plans produits en modifiant le type de mouvements de caméra et de transitions entre les plans utilisés.

L'émotion de *CameraCreature* gère aussi l'illumination globale de la scène alors que celle des personnages leur illumination personnelle, en modifiant l'intensité, la position et la couleur des lumières.

La composante motivation détermine plus directement le choix des plans en modulant sa tendance entre plan de situation, plan de conjonction entre deux personnages et plan rapproché sur un personnage donné. Les actions effectuées par les différents personnages manipulés interactivement influence la valeur de motivation de *CameraCreature*. Les personnages peuvent aussi utiliser cette voie pour demander l'attention de la caméra. Les valeurs d'émotions et de motivation sont influencées par des taux de croissance, de décroissance et d'inhibition.

L'angle utilisé pour montrer le personnage principal dépend de l'action effectuée

par celui-ci. Un déplacement du personnage favorise un plan pris de dos, alors qu'une manipulation appelle un plan pris de face.

Dans la thèse "*Intelligent Planning of Communicative Acts for Computer Animation Cinematography*" [Ken02] et l'article correspondant [KM02], Kennedy et Mercer décrivent une approche rhétorique, la *Rhetorical Structure Theory* (*RST*), qui analyse les intentions de communication fournies avec l'animation afin de sélectionner un ensemble de procédés cinématographiques transmettant ces intentions de communication. L'assistant animateur proposé reçoit en entrée une description physique de l'animation et une série d'intentions de communication fournies par l'utilisateur. La description de l'animation spécifie les caractéristiques visuelles, la disposition des lieux et des personnages, ainsi que les actions effectuées par ces derniers. Les intentions de communication peuvent identifier l'émotion associée à un personnage, l'ambiance à insuffler à la scène ou viser un renforcement de la narration. Les procédés cinématographiques utilisés pour exprimer les intentions de communication comprennent le choix de cadrage, le type d'illumination de la scène ainsi que les caractéristiques des couleurs utilisées pour le rendu.

Originalement développé pour l'analyse de textes, *Rhetorical Struture Theory* en extrait la structure en exprimant les différents énoncés sous forme d'une série de relations sémantiques. L'application du *RST* aux intentions de communication à représenter s'effectue en associant les énoncés des intentions de communication avec des procédés cinématographiques. Le système construit un arbre avec les relations sémantiques *RST* et résout les contradictions entre les procédés cinématographiques correspondants, de façon à obtenir un ensemble de procédés consistants pour chaque plan.

Le système réfère alors à sa base de connaissances cinématographiques pour traduire les procédés cinématographiques en options de rendu concrètes. Cette base de connaissances est encodée avec *LOOM*, un langage associé aux systèmes à base de connaissances qui permet de définir hiérarchiquement les concepts cinématographiques.

La combinaison de procédés cinématographiques de natures différentes permet d'obtenir une grande diversité au niveau de la composition des plans.

Cet assistant animateur, qui permet de produire de courtes séquences composées de quelques plans, ne fonctionne cependant pas en temps réel.

Afin de simuler le comportement de poissons dans un environnement naturel, Tu et Terzopoulos [TT94] utilisent, conjointement avec un ensemble de paramètres représentant certaines habitudes, les variables “frayeur”, “faim” et “libido”, indiquant l'état mental actuel de chaque poisson. Les valeurs de ces variables sont utilisées pour déterminer les choix de comportement du poisson au cours de la simulation.

Ils ne traitent pas de la problématique de positionnement de caméra, mais la Caméra *SIT* utilise des variables contextuelles simulant la perception du spectateur, dont la fonction est semblable à celles décrites dans leur article. Comme dans cet article, les variables contextuelles synthétisent leurs valeurs à partir du contexte courant du système. Tu et Terzopoulos établissent une hiérarchisation de leurs variables contextuelles pour la détermination du comportement de leur poisson.

3.4 Discussion

Tout comme le projet présenté dans le présent mémoire, la plupart des articles référés décrivent des solutions s'appliquant à une animation contrôlée en temps réel par l'utilisateur. Cette spécification s'applique particulièrement aux systèmes de fiction virtuelle dont les jeux vidéo. Seuls ceux de Kennedy et Mercer [Ken02, KM02] qui doivent connaître la totalité de l'animation à traiter pour effectuer l'analyse *Rhetorical Structure Theory* des intentions de communication, et ceux de Drucker et Zelter [DZ95] où l'association entre l'action en cours et le module de caméra est définie manuellement, ne décrivent pas des systèmes travaillant en temps réel.

On remarque aussi que les travaux entrant dans la catégorie émotionnelle n'effectuent qu'un traitement très sommaire de l'occultation des personnages impliqués dans les animations. Inversement, les travaux de Bares *et al.* [BGL98] et de Halper *et al.* [HHS01] principalement basés sur la géométrie de la scène et la gestion de l'occultation n'utilisent pas de valeurs émotionnelles pour le choix des plans. La solution documentée au chapitre 4 allie ces deux approches en utilisant un algorithme de détermination des zones d'occultation qui restreint le domaine des positions de caméra disponibles pour la sélection par les variables émotionnelles.

Le traitement des variables émotionnelles utilisées comme entrée à l'algorithme de choix de plans diffère aussi entre ces travaux et la solution proposée par la Caméra *SIT*.

Hornung [Hor03] et Tomlinson *et al.* [TBN00] utilisent plusieurs émotions distinctes, mais en les combinant avec une somme pondérée. Les trois variables émotionnelles S (perception spatiale), I (identification avec le personnage) et T (tension perçue dans la scène) utilisées dans notre Caméra *SIT* restent indépendantes et sont mises en relation lors d'une optimisation par moindres carrés. L'utilisation de ces variables émotionnelles par Hornung [Hor03] aboutit à la sélection d'un type de plan, parmi un ensemble fini de types, dans lequel la position de la caméra est prédéfinie. L'éventail des positions potentielles est beaucoup plus large dans notre Caméra *SIT* puisque chaque composante de ces positions (échelle de plan, position et élévation), est déterminée à partir de ces moindres carrés. Du côté de Tomlinson *et al.* [TBN00], les variables émotionnelles ne servent qu'à déterminer le type de transition entre les plans, dont les mouvements de caméra. Les caractéristiques des plans sont davantage reliées aux actions des personnages.

De plus, contrairement aux travaux cités où les valeurs des variables émotionnelles sont déterminées arbitrairement par l'utilisateur, notre Caméra *SIT* utilise la rétroaction de ces valeurs provenant des derniers plans présentés pour établir celles à associer au prochain plan produit.

Chapitre 4

Implantation de la Caméra *SIT*

4.1 Fonctionnalités de la Caméra *SIT*

La Caméra *SIT* contient l'implantation des principes illustrés dans ce mémoire. Elle intègre les procédés cinématographiques documentés au chapitre 2, afin de générer des séquences de plans qui communiquent trois composantes émotionnelles : la perception spatiale S, l'identification avec le personnage I, et un niveau de tension T. Ses fonctionnalités sont intégrées dans un engin de jeu, permettant de disposer d'un environnement de jeu complet pour l'évaluation de la caméra.

La distinction entre les fonctions des différentes personnes interagissant avec un système de jeu n'est pas pertinente dans le cadre du présent projet. Le terme utilisateur, employé dans la suite de ce mémoire, cumule donc les rôles d'opérateur et de joueur. L'utilisateur détermine la configuration du système, manipule le personnage dans l'environnement virtuel et devient le spectateur des séquences présentées.

Les séquences générées par la Caméra *SIT* représentent un personnage unique en déplacement dans un environnement virtuel inconnu. Les déplacements de ce personnage sont arbitrairement contrôlés par l'utilisateur. La génération des plans n'utilise aucune composante prédictive et est effectuée en temps réel.

Seul le cadrage, déterminé par la position et l'orientation de la caméra relatives au personnage, est utilisé pour exprimer les composantes émotionnelles des séquences. La séquence de plans doit équilibrer la connaissance spatiale de l'environnement du personnage et le sentiment d'identification de l'utilisateur avec ce dernier. Le niveau de tension moyen communiqué par ces plans doit correspondre à une tension désirée

arbitrairement déterminée par l'utilisateur.

La problématique de la visibilité du personnage, qui vérifie que le personnage n'est pas occulté dans l'image, est traitée à même la résolution des contraintes émotionnelles pour la sélection des paramètres de la caméra.

4.2 Intégration dans *Crystal Space*

De nos jours, plusieurs trousses de développement logiciel (*Software Development Kit*) d'engins de jeu sont disponibles gratuitement. Notre choix s'est porté sur *Crystal Space* [Cry07].

Développé en C++, *Crystal Space* possède une interface de programmation (*API*) dont les classes font l'objet d'une documentation complète, claire et facile à consulter. La disponibilité du code source permet d'ajuster certaines fonctionnalités aux besoins du projet.

En plus des fonctionnalités de base, nécessaires au rendu et à l'animation des scènes 3D, ainsi qu'à l'interaction avec celles-ci, *Crystal Space* possède des modules facilitant l'utilisation de fichiers de configuration, la saisie d'images et l'enregistrement des déplacements du personnage.

Un critère important de sélection était la compatibilité du système avec les formats courants pour le chargement des environnements de jeu et des personnages. *Crystal Space* permet l'utilisation de personnages en format *MD3*, utilisé pour le jeu *Quake 3* (*id Software*, 1999), qui sont très répandus. Les environnements de jeu de *Crystal Space* sont codés en *XML*, ce qui facilite grandement leur modification.

Secteurs *Crystal Space* organise son environnement de jeu en secteurs. Tous les objets de l'environnement appartiennent à un secteur donné. Ainsi, l'environnement est constitué de plusieurs secteurs, dans lesquels sont définis les maillages, lumières et autres composantes.

Un portail est un polygone inclus dans un maillage, qui donne accès à un secteur adjacent au secteur auquel il appartient. Un portail n'est pas bilatéral. Deux portails distincts doivent donc être définis dans leur secteur respectif pour permettre la circulation dans les deux sens entre ces secteurs. Le portail simule la transparence, permettant de voir les maillages appartenant au secteur adjacent, comme par une porte ouverte.

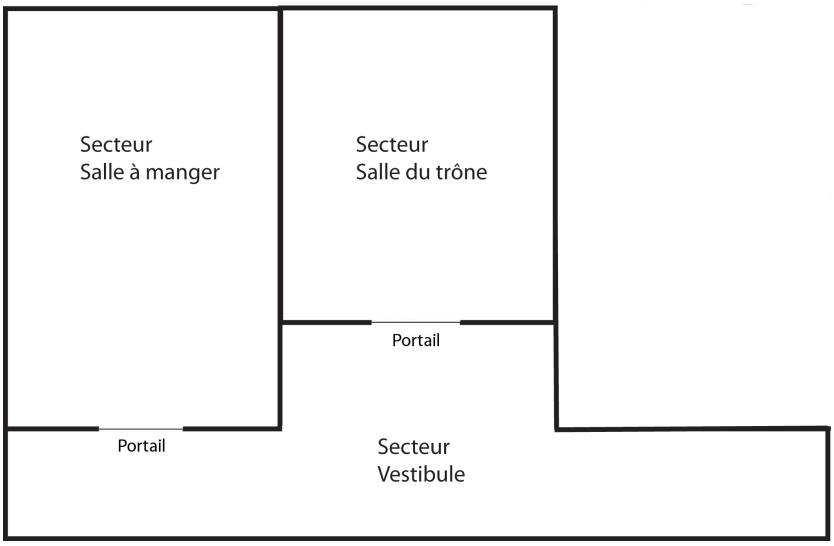


FIG. 4.1 – Environnement de jeu, château médiéval.

Dans le château, sélectionné pour ce mémoire comme environnement de jeu pour faire évoluer le personnage, chaque secteur représente une salle distincte (figure 4.1). La Caméra *SIT* peut ainsi savoir quand le personnage passe d'une salle à une autre, grâce au secteur relié à son emplacement.

Maillages Les maillages contiennent toute la géométrie constituant l'environnement virtuel. Ils représentent tant les murs, planchers et plafonds définissant les limites d'une pièce, que les meubles y étant disposés. Ils peuvent être organisés en hiérarchies, lorsque la position d'un objet dépend physiquement de celle d'un autre.

Le traitement des maillages par *Crystal Space* est généralement invisible pour la Caméra *SIT*. Les maillages ne doivent être utilisés par celle-ci que pour la gestion de la visibilité du personnage. Ils sont alors accessibles par une méthode de *Crystal Space* qui retourne un ensemble de maillages manipulables à partir d'un itérateur.

Systèmes de coordonnées *Crystal Space* fonctionne avec plusieurs systèmes de coordonnées. Un système de coordonnées global est géré au niveau de l'environnement de jeu. La caméra ainsi que tous les maillages inclus dans cet environnement possèdent leurs systèmes de coordonnées propres. Ces systèmes sont reliés au système de coordonnées global par des matrices de transformation qui positionnent les divers éléments dans l'environnement de jeu.

Dans ces systèmes de coordonnées cartésiens, les axes X et Z définissent le plan

horizontal alors que l'axe Y représente la hauteur. L'axe Z positif est aligné avec la direction du personnage ou de la caméra. Les systèmes de coordonnées utilisent la convention de la main gauche pour leur orientation.

Animation du personnage Les personnages en format *MD3* sont implantés comme des hiérarchies de maillages texturés dans *Crystal Space*. Ces maillages supportent une animation précalculée du personnage, selon des actions prédéfinies. Dans le système actuel, seule l'action *marcher* est utilisée.

Les maillages du personnage sont reliés entre eux par des matrices de transformation déterminant leurs positions relatives. La matrice de transformation du maillage principal indique la position courante du personnage dans l'environnement virtuel.

Les maillages représentant le personnage sont associés à une classe de *Crystal Space* qui en effectue les déplacements, selon la vitesse directionnelle et la vitesse angulaire désirées. Cette classe gère la détection de collisions et la gravité. Les vitesses directionnelle et angulaire peuvent être contrôlées par des clés du clavier ou par la souris. Les états du personnage, fixe ou en mouvement, en rotation ou en déplacement, dépendent des valeurs de ces vitesses et sont accessibles par la Caméra *SIT*.

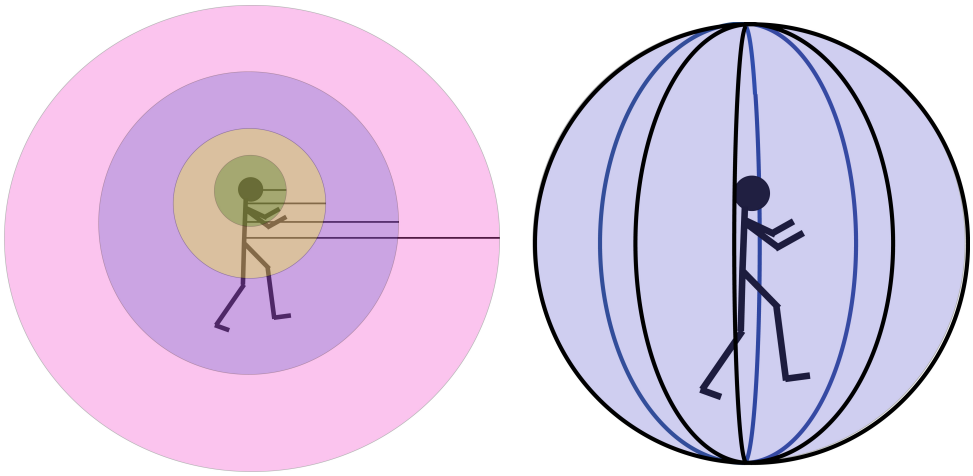
Module de caméra La caméra de *Crystal Space* utilise une matrice de transformation orthonormale qui définit son orientation, et un point 3D pour sa position. L'angle de champ de la caméra, définissant l'amplitude angulaire de la scène incluse dans l'image, ainsi que le rapport entre la hauteur et la largeur de l'image, peuvent être modifiés et consultés.

La caméra *SIT* intègre ses fonctionnalités à l'engin de jeu de *Crystal Space* en calculant la position de la caméra ainsi que la matrice de transformation de son orientation.

Variables de configuration Des variables de configuration sont utilisées pour ajuster le comportement de la Caméra *SIT*. Les identificateurs de ces variables prennent la forme *ident1.ident2.ident3*, ce qui permet de créer une hiérarchie et de regrouper les variables selon leur fonction. Par exemple, toutes les variables associées au plan d'ensemble seront identifiées par Application.PE.*ident*. Les variables de configuration développées pour la Caméra *SIT* sont regroupées par échelle de plan, angle de vue, personnage, ou variable émotionnelle.

Interface utilisateur Une interface utilisateur, permettant d'effectuer le choix de la caméra et de visualiser les valeurs courantes des variables émotionnelles, est superposée à l'image du jeu. Le moteur de rendu 2D de *Crystal Space* fournit les utilitaires nécessaires au dessin de cette interface.

4.3 Fonctionnement de la Caméra *SIT*

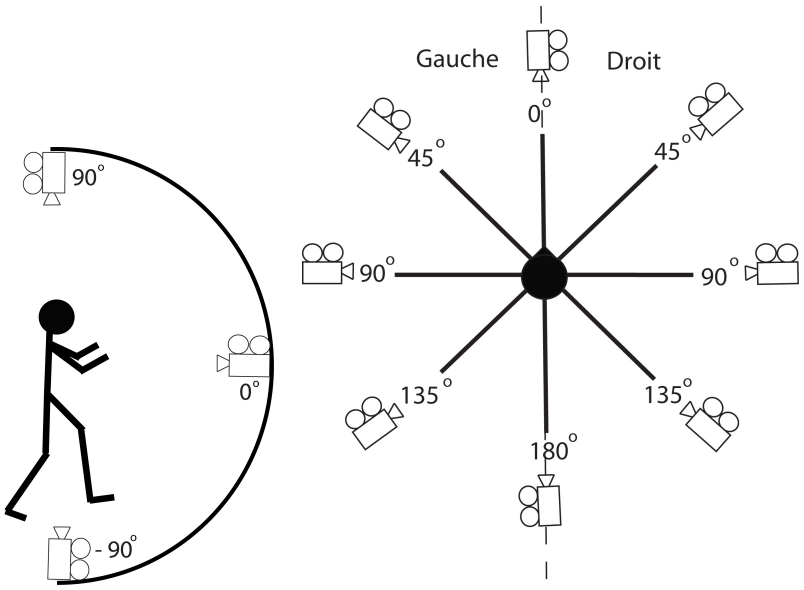


(a) (b)

FIG. 4.2 – (a) Échelles de plan, (b) Angles de vue appliqués à une échelle de plan.

Paramètres de la caméra Dans notre Caméra *SIT*, le cadrage du plan, et conséquemment la position et l'orientation de la caméra, se compose de trois paramètres indépendants : l'échelle de plan, l'angle de vue du personnage et l'élévation. Les plans élaborés par la Caméra *SIT* étant basés sur le principe classique de la centralité du personnage (section 2.6), ces paramètres s'expriment par rapport à la position courante du personnage dans l'environnement virtuel, formant ainsi un système de coordonnées sphériques autour de celui-ci.

Bien que la liberté de la caméra soit une prémisse de conception de la Caméra *SIT*, l'échelle de plan et l'angle de vue sont représentés par des valeurs discrétisées. Cette discrétisation permet d'intégrer les échelles et angles conventionnels du cinéma et de leur attribuer des pondérations émotionnelles. L'élévation ne faisant pas l'objet d'une gradation établie en cinématographie, sa valeur reste continue.



(a) (b)

FIG. 4.3 – (a) Angles d'élévation, (b) Angles de vue.

