



## MASTER RESEARCH INTERNSHIP



## BIBLIOGRAPHIC REPORT

---

# Drones cinématographiques autonomes

---

**Domain: Emerging Technologies - Human-Computer Interaction - Multiagent Systems - Multimedia - Robotics**

*Author:*  
Gaétan GILSON

*Supervisor:*  
Marc CHRISTIE  
Quentin GALVANE  
Christophe LINO  
EPI Mimetic

**Abstract:** L'intérêt croissant porté aux drones dans divers secteurs d'activité ainsi que leur forte popularité auprès du grand public ont fortement favorisé leur développement aux cours des dernières années. Plus particulièrement, l'industrie du cinéma semble s'intéresser de plus en plus à cette nouvelle forme de technologie qui leur offre de nombreuses possibilités tout en leur permettant de réduire significativement les coûts de production. D'un point de vue cinématographique cependant, les problématiques liées au pilotage des drones et au contrôle des caméras embarquées sont nombreuses et rendent particulièrement difficile l'utilisation de ces derniers. La création d'outils dédiés à ces tâches semble alors s'imposer comme une nécessité. Bien qu'il n'existe pour le moment pas de solution complète, la communauté scientifiques s'est déjà penchée sur plusieurs de ces problématiques. L'objet de cette étude bibliographique est de faire un point sur les enjeux, les grands challenges et les résultats récents dans ce domaine. Cette étude s'appuiera à la fois sur les techniques de contrôle de caméra dans le virtuel, et celles mises en place pour le contrôle de drones dédiés à la prise de vue. Dans ce cadre, nous nous focaliserons sur 3 problématiques : (i) la génération de trajectoires de caméra, (ii) le respect des propriétés cinématographiques et (iii) la gestion de l'environnement (collisions, occultations). Nous ferons enfin le lien entre ces travaux et l'objectif de ce stage qui est de proposer un langage permettant de décrire différents comportements de drone et de proposer de nouvelles méthodes permettant la coordination d'un ensemble de drones à partir de ces comportements.

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>La cinématographie virtuelle</b>	<b>1</b>
2.1	Connaissances cinématographiques . . . . .	2
2.1.1	Contraintes de composition visuelle . . . . .	3
2.1.2	Contraintes de trajectoire de caméra . . . . .	4
2.1.3	Contraintes de montage . . . . .	4
2.2	Caméra virtuelle . . . . .	4
2.3	Transposition de la connaissance cinématographique aux mondes virtuels . . . . .	4
2.3.1	Cadrage automatique pour le calcul de points de vue . . . . .	5
2.3.2	Trajectoires automatiques de caméra . . . . .	5
2.3.3	Contraintes de composition visuelle . . . . .	6
2.3.4	Gestion d'un ensemble de caméra . . . . .	9
<b>3</b>	<b>Le contrôle de drones cinématographiques</b>	<b>9</b>
3.1	Spécificités d'un drone . . . . .	9
3.2	Filmer des scènes réelles avec des drones . . . . .	10
3.3	Transposition de la connaissance cinématographique aux drones . . . . .	12
<b>4</b>	<b>Conclusion et perspectives</b>	<b>14</b>

# 1 Introduction

Les évolutions technologiques, fonctionnelles et tarifaires des drones civils, et en particulier des quadrotors, ont développé leur intégration dans de nombreux domaines d'application (surveillance, inspection, prises de vue, loisirs). Une large gamme de prix et de fonctionnalités est désormais accessible. L'industrie du cinéma et du documentaire s'est largement équipée de ces nouvelles modalités de prise de vue, permettant à la fois des réductions de coûts et une liberté créatrice nouvelle. Ces drones sont souvent contrôlés manuellement (en position ou en vitesse). Dans les cas où une caméra est embarquée sur le drone, ce contrôle manuel reste difficile. En effet le pilote doit être capable de contrôler les 10 degrés de liberté (ddl) de la caméra et du drone: orientation et position du drone dans son environnement (6 ddl), l'orientation de la caméra (3 ddl) et sa focale (zoom - 1 ddl). De plus, certaines applications demandent de respecter des contraintes particulières. Par exemple, dans le cadre de prises de vue cinématographiques, il est nécessaire de respecter différentes règles cinématographiques, par exemple différentes orientations et positions des acteurs sur les images filmées par la caméra (composition visuelle) (1), différents mouvements de caméra pour éviter de perdre de vue les acteurs ou d'effectuer des mouvements saccadés (2), et, pour finir, respecter le montage que l'on souhaite effectuer, dans une séquence de différentes prises de vue (3). Dans ce contexte cinématographique, les différentes difficultés à contrôler un drone et à respecter ce domaine amènent la création d'outils dédiés de contrôle de drones dits "cinématographiques". On définit un drone cinématographique comme un drone ayant un comportement similaire à une caméra utilisée en cinéma (recadre les acteurs, comportement similaire à une caméra portée à l'épaule ou simulant un travelling, etc.). Dans le cas où l'on souhaite réaliser le montage en temps réel, il est possible de coordonner un ensemble de drones cinématographiques pour avoir les différents cadrages nécessaires à celui-ci.

Nous allons donc, en voyant l'intérêt de créer des outils permettant de manipuler ou aidant à manipuler les drones cinématographiques, nous concentrer sur les problématiques et avancées scientifiques dans ce rapport bibliographique. Il existe actuellement de nombreuses techniques permettant de transposer les connaissances cinématographique dans un environnement virtuel, que nous allons voir dans un premier temps, mais ils ne sont pas directement utilisables pour contrôler un drone. En effet, le contrôle d'un drone demande l'ajout de contraintes supplémentaire pour respecter les différentes contraintes physiques. Il ne faut en aucun cas rentrer en collision avec des objets ou des acteurs et il faut respecter les limitations physiques du drone (vitesse, posture, continuité...). Par ces diverses limitations, des techniques propres à l'environnement physique, que nous allons voir dans un second temps, ont été proposées. Ces procédés se basent généralement sur un assemblage de techniques classiques d'animation (points clés, interpolations entre les points clés avec des courbes splines, etc.). D'autres procédés prenant en compte les différents aspects cinématographiques seront présentés.

## 2 La cinématographie virtuelle

Le domaine de la cinématographie virtuelle s'intéresse à la réutilisation des connaissances et outils du cinéma dans les mondes virtuels. Dans cette section nous présentons tout d'abord ces connaissances et outils du cinéma, avant d'étudier les approches qui ont été proposées pour transposer cette connaissance aux caméras virtuelles.