(a) (b)

(c) (d)

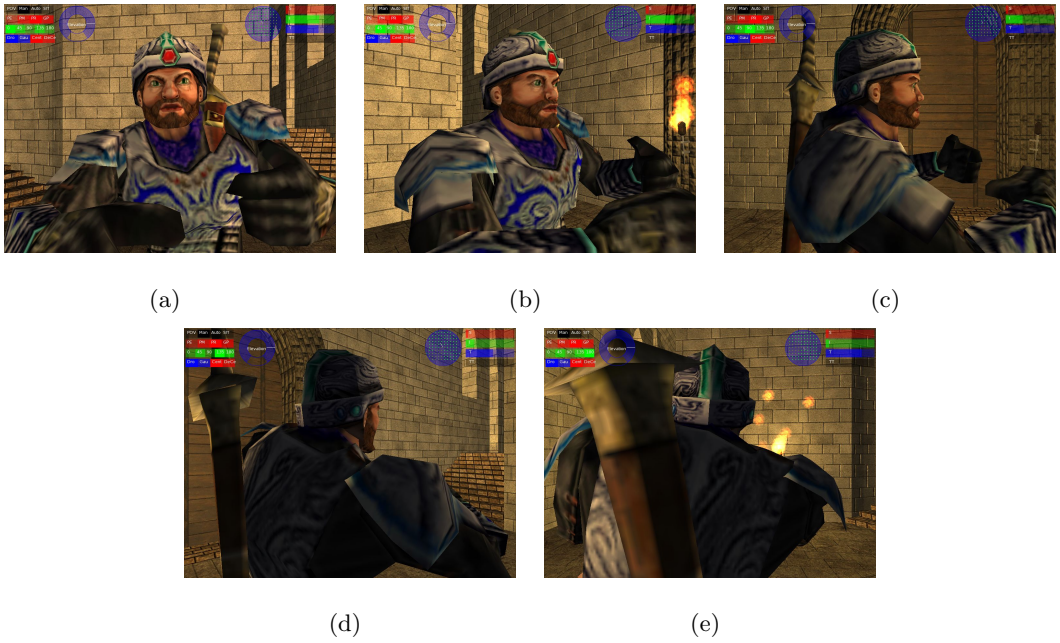
FIG. 4.4 – (a) Plan d'ensemble, (b) Plan moyen, (c) Plan rapproché, (d) Gros plan.

L'échelle de plan représente la proportion de l'image occupée par le personnage (section 2.3). Quatre échelles de plan canoniques ont été retenues pour l'implantation : le plan d'ensemble, le plan moyen, le plan rapproché et le gros plan (figure 4.4). Ces

échelles couvrent l'ensemble des motivations cinématographiques à communiquer dans les plans.

Une échelle de plan est définie par la distance entre le personnage et la caméra. Chaque échelle de plan délimite ainsi une sphère entourant le personnage, dont le rayon correspond à cette distance (figure 4.2(a)). La hauteur du centre de la sphère varie en fonction du cadrage du personnage utilisé pour cette échelle de plan. Le centre de la sphère devient le point de visée de la caméra.

La surface de cette sphère contient l'ensemble des positions de caméra valides pour une échelle de plan.



(a) (b) (c)

(d) (e)

FIG. 4.5 – (a) Angle 0 °, (b) Angle 45 °, (c) Angle 90 °, (d) Angle 135 °, (e) Angle 180 °.

L'angle de vue définit l'angle dans le plan horizontal avec lequel est montré le personnage (section 2.4). Les principaux angles cinématographiques sont implantés. Les valeurs d'angle sont réparties relativement à la frontalité du personnage. L'angle 0 ° désigne un plan de face, l'angle 45 ° un plan de trois-quarts face, l'angle 90 ° un plan de profil, l'angle 135 ° un plan de trois-quarts de dos, et l'angle 180 ° un plan de dos (figures 4.3(b) et 4.5).

Afin de faciliter la gestion de l'axe de jeu dans l'enchaînement des plans (section 2.8), les angles de vue sont considérés symétriquement pour les deux côtés du personnage.

Ainsi, un angle de vue s'identifie par son angle et le côté du personnage auquel il appartient.

Les angles de vue découpent huit demi-cercles sur la surface de la sphère de l'échelle de plan, représentant les positions de caméra valides (figure 4.2(b)).

(a) (b) (c)

FIG. 4.6 – (a) Élévation neutre, (b) Élévation positive, (c) Élévation négative.

L'élévation de la caméra constitue l'angle de cadrage vertical de l'image (section 2.5). Sa valeur continue recouvre les positions du demi-cercle représentant un angle de vue et une échelle de plan, et s'étend de $-90\,^\circ$ à $90\,^\circ$ (figure 4.3(a)). Une élévation nulle résulte en une vue régulière du sujet, une valeur positive en une vue en plongée, et une valeur négative en une vue en contre-plongée (figure 4.6).

Sélection du prochain plan La combinaison des trois paramètres, échelle de plan, angle de vue et élévation, construit l'ensemble de toutes les configurations de plans existantes. Chaque configuration potentielle est exprimée comme un intervalle d'élévations, défini pour un couple échelle de plan et angle de vue.

La Caméra *SIT* cherche, parmi cet ensemble, la configuration qui exprime le mieux le contexte émotionnel courant de la scène, représenté par les variables émotionnelles S, I et T, considérant les contraintes de visibilité liées à l'emplacement actuel du personnage dans les lieux virtuels. Les variables S, I et T traduisent respectivement la connaissance de l'environnement spatial du personnage, l'identification psychologique avec le personnage et le niveau de tension ressentie dans la séquence (section 1.3). Un ensemble de règles visant à conserver une cohérence visuelle entre les plans, concernant principalement l'angle de vue, et à éliminer certains cadrages inesthétiques est aussi utilisé afin de réduire le nombre de configurations de plan candidates.

Afin de conserver un temps de traitement raisonnable pour la sélection du prochain plan, la vérification de la visibilité du personnage est effectuée de façon approximative. La zone de visibilité ainsi établie délimite l'intervalle d'élévations pour l'échelle de plan et l'angle de vue correspondants.

La Caméra *SIT* attribue des pondérations émotionnelles pour chacun des trois axes *S*, *I* et *T* aux configurations de l'ensemble, préalablement restreint par les règles d'esthétique et l'établissement des zones de visibilité. Ces valeurs émotionnelles sont comparées avec les valeurs contextuelles obtenues à partir des caractéristiques des plans précédemment présentés, pour identifier la configuration du plan qui complète le mieux la séquence en cours.

4.4 Implantation des paramètres de la caméra

Chaque échelle de plan est représentée par deux variables de configuration : la distance entre la caméra et le personnage et la hauteur du point de visée correspondant au cadrage du personnage pour cette échelle. Le tableau 4.1 fournit les valeurs associées aux quatre échelles de plan implantées. Le point de visée est défini comme un déplacement vertical par rapport au centre de la tête du personnage. Comme toutes les dimensions utilisées dans *Crystal Space*, les distances documentées dans ce tableau sont exprimées en mètres. Ces valeurs servant au cadrage du personnage, elles sont établies en proportion de sa taille qui est d'environ 1.80 m.

Des valeurs d'élévations minimum et maximum sont configurées pour chaque échelle de plan afin d'éviter des vues indésirables, comme un gros plan en plongée du dessus de la tête du personnage ou une contre-plongée montrant le dessous de son menton. Ces valeurs sont présentées dans le tableau 4.1. Ces élévations ont été évaluées empiriquement à partir de la posture du personnage. Les élévations maximum ne sont limitées que pour les gros plans, de façon à permettre une prise de vue à vol d'oiseau pour les autres échelles de plan.

Calcul de la position de caméra La position et l'orientation de la caméra dans *Crystal Space* sont calculées à partir des valeurs d'échelle de plan, d'angle de vue et d'élévation sélectionnées pour le plan.

Pour effectuer le calcul de la position de la caméra, l'élévation, dont les valeurs sont définies entre $-90\,^\circ$ et $90\,^\circ$, doit être transformée pour l'intervalle entre $0\,^\circ$ et $180\,^\circ$, où $0\,^\circ$ est situé au-dessus du personnage, et $180\,^\circ$ en-dessous. De son côté, l'angle de vue conserve sa valeur mais les angles situés à la gauche du personnage deviennent négatifs. Les composantes de la position de caméra sont calculées ainsi

$$\begin{aligned} X_{pos} &= Rayon \times \sin\phi_e \sin\theta_a \\ Y_{pos} &= Rayon \times \cos\phi_e \\ Z_{pos} &= Rayon \times \sin\phi_e \cos\theta_a \end{aligned}$$

où ϕ_e représente l'élévation, θ_a l'angle de vue et *Rayon* le rayon associé à l'échelle de plan.

L'orientation de la caméra est obtenue par la multiplication des matrices de rotation correspondant aux trois axes.

La rotation selon l'axe X dépend de l'élévation verticale de la caméra. Elle est définie par la matrice suivante

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi_e & -\sin\phi_e \\ 0 & \sin\phi_e & \cos\phi_e \end{pmatrix}$$

où ϕ_e représente la valeur d'élévation inversée, $-90\,^\circ$ étant situé au-dessus du personnage, alors que $90\,^\circ$ est situé au-dessous.

La rotation selon l'axe Y, vertical, dépend de l'angle de vue. Elle est définie à la matrice suivante

$$\begin{pmatrix} \cos\theta_a & 0 & -\sin\theta_a \\ 0 & 1 & 0 \\ \sin\theta_a & 0 & \cos\theta_a \end{pmatrix}$$

Échelle de plan	Rayon	Point de visée	Élévation minimum	Élévation maximum
Plan d'ensemble	4.50	-0.45	$-55\,^\circ$	$90\,^\circ$
Plan moyen	2.00	-0.45	$-65\,^\circ$	$90\,^\circ$
Plan rapproché	0.76	-0.14	$-55\,^\circ$	$90\,^\circ$
Gros plan	0.40	-0.04	$-35\,^\circ$	$35\,^\circ$

TAB. 4.1 – Configuration des échelles de plan.

où θ_a représente l'angle de vue dont la valeur a été inversée, l'angle 180 ° correspondant à la vue de face, et l'angle 0 ° à un plan de dos. Les angles situés à la gauche du personnage prennent des valeurs négatives.

La rotation selon l'axe Z représente l'inclinaison latérale du cadre de l'image et n'est pas utilisée dans l'application actuelle.

4.5 Règles d'esthétique

Les règles d'esthétique visent à éliminer de l'ensemble des configurations de plan potentielles celles qui résulteraient en un cadrage indésirable ou violeraient un principe de continuité établi entre deux plans consécutifs.

La première règle correspond à l'application de la *règle des* 180 °, caractéristique du montage par continuité (section 2.8). Cette règle contribue à l'orientation de l'utilisateur dans l'espace. Elle stipule que, dans une séquence, l'ensemble des plans doivent être pris du même côté de l'axe de jeu. Comme dans le cas d'un personnage en déplacement, l'axe de jeu est représenté par la direction du mouvement, seuls les angles de vue situés du même côté du personnage que celui du plan précédent sont conservés. Le saut d'axe, qui consiste à placer la caméra de l'autre côté de l'axe de jeu, peut aussi être nécessaire afin d'étendre la connaissance spatiale à l'ensemble de l'environnement courant. La gestion de l'axe de jeu est décrite dans la section 4.10.

Une seconde règle favorise la similarité des directions apparentes du déplacement dans les images de plans successifs. Le personnage étant dirigé par des touches du clavier ou la souris, une grande disparité dans ces directions complexifierait sa manipulation. L'angle de vue du plan précédant, ainsi que les angles adjacents à celui-ci, sont donc considérés comme les angles de vue valides pour le prochain plan. L'angle de vue du personnage à sa sortie de champ est utilisé pour l'application de cette règle. Par exemple, l'angle de vue suivant un plan pris avec un angle de 45 ° pourrait prendre les valeurs 0 °, 45 ° ou 90 °. Par contre, comme le côté de l'axe de jeu doit aussi être respecté, un plan pris avec un angle de 0 ° devrait être suivi par un plan ne pouvant utiliser que les angles de vue 0 ° ou 45 °. La sélection des angles de vue doit tenir compte de la visibilité du personnage. Si l'angle de vue du plan précédent et les angles lui étant directement

adjacents ne possèdent pas un intervalle de visibilité valide, les angles adjacents à ces derniers sont considérés, de façon à trouver l'angle le plus rapproché de l'angle du plan précédant ayant un intervalle de visibilité valide. Dans le premier exemple mentionné, l'angle 135 ° serait alors utilisé.

Bien que la *règle des* 180 ° soit conservée dans le cas d'un personnage fixe, afin de permettre à l'utilisateur de se construire une image continue de l'espace, la règle concernant la direction apparente du déplacement peut être relâchée. Tous les angles de vue situés d'un même côté de l'axe de jeu sont donc valides pour un personnage fixe.

Alors que deux plans ayant un même angle de vue peuvent renforcer le point de vue utilisé sans tomber dans la répétition, un plus grand nombre de plans consécutifs adoptant ce même angle produit une séquence monotone. Afin de dynamiser la séquence, une règle exclut un angle de vue lorsque celui-ci a déjà été utilisé dans deux plans successifs.

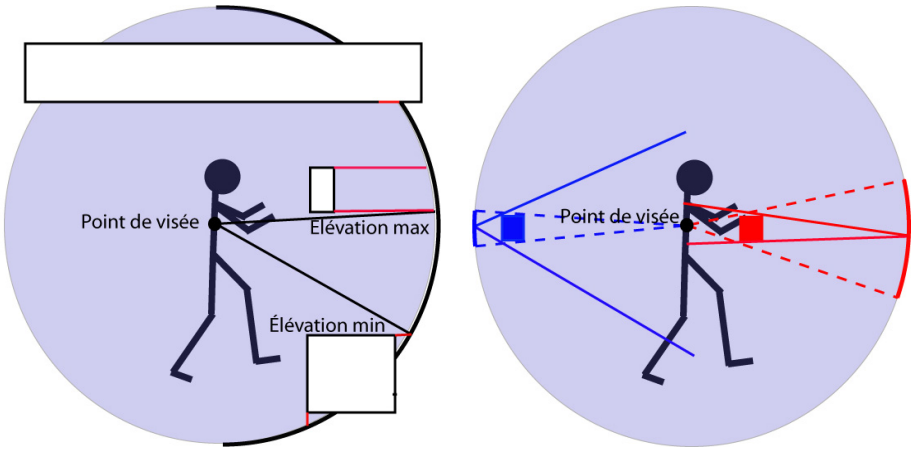
Alors que les règles précédentes adressaient l'articulation des plans, les prochaines règles régissent l'esthétique de l'image d'un plan.

La principale fonction d'un gros plan ou d'un plan rapproché est de favoriser l'identification avec le personnage. Lorsque ce plan est pris avec un angle arrière, cependant, sa fonction se déplace vers la vue subjective, l'éloignement du personnage permettant de découvrir le décor avec un point de vue semblable au sien. La représentation de l'environnement est centrale pour ce type de plan, et l'élévation de la caméra ne doit pas la restreindre. La valeur d'élévation est donc bornée à l'élévation neutre dans le cas d'une vue à 180 °, et étendue aux valeurs positives permettant une plongée pour une vue à 135 °. Les plongées sont permises pour les angles de vue à 135 ° puisque ceux-ci procurent une certaine reconnaissance du personnage, tout en privilégiant le décor.

Comme cette fonction de plan subjectif ne s'applique que lorsque le personnage est en déplacement, les combinaisons d'échelles de plan rapprochées et d'angles de vue arrières sont entièrement proscrites pour les plans montrant un personnage fixe.

La structure subjective fournit une bonne perception de l'environnement spatial, en empruntant le point de vue du personnage. Elle est ainsi indiquée lorsque le personnage entre dans un nouveau secteur de l'environnement de jeu. L'angle de vue du premier

plan effectué dans un nouveau secteur est donc forcé à 180 °.



(a) Calcul de la visibilité approximative du personnage (b) Problématique de la visibilité partielle

FIG. 4.7 – Visibilité du personnage.

4.6 Vérification de la visibilité

L'environnement virtuel dans lequel évolue le personnage est entièrement constitué de maillages représentant les divisions (murs, planchers, plafonds, escaliers, colonnes, etc.) et les objets meublant (tables, chaises, torches, etc.). Parmi l'ensemble des configurations de plan potentielles, composé de la combinaison de tous les échelles de plan, angles de vue et élévations définis, certaines impliquent une position de caméra située derrière une division ou un objet, relativement à la position du personnage. Afin d'écarter ces configurations qui engendreraient une image représentant un personnage occulté, la Caméra *SIT* intègre la visibilité du personnage dans son algorithme de sélection de plan.

La gestion de la visibilité du personnage est donc nécessaire à l'obtention d'une séquence de plans valide, lorsque cette dernière se déroule dans un environnement virtuel réaliste. Le problème de la visibilité ne constituant pas le sujet central de ce mémoire, sa résolution est alors davantage considérée comme une fonctionnalité complémentaire de la Caméra *SIT*.

L'algorithme de vérification de visibilité établit pour chaque couple échelle de plan et angle de vue appartenant à l'ensemble des configurations de plan potentielles, un

intervalle d'élévations définissant la zone de visibilité du personnage. Cet intervalle d'élévations restreint l'étendue des configurations soumises à la sélection, à l'aide des variables *SIT*, des paramètres de caméra. Si l'intervalle de visibilité est nul pour un couple, les configurations correspondantes sont complètement éliminées de l'ensemble des configurations candidates.

L'occultation partielle constitue un indice de la profondeur de l'image, souvent exploité par les décors de films [Cut05]. Une occultation partielle du personnage améliorant la perception spatiale de son environnement, il serait intéressant qu'elle soit préservée par l'algorithme de vérification de la visibilité. Le critère de visibilité du personnage doit ainsi être défini. Pour simplifier le problème, un point de visibilité unique, correspondant au point de visée de la caméra, est utilisé pour le personnage. La position du point de visée sur le personnage dépend de l'échelle de plan. Ainsi, seules sont valides les positions de caméra pour lesquelles le point de visée n'est pas occulté.

L'intervalle d'élévations est évalué à partir de l'intersection entre le demi-cercle représentant l'échelle de plan et l'angle de vue sur le personnage, et la projection parallèle des boîtes englobantes des maillages situés entre le personnage et ce demi-cercle (figure 4.7(a)).

L'utilisation de la projection parallèle fournit une estimation approximative et inexacte de la visibilité. Le recours à la projection radiale procurerait une carte exacte de la visibilité du point de visée à partir de la position de la caméra. Elle comporte cependant certaines difficultés pour reproduire la visibilité partielle réelle du personnage.

La projection radiale d'un petit maillage situé près du point de visée peut éliminer une grande région pour la position valide de la caméra, alors qu'il n'occulte qu'une petite portion du personnage. Inversement, un maillage près de la caméra peut cacher une grande proportion du personnage, alors que sa projection radiale ne restreint qu'une zone étroite pour la caméra (figure 4.7(b)). La résolution de cette problématique impliquerait l'élaboration d'une méthode plus complexe qui, compte tenu de la position imprévisible du personnage parmi l'ensemble des maillages, pourrait demander un temps de traitement prohibitif. La gestion de la visibilité ne constituant pas le thème central de ce mémoire, l'implantation de la projection parallèle a été jugée adéquate pour les besoins de la Caméra *SIT*, tout en permettant de conserver un algorithme simple et peu coûteux en temps de calcul.

Afin de pallier l'inexactitude de la solution, la vérification de la visibilité est séparée en deux procédures distinctes. La première utilise l'évaluation approximative pour déterminer l'intervalle de visibilité de toutes les combinaisons d'échelles de plan et d'angles de vue en prévision de la sélection des paramètres de caméra correspondant aux valeurs des variables émotionnelles *SIT*. La seconde établit exactement la visibilité du point de visée du personnage pour la position de caméra ainsi déterminée.

Estimation approximative de la visibilité La Caméra *SIT* ne génère qu'un seul intervalle d'élévations pour chaque couple échelle de plan et angle de vue. Cet intervalle est obtenu par l'agglomération des segments d'occultation provenant de la projection des maillages. Les maillages constituant l'environnement immédiat du personnage sont rattachés soit à la limite supérieure ou inférieure du demi-cercle utilisé, selon leur proximité à l'une ou l'autre de ces hauteurs. Cette règle est appropriée dans notre contexte, puisque les objets représentés par les maillages sont accrochés au plafond ou s'empilent à partir du plancher, ce qui est le cas des environnements composés d'objets inanimés et soumis à la gravité. Cette simplification apportée à l'implantation de la Caméra *SIT* élimine cependant certaines configurations de plan, empêchant ainsi la considération d'élévations valides mais situées dans des fenêtres de visibilité comblées par l'agglomération des segments d'occultation.

Les maillages considérés pour l'établissement de la zone de visibilité proviennent d'une méthode de *Crystal Space*. Cette méthode fournit la liste des maillages touchés par un plan vertical dont l'orientation correspond à l'angle de vue sur le personnage, et pour lesquels la distance au point de visée est inférieure au rayon de la sphère représentant l'échelle de plan. Cette méthode de *Crystal Space* retourne les maillages définis dans le secteur actuellement occupé par le personnage et ceux appartenant à un secteur adjacent si ceux-ci sont rejoints par le plan au travers d'un portail.