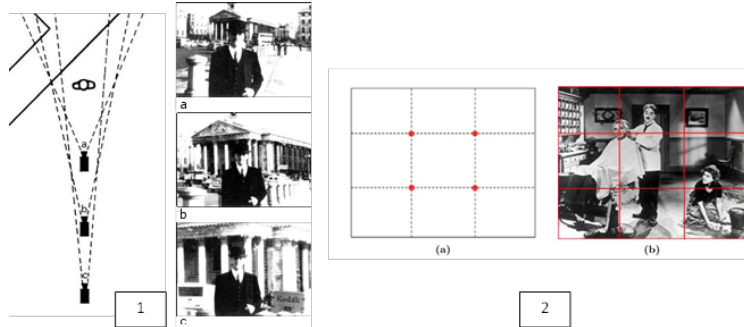


Figure 1: Différentes compositions visuelles — 1 - Exemple de différentes profondeurs de champ et focale : a) une focale courte (grand angulaire) et une distance entre la caméra et le personnage faible. b) une focale normale et une distance moyenne entre la caméra et le personnage. c) une longue focale et une distance importante entre la caméra et le personnage. 2 - La règle du tiers: Pour respecter le cadrage, on découpe le cadre en trois parties horizontales et trois parties verticales. Les quatre points d'intersections constituent les points importants de la scène ou points "fort". Les éléments importants sont placés sur une ligne créée par ce découpage du cadre. Par exemple si l'on souhaite filmer un personnage, il est placé sur une ligne verticale et sa tête sur un point fort.

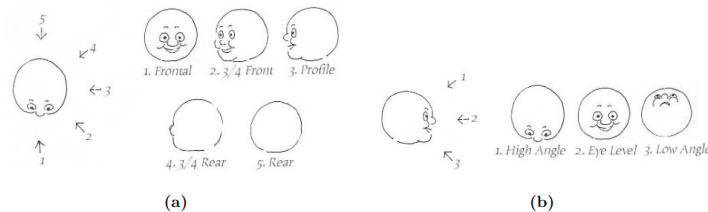


Figure 2: Illustration des différentes orientations de caméra possible selon l'axe horizontal (a) ou vertical (b)

## 2.1 Connaissances cinématographiques

En cinéma on décompose un film en scènes, qui elles-mêmes peuvent être décomposées en séquences de plans agencés dans le temps. Cette séquence de plans doit respecter un ensemble de règles permettant de maintenir une continuité dans la narration. Un film ou documentaire suit une narration développée par un enchaînement de scènes compréhensibles et logiques, dans leur globalité, pour un spectateur. Cependant, il est tout à fait possible de réaliser des faux raccords entre scènes (création d'une rupture dans la narration), lorsque l'on introduit par exemple une discontinuité (temporelle ou spatiale) entre deux plans successifs ou bien lorsque les acteurs ne sont pas suffisamment visibles à l'écran. De manière à ce que le film respecte une certaine logique, les réalisateurs ont progressivement construit un ensemble de règles. Ces règles peuvent être regroupées en 3 catégories:

- Les contraintes sur le cadrage ou la composition visuelle
- Les contraintes sur la trajectoire empruntée par la caméra
- Les contraintes sur le montage ou comment enchaîner les différentes prises de vue.

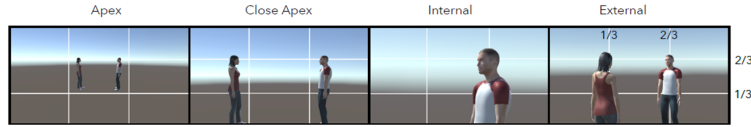


Figure 3: Différentes compositions visuelles utilisées en cinématographie

### 2.1.1 Contraintes de composition visuelle

Le choix du point de vue sur la scène est un élément central, permettant de choisir quelles informations seront présentées au spectateur, et de quelle manière certains éléments de l'histoire seront mis en valeur. Pour ce faire, il existe un ensemble de règles liées à la composition visuelle (*i.e.* l'arrangement des éléments de la scène à l'écran). On peut notamment citer trois critères visuels importants : la position, la taille et l'angle de vue des éléments à l'écran. Ces différents critères permettent de contrôler l'apparence de ces éléments.

Pour respecter ces critères, il faut pouvoir positionner correctement les acteurs à l'écran. On utilise donc, dans ce but, la règle du tiers que l'on illustre dans la partie 1 de la figure 1 qui consiste à subdiviser l'écran en trois parties égales horizontalement et verticalement, puis à placer les éléments importants le long d'une droite (appelée "power-line") ou à l'intersection des deux droites (appelée "power-point"). Cette règle permet de faire ressortir les éléments importants/les acteurs de la scène.

Il est aussi important de respecter les différentes apparences que l'on souhaite donner à un personnage ou objet à l'image. Pour cela, il est possible de faire varier sa taille ou son orientation sur l'enregistrement. Pour varier sa taille, on place la caméra à différentes distances du personnage allant d'un très gros plan où l'on perçoit une seule partie du corps (main, tête...) à un plan lointain où l'environnement occupe plus d'espace que le personnage. On détermine aussi plusieurs orientations par l'angle entre le regard de l'acteur et le sens de la caméra. Ces orientations peuvent être effectuées selon l'axe vertical ou horizontal comme l'illustre la figure 2.

Nous allons par la suite utiliser certaines compositions 3:

- L'Apex: une prise de vue où les acteurs sont centrés sur l'image et où leurs yeux sont placés horizontalement au 2/3 de l'image.

- L'internal: Un plan rapproché d'un sujet, orienté par rapport au regard pour garantir une vue partiellement frontale. Le sujet est placé verticalement au 1/3 ou 2/3 de l'image vérifiant ainsi la règle du tiers.

- L'external: Un plan où un seul sujet est ciblé mais où la seconde cible est visible. Si le sujet principal est dans le premier tiers, la cible secondaire est placée dans le tiers suivant et réciproquement (pour respecter la règle du tiers).

Après avoir fixé la proportion et l'orientation que l'on souhaite pour un personnage ou objet, on peut faire varier la focale et la profondeur de champ (zone de l'espace dans laquelle doit se trouver le sujet pour qu'il soit nette) pour agrandir l'apparence des objets en arrière plan comme l'illustre la deuxième partie de la figure 1. Pour cela, lorsqu'on s'approche du personnage on diminue la focale (On dézoom) et lorsqu'on s'en éloigne on augmente la focale (On zoom). Cela permet de faire varier l'importance de l'environnement du personnage ou de l'objet.

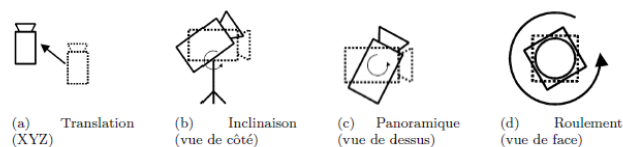


Figure 4: Les différents degrés de liberté (ddl) de la caméra. L'angle d'ouverture de la caméra est parfois ajouté comme 7ème degré de liberté.

### 2.1.2 Contraintes de trajectoire de caméra

Il est aussi possible d'effectuer une variation de la focale et de la profondeur de champ en bougeant la caméra. Cet effet se nomme le travelling compensé et contraint la caméra de bouger d'une certaine manière.

Il existe d'ailleurs de nombreuses contraintes sur la trajectoire empruntée par la caméra. On peut soit la faire pivoter sur un axe ("panoramique") ou la déplacer dans l'espace ("travelling"). Dans le cadre de la cinématographie, le travelling est effectué grâce à une caméra embarquée dans un chariot, se déplaçant sur des rails. ce dispositif permet de déplacer la caméra de manière fluide. Il existe deux types de travellings:

- les travellings perpendiculaires au champ. La caméra se déplace dans la direction ou elle est orientée.
- les travellings parallèles au champ. La caméra se déplace perpendiculairement à la direction ou elle est orientée.

### 2.1.3 Contraintes de montage

Pour finir il existe différentes contraintes sur le montage. Le montage doit prendre en compte les raccords entre les différentes scènes. Par exemple, une règle typique, est que lorsqu'un élément quitte l'écran par un des côtés, alors on s'attend à le voir re-rentrer par l'autre côté pour assurer la continuité du mouvement.

## 2.2 Caméra virtuelle

Une caméra virtuelle définit le point de vue de l'utilisateur sur la scène. Elle comporte 7 degrés de libertés (DDL) qu'il est important de déterminer pour contrôler celle-ci. En effet on a besoin de positionner et d'orienter cette caméra dans un environnement 3D. Cela définit donc 3 DDL pour la positionner et 3 DDL pour l'orienter. De plus, il est possible d'ajouter l'angle d'ouverture (ou distance focale) de la caméra comme un ddl et faire passer le nombre de ddl de 6 à 7. Certains travaux utilisent une focale fixe. Dans ce cas, seul 6 DDL doivent être déterminés. Effectivement, le contrôle d'une caméra virtuelle implique de déterminer ses DDL à chaque pas de temps.

## 2.3 Transposition de la connaissance cinématographique aux mondes virtuels

Transposer les différentes connaissances cinématographiques aux mondes virtuels pose différents problèmes, liés à chacune des catégories définies plus haut. Tout d'abord, concernant la composition visuelle, il faut, pour chaque point de vue, trouver une configuration de caméra (de dimension 7)

satisfaisant une liste de critères cinématographiques. Ensuite il faut être capable de générer la trajectoire entre chaque point de vue tout en respectant les contraintes de trajectoires:

- les délais définis entre les deux points de vue doivent être fixés
- il faut (généralement) garder visible les objets d'intérêt/acteurs que l'on filme
- il ne faut pas rentrer en collision avec l'environnement ou les acteurs
- il faut respectée une certaine composition visuelle, conforme aux règles de cadrage.

Et finalement, il faut conserver la continuité du film lors des changements de plan et déterminer quand changer de plan et quelle caméra choisir pour ce changement.

Dans le cadre de ce stage, nous nous concentrerons essentiellement sur les deux premiers aspect (composition visuelle et génération de trajectoires).

Ces différents problèmes ont longtemps été étudiés par la communauté scientifique et des solutions techniques ont été proposées pour s'attaquer à ces problèmes.

### **2.3.1 Cadrage automatique pour le calcul de points de vue**

Concernant la composition visuelle, Blinn a proposé une solution permettant de résoudre le problème de positionnement et d'orientation d'une caméra virtuelle pour satisfaire un cadrage défini de deux acteurs, tout en gardant la caméra à une certaine distance d'un des acteurs [Bli88]. Cette méthode se base sur l'extraction de la posture de la caméra (position et orientation) à partir de la projection des acteurs sur l'écran. Il en résulte une méthode itérative calculant algébriquement une approximation du problème, si elle existe.

Ce premier modèle de cadrage automatique proposé comporte néanmoins de nombreuses limitations. Tout d'abord, il est limité à 2 cibles et 2 contraintes (position des acteurs sur l'écran et distance entre un acteur et la caméra) et ne gère donc pas les autres critères visuels que l'on peut considérer, comme l'angle de vue, la taille des acteurs, la profondeur de champ, le zoom et la visibilité des cibles. Et deuxièmement, ce modèle calcule seulement un point de vue à un instant donné, il ne permet donc pas de calculer une trajectoire.

### **2.3.2 Trajectoires automatiques de caméra**

Oskam propose donc une solution en 2009 [OSTG09] pour déterminer les trajectoires d'une caméra dans un environnement virtuel. Le but de l'approche proposée est de générer une trajectoire de caméra à partir d'un point de départ, d'un point d'arrivée et d'un point de "focus". Ce dernier point indique le point d'intérêt que la caméra doit garder en visuel. Pour générer une trajectoire il faut donc que la caméra évite les obstacles présents dans l'environnement et présents entre la caméra et le point de focus. La solution proposée pour respecter cette problématique est de calculer une "roadmap" (une carte des routes possibles) que l'on illustre en partie (1) de la figure 5. Cette roadmap discrétise l'espace vide de la scène 3D sous forme de sphères et représente la visibilité (par une probabilité de voir le point de focus entre 0 et 1) pour chaque couple de sphère remplissant l'espace vide. Elle permet de calculer, en temps réel, la trajectoire de caméra après sa génération et assure la caméra d'éviter les obstacles. Son calcul définit des disques par chevauchement des sphères. chaque point de passage est choisi dans l'ensemble des points de la surface d'un disque (un point de passage par disque). Ils seront emprunté par la caméra pour passer d'une sphère à une autre. Le calcul de cette trajectoire se fait en choisissant les sphères de passage (sphères par lesquelles la caméra va passer en suivant sa trajectoire) de la caméra selon leurs visibilité en utilisant l'algorithme A\*. L'algorithme A\* est un algorithme heuristique permettant de trouver un

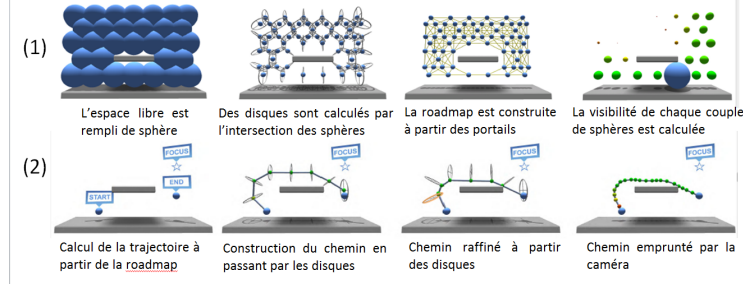


Figure 5: Les différentes étapes pour générer une trajectoire. On crée dans un premier temps la roadmap (1) puis on calcule la trajectoire de la caméra à partir de cette roadmap (2)

chemin en minimisant un coût. Il procède en choisissant, à une configuration donnée, la sphère de passage ayant le coût minimal. Ce coût est défini en fonction de la distance et de la visibilité. Plus la distance est élevée et plus la visibilité est faible, plus le coût est élevé et moins il y a de chance de passer par cette sphère.