L'algorithme 1 décrit le traitement itératif des maillages pour l'établissement de l'intervalle de visibilité. L'algorithme procède à l'intersection du demi-cercle représentant l'échelle de plan et l'angle de vue sur le personnage avec chacune des faces des boîtes englobantes des maillages. Ces intersections produisent des segments, tronqués par la frontière du demi-cercle, qui délimitent les zones d'occultation du personnage. Si un maillage contient un portail, la position de celui-ci crée une ouverture dans le segment d'occultation résultant, établissant ainsi une zone de visibilité qui permet à la caméra de

voir le personnage au travers du portail. L'occultation provenant des autres maillages, situés devant ou derrière le portail, est quant à elle traitée normalement. Chaque segment, dont seules les coordonnées verticales sont considérées, est ensuite rattaché à la région d'occultation supérieure ou inférieure, selon les distances entre ses extrémités et les hauteurs extrêmes du demi-cercle. Les hauteurs des extrémités du segment vont alors prolonger la région d'occultation à laquelle il appartient, réduisant ainsi l'étendue de l'intervalle de visibilité, préalablement initialisée avec les valeurs délimitant l'intervalle d'élévation esthétiquement valide pour l'échelle de plan utilisée (section 4.4). À la fin du traitement des maillages, l'intervalle de visibilité, jusqu'alors exprimé en coordonnées verticales, est converti en angles d'élévation.

```
Intervalle.YMin ← Échelle.ÉlévationMin;
Intervalle.YMax ← Échelle.ÉlévationMax;
for chaque Maillage do
    obtenir la BoîteEnglobante;
    diviser la BoîteEnglobante en Faces;
    for chaque Face do
        SegmentOccultation ← Intersection(Face, DemiCercle);
        if (DemiCercle.YMax − SegmentOccultation.YMax) <
        (SegmentOccultation.YMin − DemiCercle.YMin) then
            if (SegmentOccultation.YMin < Intervalle.YMax) then
                Intervalle.YMax ← SegmentOccultation.YMin;
            end
        end
        else
            if (SegmentOccultation.YMax > Intervalle.YMin) then
                Intervalle.YMin ← SegmentOccultation.YMax;
            end
        end
    end
end
```

Algorithme 1 : Traitement itératif de l'occultation des maillages.

Vérification exacte de la visibilité Lorsque la Caméra *SIT* a sélectionné l'échelle de plan, l'angle de vue et l'élévation à utiliser pour le plan, la visibilité du personnage est

vérifiée à partir de la position de caméra correspondante. Un rayon est alors tracé entre cette position et le point de visée du personnage. Si un maillage est touché, la procédure cherche l'élévation résultant en une vue dégagée du point de visée du personnage.

Cet algorithme sépare dichotomiquement l'intervalle d'élévations correspondant à l'échelle de plan et l'angle de vue, pour trouver l'élévation produisant une position de caméra à partir de laquelle le rayon lancé vers le point de visée n'intersectera aucun maillage. La séparation dichotomique s'effectue de façon à trouver une élévation valide rapprochée de l'élévation initiale.

4.7 Cadrage du personnage

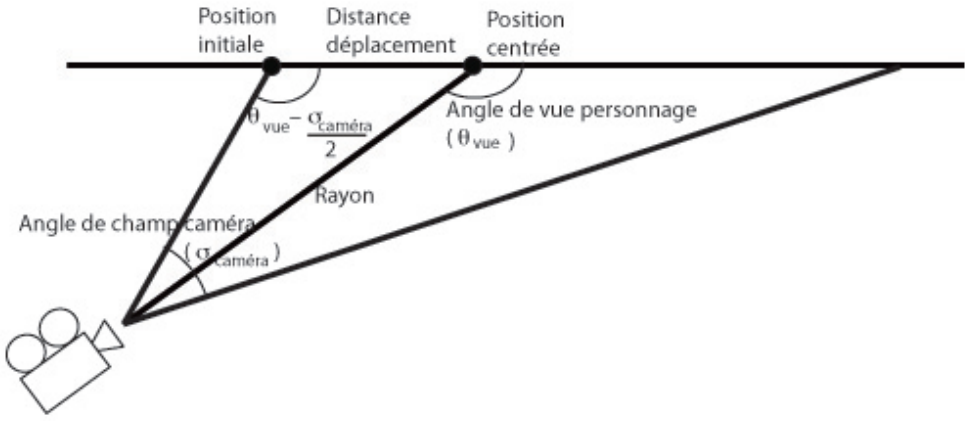


FIG. 4.8 – Calcul de la distance entre la position initiale et la position centrée du personnage.

En cinématographie, un mouvement peut couvrir l'image complète d'un plan ou se concentrer dans une de ses moitiés. Ainsi, le déplacement d'un personnage peut soit traverser l'image entière, soit s'étendre d'un côté du cadre jusqu'au centre de l'image, ou inversement du centre vers un des côtés.

Le *frame cut* (section 2.8) est obtenu lorsque le personnage sort d'un plan par un des côtés du cadre pour débuter le plan suivant par le côté opposé ; il implique généralement que le mouvement couvre l'ensemble de l'image.

Pour exécuter ce type de plan, le cadrage initial du plan ne doit pas être centré sur la position initiale du personnage, mais doit plutôt positionner ce dernier près du cadre.

La position médiane anticipée du déplacement à l'intérieur de l'image est donc calculée, afin de servir de centre au cadre du plan. Cette position est obtenue en ajoutant la distance correspondant à la moitié du déplacement complet effectué dans le plan à la position initiale du personnage, selon l'orientation de celui-ci.

La distance de déplacement entre la position initiale du personnage et sa position centrée dans l'image est calculée avec la loi des sinus, en utilisant l'angle de champ de la caméra et l'angle de vue sur le personnage (figure 4.8). L'angle de champ de la caméra définit l'amplitude angulaire de la scène incluse dans l'image. La position centrée correspond au point où la ligne de trajectoire du personnage intersecte la ligne médiane de l'angle de champ, selon la position de la caméra correspondant à l'angle de vue utilisé.

$$Distance \quad = \quad Rayon \times \frac{\sin\left(\frac{\sigma_{caméra}}{2}\right)}{\sin\left(\theta_{vue} - \frac{\sigma_{caméra}}{2}\right)}$$

où $\sigma_{caméra}$ désigne l'angle de champ horizontal de la caméra, θ_{vue} l'angle de vue sur le personnage et *Rayon* le rayon de la sphère correspondant à l'échelle de plan. Ce rayon représente la distance entre la caméra et la position centrée du personnage dans l'image.

Cette formule s'applique pour les angles de vue de 45 °, 90 ° et 135 ° seulement. La caméra est placée directement derrière le personnage pour les plans à 180 °, afin de lui laisser un maximum de distance d'éloignement dans le plan. Elle est placée au double du rayon correspondant à l'échelle de plan pour les plans à 0 °, de façon à positionner le personnage au point diamétralement opposé à la caméra, dans la sphère correspondant à l'échelle de plan.

Un rayon est tracé entre les positions initiale et centrée du personnage, afin de signaler la présence d'un maillage sur sa trajectoire. Comme une collision avec un maillage immobiliserait le personnage, le point de collision le plus proche de la position initiale est utilisé comme position centrale de l'image.

Afin de vérifier approximativement que le personnage dispose d'une région de visibilité minimale le long de son déplacement dans l'image, la visibilité du personnage est évaluée à la fois pour sa position initiale dans le plan et pour sa position centrée, par l'algorithme documenté à la section 4.6. L'intervalle d'élévations résultant correspond à l'intersection des deux intervalles d'élévations ainsi produits. Si la visibilité consolidée de ces deux

positions est nulle, le plan peut débuter avec le personnage initialement centré. Le déplacement couvrira alors la moitié de l'image.

Ce cadrage, possible pour un plan d'ensemble ou d'un plan moyen seulement, évite l'élimination de certaines configurations de plan candidates. Pour les plans rapprochés et les gros plans, la zone d'image à couvrir par le déplacement devient alors trop courte, provoquant des coupures de plans trop rapides, ce qui proscrit l'utilisation d'un cadrage centrant initialement le personnage pour ces échelles.

4.8 Variables émotionnelles *SIT*

Le rôle des variables *SIT* est de représenter le contenu émotionnel d'un plan ou d'une série de plans afin de tenter de simuler la perception de l'utilisateur. Elles permettent ainsi d'établir les caractéristiques visuelles des plans subséquents en y équilibrant les éléments de perception spatiale, d'identification au personnage et de communication de tension.

L'attribution de valeurs émotionnelles à la composition de l'image résulte de notre formalisation de principes cinématographiques largement documentés, et de la quantification de leur impact cognitif sur l'utilisateur. Cette formalisation, prérequise à la systématisation de ces techniques, est peu établie, bien que l'esthétique filmique soit abondamment analysée.

Les variables *SIT* servent à la fois à quantifier le contexte émotionnel résultant des plans de la séquence précédemment présentés, et à établir la pondération des configurations candidates à la sélection du prochain plan. La Caméra *SIT* cherche parmi les valeurs *SIT* potentielles celles qui équilibrent le mieux l'ambiance actuelle de la séquence, représentée par les variables *SIT* contextuelles. Les variables contextuelles étant réévaluées après chaque plan produit, la rétroaction des valeurs *SIT* sélectionnées permet à la Caméra *SIT* de s'adapter aux niveaux émotionnels atteints. Ainsi, plusieurs plans favorisant une perception de tension élevée seront suivis de plans permettant de la relâcher. De même, une série de plans orientés vers l'identification du personnage feront place à des plans permettant une meilleure perception de son environnement.

Les éléments de la composition de l'image considérés pour l'évaluation émotionnelle, se réduisent aux trois paramètres définissant la position de la caméra : l'échelle de plan, l'angle de vue et l'élévation. Des pondérations, reliées à la perception spatiale, à l'iden-

tification avec le personnage et à la perception de la tension, sont attribuées à ces trois composantes et sont amalgamées pour produire respectivement les valeurs des variables *S*, *I* et *T*. Ces pondérations ont été déterminées empiriquement par nous à partir de règles interprétant les principes cinématographiques. Toutes les valeurs impliquées dans l'évaluation des variables *SIT* appartiennent à l'intervalle continu de 0 à 1.

4.8.1 Variables émotionnelles *I* et *T*

Évaluation des variables émotionnelles *I* et *T* Les variables émotionnelles I et T utilisent la même évaluation tant pour leur déclinaison contextuelle que potentielle. Les valeurs de ces variables sont entièrement déterminées par le cadrage du plan. Elles sont obtenues par la multiplication des termes correspondant aux trois composantes de la position de la caméra.

$$\begin{aligned} I &= I_{échelle} \times I_{angle} \times I_{élévation} \\ T &= T_{échelle} \times T_{angle} \times T_{élévation}. \end{aligned}$$

Les pondérations associées à l'échelle de plan et l'angle de vue proviennent principalement de variables de configuration du système. L'élévation possède une valeur continue, contrairement aux deux autres composantes qui sont discrétisées. Elle est donc plutôt exprimée comme une fonction.

Calcul des variables contextuelles *I* et *T* Comme les variables contextuelles représentent les niveaux émotionnels atteints suite à la présentation d'une série de plans, elles sont obtenues par la moyenne effectuée sur les valeurs de I et T provenant d'un certain nombre de plans précédents.

Le nombre de plans utilisés pour cette moyenne influence le comportement de la Caméra *SIT*. Un petit nombre de plans augmente la variation des valeurs contextuelles, relativement à l'utilisation d'un plus grand nombre. Cette variation influence la diversité dans les caractéristiques des plans consécutifs. Des variables contextuelles trop stables tendent vers un équilibre qui se manifeste par une uniformité au niveau des plans produits. En revanche, une certaine variation dynamise les plans, et occasionne des changements d'échelle, d'angle de vue et d'élévation dans la séquence.

Les valeurs du nombre de plans peuvent être différentes pour les calculs de I et de T, de façon à obtenir un meilleur contrôle sur l'effet de ces deux variables contextuelles. Les résultats observables au chapitre 5 ont été générés avec cinq plans précédents pour la variable contextuelle I, et dix pour la variable T. Ces valeurs empiriques ont été retenues, car elles produisent la variation désirée dans les plans formant ces séquences.

Un personnage en déplacement modifie sa position dans l'image au cours d'un plan. Ainsi, sa distance avec la caméra et son angle de vue peuvent changer. L'évaluation des variables contextuelles doit donc utiliser les moyennes de ces composantes, observées pendant la durée d'un plan.

Règles pour l'évaluation de la variable émotionnelle *I* La variable émotionnelle I vise à quantifier le niveau de sentiment d'identification psychologique avec le personnage dans un plan donné. De façon générale, l'identification est proportionnelle à la reconnaissance faciale du personnage. Conséquemment, la valeur du terme $I_{échelle}$ est opposée à la distance avec le personnage (tableau 4.2), un gros plan possédant un pointage I élevé, alors que celui d'un plan d'ensemble est faible (section 2.3).

De même, le niveau d'identification I_{angle} est élevé pour les angles frontaux et bas pour les plans de dos (tableau 4.3).

$I_{élévation}$ est inversement proportionnel à la valeur absolue de l'élévation utilisée. Ainsi, elle est maximale pour l'élévation $0\,^{\circ}$, où la caméra est située au niveau du personnage, et décroît linéairement avec l'augmentation ou la diminution de l'élévation. Sa valeur devient nulle lorsque l'élévation atteint $\pm 90\,^{\circ}$, puisque cette élévation extrême empêche la reconnaissance du personnage.

Règles pour l'évaluation de la variable émotionnelle *T* La variable émotionnelle T exprime le niveau de tension ressentie lors d'un plan donné.

L'échelle de plan détermine la proximité psychologique de l'utilisateur avec le personnage. Celle-ci participant à l'effet de tension ressentie, la contribution de l'échelle de plan au terme $T_{échelle}$ suit une règle similaire à celle utilisée pour le terme $I_{échelle}$ (tableau 4.2), les plans rapprochés augmentant la tension, et les plans éloignés la tempérant.

Les valeurs attribuées à T_{angle} pour les angles de vue frontaux suivent aussi la règle de l'identification (tableau 4.3). Par contre, la tension augmente aussi pour les plans de dos puisque ces angles de vue simulent un plan subjectif qui induit un sentiment

d'anticipation (section 2.4).

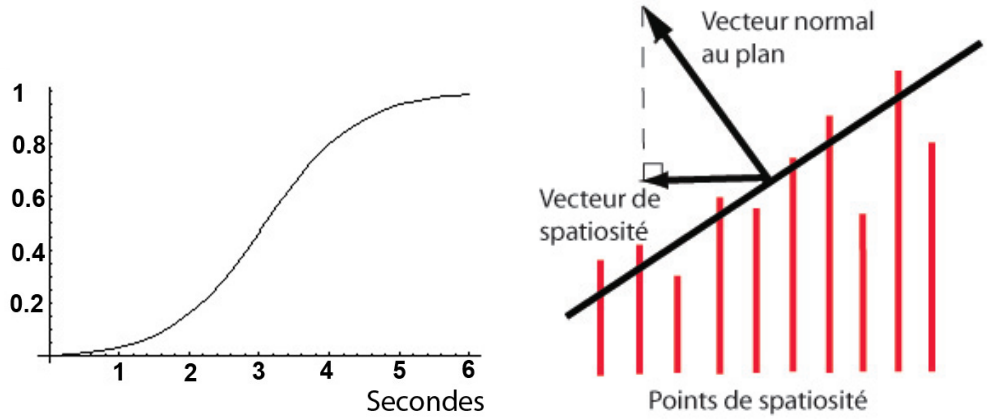
L'élévation contribue directement à la tension insufflée à un plan (section 2.5). $T_{élévation}$ est implanté comme une fonction sous-linéaire de l'élévation en valeur absolue. Un plan dont l'élévation est nulle communique une tension plus faible, alors que la tension augmente rapidement avec la valeur de l'élévation, en plongée comme en contre-plongée.

La tension ressentie est principalement influencée par une élévation haute, celle-ci diminuant les parties provenant de l'échelle de plan et de l'angle de vue. Les valeurs de ces composantes sont interpolées linéairement en fonction de la valeur absolue de l'élévation. L'apport de l'angle de vue devient nul à $\pm 90\,^\circ$, puisque la différence entre les angles y est alors imperceptible. Celle de l'échelle de plan, de son côté, est réduite à $\frac{1}{5}$ pour les élévations $\pm 90\,^\circ$, lui conservant une légère influence sur la valeur globale de T.

La jonction entre les plans peut aussi influencer la tension perçue de la séquence. L'écart entre les échelles de plans consécutifs amplifie la contribution de l'échelle à la tension. Ainsi, passer directement d'un plan d'ensemble à un plan rapproché ou un gros plan provoque un accroissement de la tension ressentie. Inversement, le passage rapide vers une échelle éloignée permet de la relâcher (voir la règle des $30\,^\circ$ à la section 2.8). La portion de la tension provenant de l'échelle est ajustée avec l'écart d'échelles entre le plan précédent et le plan courant, multipliée par un poids configurable, actuellement initialisé à 0.2. L'écart d'échelles est positive si la caméra s'approche du sujet, et négative si elle s'en éloigne.

Un changement d'échelle de plan engendre une discontinuité au niveau des tailles des objets dans l'image, habituellement camouflée par un changement d'angle. Négliger le changement d'angle de vue entre deux plans provoque ainsi un sentiment de tension, en vertu de la *règle des* $30\,^\circ$ (section 2.8). Lorsque deux plans consécutifs sont pris à partir d'un même angle de vue, la tension relative à l'angle de vue est directement posée à 1, au lieu d'être calculée comme suit.

$$
\begin{aligned}
T_{élévation} &= \frac{\frac{1}{2}\sqrt{|Élévation|}}{\sqrt{3\pi/2}} + \frac{1}{2} \\
T_{angle} &= \left[Pondération\,(Angle) \times \left(\frac{\pi}{2} - |Élévation|\right) + |Élévation|\right] / \frac{\pi}{2} \\
T_{échelle} &= \left[Pondération\left(Échelle\right) \times \left(\frac{5\pi}{8} - |Élévation|\right) + |Élévation|\right] / \frac{5\pi}{8} \times \\
&\quad \left(1 + DeltaÉchelle \times DeltaÉchellePoids\right).
\end{aligned}
$$



(a) (b)

FIG. 4.9 – (a) Fonction de Verhulst, (b) Vecteur de spatiosité.

Échelle de plan	S	I	T
Plan d'ensemble	0.95	0.10	0.10
Plan moyen	0.70	0.40	0.60
Plan rapproché	0.40	0.70	0.75
Gros plan	0.20	0.90	0.99

TAB. 4.2 – Pondérations de S, I et T par échelle de plan.

Angle de vue	I	T
0 °	0.99	0.99
45 °	0.85	0.85
90 °	0.50	0.50
135 °	0.25	0.75
180 °	0.10	0.99

TAB. 4.3 – Pondérations de I et T par angle de vue.

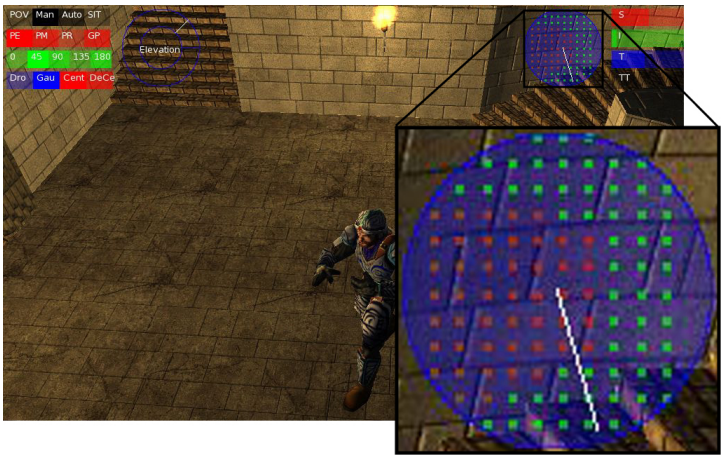


FIG. 4.10 – Points de spatiosité.

4.8.2 Variable émotionnelle *S*

Points de spatiosité À l'opposé des variables émotionnelles I et T, dont les valeurs se basent entièrement sur les caractéristiques de l'image, la variable S dépend aussi de la connaissance spatiale acquise autour de la position courante du personnage. La connaissance spatiale s'exprime par l'exposition de portions spécifiques de l'environnement dans les plans précédemment présentés.

Les points de spatiosité permettent de quantifier cette connaissance en partitionnant l'espace et en conservant pour chaque partie la mesure de son exposition dans les plans. Les points de spatiosité sont disposés selon une grille en deux dimensions, couvrant la surface des secteurs de l'environnement de jeu. Cette grille établit une carte de la connaissance spatiale des lieux. La bidimensionnalité de la grille représente la perception spatiale comme étant constante à la verticale d'une position donnée, ce qui modélise le champ visuel d'un individu regardant devant lui, tout comme les plans pris avec une élévation neutre. Bien que cette représentation soit suffisante pour orienter l'utilisateur dans l'environnement, l'utilisation d'une grille en trois dimensions occupant l'espace du secteur permettrait d'encore mieux mesurer la connaissance d'éléments spécifiques le composant, moyennant un coût plus élevé en mémoire et en temps de traitement.

Les points de spatiosité fonctionnent à la manière d'accumulateurs. Lorsque l'emplacement d'un point est inclus dans un plan, sa valeur est augmentée. Cette valeur diminue avec le temps, simulant ainsi la mémoire cognitive des lieux. Cette diminution induit une variation temporelle à la variable émotionnelle S. Une diminution rapide des

valeurs implique une grande variation de la valeur de la variable émotionnelle, alors qu'une diminution s'échelonnant sur une durée plus longue résulte en une valeur plus stable. Comme le nombre de plans utilisé pour calculer les variables contextuelles I et T, cette durée de vie participe à la dynamisation du système de variables *SIT*, qui crée une diversification dans les plans d'une séquence.

La variable contextuelle S provient de la moyenne des valeurs associées aux points de spatiosité localisés à l'intérieur d'un certain rayon autour du personnage. Ainsi, lorsque le personnage se situe dans une région peu connue, la Caméra *SIT* privilégiera les plans ayant un fort contenu spatial. Inversement, lorsqu'il se trouve dans une région abondamment montrée, les plans seront davantage sélectionnés pour leurs caractéristiques identificatoires. Grâce à la durée de vie de la valeur des points de spatiosité, une région précédemment visitée, mais délaissée depuis un moment, sera traitée comme un lieu nouvellement découvert. La durée de vie favorise aussi le recours régulier aux plans de situation, qui permettent de consolider la compréhension de l'environnement.

Les points de spatiosité pourraient se voir attribuer une valeur résiduelle non nulle qui permettrait d'obtenir un comportement différent de la Caméra *SIT* selon que le personnage entre dans une région n'ayant jamais été visitée, ou retourne dans une région antérieurement explorée. Cette valeur résiduelle représenterait la persistance de la mémoire des lieux connus. L'implantation actuelle de la Caméra *SIT* laisse les points de spatiosité redescendre à zéro, afin d'exprimer davantage de spatialisation dans les plans.

Autre mesure de perception spatiale, le vecteur de spatiosité est aussi extrait des points de spatiosité. Alors que la variable contextuelle S contient le niveau de connaissance spatiale pour la portion de l'environnement entourant le personnage, le vecteur de spatiosité indique l'écart de connaissance entre les différentes zones du secteur entier. L'évaluation du vecteur de spatiosité s'effectue sur l'ensemble des points du secteur courant, afin de diminuer la variabilité du vecteur au cours de la séquence. Le vecteur de spatiosité est composé à partir du vecteur normal au plan approximé par les points de spatiosité, dont leur hauteur provient de leur valeur. Ce vecteur normal est ensuite projeté sur le plan horizontal XZ pour former le vecteur de spatiosité (figure 4.9(b)). Alors que la direction du vecteur de spatiosité indique la zone ayant la moins été montrée récemment, sa norme donne l'amplitude des différences de connaissance pour le secteur.

Le vecteur de spatiosité influence la sélection des angles de vue de la caméra de façon à favoriser ceux montrant les régions les moins connues. Elle établit aussi le moment où une transgression de l'axe de jeu est désirable afin de révéler un côté négligé de l'environnement.

La figure 4.10 montre les points de spatiosité tels que représentés dans l'interface du jeu. Le cadran affiche la région circulaire entourant la position actuelle du personnage et les points de spatiosité y étant définis. Les points suivent un code de couleur dépendant de leur valeur. Un point est indiqué en vert lorsque sa valeur est nulle, et décale vers le rouge, proportionnellement à sa valeur. La forme de la pyramide de vue de la caméra intersectant les points est reconnaissable. La ligne blanche représente le vecteur de spatiosité obtenu à partir de cet ensemble de points. Sa norme est élevée puisque cette image provient d'un plan individuel et non d'une séquence de plans.

Distribution des points de spatiosité La distribution des points de spatiosité, utilisée pour l'implantation courante de la Caméra *SIT*, suit une grille régulière déployée sur l'aire de chaque secteur de l'environnement de jeu. Les points, actuellement espacés de 1 m, pourraient être rapprochés pour augmenter la précision de la variable S contextuelle et du vecteur de spatiosité. La répartition des points pourrait aussi être ajustée selon les zones d'intérêt de l'environnement de jeu [GSHG98] afin d'y attirer la direction de la caméra, et par conséquent l'attention de l'utilisateur.

Le traitement des points de spatiosité s'effectue par secteur. Cette gestion permet l'utilisation des maillages délimitant les pièces correspondant aux secteurs pour déterminer les positions des points à l'intérieur de ces pièces. Les secteurs définissant des aires connexes, toute la surface de l'environnement de jeu se trouve couverte par les points de spatiosité. La notion de secteur est donc transparente pour le fonctionnement des points, entièrement régi par leur valeur et la durée de vie de celle-ci.

Valeurs d'un point de spatiosité Un point de spatiosité possède deux valeurs distinctes pour mesurer la connaissance spatiale : sa charge et le nombre de plans ayant montré son emplacement. La charge ajoutée à un point dépend de la distance entre ce point et la caméra, et de certaines caractéristiques reliées au plan, comme son échelle, son élévation et sa durée. Le nombre de plans, de son côté, comptabilise simplement le nombre de fois où le point aura été vu dans l'image.

La charge, dont l'évaluation prend en compte le contexte entourant l'exposition du point dans un plan, est utilisée pour le calcul de la variable S contextuelle. Le nombre de plans est, par contre, plus indiqué pour le calcul du vecteur de spatiosité. Cette valeur étant moins nuancée que la charge, elle produit un vecteur plus prononcé.

Les valeurs de la charge et du nombre de plans d'un point de spatiosité sont incrémentées lorsque ce point se trouve inclus dans le champ de la caméra lors d'un plan. Les points de spatiosité composant une grille bidimensionnelle, un point est inclus lorsque son emplacement se situe à l'intérieur de la projection de la pyramide de vue de la caméra sur le plan horizontal XZ. La forme de cette projection diffère selon l'élévation de la caméra. Elle passe d'un triangle, lorsque l'élévation est nulle, à un quadrilatère, pour une élévation verticale. Afin de respecter les dimensions de la salle dans laquelle évolue le personnage, la pyramide de vue est tronquée par les limites (murs, planchers et plafond) de celle-ci.

La valeur de la charge associée à un plan prend la forme :

$$Charge \quad = \quad S_{échelle} \times S_{élévation} \times f(durée) \times \frac{1}{distance^2}.$$

$S_{échelle}$ exprime la proportion de l'image occupée par le décor relativement au personnage, en fonction de l'échelle de plan. Alors que les échelles larges, comme le plan d'ensemble ou le plan moyen, laissent une grande place à l'environnement, les plans rapprochés et les gros plans, se concentrant sur le personnage, ne permettent que de l'apercevoir (section 2.3). Ces proportions, exprimées sous forme de variables de configuration, sont présentées dans le tableau 4.2.

L'élévation de la caméra influence la perception spatiale dans l'image. Une plongée modérée améliore cette perception, alors que la contre-plongée lui nuit (section 2.5). Une plongée trop marquée restreint cependant l'image à une vue quasi verticale des lieux, limitant ainsi sa portée. Il a été établi empiriquement que l'évolution de la connaissance spatiale en fonction de l'élévation suivrait une fonction sous-linéaire, dont la valeur optimale se situerait à 45 °.

La distance d'un point de spatiosité dans l'image diminue la capacité de distinguer la région correspondant à son emplacement. Cette capacité est évaluée proportionnellement à la taille de la projection de cette région dans l'image, en fonction de la distance entre

le point de spatiosité et la caméra.

La durée d'un plan est reliée au processus de cognition de l'image. La fonction de croissance de Verhulst a été choisie pour représenter le niveau de déchiffrage d'une image en fonction du temps d'observation. Cette fonction logistique est souvent utilisée en biologie pour modéliser la croissance d'une population. Le taux d'acquisition de l'information croît lentement au début de l'observation, le système visuel effectuant un balayage préliminaire de l'image. Cette acquisition s'accélère par la suite pour plafonner après un certain temps lorsque l'essentiel de l'image a été analysé. La fonction de Verhulst s'exprime comme suit :

$$f(durée) = \frac{K \times P_0 \times e^{t \times durée}}{K + P_0 \times (e^{t \times durée} - 1)}$$

où *durée* représente la durée du plan en secondes, t le taux de croissance de la fonction, actuellement établi à 1.5, K la valeur maximale vers laquelle converge la fonction, fixée à 100, et P_0 la valeur initiale à partir de laquelle est calculée la croissance. Une croissance ne pouvant s'établir à partir d'une valeur nulle, P_0 est initialisée à 1. Le résultat de la fonction appartient à l'intervalle [1..100[. Ce résultat est transposé dans l'intervalle [0..1[, commun aux termes servant à l'évaluation de la charge. La figure 4.9(a) montre la croissance de la fonction avec ce taux. L'acquisition de l'information de l'image y étant représentée est surtout effectuée entre la première et la cinquième seconde d'observation.

La valeur cumulative des charges provenant de la série de plans produits est limitée à 1.

La charge d'un point de spatiosité et le nombre de plans l'incluant sont tous deux soumis à une diminution en fonction du temps. Ces valeurs disposent ainsi d'une durée de vie durant laquelle elles diminuent linéairement.

Cette altération est appliquée à la charge et au nombre de plans d'un point lorsque ceux-ci sont consultés, ou mis à jour avec les valeurs provenant d'un nouveau plan. Le temps correspondant à la dernière mise à jour, conservé dans le point de spatiosité, est alors soustrait du temps actuel et la valeur est ajustée selon la proportion de cette différence sur la durée de vie maximum établie.

La durée de vie des points de spatiosité s'allie au déplacement du personnage pour causer la variation de la variable contextuelle S, ce qui favorise les changements d'échelle,

d'angle de vue et d'élévation entre les plans constituant une séquence. Dans le cas d'un personnage fixe, seule la durée de vie joue ce rôle. La durée de vie des points doit donc être plus courte dans ce cas, afin de converser la diversité des plans. Les points de spatiosité possèdent une durée de vie de 20 s lorsque le personnage est en déplacement, mais cette valeur descend à 10 s lorsqu'il reste fixe.

Variable contextuelle *S* La variable contextuelle S exprime le niveau de connaissance de la région entourant la position courante du personnage, au moment d'un changement de plan. Les plans produits étant centrés sur le personnage, cette région correspond à la portion de l'environnement représentée dans le prochain plan, et la variable contextuelle S en représente ainsi la connaissance.

La variable contextuelle S est obtenue en calculant la moyenne des charges des points de spatiosité contenus à l'intérieur d'un cercle entourant le personnage. Le rayon choisi pour ce cercle est de 10 m, ce qui couvre une surface significative autour du personnage. Cette surface contient aussi un nombre significatif de points de spatiosité, ceux-ci étant distancés de 1 m.

Calcul du vecteur de spatiosité Le vecteur de spatiosité indique la région de l'environnement pour laquelle la connaissance spatiale est moindre. Il favorise, pour le prochain plan, l'angle de vue donnant accès à cette région et peut déterminer que la caméra doit transgresser l'axe de jeu afin de montrer une portion de l'espace négligée par les derniers plans confinés à un même côté de l'axe.

Le vecteur de spatiosité provient du vecteur normal à un plan qui passerait par les points 3D correspondant aux points de spatiosité, dont les emplacements en 2D sont augmentés par leur nombre de plans respectif, comme troisième dimension (figure 4.9(b)). La valeur du nombre de plans utilisée est altérée par le temps écoulé depuis la dernière vue du point. La hauteur des points étant moins élevée à l'emplacement de la zone la moins connue, la pente du plan représente la différence de connaissance enregistrée dans ce secteur. L'orientation de la projection du vecteur normal au plan, sur le plan horizontal XZ pointe vers cette zone. Le vecteur normal étant normalisé, la norme de sa projection est proportionnelle à la pente du plan, et informe sur le degré de décalage cognitif, relatif au reste du secteur. Cette norme détermine l'importance à attribuer à l'orientation du vecteur de spatiosité lors de la sélection de l'angle de vue du prochain

plan ou du changement de côté de l'axe de jeu.

Le plan correspondant aux points de spatiosité est approximé par moindres carrés, appliqués avec la décomposition en valeurs singulières (SVD). Le système utilise les bibliothèques de *GNU Scientific Library* pour effectuer la décomposition SVD.

L'équation du plan

$$Ax + By + Cz + D = 0$$

où D représente une valeur constante correspondant à la distance du plan à l'origine, s'exprime de la façon suivante,

$$\begin{pmatrix} P_1.x & NbPlans\,(P_1) & P_1.y \\ P_2.x & NbPlans\,(P_2) & P_2.y \\ P_3.x & NbPlans\,(P_3) & P_3.y \\ \vdots & \vdots & \vdots \end{pmatrix} \times \begin{pmatrix} A \\ B \\ C \end{pmatrix} = \begin{pmatrix} -D \\ -D \\ -D \\ \vdots \end{pmatrix}.$$

P_i représente un point de spatiosité dont l'emplacement est défini par ses coordonnées x et y. Le nombre de plans l'ayant inclus est donné par la fonction $NbPlans$.

Comme la normale au plan ne dépend pas de la distance de ce dernier à l'origine, D est remplacé par une constante arbitraire dans le système d'équations. La résolution du système par (SVD) produit les coefficients A, B et C définissant le plan. La normale au plan, définie par ces mêmes coefficients, est projetée sur le plan horizontal XZ pour obtenir le vecteur de spatiosité.

Évaluation de S pour les plans potentiels Les points de spatiosité conservent la connaissance spatiale acquise lors des derniers plans présentés et constituent ainsi la variable S contextuelle. Afin d'effectuer la sélection du prochain plan, il faut pouvoir comparer cette valeur contextuelle avec les pondérations des différentes configurations potentielles. L'évaluation de cette valeur potentielle de S est semblable à celle de la charge d'un point de spatiosité.

$$S_{potentiel} = S_{échelle} \times S_{élévation} \times Const_{dist} \times S_{angle}.$$

Les termes $S_{échelle}$ et $S_{élévation}$ sont directement liés aux paramètres de la caméra pour le plan potentiel et utilisent les mêmes valeurs que dans l'évaluation de la charge.

Le terme $Const_{dist}$ représente une estimation de la capacité moyenne à distinguer la position d'un point de spatiosité, en fonction de la distance entre ce point et la caméra. Cette constante a été empiriquement établie à $\frac{1}{3}$, pour la production des séquences présentées au chapitre 5.

Difficilement évaluable pour un plan potentiel, la notion de durée est éliminée du calcul. Son retrait est cependant équilibré par l'ajout d'un nouveau terme, absent de l'évaluation de la charge. S_{angle} augmente la pondération des angles de vue pour lesquels l'orientation de la caméra est semblable à la direction du vecteur de spatiosité.

$$S_{angle} = |VecteurSpatiosité| \times \left(\frac{VecteurSpatiosité \cdot AngleVue}{2 \times |VecteurSpatiosité| \times |AngleVue|} + \frac{1}{2} \right) + (1 - |VecteurSpatiosité|) .$$

$AngleVue$ représente l'orientation de la caméra résultant de l'angle de vue sur le personnage. La valeur de S_{angle} provient du cosinus du plus petit angle formé entre l'orientation de la caméra et le vecteur de spatiosité. Cette valeur est transposée entre 0 et 1. Elle est élevée lorsque l'angle entre les deux vecteurs est petit et faible lorsque cet angle est plus grand. La norme du vecteur de spatiosité pondère linéairement la contribution de cet angle à la valeur de $S_{potentiel}$. Le vecteur de spatiosité provenant de la projection d'un vecteur de longueur unité, sa norme est incluse dans l'intervalle [0..1]. Lorsque la norme est faible, les valeurs de S_{angle} sont semblables pour tous les angles de vue, alors qu'elles prennent la valeur du produit scalaire lorsque la norme est élevée, et varient selon l'écart entre l'angle de vue et le vecteur de spatiosité.

4.9 Sélection du prochain plan

L'étape de sélection du prochain plan constitue le coeur fonctionnel de la Caméra *SIT*. Elle détermine l'échelle de plan, l'angle de vue et l'élévation qui constitueront le prochain plan.

La sélection vise à choisir, parmi l'ensemble des configurations potentielles, préalablement réduit par l'application des règles d'esthétique (section 4.5) et la vérification de la visibilité du personnage (section 4.6), celle qui permet le mieux d'atteindre le contexte émotionnel désiré. Elle tente de minimiser les différences entre les valeurs des variables *SIT* potentielles, spécifiques à chaque configuration, et les valeurs

cibles à exprimer dans le prochain plan. Les valeurs cibles proviennent des variables *SIT* contextuelles.

Comme les paramètres de caméra possèdent des pondérations S et I généralement opposées (section 4.8), les valeurs S et I cibles cherchent à compenser les valeurs courantes afin d'obtenir des composantes émotionnelles S et I équivalentes dans la séquence.

Les variables S et I étant traitées en opposition et toujours positives, elles peuvent s'exprimer une par rapport à l'autre et être combinées en une variable unique. SI représente la proportion de la variable S dans la somme de ces deux valeurs.

$$SI = \frac{S}{S+I}.$$

Cette combinaison permet de réduire le problème de moindres carrés à deux variables pour trouver la configuration appropriée.

Pour obtenir des valeurs $S_{contextuelle}$ et $I_{contextuelle}$ équivalentes, la proportion $\frac{S}{S+I}$, représentée par $SI_{contextuelle}$, doit tendre vers 0.5. Ainsi la valeur de SI_{cible}, pour le prochain plan, et $SI_{contextuelle}$ des plans précédents doivent se compléter.

$$SI_{cible} = 1 - \frac{S_{contextuelle}}{S_{contextuelle} + I_{contextuelle}}.$$

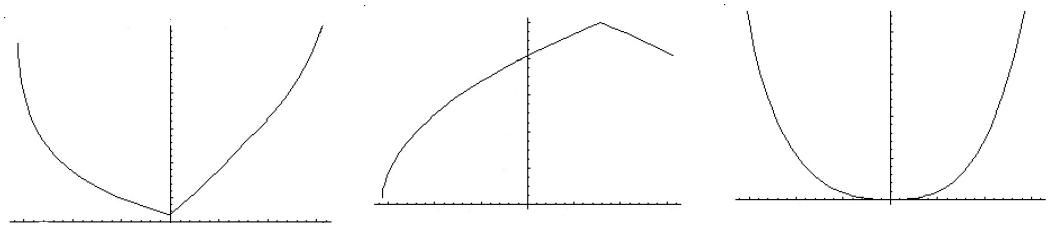
La proportion de 0.5, traitant les éléments spatial et identificatoire en parties égales, pourrait être changée pour favoriser une ou l'autre des composantes.

Au tout début de la séquence, les valeurs de $S_{contextuelle}$ et $I_{contextuelle}$ sont toutes deux nulles, et croissent avec l'évaluation des premiers plans. Afin de donner préséance à la perception spatiale, pour des raisons d'orientation et de manipulation du personnage, $I_{contextuelle}$ est ignorée et $SI_{contextuelle}$ est remplacée par $S_{contextuelle}$, tant que cette dernière valeur n'ait pas atteint un seuil minimum. S_{cible}, remplaçant SI_{cible}, prend alors la valeur 1, puisque les plans viseront alors à maximiser la spatialisation au détriment de l'identification. Le seuil de connaissance spatiale minimum a été arbitrairement établi à 0.15. Cette valeur permet la production de quelques plans entièrement axés sur la connaissance spatiale en début de séquence.