Après avoir calculé un chemin gardant au maximum la visibilité, on raffine le résultat en travaillant sur la visibilité des sphères. La visibilité des sphères pouvant varier entre 0 et 1, on cherche à minimiser la distance parcouru dans des lieux où le point de focus n'est pas visible. Cette minimisation est réalisée dans les disques, en réutilisant l'algorithme  $A^*$ . Après ce raffinement, une trajectoire de caméra est finalement calculée. Les différentes étapes du calcul sont illustrées partie (2) de la figure 5.

### 2.3.3 Contraintes de composition visuelle

L'approche proposée par Blinn permet donc de calculer une trajectoire gardant un point en visuel mais ne permet pas de rajouter diverses contraintes cinématographiques. Afin de pouvoir respecter ces contraintes, une nouvelle représentation des configurations de caméra est présentée [LC12]. Cette représentation, sous forme d'une surface torique, permet d'avoir un ensemble de points respectant une certaine composition visuelle. Elle permet aussi de trouver une solution algébrique au problème de Blinn. Pour démontrer cette solution, on simplifie tout d'abord le problème en 2 dimensions et 2 cibles. On représente dans cette situation la caméra et les deux cibles vue de dessus. La caméra apparaît donc comme un segment. En utilisant le théorème des angles inscrit, on s'aperçoit que si la caméra appartient à l'arc de cercle formé par les 2 cibles, alors l'angle réalisé par la 1ère cible, la caméra et la 2ème cible est identique le long de l'arc de cercle comme l'illustre la partie gauche de la figure 6. Il est possible d'étendre, par la suite, ce problème dans un environnement 3D en effectuant une révolution autour de l'axe formé par les deux cibles. La surface ainsi créée garantit, pour une distance entre la caméra et un acteur donné, de contenir toutes les solutions pour respecter la composition visuelle donnée. On peut donc naviguer entre les solutions grâce aux paramètres  $\theta$  et  $\phi$  (partie droite de la figure 6) pour satisfaire les contraintes que l'on souhaite, réduisant ainsi l'espace de recherche de 6 (ou 7 si l'on ajoute la focale) à 2 paramètres (ou 3).

En plus de réduire le nombre des degrés de liberté et donc, de réduire le nombre de paramètres à définir pour chaque composition visuelle, cette approche permet de prendre en compte un nombre plus important de contraintes cinématographique que Blinn. Mais elle ne propose pas de méthode



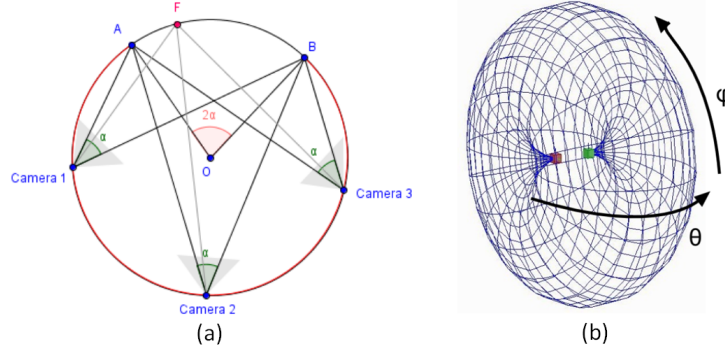


Figure 6: Ensemble des solutions pour une composition visuelle donnée dans un environnement 2D et 3D. (a) Dans un environnement 2D, les solutions sont représentées par l'arc de cercle rouge. (b) Dans un environnement 3D, les solutions sont sous forme d'une surface torique.

pour gérer la trajectoire de la caméra. Pour pouvoir contrôler la trajectoire de la caméra Lino [LC15] propose une application à des tâches classiques de contrôle de caméra. Cette approche utilise le modèle de représentation torique [LC12] (2 cibles), en lui rajoutant comme 3ème paramètre  $\alpha$ , l'angle entre la caméra et les deux cibles.

Dans le but de générer une trajectoire, il est défini un ensemble de points clés où l'on précise la composition visuelle que l'on souhaite à un certain temps. Pour définir les points clés, un outil a été développé (haut de la figure 7). Il propose de modifier le positionnement des cibles (en créant la surface torique pour un  $\alpha$  fixé), de modifier la taille des cibles sur l'écran (en changeant la distance entre la cible et la caméra), de changer l'angle de vue des cibles (en modifiant les paramètres  $\theta$  et  $\phi$ ) et de réaliser un travelling compensé (en conservant  $\alpha$ , tout en gardant les propriétés visuelle des cibles, mais en faisant varier distance focale et distance aux objets.).

Une trajectoire est ensuite générée (figure 7.e) entre deux points clés en:

- exprimant les deux configurations de la caméra indépendamment
- exprimant une première trajectoire  $\tau$  favorisant les cibles du 1er point clé
- exprimant une seconde trajectoire  $\tau'$  favorisant les cibles du 2ème point clé
- fusionnant les 2 trajectoires précédemment calculées

Les premières et seconde trajectoires sont calculées par interpolation des différentes propriétés visuelles de chaque points clés. Elles sont par la suite assemblées par une pondération devant être réglées par l'utilisateur (figure 7.f)).

Cette approche donne donc une application plus robuste et efficace comparé aux approches itératives ou heuristique. En effet les techniques itératives ne permettent pas d'assurer que la position de la caméra reste la même entre chaque exécution.