Le niveau de tension exprimée par une séquence est manuellement déterminé par l'utilisateur. La variable émotionnelle T doit donc converger vers cette valeur de tension

désirée, au cours de la séquence. $T_{désirée}$ étant la moyenne entre les valeurs $T_{contextuelle}$, correspondant aux plans précédents, et T_{cible} du prochain plan,

$$T_{cible} \quad = \quad 2 \times T_{désirée} - T_{contextuelle}.$$



(a) (b) (c)

FIG. 4.11 – (a) Valeurs de SI en fonction de l'élévation, (b) Valeurs de S en fonction de l'élévation, (c) Valeurs de T en fonction de l'élévation.

Chaque configuration potentielle étant représentée sous la forme d'un intervalle d'élévations, associé à un couple échelle de plan et angle de vue, les valeurs SI et T correspondantes sont exprimées comme des fonctions de l'élévation pour ce couple. Les figures 4.11(a) et 4.11(c) montrent respectivement les fonctions définies entre les valeurs SI et T et l'élévation, pour une échelle de plan et un angle de vue donnés. Lorsque seule la valeur S est considérée, la fonction représentée à la figure 4.11(b) remplace celle de SI.

La Caméra SIT cherche parmi les paires de fonctions SI et T potentielles l'élévation, l'échelle de plan et l'angle de vue pour lesquels la somme des carrés des différences avec les valeurs cibles est minimisée

$$\min \left((SI_{cible} - SI_{potentielle})^2 + (T_{cible} - T_{potentielle})^2 \right).$$

Si la valeur de $S_{contextuelle}$ est sous son seuil minimum, la Caméra SIT cherche à minimiser en fontion de la variable S indépendamment

$$\min \left((1 - S_{potentielle})^2 + (T_{cible} - T_{configuration})^2 \right).$$

La Caméra *SIT* résout ces moindres carrés pour chaque couple échelle de plan et angle de vue, afin de trouver l'élévation optimale et son pointage correspondant. Les pointages ainsi obtenus sont comparés entre eux pour déterminer l'échelle de plan et l'angle de vue complétant la solution.

L'application des moindres carrés fournit une solution intermédiaire entre les configurations remplissant les objectifs des valeurs *SI* et *T*. Elle favorise les procédés cinématographiques communiquant le niveau de tension désiré, qui équilibrent le mieux les pôles spatial et identificatoire.

L'exactitude de la solution aux moindres carrés n'est toutefois pas cruciale. Les variables *SIT*, sur lesquelles se base cette évaluation, possèdent déjà des pondérations empiriques et arbitraires, et les caractéristiques émotionnelles des plans se régularisent sur l'ensemble de la séquence, grâce à la rétroaction des variables *SIT*. La Caméra *SIT* peut donc implanter un algorithme approximatif plus rapide pour cette résolution.

Cet algorithme détermine de manière dichotomique pour chacune des fonctions *SI* et *T* la valeur d'élévation pour laquelle la fonction s'approche le plus de sa valeur cible. Les moindres carrés sont ensuite échantillonnés systématiquement, selon un incrément établi, entre ces deux élévations jusqu'à l'obtention d'un minimum.

Comme la composante spatiale est moins essentielle lorsque le personnage est fixe, la Caméra *SIT* peut alors chercher à augmenter l'identification au personnage. Ce décalage émotionnel s'effectue en diminuant la valeur de SI_{cible}, et conséquemment la proportion de la composante spatiale, par un facteur configurable, afin de favoriser les plans possédant un plus haut niveau d'identification.

4.10 Gestion de l'axe de jeu

La règle de montage des 180° (section 2.8) demande que les plans constituant une séquence soient exécutés d'un même côté de l'axe de jeu. Cette règle permet la construction d'une représentation mentale cohérente de l'espace, à partir des portions continues de l'environnement montrées dans les différents plans, où les éléments de la scène sont mis en conjonction de façon consistante. Elle fournit aussi une direction constante du personnage dans l'image, ce qui procure les repères pour son orientation.

La Caméra *SIT* utilise le côté du personnage (figure 4.3(b)) pour identifier le côté

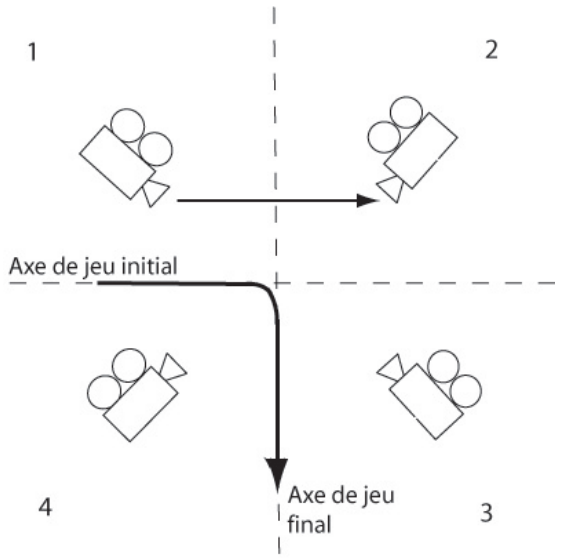


FIG. 4.12 – Côté de la caméra suite à un changement de direction d'un personnage.

de l'axe de jeu permettant de positionner la caméra. Dans le cas d'un personnage en mouvement, l'axe de jeu est défini par la trajectoire de son déplacement. Ainsi, le côté de l'axe de jeu devra être ajusté selon l'orientation courante du personnage, afin d'obtenir une vue constante malgré ses changements de direction.

Le côté d'axe de jeu est calculé à partir des vecteurs des directions initiale et finale du personnage dans le plan. Le côté d'axe de jeu reste constant entre deux plans si la direction du personnage reste sensiblement la même, c'est-à-dire si le plus petit angle formé entre ces deux directions est inférieur à 45 °.

$$DirectionInitiale \cdot DirectionFinale > 0.5,$$

assumant que les deux vecteurs de direction sont des vecteurs unitaires.

Afin de conserver la même région pour la position de la caméra, le côté de l'axe de jeu doit être inversé si la direction finale du personnage est opposée à sa direction initiale, c'est-à-dire, si le plus petit angle entre ces deux directions est supérieur à 135 °.

$$DirectionInitiale \cdot DirectionFinale < -0.5.$$

Si l'angle entre les directions initiale et finale est entre 45 ° et 135 °, la Caméra *SIT* conclut que le personnage a effectué un virage à gauche ou à droite pendant le plan.

Afin de favoriser la compréhension du mouvement, un changement de direction doit être visible dans l'image. Si la caméra se trouve à l'intérieur d'un virage à la fin d'un plan, elle doit y demeurer pour le plan suivant. Pareillement, un plan se terminant sur un virage montré de l'extérieur, doit être suivi par une vue extérieure de la suite de ce mouvement.

La figure 4.12 montre les différentes configurations de la position de la caméra pour un changement de direction du personnage. Lorsque la caméra est située dans les cadrans 2 et 3, les régions correspondant à la position de la caméra relativement aux axes de jeu définis au début et à la fin du mouvement montrent toutes deux l'extérieur du virage. De façon similaire, une caméra située dans le cadran 4 demeure à l'intérieur du virage à la fin du mouvement. Dans ces trois cas, le côté de l'axe de jeu provient du côté du personnage où se trouve la caméra à la fin du mouvement.

Dans le cadran 1, cependant, la caméra est placée à l'extérieur du virage au début du mouvement, mais la région située de son côté de l'axe de jeu défini à la fin du mouvement correspond à l'intérieur du virage. La caméra doit donc être déplacée vers la région représentée par les cadrans 2 et 3 afin de rester à l'extérieur du virage dans le prochain plan. Le côté de l'axe de jeu et donc inversé par rapport au côté du personnage où se trouve la caméra.

Comme le côté d'axe varie peu durant une séquence, sa valeur initiale doit être significative. Le nombre de points de spatiosité disposés de chaque côté de la position du personnage en début de séquence permet d'identifier la région du secteur possédant l'aire la plus importante. Le côté d'axe de jeu est initialisé au côté opposé à cette région afin qu'elle puisse être montrée dans les plans.

Il peut être désirable de transgresser l'axe de jeu afin d'exposer une portion de l'environnement actuellement inaccessible et négligée dans les plans antérieurs. Cette transgression peut s'effectuer par l'utilisation d'un angle de vue neutre intermédiaire. Les angles 0° et 180° sont neutres, puisqu'ils ne sont associés à aucun côté de l'axe de jeu. Le recours à un angle neutre évite la désorientation en éliminant les vues opposées consécutives et l'inversion subite de la direction du personnage. Lors de la sélection du premier plan suivant une inversion du côté de l'axe de jeu, les angles de vue potentiels sont donc restreints à 0° et 180° afin de faciliter la transition.

L'inversion du côté de l'axe de jeu est effectuée lorsque le vecteur de spatiosité indique la région opposée à celle étant actuellement montrée, et que sa norme implique une différence de connaissance spatiale élevée entre ces deux régions. Un certain nombre de plans consécutifs doivent être montrés à partir d'un même côté, afin de préserver la cohérence entre les plans. Les seuils actuellement configurés dans le système sont de 0.4 pour la norme du vecteur de spatiosié et de 5 plans consécutifs.

4.11 Transitions

Les transitions relient les prises de vue de lieux distincts afin de les combiner en une séquence continue. Il existe plusieurs types de transitions, dont la coupe franche, les mouvements de caméra et les ponctuations (fondus au noir ou au blanc, plan filé etc.), possédant leurs qualités expressives respectives.

Deux transitions ont été implantées dans la version actuelle de la Caméra *SIT*, la coupe franche et le panoramique de recadrage. Alors que la coupe franche se distingue des mouvements de caméra en localisant directement la caméra à sa prochaine position, le panoramique effectue une brève rotation de la caméra.

La coupe franche est la principale transition utilisée par la Caméra *SIT* entre les plans. Elle constitue la figure de transition la plus répandue dans la production cinématographique, et y est observable à toutes les époques. Elle est caractéristique des séquences montées, telles celles produites par la Caméra *SIT*, mais permet aussi de joindre des plans séquences.

Contrairement aux mouvements de caméra, la coupe franche peut relier des plans pour lesquels les positions de caméra ne possèdent aucun paramètre commun. Elle s'applique ainsi facilement à tous les contextes de transition, dont ceux impliquant des conditions restreintes dues à la visibilité du personnage.

Le panoramique de recadrage permet de conserver le personnage centré dans le cadre du plan. Appliquant le principe de centralité du style classique Hollywoodien, il en souligne l'importance. Ce mouvement de caméra permet aussi une transition plus douce que la coupe franche. Il est utilisé lorsque le personnage s'éloigne du centre de l'image sans sortir du cadre.

Bien que l'intégration d'autres mouvements de caméra résulterait en une plus grande variété dans les effets cinématographiques produits, la coupe franche et le panoramique sont suffisants pour démontrer le fonctionnement du modèle *SIT* exposé dans ce mémoire.

Les conditions de changements de plan appelant une coupe franche visent à constamment maintenir le personnage visible dans l'image. Ainsi, dans le cas d'un personnage en déplacement, les sorties de champ provoquent la majorité des changements de plan. Six zones de hors champ sont formées des quatre régions définies par les limites du cadre de l'image, de celle à l'arrière de la caméra, et de celle située dans le fond de l'image.

Un changement de plan est initié lorsque le personnage traverse un des bords du cadre. Cette transition résulte en un *frame cut* (section 2.8). Une coupe franche est aussi effectuée lorsque la position du personnage correspond aux coordonnées Z négatives de la caméra ou que sa distance avec la caméra est trop élevée. Une distance inférieure à 10 m, soit deux fois la distance utilisée pour le plan d'ensemble, accorde au personnage une certaine liberté de déplacement dans l'image, tout en évitant qu'il en constitue une proportion trop faible.

Un nouveau plan est aussi commandé lorsque le personnage n'est plus visible, dû à son occultation par un élément de l'environnement. Comme l'occultation temporaire d'un personnage en déplacement est un indice fort de la perception de la profondeur dans l'image cinétique [Cut05], seules les occultations plus longues résultent en un changement de plan. L'occultation est évaluée par un rayon tracé entre la caméra et le centre du personnage. Si ce rayon intersecte un maillage, l'occultation est signalée.

Comme les sorties de champ du personnage gèrent les changements de plan, la durée d'un plan se trouve naturellement reliée à l'échelle de plan utilisée, l'image d'une échelle rapprochée étant traversée plus rapidement que celle d'une échelle large. Cette variation dans la durée des plans transmet un rythme à la séquence, qui contribue à la tension ressentie.

Pour un personnage fixe, les changements de plan permettent de varier les points de vue. Ils sont alors entièrement déclenchés par la durée des plans, proportionnelle à l'échelle de plan. Un plan large requérant un plus grand temps d'analyse visuelle qu'un

plan rapproché (section 2.8), il est présenté plus longtemps.

Contrairement à la coupe franche qui établit immédiatement la nouvelle position de la caméra, le panoramique doit s'effectuer progressivement au cours des images, afin d'obtenir un aspect continu.

À chaque image, la matrice d'orientation de la caméra est transformée par une matrice de rotation selon l'axe Y, représentant l'angle de rotation mesuré depuis la dernière image présentée. La vitesse angulaire, exprimée en $^\circ/s$, est multipliée par la différence de temps entre deux images consécutives pour produire cet angle de rotation. La vitesse angulaire est une variable de configuration du système. La valeur de $45\,^\circ/s$ a été utilisée pour l'application.

Chapitre 5

Présentation des résultats

5.1 Méthode utilisée

Analyse d'une séquence Tel qu'exposé au chapitre 1, la finalité de la Caméra *SIT* comprend l'amélioration de la compréhension spatiale des lieux virtuels visités, une plus grande identification psychologique de l'utilisateur avec le personnage et la communication d'un niveau déterminé de tension dans la scène présentée.

La perception émotionnelle et cognitive appartient au domaine subjectif et est variable selon les individus. La quantification de cette perception nécessite une étude menée sur un échantillonnage d'utilisateurs et une interprétation statistique des résultats s'en dégageant. Nous laissons en travail futur la réalisation de cette étude (chapitre 6) et utilisons plutôt certaines mesures indicatives afin d'analyser les séquences produites. Les séquences présentées dans ce chapitre feront l'objet d'une analyse qualitative descriptive et d'une analyse quantitative utilisant ces mesures.

Une analyse filmique étudie la construction d'une séquence, en se basant sur la composition des plans et de leur combinaison, afin d'en dégager une signification qui, s'ajoutant à la narration, contribue à l'expression cinématographique de l'oeuvre. L'analyse d'une séquence cherche à identifier l'observateur dont le point de vue est représenté par le plan, la distance établie entre cet observateur et l'action montrée, ainsi que les liens symboliques et les raccords visuels établissant des liaisons entre personnages ou avec les événements.

La Caméra *SIT* ne communiquant que les trois orientations émotionnelles définies,

l'analyse qualitative des séquences produites se limite à couvrir leur représentation, s'attardant particulièrement à la continuité de l'espace représenté, essentielle à l'orientation de l'utilisateur manipulant l'avatar.

Les mesures retenues pour quantifier les contributions émotionnelles et cognitives des séquences comprennent les variations d'échelles de plan, les choix d'élévations utilisées dans les plans, l'évolution des variables émotionnelles S et I, et les valeurs de la variable émotionnelle T mises en relation avec le niveau de tension déterminé. Une évaluation de la connaissance spatiale développée au cours de la séquence est obtenue par les points de spatiosité (section 4.8).

Les points de spatiosité discrétisent l'environnement dans lequel évolue le personnage, chaque point de spatiosité contenant le niveau de connaissance spatiale de la région correspondant à son emplacement. Un point de spatiosité possède deux valeurs, sa charge et le nombre de plans dans lesquels il apparaît. La charge d'un point dépend de l'échelle, l'élévation et la durée du plan montrant l'emplacement du point, ainsi que de la distance du point à la caméra. Comme nous cherchons à mesurer la distribution des prises de vue dans l'environnement, le nombre de plans est mieux indiqué pour estimer la connaissance spatiale transmise par les séquences produites par la Caméra *SIT*. Contrairement aux valeurs servant au calcul du vecteur de spatiosité, l'accumulation du nombre de plans utilisée pour cette quantification n'est pas altérée en fonction du temps.

(a) Vestibule (b) Salle du trône (c) Salle à manger

FIG. 5.1 – Secteurs de l'environnement de jeu.

Production des séquences Les séquences analysées dans ce chapitre sont produites à partir d'un même itinéraire du personnage à l'intérieur de l'environnement virtuel.

L'enregistrement de cet itinéraire s'effectue en conservant les positions et orientations consécutives du personnage, synchronisées avec l'horloge virtuelle du jeu.

L'environnement virtuel utilisé, un château médiéval, contient trois secteurs, chacun correspondant à une salle avec des caractéristiques distinctes de visibilité. Le secteur correspondant au vestibule (figure 5.1(a)) se compose d'une aire ouverte avec peu de potentiel d'occultation, prolongée de deux corridors étroits qui imposent leurs restrictions sur les prises de vue. Les escaliers menant à la salle du trône constituent un environnement mi-clos. Dans la salle du trône (figure 5.1(b)), des colonnes régulièrement disposées fournissent des limitations éparses de visibilité. La salle à manger (figure 5.1(c)) contient, pour sa part, plusieurs maillages représentant les tables et chaises, et fournit un environnement d'occultation assez dense, bien que concentrée vers le bas.

Plusieurs séquences ont été produites afin d'observer les différentes solutions proposées par la Caméra *SIT*. Tous les plans constituant les séquences ont été automatiquement sélectionnés par la Caméra *SIT*. La séquence de base correspond à une tension nulle, reliée à une vision neutre de la réalité. Elle est utilisée pour analyser la représentation de l'espace et l'équilibre entre la perception spatiale et l'identification avec le personnage. Les caractéristiques de cette séquence sont comparées à celles provenant d'une séquence générée semi-aléatoirement, reproduisant des règles similaires à celles de la Caméra *SIT*, mais sans le recours aux variables émotionnelles *SIT*. Cette comparaison vise à mesurer l'apport de l'heuristique émotionnelle de la Caméra *SIT* sur la perception de l'utilisateur.

Des séquences communiquant respectivement des tensions de 0.2 et 0.4 sont aussi comparées à cette séquence neutre, afin de constater les choix de plans effectués en fonction de la tension à représenter. Avec la configuration actuelle de la pondération émotionnelle du système, la tension des séquences produites peut s'étendre de 0.1 à 0.4. L'examen de deux séquences provenant de films connus contribue à l'évaluation de l'impact de ces choix de positions de caméra.

La Caméra *SIT* ne contrôle pas l'illumination du personnage. Ainsi, dans les séquences présentées, celle-ci est entièrement dépendante de l'illumination créée pour l'environnement virtuel utilisé. L'aspect de certains plans, particulièrement les gros plans et les plans rapprochés, souffre d'une sous-exposition du personnage. Une illumination liée au personnage, laissée en travail ultérieur, permettrait d'améliorer l'esthétique des

plans.

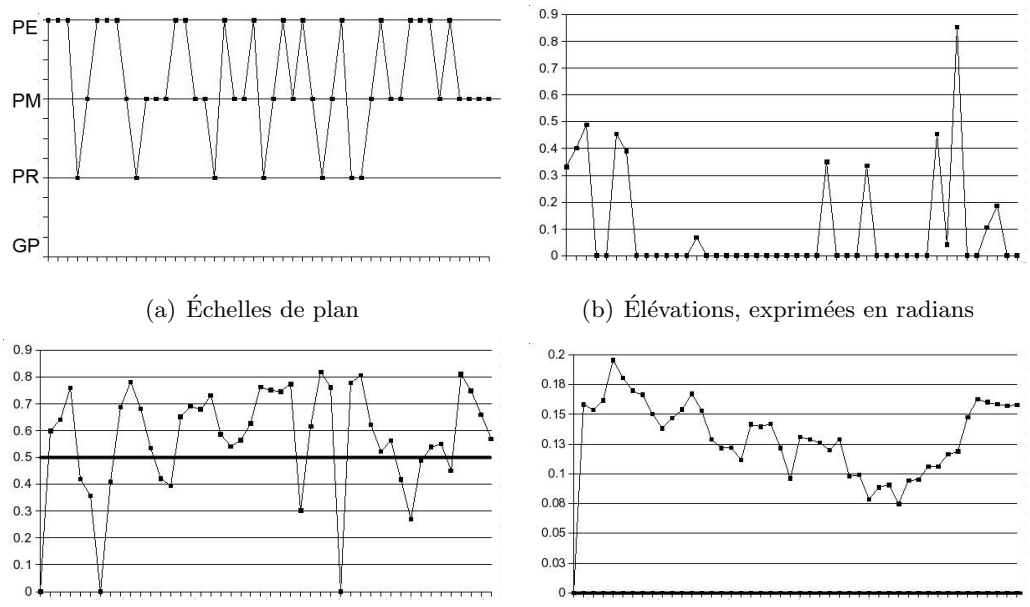
5.2 Analyse d'une séquence communiquant une tension nulle

Description de la séquence La première séquence présentée a été produite avec une tension désirée de 0. La séquence débute avec le personnage traversant le corridor menant à la salle à manger en direction du vestibule. Il se dirige alors vers la salle du trône qu'il explore pendant un certain temps. Sortant de la salle du trône, il prend le corridor menant aux escaliers du donjon, puis revient sur ses pas pour se rendre à la salle à manger, située à l'autre bout du vestibule. Il parcourt la salle à manger entre les tables et les chaises. La séquence se termine alors que le personnage, sortant de la salle à manger, regagne sa position initiale.