Mais, malgré son efficacité, cette méthode est limitée à deux cibles, les trajectoires ne prennent pas en compte la visibilité ou les obstacles et le réglage de la vitesse de la caméra doit s'effectuer manuellement pour créer un mouvement réaliste. Une amélioration a donc été proposée par Galvane [GCLR15]. Dans l'approche proposée, les trajectoires des cibles que l'on souhaite filmer ainsi que les point de départ et d'arrivée de la caméra doivent être connues. A partir des différentes trajectoires des cibles, une trajectoire "brut" est calculée en interpolant les compositions visuelles des cibles pour les caméras de départ et d'arrivée. On constate que cette trajectoire "brute" n'est pas fluide ou naturelle de part les nombreuses variations de vitesse qu'effectue la caméra en suivant

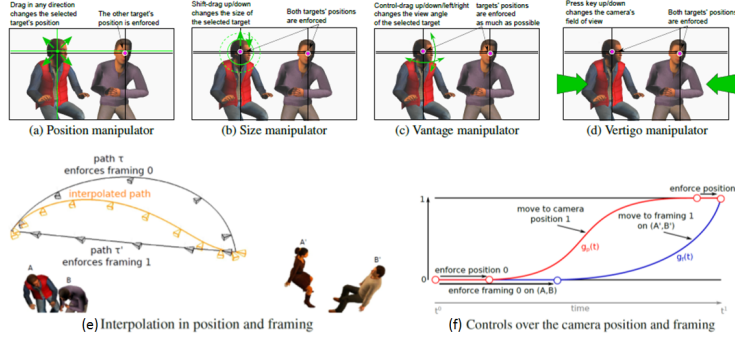


Figure 7: Différentes étapes pour déterminer la trajectoire de la caméra avec l'espace Torique [LC15]. La composition visuelle est contrôlée par 4 actions : (a) repositionner une cible sur l'écran en gardant l'autre cible à la même position. (b) changer la taille d'une cible en gardant l'autre cible de même taille. (c) changer l'angle de vue d'une cible en gardant le plus possible la position de l'autre cible. (d) Changer le focus de la caméra en gardant la même composition visuelle. (e) pour chaque point clé une trajectoire  $\tau$  et  $\tau'$  est calculée favorisant respectivement les cibles du 1er point clé puis du 2ème point clé. Ces trajectoires sont interpolées en les reliant grâce à la courbe d'interpolation. (e) Cette courbe permet à l'utilisateur de contrôler l'interpolation entre les deux courbes en réglant le temps où la trajectoire passe du 1er au 2ème point clé et la vitesse à laquelle bouge la caméra.

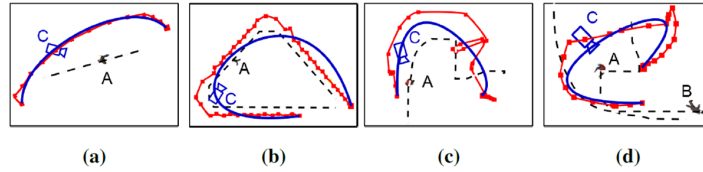


Figure 8: Exemple de rails de caméra (courbe de Bézier en bleu) calculés pour approximer les trajectoires "brutes" (en rouge) en suivant un ou plusieurs acteurs effectuant des mouvements de plus en plus complexes (de (a) jusqu'à (b)).

cette trajectoire. Pour pouvoir éviter ce problème, les auteurs définissent une nouvelle trajectoire en cherchant une courbe de Bézier à 4 points de contrôle se rapprochant le plus possible à la trajectoire "brute". Cette trajectoire peut ainsi être assimilée aux rails utilisés en cinématographie.

Après avoir minimisé l'écart entre la courbe de Bézier et la trajectoire brute, les auteurs limitent les variations de vitesse et accélération sur la position et l'orientation de la caméra. Cela assure donc un mouvement fluide de celle-ci (figure 8).

Ces différentes approches proposent donc diverses manières pour gérer la composition visuelle et la trajectoire d'une seule caméra, mais ne considèrent pas la coordination de plusieurs caméras pour filmer une scène.

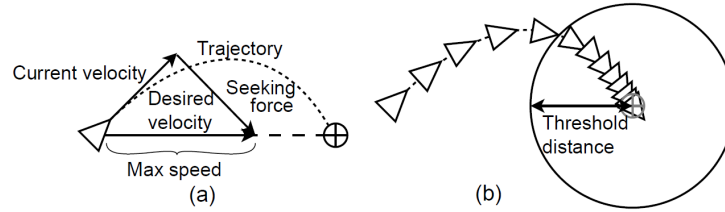


Figure 9: Illustration de deux comportements: (a) comportement lors d’une recherche. (b) comportement lors d’une arrivée à une cible

### 2.3.4 Gestion d’un ensemble de caméra

Il est intéressant, si l’on souhaite avoir plusieurs plans d’une même scène ou si plusieurs événements que l’on souhaite filmer se réalisent simultanément, de disposer de plusieurs caméras contrôlables. Une méthode, proposée par Galvane [QGC<sup>+</sup>13], permet de filmer les scènes que l’on juge intéressantes dans un environnement complexe (une foule en mouvement). Dans cet article, deux types d’événements influencent les caméras:

- Un face à face: Une scène où 2 acteurs se regardent mutuellement pendant un certain temps
- Un mouvement de groupe: Une scène où un groupe d’acteurs bougent dans la même direction

On souhaite donc manipuler les caméras pour qu’ils puissent filmer ces scènes. Or, calculer une trajectoire pour chaque caméra est coûteux en temps de calcul. De plus, l’environnement est dynamique. Et donc, une trajectoire calculée à un certain temps peut ne plus être viable au temps suivant. Pour palier à ces problèmes, les caméras sont contrôlées par une force de "steering" (pilotage). Les trajectoires des caméras varient en fonction des forces qui leurs sont appliquées. Il est donc possible, comme le montre l’article de Reynolds [REY99] d’effectuer un ensemble varié de comportement. Par exemple, il est possible de simuler la fuite, la poursuite, l’errance, d’éviter des obstacles... Un exemple montrant un comportement lors d’une recherche et lors d’une arrivée est illustré figure 9. Dans cet article, les forces de "steering" sont utilisées pour maintenir une bonne composition visuelle de l’évènement, de maintenir la visibilité de la caméra sur l’évènement, de maintenir une distance donnée de l’évènement et d’assurer que plusieurs caméras filment l’évènement sous différents angles.

Finalement, ces diverses approches proposent des méthodes pouvant être appliquées principalement aux environnements virtuels. En effet, contrôler un drone demande de respecter de nombreuses contraintes supplémentaires.