Les figures 5.9 à 5.13 montrent l'intégralité des plans constituant cette première séquence. La figure 5.2 présente les variations des échelles de plan utilisées dans la séquence, les élévations avec lesquelles chaque plan est vu, la progression des variables émotionnelles contextuelles S et I au cours de la séquence, et les niveaux de tension contextuels calculés.

Analyse de la séquence Les premiers plans marquant l'arrivée du personnage dans une salle sont présentés en plan d'ensemble avec un angle de vue de 180 ° relativement au personnage (figures 5.9(a) à 5.9(c), 5.9(h), 5.11(g) et 5.11(m)). Cette composition s'assimile à un plan subjectif puisqu'elle fournit à l'utilisateur une vue ayant la même orientation que le personnage. Ces plans ont une fonction de plans de situation, permettant à l'utilisateur de positionner d'entrée de jeu certains éléments de la pièce. L'élévation haute de ces plans augmente la perception de la disposition des lieux.

Les plans subséquents découpent l'espace avec des échelles plus rapprochées. L'échelle médiane utilisée dans cette séquence est le plan moyen qui, montrant à la fois le personnage et son environnement immédiat, est indiquée pour représenter les déplacements du personnage. Les plans moyens sont entrecoupés par des plans rapprochés qui permettent une vue plus intime du personnage (figures 5.9(d), 5.9(n), 5.10(j), 5.10(o), 5.11(k), 5.11(o) et 5.12(a)). Comme ils nuisent à la manipulation du personnage, la Caméra *SIT* évite de présenter plusieurs plans rapprochés consécutifs. La reprise



(a) Échelles de plan

(b) Élévations, exprimées en radians

(c) $\frac{S}{S+I}$

(d) Tension

FIG. 5.2 – Séquence *SIT*, tension=0. L'abscisse des graphes représente la séquence de plans.

régulière de plans d'ensemble permet à l'utilisateur de mieux situer le personnage dans l'environnement spatial.

Les plans d'ensemble établissent aussi une composition d'image en profondeur qui permet de visualiser les déplacements du personnage à l'intérieur du cadre de l'image (figures 5.9(c), 5.9(h) à 5.9(k), 5.11(g), 5.11(h), 5.13(b) et 5.13(c)). Les panoramiques, particulièrement celui effectué au plan 37 (figures 5.12(k) à 5.12(m)), étendent la portée de ces plans en augmentant la surface couverte par la prise de vue. Le plan 30 (figures 5.11(m) et 5.11(n)) renforce l'impression de profondeur de son image par une composition faite au travers d'une ouverture. Les scènes filmées en profondeur de champ présentent l'action en temps réel et suscitent ainsi l'attention de l'utilisateur. Elles installent cependant une distance émotionnelle entre l'utilisateur et l'action en cours.

La figure 5.2(a) permet de constater l'équilibre établi entre les échelles de plan au cours de la séquence. Les variations des échelles de plan impliquent, dans la majorité des cas, une seule taille de plan, une variation de plusieurs tailles étant pénalisée par une tension plus élevée.

Un changement de plan s'effectuant lorsque le personnage sort du champ, la durée d'un plan est directement proportionnelle à son échelle. Le recours pour cette séquence

aux plans d'ensemble et moyens impose un rythme assez lent, ponctué par la distribution des plans rapprochés.

La majorité des plans formant cette séquence utilisent une élévation nulle (figure 5.2(b)). Les plans, dont l'élévation produit une image semblable à la vision usuelle du spectateur, induisent une perception de faible tension. Une élévation élevée est associée aux plans ayant une forte composante spatiale (figures 5.9(a) à 5.9(c), 5.9(h), 5.11(g), 5.11(m) et 5.12(o)). L'élévation sert aussi parfois à éviter l'occultation du personnage par un élément du décor (figures 5.9(f) et 5.12(k)). Bien que la contre-plongée puisse aussi être utilisée pour résoudre la visibilité, aucun plan constituant cette séquence ne l'utilise dans ce but. Cette préférence pour les plongées s'explique par la présence d'un plus grand nombre de maillages dans la partie inférieure des secteurs dans l'environnement utilisé.

Les changements d'angle de vue entre les plans dynamisent la séquence et permettent de synthétiser l'espace morcelé par le cadre de l'image. On constate un équilibre entre les angles arrière au personnage, facilitant sa direction, et les vues avant ou de profil qui favorisent sa reconnaissance.

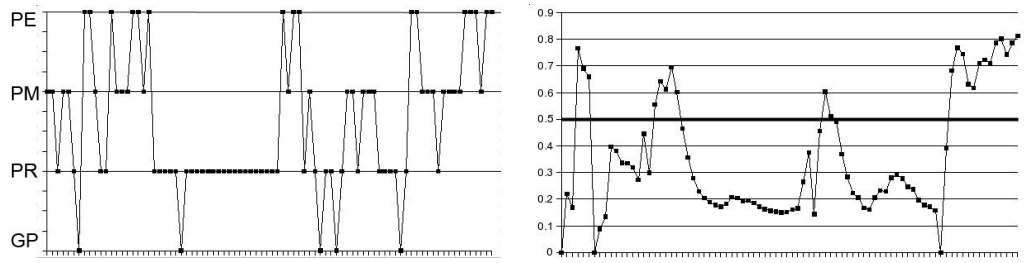
Les figures 5.11(m) à 5.12(n) montrent comment la salle à manger est révélée par les plans. Le plan 30 introduit la pièce et fournit une vue sommaire des tables et chaises la meublant. Le plan 31, bien qu'en plan rapproché, montre le mur extérieur de la pièce. Les plans 32 et 33, quant à eux, décrivent le mur opposé, sur lequel on remarque deux tapisseries. Le plan 32 situe ces tapisseries l'une par rapport à l'autre. Les trois panoramiques des plans 34, 35 et 37 lient respectivement le mur extérieur et le mur du fond, le mur du fond avec le mur intérieur, et le mur intérieur avec le mur d'entrée. Le plan 38 établit la conjonction entre le mur d'entrée et les murs adjacents. La disposition des tables et chaises de la salle est entièrement révélée aux plans 37 et 38. Idéalement, cette information aurait dû être disponible dès le plan 30, mais était partiellement dissimulée par la porte.

La position de la caméra par rapport à l'axe de jeu permet de communiquer un espace cohérent en suivant le personnage autour de la pièce. Initialement située à la gauche du personnage (plans 31 à 34), elle traverse du côté droit à la faveur d'un changement de direction du personnage, visible au plan 34. Elle conserve cette position lors du tournant

effectué au plan 35, de façon à rester à l'intérieur du mouvement. La caméra demeure aussi de ce même côté de l'axe de jeu suite au virage à 180 ° du personnage, montré au plan 41 (figures 5.13(b) et 5.13(c)).

Une gestion similaire de l'axe de jeu, subordonnée à la direction du personnage, peut être constatée dans le reste de la séquence. Les plans 26 à 28 (figures 5.11(h) à 5.11(k)) utilisent un angle de vue neutre (180 °) afin d'initier une transgression de l'axe de jeu permettant de voir la partie du vestibule opposée à la porte d'entrée. Un changement de côté de l'axe de jeu similaire peut aussi être observé aux plans 17 à 19 (figures 5.10(j) à 5.10(l)).

Dans les plans 27 et 29 (figures 5.11(i), 5.11(j) et 5.11(l)), la Caméra *SIT* favorise l'angle de vue arrière plutôt qu'une échelle de plan rapprochée pour traiter l'espace exigu du corridor.



(a) Échelles de plan (b) $\frac{S}{S+I}$

FIG. 5.3 – Séquence semi-aléatoire.

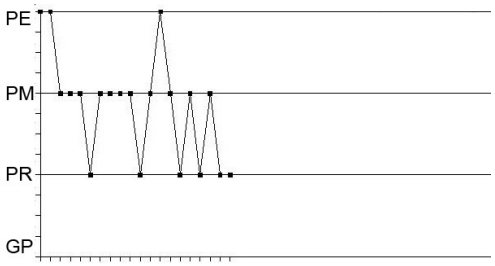
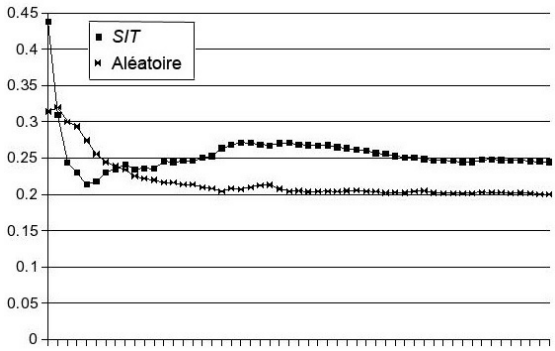
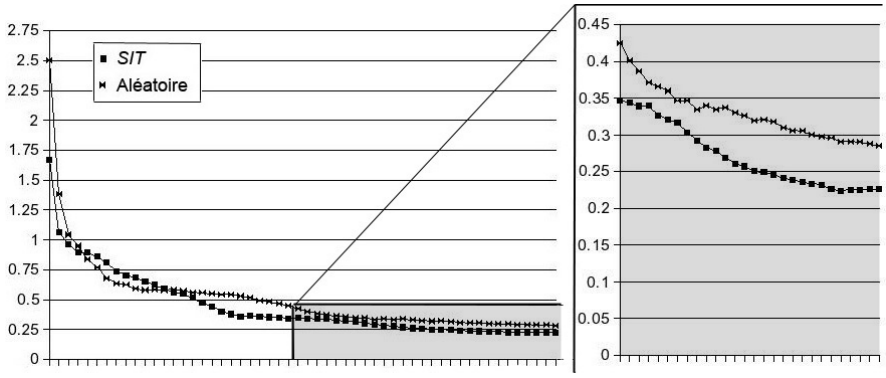


FIG. 5.4 – Échelles de plan, séquence "The Birds".

Comparaison avec une séquence semi-aléatoire La génération semi-aléatoire de séquences implante les mêmes règles que la Caméra *SIT*, en remplaçant l'algorithme basé sur les variables émotionnelles *SIT* par une sélection aléatoire de la position de la caméra. L'utilisation de règles semblables permet de reproduire un comportement simi-



(a) Proportion du nombre de plans



(b) Coefficient de variabilité du nombre de plans

FIG. 5.5 – Statistiques de connaissance spatiale.

laire à celui de la Caméra *SIT* et d'isoler ainsi la contribution des variables émotionnelles *SIT* dans l'établissement des caractéristiques des plans.

Dans la séquence semi-aléatoire, l'échelle d'un plan est choisie au hasard entre l'échelle du plan précédent, et les deux échelles adjacentes. La même règle s'applique pour le choix de l'angle de vue. L'axe de jeu est traversé après un certain nombre de plans consécutifs effectués d'un même côté. L'élévation reste la plus près possible de la valeur neutre, où la caméra est placée à la même hauteur que le personnage, en tenant compte de la visibilité de ce dernier. Cette élévation produit une tension faible sur l'ensemble de la séquence.

Les figures 5.5(a) et 5.5(b) fournissent les données statistiques représentant la quantification de la connaissance spatiale obtenue par les points de spatiosité. Ces données sont extraites des points de spatiosité situés dans un même secteur de l'environnement virtuel.

La figure 5.5(a) montre l'évolution au cours de la séquence de la proportion des plans précédemment générés dans laquelle est inclus, en moyenne, un point de spatiosité donné. Cette donnée mesure l'exposition moyenne de chaque point durant une séquence.

$$\mu = \frac{\sum_{point=1}^{NbPoints} NbPlans\,(point)}{NbPoints}$$

$$Proportion = \frac{\mu}{NbPlansGénérés}$$

La figure 5.5(b) trace la progression du coefficient de variabilité correspondant au nombre de plans incluant un point. Cette donnée, provenant de l'écart type divisé par la moyenne, représente une mesure de dispersion sur l'ensemble des points de spatiosité du secteur. Elle indique la répartition des surfaces couvertes par les différents plans sur l'aire du secteur observé, les points de spatiosité y étant distribués uniformément.

$$C_v = \frac{\sqrt{\frac{\sum_{point=1}^{NbPoints}(NbPlans(point)-\mu)^2}{NbPoints}}}{\mu}$$

Ces deux graphes comparent les données de séquences produites par la Caméra *SIT*, à l'aide des variables émotionnelles *SIT*, et aléatoirement. Ces graphes utilisent des données moyennes obtenues à partir de dix séquences, pour chacun de ces deux types.

On remarque que ces deux données convergent vers leurs valeurs d'équilibre respectives lorsque le nombre de plans devient suffisamment élevé pour générer des résultats significatifs.

La proportion des plans incluant un point donné correspondant aux séquences issues de la Caméra *SIT* est alors supérieure à celle des séquences aléatoires de 22%. Cette donnée indique que la couverture des plans générés par la Caméra *SIT* est plus élevée en moyenne, ce qui implique que ces plans présentent une plus grande proportion de l'environnement que ceux générés aléatoirement. Cette différence peut être due tant à la sélection des échelles de plan que des angles de vue utilisés.

Le coefficient de variabilité provenant des séquences *SIT* converge vers une valeur 21% inférieure au coefficient de variabilité des séquences aléatoires, ce qui indique que

les prises de vue y sont mieux distribuées dans l'environnement. Cet avantage est attribuable au rôle du vecteur de spatiosité dans les choix d'orientation de la caméra et dans les changements de côté de l'axe de jeu. L'influence du vecteur de spatiosité est cependant diminuée par les autres règles régissant le choix de l'angle de vue, comme la restriction portant sur les angles de vue adjacents à l'angle de vue du plan précédent.

Les données précédemment analysées ne comparent que l'exposition des points de spatiosité sans considérer la cohérence spatiale des plans présentés, qui augmente la compréhension cognitive de l'environnement de jeu.

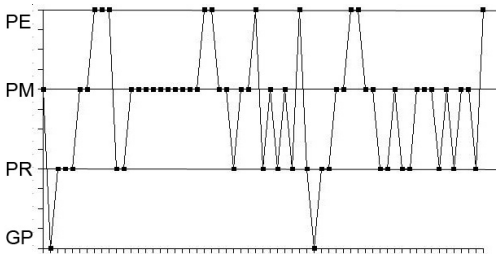
Les figures 5.2(a) et 5.3(a) permettent de comparer les variations d'échelles de plan des séquences *SIT* et aléatoire. Bien que les variations d'échelles de plan de la séquence aléatoire aient une allure générale assez proche de celles de la séquence *SIT*, principalement due à la règle qui restreint la majorité des variations à une échelle, leur distribution est très différente. La séquence *SIT* s'articule autour du plan moyen, avec un recours régulier aux plans d'ensemble, surtout lorsque le personnage entre dans un nouveau secteur, et l'insertion intermittente de plans rapprochés. Les échelles de la séquence aléatoire suivent une distribution plus erratique qui ne répond pas à la connaissance spatiale de l'environnement. La série de plans rapprochés observée provient de la présence du personnage dans un lieu confiné comme le corridor.

La figure 5.4 montre la variation d'échelles de plan d'une séquence issue du film "The Birds" (Alfred Hitchcock, 1963) [Bel79]. Cette séquence se situe vers le début du film, lorsque le personnage de Mélanie Daniels livre un couple d'inséparables à la maison de Mitch Brenner. Cette séquence décrivant les déplacements du personnage seul, sa dynamique s'assimile au contexte des séquences produites par la Caméra *SIT*. Cette séquence est construite selon un mode personnage vu, personnage voyant, par l'alternance des plans décrivant la progression du personnage et des plans subjectifs montrant ce que le personnage de Mélanie regarde durant son trajet, notamment sa destination, la maison de Mitch. Les plans appartenant à la partie subjective de la séquence remplissant une fonction distincte des plans portant sur le personnage, ils ont été retirés du graphe de variation scalaire.

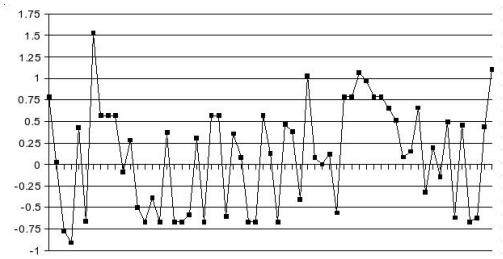
L'agencement des variations d'échelles de plan pour cette courte séquence montre une parenté avec la séquence *SIT*. Trois plans d'ensemble y sont observables, dont deux marquant le début de la séquence, alors que Mélanie amorce sa traversée du lac en ba-

teau. Le troisième plan d'ensemble se situe lors de l'entrée de Mélanie dans la maison de Mitch. Tout comme dans les séquences *SIT*, ces plans d'ensemble servent de plans de situation pour présenter un nouveau lieu. Comme les séquences générées, cette séquence utilise le plan moyen comme échelle de plan médiane. Alors que les plans moyens permettent de suivre la position de Mélanie, l'insertion de plans rapprochés fournit une image plus intime, permettant au spectateur de développer un sentiment d'empathie envers elle. La séquence de "The Birds" utilise, davantage de plans rapprochés que de plans d'ensemble, contrairement à la séquence *SIT* où la situation du personnage est nécessaire à sa manipulation.

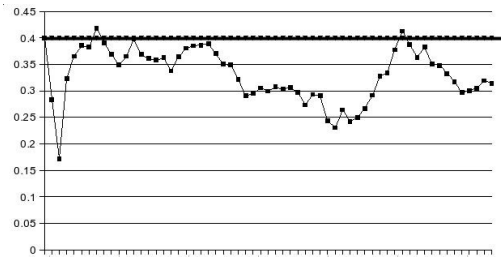
La séquence produite par la Caméra *SIT* et la séquence aléatoire peuvent aussi être comparées par la progression de leurs variables émotionnelles S et I contextuelles respectives. Les figures 5.2(c) et 5.3(b) montrent la proportion $\frac{S}{S+I}$ pour ces deux séquences. Les trois plans ayant $\frac{S}{S+I} = 0$ présents dans les deux graphes, indiquent l'entrée du personnage dans un nouveau secteur de l'environnement de jeu. Les valeurs provenant de la séquence *SIT* convergent davantage vers la valeur d'équilibre entre les variables S et I, établie à 0.5, que celles de la séquence aléatoire.



(a) Échelles de plan.

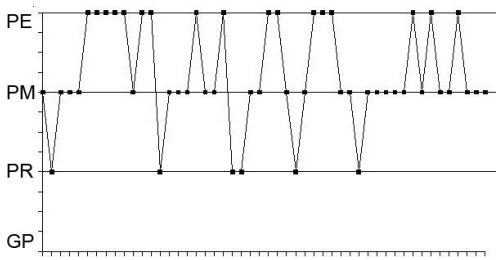


(b) Élévations, exprimées en radians.

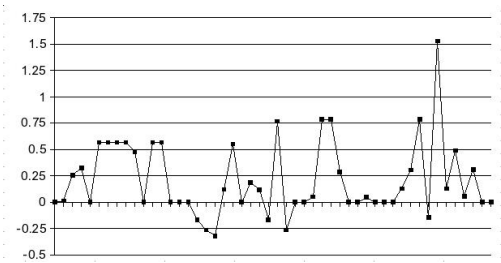


(c) Tension.

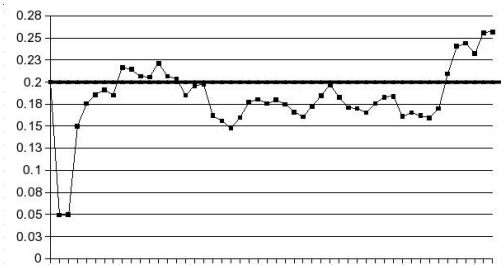
FIG. 5.6 – Séquence *SIT*, tension=0.4.



(a) Échelles de plan.



(b) Élévations, exprimées en radians.



(c) Tension.

FIG. 5.7 – Séquence *SIT*, tension=0.2.

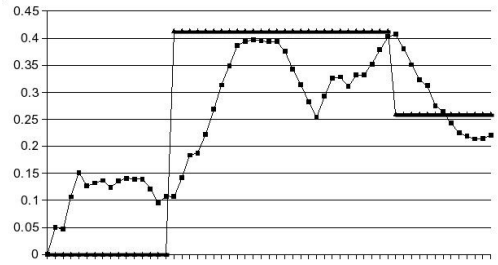


FIG. 5.8 – Séquence *SIT*, tension variable.

5.3 Analyse d'une séquence communiquant une tension élevée

Les images provenant de séquences tirées des films "Psycho" (Alfred Hitchcock, 1960) et "Taxi driver" (Martin Scorsese, 1976), reconnus pour leur représentation esthétique du suspence et de la violence, sont exposées dans les figures 5.14 et 5.15. On y remarque des choix visuels basés sur l'élévation de la caméra et les plans rapprochés, deux techniques utilisées par la Caméra *SIT* pour exprimer la tension dans l'image.

Hitchcock traduit dans cette scène de "Psycho" l'anticipation du drame à l'aide de l'augmentation de la plongée dans les plans. Cette plongée croissante permet de passer progressivement d'une ambiance normale, montrée avec un angle neutre, à l'assassinat

du déctective, filmé à vol d'oiseau. Cette dernière élévation extrême, bien qu'ayant instauré une figure de style largement réutilisée par la suite, avait pour fonction originale de masquer l'identité réelle de l'agresseur [TS83].

La séquence du film "Taxi driver", durant laquelle le protagoniste procède à un massacre, marque l'apogée de la violence latente du film. Scorsese y alterne les plans en plongée et contre-plongée qui soulignent la brutalité de la scène, alors que les plans rapprochés en renforcent l'impact émotif [Pic03]. Le montage très serré, contribuant à brusquer le spectateur, compte près d'une soixantaine de plans pour environ une minute de film.

La figure 5.16 contient le début de la scène décrite à la section 5.2, mais produite avec une tension ciblée à 0.4, qui correspond à une tension très élevée. Les plans générés reprennent les techniques de plongée et de contre-plongée marquées ainsi que de plans rapprochés, illustrées par les séquences de film précédentes, afin d'induire un sentiment de tension et d'inconfort chez l'utilisateur. Cette séquence serait appropriée pour précéder un événement inattendu dans le jeu.

Le premier plan de la séquence (figure 5.16(a)) est très semblable au plan correspondant de la séquence neutre (figure 5.9(a)), mais utilise une échelle plus rapprochée et une élévation plus accentuée. Le plan suivant (figures 5.16(b) à 5.16(d)) débute avec un gros plan du personnage pris avec un angle arrière. Le personnage s'éloignant de la caméra convertit ce plan en un plan d'ensemble, s'ouvrant sur une profondeur de champ du vestibule. Les deux plans rapprochés subséquents (figures 5.16(e) et 5.16(f)), alternant une contre-plongée marquée et une plongée, suivis d'un plan pris avec une caméra au plancher (figure 5.16(g)) et d'une prise de vue à vol d'oiseau (figure 5.16(h)) forment un enchaînement dont l'effet d'angoisse est augmenté par le cloisonnement du personnage dans l'image, contrastant avec le plan ouvert précédent. Par leur rythme rapide, ces quatre plans s'opposent au long plan précédent et aux plans de situation suivants. Ces plans de situation (figures 5.16(i) et 5.16(j)), qui introduisent la salle du trône, sont semblables à celui créé pour la séquence neutre (figures 5.9(h) et 5.9(i)). Les deux plans suivants (figures 5.16(k) et 5.16(l)), bien qu'utilisant une élévation normale, frappent par le passage direct des plans d'ensemble aux plans rapprochés. Les trois derniers plans présentés (figures 5.16(m) à 5.16(o)) utilisent la plongée et la contre-plongée, plutôt que la proximité avec le personnage, pour maintenir la sensation d'inconfort. De

façon générale, les angles de vue frontaux sont favorisés pour ce niveau de tension, ce qui augmente le lien psychologique avec le personnage.

Évidemment, les séquences produites par la Caméra *SIT* ne traduisent aucune intension d'auteur, ni ne supportent le contexte diégétique, contrairement aux séquences de films présentées. La Caméra *SIT* ne fournit qu'une direction générale pour les choix de cadrage. Ainsi, certains des effets décrits dans cette analyse proviennent du hasard de la composition d'image. Ils démontrent cependant la latitude de la Caméra *SIT*, qui ne restreint pas le potientiel expressif des plans générés.

Contrairement à la séquence produite avec une tension basse, où les élévations, toutes positives, expriment une grande perception spatiale ou accommodent la visibilité du personnage (figure 5.2(b)), cette séquence favorise des élévations excentriques, tant positives (plongée) que négatives (contre-plongée) (figure 5.6(b)).

Elle utilise aussi davantage d'échelles de plan rapprochées que la séquence neutre (figures 5.6(a) et 5.2(a)). On y remarque quelques gros plans, qui étaient absents de la première séquence. Ces graphes ne montrent cependant pas que la durée d'un plan est directement proportionnelle à son échelle. La séquence tendue possède ainsi un rythme plus élevé dans l'enchaînement de ses plans que la séquence neutre. Ces coupures rapides entre les plans contribuent à l'augmentation de la fébrilité de l'utilisateur.

Le recours fréquent à des contre-plongées et à des plans rapprochés nuit à la perception spatiale et complique la manipulation du personnage. Cette difficulté peut, cependant, s'ajouter à l'anxiété de l'utilisateur.

La figure 5.6(c) trace la tension évaluée au cours de la séquence. On voit que la tension se dégageant des plans parvient rarement à atteindre la valeur cible de 0.4. La pondération émotionnelle actuellement attribuée aux différents éléments du plan plafonne la tension à cette valeur. De son côté, la tension de la séquence neutre oscille entre 0.1 et 0.2, puisque les configurations de plans ne possèdent jamais une tension complètement nulle (figure 5.2(d)). Dans les graphes 5.2(d), 5.6(c), 5.7(c) et 5.8, la tension à exprimer est indiquée en gras.

La figure 5.17 montre une portion de séquence produite avec une tension de 0.2. Cette séquence communique une tension davantage assimilable à une inquiétude qu'à une angoisse. Les échelles de plan utilisées pour cette séquence conservent une certaine

distance entre l'utilisateur et l'action. Ces échelles de plan sont d'ailleurs semblables à celles de la séquence neutre (figure 5.7(a)). Seule l'élévation, dont l'amplitude est moins marquée que pour la séquence à tension élevée (figure 5.7(b)), vient déstabiliser l'utilisateur. La figure 5.7(c) montre que cette séquence réussit à maintenir une tension voisine de la valeur désirée.

Il est possible de varier la tension engendrée au cours d'une séquence pour souligner certains moments forts ou, au contraire, diminuer l'intensité d'une action présentée. La figure 5.8 montre comment la Caméra *SIT* modifie la tension de ses plans pour s'adapter aux changements de la tension à communiquer.

5.4 Temps de calcul

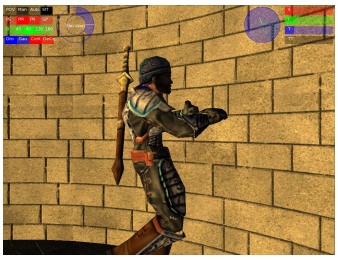
Comme la Caméra *SIT* doit générer les plans en temps réel lors du déplacement du personnage, le temps pris pour calculer la position de la caméra est crucial pour conserver l'apparente fluidité du mouvement à l'écran. Sur un échantillon de 45 plans produits pour une séquence, le temps de calcul varie entre 2394 μs et 29336 μs pour une moyenne de 11985 μs. Avec un temps de calcul de moins 30 ms, le temps de coupure entre deux plans passe donc inaperçu pour l'utilisateur. Ainsi, la vérification de la visibilité, effectuée approximativement pour diminuer les temps de traitement, pourrait être résolue de façon plus complète, tel que mentionné dans la section 6.3.

(a) Plan 1

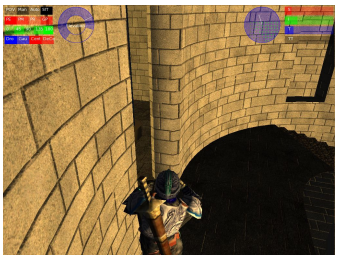
(b) Plan 2

(c) Plan 3

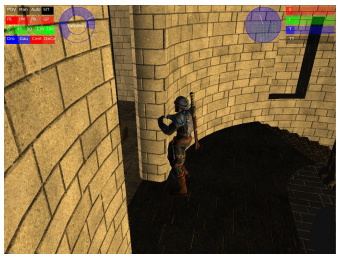
(d) Plan 4



(e) Plan 5



(f) Plan 6



(g) Plan 6

(h) Plan 7

(i) Plan 7

(j) Plan 8

(k) Plan 8 panoramique

(l) Plan 9

(m) Plan 9

(n) Plan 10

(o) Plan 11

FIG. 5.9 – Séquence *SIT*, tension=0.

(a) Plan 12

(b) Plan 13

(c) Plan 13

(d) Plan 14

(e) Plan 14

(f) Plan 15

(g) Plan 15

(h) Plan 15

(i) Plan 16

(j) Plan 17

(k) Plan 18

(l) Plan 19

(m) Plan 20

(n) Plan 21

(o) Plan 22

FIG. 5.10 – Séquence *SIT*, tension=0.

(a) Plan 23 (b) Plan 23 (c) Plan 23 panoramique

(d) Plan 24 (e) Plan 25 (f) Plan 25

(g) Plan 26 (h) Plan 26 (i) Plan 27

(j) Plan 27 panoramique (k) Plan 28 (l) Plan 29

(m) Plan 30 (n) Plan 30 (o) Plan 31

FIG. 5.11 – Séquence *SIT*, tension=0.

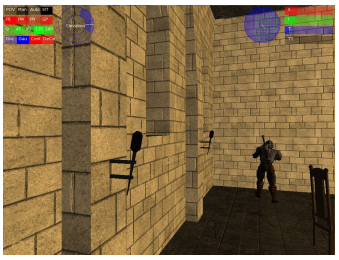
(a) Plan 32

(b) Plan 32

(c) Plan 33

(d) Plan 33

(e) Plan 34



(f) Plan34

(g) Plan 34 panoramique

(h) Plan 35

(i) Plan 35 panoramique

(j) Plan 36

(k) Plan 37

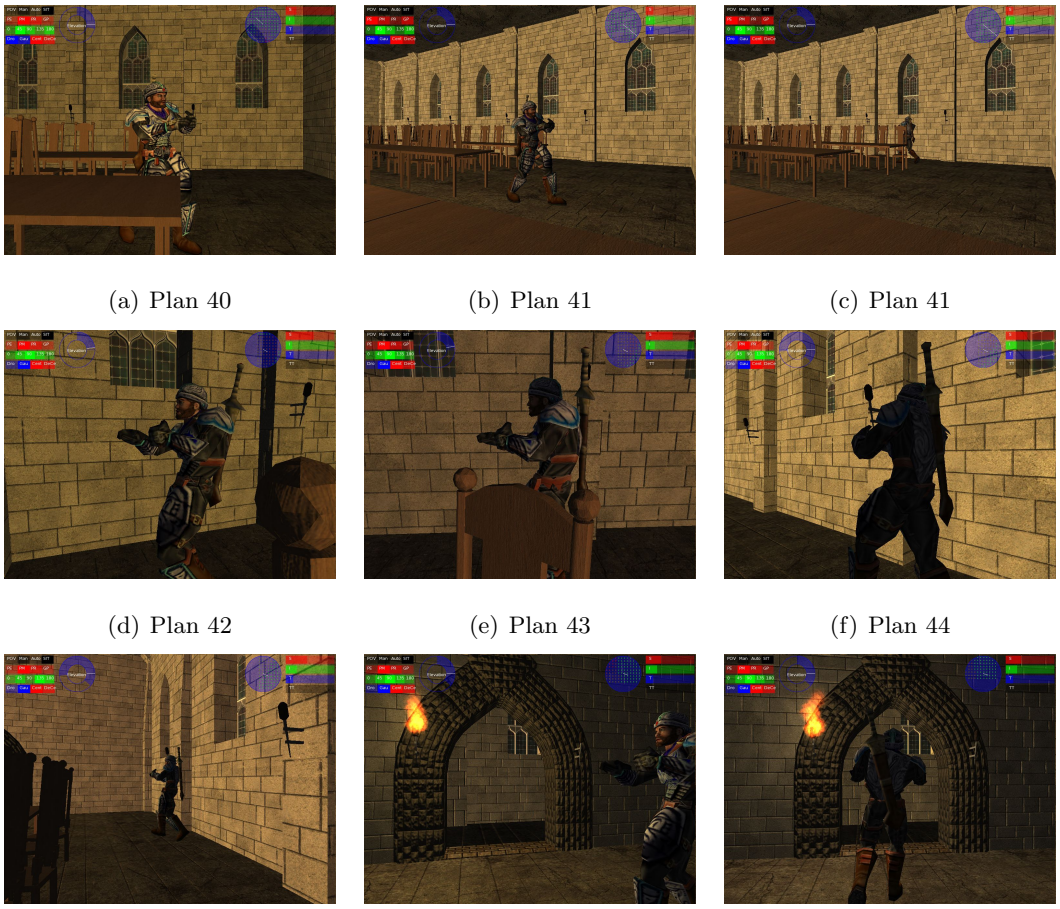
(l) Plan 37 panoramique

(m) Plan 37

(n) Plan 38

(o) Plan 39

FIG. 5.12 – Séquence *SIT*, tension=0.



(a) Plan 40 (b) Plan 41 (c) Plan 41

(d) Plan 42 (e) Plan 43 (f) Plan 44

(g) Plan 44 panoramique (h) Plan 45 (i) Plan 45 panoramique

FIG. 5.13 – Séquence *SIT*, tension=0.

(a) (b)

(c) (d)

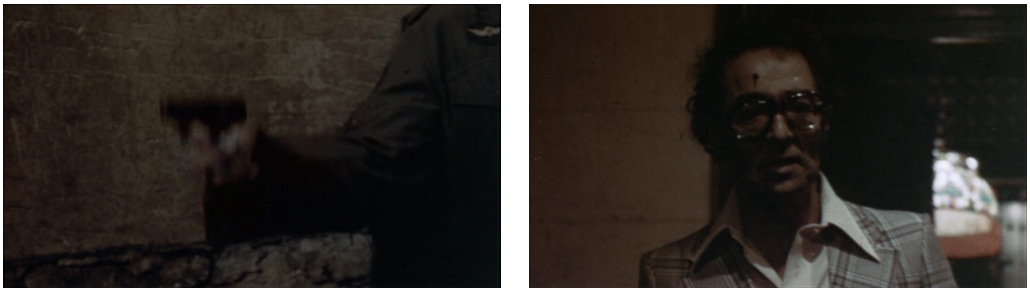
(e) (f)

FIG. 5.14 – “Psycho” (Alfred Hitchcock, 1960).

(a) (b)

(c) (d)

(e) (f)



(g) (h)

(i) (j)

FIG. 5.15 – "Taxi driver"(Martin Scorsese, 1976).

(a) Plan 1 (b) Plan 2 (c) Plan 2

(d) Plan 2 (e) Plan 3 (f) Plan 4

(g) Plan 5 (h) Plan 6 (i) Plan 7

(j) Plan 8 (k) Plan 9 (l) Plan 10

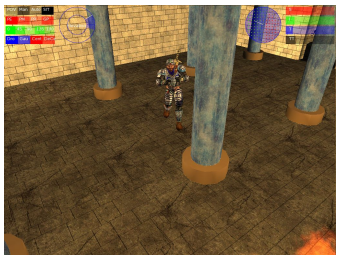
(m) Plan 11 (n) Plan 12 (o) Plan 13

FIG. 5.16 – Séquence *SIT*, tension=0.4.

(a) Plan 1

(b) Plan 2

(c) Plan 3



(d) Plan 4

(e) Plan 5

(f) Plan 6

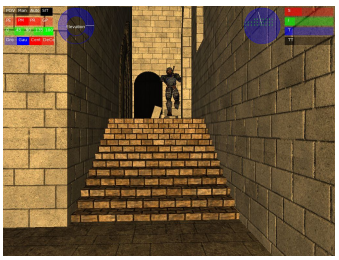
(g) Plan 7

(h) Plan 7 panoramique

(i) Plan 8

(j) Plan 9

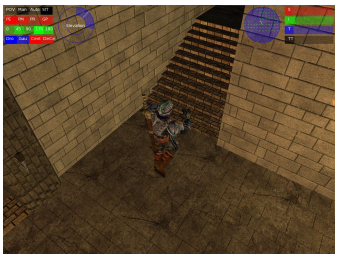
(k) Plan 10



(l) Plan 10

(m) Plan 11

(n) Plan 12



(o) Plan 13

FIG. 5.17 – Séquence *SIT*, tension=0.2.

Chapitre 6

Conclusion

6.1 Résumé de la méthode implantée

L'objectif premier de ce mémoire était l'application de principes cinématographiques à l'image du jeu vidéo afin d'en augmenter le contenu émotionnel. L'approche vise à remplacer la caméra subjective, dont l'utilisation est prédominante dans dans les jeux, par des séquences de plans non subjectifs portant sur le personnage. Alors que la caméra subjective est mieux adaptée pour des situations d'interactions particulières, comme viser une cible, plusieurs situations peuvent avantageusement être représentées par une séquence non subjective. Le choix effectué pour ce mémoire porte sur les déplacements d'un personnage unique dans son environnement virtuel. Les séquences produites visent à induire un sentiment d'identification avec le personnage, entraînant de l'empathie pour celui-ci, à favoriser une perception consistante de l'environnement de jeu dans lequel il évolue, et à créer une atmosphère de tension qui communique l'intention de la scène. Parmi l'ensemble des procédés existants en cinéma, seuls ceux concernant le cadrage de l'image sont utilisés pour exprimer les trois émotions ciblées.

La Caméra *SIT* développée est intégrée à l'engin de jeu *Crystal Space*. Le déplacement du personnage, contrôlé par l'utilisateur, est imprévu, et sa position dans l'environnement de jeu doit être traitée en temps réel par la Caméra *SIT* lors de la génération des plans. La latitude expressive de la caméra est maximisée dans le cadre de son fonctionnement.

Les variables S, I et T traduisent le contenu émotionnel transmis par une séquence de plans et permettent de relier les procédés cinématographiques aux trois paramètres définissant la position et l'orientation de la caméra relatives à la position du personnage : l'échelle de plan, l'angle de vue et l'élévation. Les valeurs des variables émotionnelles *SIT*, alliées à une vérification approximative de la visibilité du personnage dans l'image et au respect de règles régissant la cohérence visuelle des plans, comme la *règle des* 180 °, établissent les caractéristiques des plans générés. La coupe franche et le panoramique de recadrage servent de transitions entre les plans des séquences provenant de la Caméra *SIT*.

Alors que les composantes identification avec le personnage et tension ressentie sont entièrement constituées par des pondérations empiriques associées aux paramètres de la caméra, l'évaluation de la connaissance spatiale a recours à une grille de points de spatiosité qui conservent la quantification de cette connaissance pour chaque portion de l'environnement de jeu. Les points de spatiosité fournissent une évaluation du niveau de connaissance spatiale atteint pour la zone entourant la position du personnage. Un vecteur de spatiosité indique l'emplacement de la région ayant bénéficié de moins d'exposition dans les plans.