## 3 Le contrôle de drones cinématographiques

### 3.1 Spécificités d’un drone

Il est possible d’embarquer une caméra dans un drone pour filmer des scènes réel. Or, manipuler un objet réel est beaucoup plus contraignant que de manipuler une caméra virtuelle. En effet, cela demande :

- De connaître la position et l’orientation du drone en temps réel
- De découvrir son environnement au fur et à mesure ( La connaissance de l’environnement n’est pas acquise, le drone doit apprendre la position des objets)



Figure 10: (A gauche) L’outil interactif permettant de gérer la caméra. L’utilisateur spécifie les images clés dans un environnement virtuel en 3D (a) ou 2D (b). L’utilisateur peut aussi contrôler le timing de ces images clés en éditant les courbes (d). Le résultat peut être prévisualisé dans l’environnement 3D (c). L’utilisateur peut régler la focale de la caméra pour s’approcher le plus possible de la caméra embarquée par le drone (e). L’outil propose aussi un retour visuel permettant notifier l’utilisateur si la trajectoire qu’il veut réaliser respecte les limites physiques du drone (f). Quand l’utilisateur est satisfait du résultat, il peut lancer l’enregistrement dans le monde réel en appuyant sur le bouton ”Start Capture” (g). (A droite) Exemples de divers images extraient d’un enregistrement réalisé dans le monde réel, grâce à cet outil (h)

- De prendre en compte que l’environnement est dynamique (peut changer à chaque instant)
- De respecter les limites physiques du drone (vitesses de déplacement et de rotation du drone (borné), vitesses de rotation des pales et limitation des angles de vue de la caméra par sa boîte et le drone.)
  - D’éviter de rentrer en collision avec les acteurs ou d’autres drones
  - D’éviter de rentrer en collision avec l’environnement
  - De commander 3 degrés de liberté supplémentaires pour orienter le drone. 1 DDL permet de faire tourner le drone sur l’axe vertical et les 2 autres permettent de déplacer le drone. En effet, Le drone doit rester parallèle au sol ou légèrement incliné. On doit donc changer l’angle de la caméra pour changer l’angle de vue.
- De créer les trajectoires des drones d’une continuité C4 (du 4ème ordre) pour respecter les équations physiques de mouvement du drone. Les différentes approches qui vont suivre se reposent sur une modélisation du drone et de sa dynamique.

### 3.2 Filmer des scènes réelles avec des drones

Des méthodes ont été proposées pour filmer des scènes réelles avec des drones. Notamment Joubert et al. [JRT<sup>+</sup>15] qui ont proposés une des 1ères techniques permettant de filmer une scène avec un drone, à partir d’images clés et d’une interpolation entre ces images. Ces images clés, ou plans visuels, sont définis par la position du drone (images clés ”regarde de”) et la position à laquelle le drone doit filmer (images clés ”regarde à”) à un certain temps. Pour les définir, un outil a été développé (figure 10). Dans cet outil, on spécifie les plans visuels que l’on souhaite dans un environnement 3D réaliste (GOOGLE EARTH). Cet environnement virtuel permet d’avoir une prévisualisation virtuelle et complète de l’enregistrement.

Après que l’utilisateur est rentré les différents plans clés qu’il souhaite, une trajectoire d’une

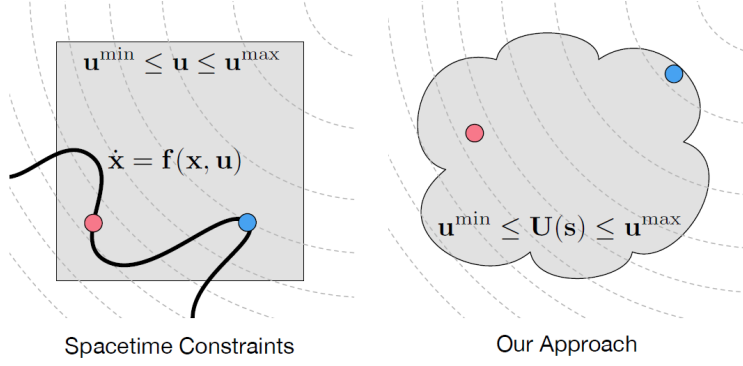


Figure 11: Illustration des différences entre l'approche par contraintes spatio-temporelles (A gauche) et l'approche proposée (A droite). Dans les contraintes spatio-temporelle, Les limitations physiques sont simples à effectuer car elles peuvent être définies par des inégalités linéaires (région grisée à gauche). Mais, dans ce cas, la dynamique du drone doit être définie par des équations non linéaires (courbe en gras à gauche). Cette non linéarité force à un solveur de réaliser un nombre important d'étape pour avoir une approximation correcte entre la solution initiale (point rouge à gauche) et la solution finale (point bleu à gauche). Dans l'approche proposée, les limitations physiques sont définies de manière non-linéaire (région grisée à droite) mais la dynamique du drone est implicite. cette approche permet à un solveur une progression rapide entre une solution initiale (point rouge à droite) et la solution optimale (point bleu à droite)

continuité C4 est générée. Pour cela, une courbe par morceaux, composée de polynôme du 7ème degrés, est calculée. En effet, de nombreuses expérimentation montrent que ce type de courbe produit la trajectoire la mieux bornée et la plus lisse. Pour trouver les différents coefficients de ce polynôme, on cherche un nombre distinct de morceaux de polynôme, qui, reliés ensemble assure la continuité C4 et interpolent parfaitement les images clés.

Par la suite, la courbe est re-paramétrisée pour produire deux fonctions permettant à l'utilisateur d'orienter et de positionner le drone en fonction du temps (1).

Finalement, l'outil indique visuellement si l'enregistrement prévu viole les limites physiques (vitesses et angles cf [Spécificités d'un drone]) de leur drone et calcul le signal de contrôle du drone. L'utilisateur peut, pour rendre la trajectoire possible, adapter ces différentes images clés et leurs timing. Mais comme il est difficile de savoir qu'est ce qu'il faut changer pour être dans les limites des contraintes que peut accepter le drone, Roberts et al. [RH16] ont proposé une solution permettant de modifier automatiquement le timing de la trajectoire pour quelle soit possible. Cette solution utilise l'outil proposé par Joubert et al. [JRT<sup>+</sup>15] en ajoutant des équations non-linéaires qui définissent les limites physiques du drone. Ces équations permettent à un solveur de trouver une solution respectant la configuration que l'on souhaite plus rapidement par rapport aux contraintes spatio-temporelles [WK88], qui propose une manière simple de définir des limites physiques (équations linéaires) mais dont les solutions pour une configuration sont laborieuses à trouver (figure 11). Cette méthode propose donc une optimisation donnant un temps de vol plus rapide qu'une simple variation du temps et plus efficace que l'approche par contraintes spatio-temporelles.

Une autre solution, proposée par Gebhardt [GHN<sup>+</sup>16]], apporte un autre outil et une autre manière d'interagir avec des trajectoires de drones. Cet outil permet, grâce à une tablette, de contrôler la trajectoire du drone avec des images clés "regarde de" et "regarde à" (figure 12). Après

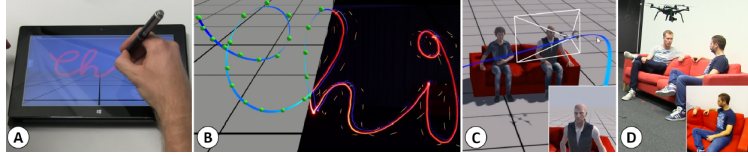


Figure 12: Outil interactif pour réaliser différentes trajectoires de drone. (A) Interface utilisateur permettant de spécifier les images clés et la dynamique du vol du drone. (B) L'algorithme d'optimisation permettant de générer les trajectoires possibles (C) et une prévisualisation 3D permettant d'éditer ceux-ci. (D) Le vol réalisé dans un environnement réel.

les entrées utilisateurs via l'outil, on minimise l'énergie dépensé par le drone. Cela se fait souvent en tentant de d'atteindre les positions souhaitées avec un minimum d'effort. Le calcul s'effectue en minimisant la somme quadratique de la différence entre l'entrée souhaitée et observée (l'erreur). Cette opération de minimisation s'effectue en respectant les limites physiques du drone. Donc, Les trajectoires générées appartiennent forcément à l'ensemble des trajectoires possibles.

Cette méthode est utilisable pour de multiples scénarios comme des vidéos-aériennes, des peintures lumineuses et des courses de drones.

Les méthodes que nous venons de présenter sont donc assez générales mais ne proposent pas de gérer la visibilité ou les différentes collisions possibles avec la scène. De plus, ils n'intègrent pas vraiment de connaissances cinématographiques. L'utilisateur doit lui-même avoir cette connaissance et manipuler les paramètres du drone et de la caméra.

### 3.3 Transposition de la connaissance cinématographique aux drones

Dans le cas où l'utilisateur cherche à respecter certaines propriétés cinématographiques de cadrage ou compositions visuelles. Il est intéressant de proposer une méthode permettant d'intégrer ces connaissances en respectant les différentes contraintes du monde réel. Joubert [JEG<sup>+</sup>16] a proposé, dans cette optique, une technique permettant de réaliser une transition fluide entre différentes prises de vue correspondant à un modèle cinématographique. Pour ce faire, la position du drone est connue grâce à un système GPS ; on se limite à un ou deux acteurs devant rester à la même position et équipés d'un système de tracking. Un tel système permet de connaître la position exacte des acteurs que l'on souhaite filmer, dans le repère utilisé par le drone. Grâce à ces informations et un réglage utilisateur permettant de définir quelle composition il souhaite (Apex, Close Apex, Internal ou External. Voir section 2.1.1), le drone se place automatiquement. Pour placer celui-ci, un modèle de la caméra et des acteurs (figure 13) est créé. Chaque composition visuelle fixe les paramètres de ce modèle et permet de déduire le lieu où doit se placer et le lieu où doit regarder la caméra.

Dans le cas où l'utilisateur souhaite passer d'une composition visuelle à une autre, une trajectoire respectant les diverses restrictions du drone (Voir section 3.1) est calculée. Premièrement, on fixe les positions et orientations de départ et d'arrivée grâce au modèle fixé précédemment. Puis on interpole linéairement ces deux points par rapport à la distance d'une cible et de la caméra ; on interpole circulairement par rapport à l'orientation de la caméra quand elle regarde cette même cible aux positions de départ et d'arrivée. De cette manière, il en résulte deux trajectoires réalisées par rapport aux deux cibles. Finalement, en pondérant ces deux trajectoires et en imposant que la trajectoire résultante ne doit pas être trop près des acteurs, on en déduit la trajectoire finale



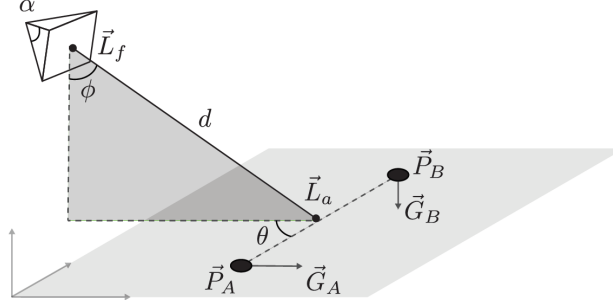


Figure 13: Aperçu du modèle de la caméra et des acteurs. On attribut à chaque acteur un vecteur  $\mathbf{P}$  et  $\mathbf{G}$ . On modélise la caméra comme un point d'où il regarde  $\mathbf{L}_f$ , un point qu'il regarde  $\mathbf{L}_g$  et un angle de vue  $\alpha$ . Des angles  $\theta$  et  $\phi$  sont définis pour décrire les angles entre le sens où est pointé la caméra et la ligne de l'action (droite passant par les positions des 2 acteurs).  $d$  indique la distance entre le point d'où regarde la caméra et le point qu'il regarde.

réalisable par le drone.

Ce principe donne une manière permettant d'inclure des connaissances cinématographiques lors d'un enregistrement avec un drone, mais le système GPS manque de précision, le système de tracking est encombrant, il peut y avoir que deux sujets, le drone n'est pas capable d'éviter les obstacles présent dans son environnement et, finalement, les acteurs doivent rester statiques. Pour palier à ce dernier problème, Galvane [GFTG16] a proposé un nouveau modèle permettant d'intégrer une connaissance cinématographique dans le comportement de drones autonomes pour se déplacer et cadrer des acteurs.

Ce modèle se base sur une adaptation du modèle de Lino [LC12] / Lino [LC15], où l'intégration de la distance minimum aux acteurs est directement incluse dans le modèle proposé. Cette adaptation propose de produire des images respectant un cadrage spécifique en utilisant uniquement des commandes cinématographiques via le PSL (prose storyboard language), un langage permettant de spécifier un scénario. Ce langage permet de décrire un film par ces différents cadrages successifs. En utilisant un système de capture de mouvements et les spécifications sur le cadrage données par l'utilisateur, ce modèle permet de générer une trajectoire fluide, tout en gardant la composition visuelle voulue (sur les cibles désirées).