La mise en relation des variables *SIT* potentielles et contextuelles, ces dernières représentant le niveau émotionnel courant, permet de dégager les plans dont les caractéristiques compléteront les qualités expressives de la séquence en cours. La sélection des plans tente d'équilibrer les propriétés spatiales et identificatoires dans l'enchaînement. Le niveau de tension à insuffler à la scène est déterminé manuellement par l'utilisateur. Il influence le choix des cadrages en privilégiant ceux qui le manifestent le mieux. La rétroaction des variables *SIT*, effectuée lorsque les valeurs émotionnelles d'un plan généré sont ajoutées aux variables *SIT* contextuelles, permet à la Caméra *SIT* d'adapter sa génération de plans au degré de connaissance du voisinage actuel du personnage et aux variations dans la tension à atteindre.

6.2 Évaluation de la solution

Les objectifs premiers de la Caméra *SIT* étant de nature perceptive, leur évaluation complète requiert une étude empirique appuyée par étude statistique des résultats. Cette étude sera élaborée à la section 6.3. La suite de cette section commente les résultats

présentés au chapitre 5.

Forces de la solution Une des forces de la solution, observée dans les séquences contenues dans les figures 5.9 à 5.13, 5.16 et 5.17, réside dans la diversité créative des plans générés. On y note plusieurs plans dont les cadrages atypiques procurent des effets visuels intéressants et variés. L'enchaînement des plans ne comporte pas, non plus, de motifs répétitifs, malgré la redondance des actions du personnage. La Caméra *SIT* peut ainsi produire une variété illimitée de représentations d'un personnage en simple déplacement. La capacité de la Caméra *SIT* à exprimer la tension dans une séquence est aussi observable dans ses choix de cadrages, dont les caractéristiques diffèrent selon cette valeur.

L'environnement spatial décrit dans l'enchaînement de plans est exhibé de façon cohérente, ce qui facilite sa compréhension. Les variations d'échelles de plan, permettant d'alterner les vue rapprochées sur le personnage avec les plans de situation, s'effectuent de façon similaire à celles observées dans les productions cinématographiques.

La configuration de l'ensemble des valeurs reliées au traitement des variables émotionnelles *SIT* et des points de spatiosité confère une grande flexibilité au comportement de la Caméra *SIT*. Cette adaptabilité facilite l'ajout de nouveaux procédés cinématographiques, comme des mouvements de caméra, la modification de l'esthétique des plans, ou une variation de l'importance relative d'un des axes émotionnels.

Faiblesses de la solution Le désavantage principal de la Caméra *SIT* réside dans la perception spatiale nécessaire au contrôle du personnage. Bien que la représentation de l'environnement soit cohérente, et que les directions des déplacements dans les images de plans successifs soient semblables, la perception spatiale est limitée par la fréquence des plans rapprochés qui restreignent la localisation du personnage dans l'environnement. Cette faiblesse complique le contrôle du personnage, comparativement à l'utilisation d'une caméra subjective (ou *point of view*).

L'implantation de la vérification de la visibilité pourrait être améliorée. L'utilisation d'un algorithme approximatif et le traitement d'une fenêtre de visibilité unique pour chaque paire échelle de plan et angle de vue, écartent certaines configurations de plan qui pourraient améliorer la qualité expressible des séquences.

6.3 Travaux ultérieurs

Les fonctionnalités développées pour la Caméra *SIT* constituent un cadre d'application qui démontre l'efficacité globale de la solution proposée. L'ajout de nouvelles fonctionnalités contribuerait à l'amélioration de cette application. La majorité des travaux proposés dans cette section concerne le contrôle de la position et de l'orientation de la caméra.

Étude empirique L'évaluation des contributions perceptuelles et cognitives de la Caméra *SIT* devrait être le sujet d'une étude empirique approfondie, suivie d'une analyse statistique. Cette étude soumettrait la manipulation du personnage dans l'environnement de jeu à un échantillonnage d'utilisateurs afin de comparer la Caméra *SIT* avec la caméra subjective classique. Malheureusement, cette étude requiert des compétences actuellement hors de notre champ d'expertise afin d'établir l'échantillonnage des sujets soumis aux essais, de définir le banc d'essai, et d'interpréter les données recueillies durant les tests. L'échantillonnage doit contenir un certain nombre d'utilisateurs et ne doit pas être homogène afin d'éviter les biais au niveau des résultats de l'étude.

Le banc d'essai soumis aux utilisateurs doit permettre la vérification des différents apports cognitifs de la Caméra *SIT* en comparaison avec la caméra subjective. Cette vérification peut passer par la réalisation d'objectifs définis sous forme de tâches à accomplir à l'intérieur de l'environnement de jeu. L'établissement d'un questionnaire permettrait aussi d'évaluer les connaissances acquises durant la session de jeu. Le banc d'essai doit tester chacun des trois axes émotionnels, soit la connaissance spatiale des lieux, la perception de l'état émotif du personnage, attendu que le personnage puisse communiquer certaines émotions facialement ou autrement, et l'identification de moments où une tension est suggérée. Ces essais pourraient être menés avec des configurations de pondérations émotionnelles *SIT* différentes, de façon à dégager l'impact réel de chaque procédé cinématographique sur les performances des participants.

Les résultats provenant de ces essais doivent être interprétés à l'aide d'analyses statistiques pour déterminer leur validité, et confirmer ou infirmer l'hypothèse de gain perceptuel attribuable au fonctionnement de la Caméra *SIT*.

Esthétique des plans Quelques ajouts simples permettraient d'améliorer l'esthétique des images produites par la Caméra *SIT*. Le personnage pourrait bénéficier

d'une illumination dédiée qui lui procurerait une illumination de base appropriée, quelle que soit sa position dans l'environnement de jeu. Cette illumination serait particulièrement profitable pour les gros plans et les plans rapprochés, pour éviter que le personnage ait le visage obscurci, lorsqu'il occupe une région plus sombre de l'environnement.

L'utilisation d'une caméra avec grande distance focale pour les gros plans éliminerait la distorsion du visage actuellement observable. Cette distorsion est causée par les distances relatives des différents éléments du visage avec la caméra, lorsque celle-ci est proche du sujet. Une longue focale utilisée avec une caméra éloignée, telle qu'utilisée dans la production cinématographique, permet d'applanir le visage et d'obtenir une image plus harmonieuse.

Visibilité La vérification de la visibilité, exposée à la section 4.6, profiterait de meilleurs algorithmes. La résolution de la visibilité exacte du personnage doit tenir compte du niveau d'occultation partielle provenant des maillages traités et être exécutée en temps réel lors de la sélection du prochain plan. Le personnage est divisé en zones d'importance permettant de déterminer des degrés d'occultation distincts pour ses différentes parties. Ces zones pourraient varier selon l'échelle de plan. Par exemple, la visibilité de la tête du personnage étant cruciale lors des gros plans, celle-ci bénéficierait d'une zone d'occultation nulle ou faible. Elle pourrait cependant tolérer une occultation moyenne dans les plans éloignés. Les autres parties du corps seraient ainsi évaluées en fonction de la nécessité de leur visibilité dans les plans.

L'établissement d'une technique de précalcul de la visibilité en fonction de la géométrie de l'environnement de jeu, basée sur la discrétisation du volume de l'espace ou des volumes englobants, serait probablement nécessaire pour conserver un temps de calcul suffisamment bas. L'article [COCSD03] propose plusieurs techniques de visibilité pouvant être appliquées dans notre contexte.

Le traitement d'intervalles de visibilité multiples plutôt que l'intervalle unique de l'application actuelle, pour un couple échelle de plan et angle de vue, fournirait une solution moins restreinte des positions de caméra pour lesquelles le personnage est visible.

Mouvements de caméra L'ajout de mouvements de caméra pourrait diversifier les transitions utilisées par la Caméra *SIT*, et compléter l'ensemble des procédés

cinématographiques disponibles pour la communication des émotions.

Les travellings se divisent en deux types : le travelling avant ou arrière où la caméra s'approche ou s'éloigne du personnage, et le travelling d'accompagnement qui suit le personnage dans son déplacement. Un travelling d'accompagnement peut être avant, arrière ou latéral, selon la position de la caméra relativement au personnage.

Les travellings avant ou arrière font office de *zoom-in* et *zoom-out* et permettent de lier des plans pris avec le même angle de vue mais d'échelles différentes. Leur vitesse peut influencer le sentiment de tension émanant de la scène.

Les travellings d'accompagnement possèdent un niveau élevé d'identification avec le personnage. Leur utilisation permettrait d'augmenter la composante I de la Caméra *SIT* en ayant moins recours aux plans rapprochés, ce qui favoriserait une meilleure perception spatiale.

La régularité du déplacement d'un travelling peut aussi modifier la perception du plan. Un travelling ajustant exactement sa vitesse sur celle du personnage sera plus neutre, alors qu'un mouvement variable dynamise davantage l'image. L'effet documentaire du style caméra à l'épaule, où le cadre de l'image est instable, communique l'agitation du personnage.

Le panoramique de recadrage, utilisé pour recentrer le personnage dans l'image après un déplacement, est actuellement implanté dans la Caméra *SIT*. Les panoramiques pourraient être étendus aux plans de conjonction, permettant de relier le personnage à un élément du décor.

Plans de conjonction Une façon de souligner un élément de l'environnement de jeu est de produire des plans de conjonction entre le personnage et cet élément. Ces plans peuvent s'effectuer en cadrant dans un même plan le personnage et l'élément. Un panoramique peut aussi servir de conjonction. Le plan débute alors sur le personnage centré, et la caméra pivote vers l'élément important, de manière à le centrer dans l'image à son tour.

Les plans de conjonction servent à lier le personnage à un élément prédéfini du décor, un portail pour lui indiquer une ouverture, ou un maillage avec lequel il serait entré en collision au cours de son déplacement. Les plans de conjonction peuvent aussi servir à l'intégration d'un second personnage dans les plans.

Spatialisation de l'environnement Les points de spatiosité, utilisés pour évaluer la connaissance spatiale par région, sont actuellement distribués uniformément sur le plan horizontal de la surface de jeu. Une modification de la concentration des points pourrait servir à attirer l'attention de l'utilisateur sur certaines zones de l'environnement de jeu. Ces zones d'attention, correspondant à des emplacements ou objets importants dans le jeu, auront une concentration plus faible de points de spatiosité que le reste de l'environnement. Elles disposeront ainsi d'une moins grande capacité d'accumulation de connaissance spatiale. La Caméra *SIT* tentera de compenser ce déficit en augmentant le nombre de plans les incluant. Un même résultat serait obtenu par l'attribution de poids variables aux points de spatiosité, selon leur position.

La grille des points de spatiosité pourrait aussi être étendue en trois dimensions, pour permettre de mieux gérer la connaissance des éléments de l'environnement de jeu selon leur hauteur. Avec la grille actuelle, distribuée sur le plan horizontal de l'environnement de jeu, le découpage de l'espace par la pyramide de vue de la caméra est traité horizontalement seulement, l'ensemble de l'espace situé à la verticale d'un point de spatiosité étant considéré de façon équivalente. L'utilisation d'une grille tridimentionnelle ajouterait le découpage vertical de l'espace par la pyramide. Ainsi la Caméra *SIT* équilibrerait les plans pour montrer les régions situées à des hauteurs négligées par les prises de vue précédentes.

Objectifs émotionnels planifiés La fluidité des séquences serait améliorée en établissant des objectifs émotionnels atteignables sur plusieurs plans, plutôt que de compenser le contexte courant à chaque nouveau plan. Le contenu émotionnel des plans pourrait converger progressivement vers une valeur établie, en évitant les variations marquées et subites des caractéristiques de plans successifs. Un exemple d'utilisation serait l'installation progressive d'une atmosphère de tension qui culminerait après un certain nombre de plans. Cette approche pourrait aussi faciliter le changement de côté de l'axe de jeu en échelonnant le déplacement de l'angle de vue vers un angle neutre sur plusieurs plans.

Rythme de la séquence La variation des durées des plans peut influencer la perception de la durée de l'action présentée. Une série de plans de durées égales induit un sentiment d'attente et de lenteur, alors que l'utilisation de plans de plus en plus courts

accélère l'action et augmente la tension.

Le moment de coupure entre deux plans peut aussi jouer sur la perception du temps. Un plan coupé avant que le personnage n'ait traversé le cadre augmente la rapidité de l'action, alors qu'une coupure survenue après que le personnage ait quitté l'image la diminue.

Pondérations émotionnelles Le secteur le moins formalisé de la Caméra *SIT* reste l'attribution des pondérations émotionnelles aux divers procédés cinématographiques. Cette attribution a été effectuée empiriquement, en suivant les principes généraux documentés dans les ouvrages d'études cinématographiques. La formalisation des effets cognitifs n'y ayant pas été trouvée, des recherches effectuées du côté de la psychologie cognitive pourraient contribuer à l'établir. Ces pondérations pourraient aussi être déterminées par un système d'apprentissage qui les établirait à partir de la rétroaction de l'utilisateur, selon les résultats obtenus.

Autres procédés cinématographiques En plus des techniques concernant le cadrage, les mouvements de caméra et les durées des plans, d'autres procédés permettent d'influencer la perception de l'utilisateur.

Les techniques d'éclairage classiques peuvent établir une atmosphère et jouer sur la tension de la scène. De façon générale, les éclairages *low key*, plutôt sombres et produisant des ombres nettes, induisent une atmosphère plus dramatique que les *high key*, plus lumineux et générant des ombres douces.

L'éclairage contribue aussi à la perception de la profondeur de l'image et de la compréhension spatiale de l'environnement en soulignant les contours des objets, et en découpant l'espace en zones d'ombre et de lumière. L'illumination sélective d'une région ou d'un élément de l'environnement contribue à influencer l'attention de l'utilisateur.

Les caractéristiques de l'objectif de la caméra sont aussi souvent utilisées en cinématographie. La variation d'une profondeur de champ étroite contrôle la portion de l'environnement qui est vue par l'utilisateur. La modification de la distance focale, de son côté, modifie la perspective de l'image, permettant de jouer sur la perception des distances entre les divers éléments constituant l'environnement.

Un amalgame de toutes les techniques proposées formerait un système de cinématographie plus complet, intégrable à un engin de jeu. Un tel système améliorerait l'expérience de l'utilisateur, en lui procurant des émotions reliées au contexte diégétique du jeu, un meilleur contrôle de son personnage, et une compréhension accrue des éléments narratifs constituant le monde fictionnel avec lequel il doit interagir.

Nos travaux établissent une base appréciable qui démontre l'intérêt d'une telle approche. Ils soulignent cependant que plusieurs développements peuvent encore contribuer à son amélioration.

Bibliographie

[Ari76] D. Arijon. *Grammar of the film language*. Hastings House, 1976.

[Bel79] R. Bellour. *L'analyse du film*. Albatros, 1979.

[BGL98] W.H. Bares, J.P. Grégoire et J.C. Lester. « Realtime constraint-based cinematography for complex interactive 3D worlds ». Dans *AAAI '98/IAAI '98 : Proceedings of the fifteenth national/tenth conference on Artificial intelligence/Innovative applications of artificial intelligence*, pages 1101–1106, Menlo Park, CA, USA, 1998. American Association for Artificial Intelligence.

[BL99] W.H. Bares et J.C. Lester. « Intelligent multi-shot visualization interfaces for dynamic 3D worlds ». Dans *IUI '99 : Proceedings of the 4th international conference on Intelligent user interfaces*, pages 119–126, New York, NY, USA, 1999. ACM Press.

[Bor85] D. Bordwell. *Narration in the fiction film*. University of Wisconsin Press, 1985.

[BT00] D. Bordwell et K. Thompson. *L'art du film : une introduction*. De Boeck Université, 2000.

[BTS85] D. Bordwell, K. Thompson et J. Staiger. *The classical Hollywood cinema : film style and mode of production to 1960*. De Boeck Université, 1985.

[CAH+96] D.B. Christianson, S.E. Anderson, L. He, D.H. Salesin, D.S. Weld et M.F. Cohen. « Declarative Camera Control for Automatic Cinematography ». Dans *AAAI/IAAI, Vol. 1*, pages 148–155, 1996.

[COCSD03] D. Cohen-Or, Y.L. Chrysanthou, C.T. Silva et F. Durand. « A survey of visibility for walkthrough applications ». *Transactions on Visualization and Computer Graphics*, volume 9, numéro 3, pages 412–431, 2003.

[Cry07] « Crystal Space ». `http://www.crystalspace3d.org`, 2007. Trousse de développement logiciel de jeu.

[Cut05] J.E. Cutting. « Perceving Scenes in Film and in the World ». Dans J.D. Anderson et B. Fisher Anderson, éditeurs. *Moving image theory : ecological considerations*, pages 9–27. Southern Illinois University Press, 2005.

[DZ95] S.M. Drucker et D. Zeltzer. « CamDroid : a system for implementing intelligent camera control ». Dans *SI3D '95 : Proceedings of the 1995 symposium on Interactive 3D graphics*, pages 139–144, New York, NY, USA, 1995. ACM Press.

[FTT99] J. Funge, X. Tu et D. Terzopoulos. « Cognitive modeling : knowledge, reasoning and planning for intelligent characters ». Dans *SIGGRAPH '99 : Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 29–38, New York, NY, USA, 1999. ACM Press/Addison-Wesley Publishing Co.

[Gro05] T. Grodal. « Film Lighting and Mood ». Dans J.D. Anderson et B. Fisher Anderson, éditeurs. *Moving image theory : ecological considerations*, pages 152–163. Southern Illinois University Press, 2005.

[GSHG98] G. Greger, P. Shirley, P.M. Hubbard et D.P. Greenberg. « The Irradiance Volume ». *IEEE Computer Graphics and Applications*, volume 18, numéro 2, pages 32–43, 1998.

[HCS96] L. He, M.F. Cohen et D.H. Salesin. « The virtual cinematographer : a paradigm for automatic real-time camera control and directing ». Dans *SIGGRAPH '96 : Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 217–224, New York, NY, USA, 1996. ACM Press.

[HHS01] N. Halper, R. Helbing et T. Strothotte. « A Camera Engine for Computer Games : Managing the Trade-Off Between Constraint Satisfaction and Frame Coherence ». Dans A. Chalmers et T.-M. Rhyne, éditeurs. *EG 2001 Proceedings*, volume 20(3), pages 174–183. Blackwell Publishing, 2001.

[Hor03] A. Hornung. « Autonomous Real-Time Camera Agents in Interactive Narratives and Games ». Mémoire de maîtrise, Department of Computer Science V, Aachen University of Technology, 2003.

[Int84] M. Intintoli. *Taking Soaps Seriously*. Praeger Publishers, 1984.

[Ken02] K. Kennedy. *Intelligent Planning of Communicative Acts for Computer Animation Cinematography*. Thèse de doctorat, Faculty of Graduate Studies, The University of Western Ontario, 2002.

[KM02] K. Kennedy et R.E. Mercer. « Planning animation cinematography and shot structure to communicate theme and mood ». Dans *SMARTGRAPH '02 : Proceedings of the 2nd international symposium on Smart graphics*, pages 1–8, New York, NY, USA, 2002. ACM Press.

[NT99] Y. Nam et D. Thalmann. « CAIAS : Camera Agent based on Intelligent Action Spotting for Real-Time Participatory Animation in Virtual Stage ». Dans *5th International Conference on Virtual Systems and Multi-Media (VSMM) 1999*, 1999.

[Per04] B. Perron. « Sign of a Threat : The Effects of Warning Systems in Survival Horror Games ». Dans *COSIGN 2004 Proceedings*, pages 132–141, 2004.

[Pic03] M. Picard. « L'esthétique de la violence dans l'oeuvre cinématographique de Martin Scorsese et Takeshi Kitano ». Mémoire de maîtrise, Université de Montréal, 2003.

[Sta07] « Stanford AI Lab ». "http ://robotics.stanford.edu/˜latombe/cs326/2004/class16/class16.htm", 2007.

[Tan96] E.S. Tan. *Emotions and the Structure of Narrative Film : Film as an Emotion Machine*. Lawrence Erlbaum Associates, 1996.

[TBN00] B. Tomlinson, B. Blumberg et D. Nain. « Expressive autonomous cinematography for interactive virtual environments ». Dans *AGENTS '00 : Proceedings of the fourth international conference on Autonomous agents*, pages 317–324, New York, NY, USA, 2000. ACM Press.

[TS83] F. Truffaut et H. Scott. *Hitchcock*. Ramsay, 1983.

[TT94] X. Tu et D. Terzopoulos. « Artificial fishes : physics, locomotion, perception, behavior ». Dans *SIGGRAPH '94 : Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, pages 43–50, New York, NY, USA, 1994. ACM Press.