L'interprétation du PSL, pour en déduire une position sur la surface sphérique (1 acteur) ou torique (2 acteurs), s'effectue selon la position voulue de la cible sur l'écran et la taille voulue de la cible sur l'écran. La position optimale pour la caméra correspond au point sur la surface qui satisfait le plus cette contrainte.

Dans le cas où l'on souhaite effectuer une transition entre deux configurations de caméra différentes, on se base sur l'interpolation des différentes compositions visuelles proposées par [caméra on rails]. Pour l'appliquer à un monde réel, on calcule la posture initiale de la caméra en prenant en compte la cible visuelle de celle-ci. Puis, en sachant le temps de trajet que l'on souhaite, on interpole chaque propriété visuelle correspondant à la surface sphérique (ou torique) pour la posture initiale et finale de la caméra. En sachant que la posture finale est spécifiée par l'utilisateur. Le résultat de la trajectoire ainsi calculé est illustré figure 14.

L'ensemble des solutions proposées dans cette partie donnent donc des moyens pour permettre à une personne novice (cad une personne ayant peu de connaissances cinématographiques et sur les drones) de réaliser un ensemble d'actions respectant divers aspects cinématographiques. Mais,

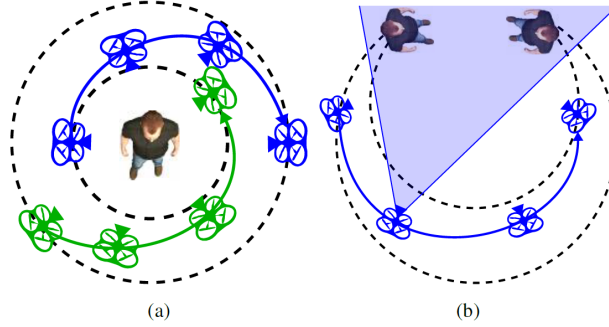


Figure 14: L’interpolation de compositions visuelles crée une trajectoire naturelle (a) et maintient les propriétés visuelles sur la cible le long de la trajectoire (b)

malgré la simplification d’utilisation que ces solutions proposent, elles sont limitées par certains aspects. En effet, elles se limitent à l’utilisation d’un seul drone et deux acteurs maximum. De plus, elles peuvent nécessiter de se munir d’un système de tracking plus ou moins encombrant. Et, finalement, ils n’intègrent pas de système permettant d’esquiver les obstacles.

## 4 Conclusion et perspectives

Au travers de cette bibliographie, nous avons présenté un ensemble de techniques et d’outils susceptibles d’aider à la production de contenus cinématographiques pour les drones. Une grande partie de ces contributions se concentrent sur l’idée de simplifier au maximum la création de contenus pour des novices. Les différentes méthodes d’interactions proposées – qu’il s’agisse de textes, de manipulations dans l’espace image, de définition de plans clés, ou même simplement via un dessin de la trajectoire – offrent toutes une grande liberté créatrice à l’utilisateur. Cependant, à travers toutes ces approches, beaucoup d’aspects n’ont pas été traités, ou l’ont été séparément. Dans le cadre du contrôle de drones, les notions de visibilité ou d’évitement d’obstacles par exemple restent peu abordées. Adapter des méthodes qui restent pour le moment dédiées à des mondes virtuels semble donc une piste intéressante pour de futures recherches. La coordination d’un ensemble de drones qui filment une même scène, sous divers angles, en respectant les diverses contraintes cinématographiques constitue également un axe de recherche intéressant et pour le moment inexploré. C’est précisément dans ce cadre que vient se situer ce stage.

## References

- [Bli88] Jim Blinn. Where am i? what am i looking at? *IEEE Computer Graphics and Applications*, 8(4):76–81, 1988.
- [GCLR15] Quentin Galvane, Marc Christie, Chrstophe Lino, and Rémi Ronfard. Camera-on-rails: automated computation of constrained camera paths. In *Proceedings of the 8th ACM SIGGRAPH Conference on Motion in Games*, pages 151–157. ACM, 2015.
- [GFTG16] Quentin Galvane, Julien Fleureau, Francois-Louis Tariolle, and Philippe Guillotel. Automated Cinematography with Unmanned Aerial Vehicles. In M. Christie, Q. Galvane,



- A. Jhala, and R. Ronfard, editors, *Eurographics Workshop on Intelligent Cinematography and Editing*. The Eurographics Association, 2016.
- [GHN<sup>+</sup>16] Christoph Gebhardt, Benjamin Hepp, Tobias Nægeli, Stefan Stevšić, and Otmar Hilliges. Airways: Optimization-based planning of quadrotor trajectories according to high-level user goals. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 2508–2519, New York, NY, USA, 2016. ACM.
  - [JEG<sup>+</sup>16] N. Joubert, J. L. E, D. B Goldman, F. Berthouzoz, M. Roberts, J. A. Landay, and P. Hanrahan. Towards a drone cinematographer: Guiding quadrotor cameras using visual composition principles. *ArXiv e-prints*, October 2016.
  - [JRT<sup>+</sup>15] Niels Joubert, Mike Roberts, Anh Truong, Floraine Berthouzoz, and Pat Hanrahan. An interactive tool for designing quadrotor camera shots. *ACM Transactions on Graphics (TOG)*, 34(6):238, 2015.
  - [LC12] Christophe Lino and Marc Christie. Efficient composition for virtual camera control. In *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA '12, Aire-la-Ville, Switzerland, Switzerland, 2012. Eurographics Association.
  - [LC15] Christophe Lino and Marc Christie. Intuitive and efficient camera control with the toric space. *ACM Transactions on Graphics (TOG)*, 34(4):82, 2015.
  - [OSTG09] Thomas Oskam, Robert W Sumner, Nils Thuerey, and Markus Gross. Visibility transition planning for dynamic camera control. In *Proceedings of the 2009 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pages 55–65. ACM, 2009.
  - [QGC<sup>+</sup>13] Quentin, Marc Galvane, Rémi Christie, Chen-Kim Ronfard, Marie-Paule Lim, and Cani. Behaviors for autonomous cameras. In *SIGGRAPH conference on Motion in Games*, pages 93–102. ACM, 2013.
  - [REY99] C REYNOLDS. Steering behaviors for autonomous characters. In *Game Developers Conference*, pages 763—782, 1999.
  - [RH16] Mike Roberts and Pat Hanrahan. Generating dynamically feasible trajectories for quadrotor cameras. *ACM Transactions on Graphics (Proc. SIGGRAPH 2016)*, 35(4), 2016.
  - [WK88] A. Witkins and M. Kass. Spacetime constraints. In *SIGGRAPH 1988*, 1988.