

# Développement d'un module de suivi de mouvements et localisation d'un patient par rapport à un ensemble d'équipements dans une salle de traitement par radiothérapie

Lyon, France - LIRIS / SAARA - **Rapport Interne**

MA PORTELA SOTELO  
miguel.portela-sotelo@liris.cnrs.fr

J-M MOREAU  
jean-michel.moreau@liris.cnrs.fr

31 septembre 2009

## Résumé

Dans ce rapport, nous présentons un module de suivi de mouvement et localisation pour un système de positionnement d'un patient, pour des traitements par radiothérapie. Ce travail se focalise sur le besoin exprimé par des médecins et des radiophysiciens d'avoir une vue globale du système du traitement pour pouvoir suivre le patient et la totalité des équipements mobiles ou non. On veut visualiser pour superviser, contrôler et finalement corriger le positionnement. Notre module est « basé modèle » et utilise des caméras vidéo pour surveiller la salle de traitement. On cherche donc à améliorer la qualité en augmentant la sécurité du traitement. Nous avons développé un prototype d'application comportant une partie de ce module, qui nous a permis de valider sa faisabilité. Nous décrivons ce prototype, et nous présentons les résultats.

Mots clefs : Réalité augmentée, suivi basé modèle, reconstruction géométrique, temps itératif

## 1 Introduction

Ce travail s'inscrit dans le cadre du centre ETOILE<sup>1</sup>, et vise à compléter un équipement de pointe qui sera utilisé dans ce centre de soin et de recherche clinique du cancer prochainement construit à Lyon. À travers ce rapport, nous présentons le travail effectué lors de 6 mois de stage de fin d'études (niveau M2), qui a débuté le 1<sup>er</sup> avril 2008. Tout d'abord, nous parlerons des besoins qui l'ont motivé. Ensuite, les objectifs seront énumérés et les solutions exploitées seront expliquées et comparées. Après avoir exposés les résultats, nous finirons par tirer des conclusions et présenter les perspectives possibles de ce travail.

---

1. pour plus d'information : [www.centre-etoile.org](http://www.centre-etoile.org)

## 1.1 Le LIRIS et le Centre Léon Bérard

Le LIRIS (Laboratoire d'InfoRmatique en Image et Systèmes d'information) est né en 2003 à la suite du regroupement de plusieurs laboratoires de recherche lyonnais (LIGIM, LISI, RFV) dans le domaine des Sciences et Techniques de l'Information et de la Communication. Il est associé au CNRS avec le label UMR 5205.

Regroupant environ 280 personnes, dont près de 110 chercheurs et enseignants-chercheurs, le LIRIS a quatre tutelles : l'INSA de Lyon, l'Université Claude Bernard Lyon 1, l'Ecole Centrale de Lyon et l'Université Lumière Lyon 2, et des sites à La Doua, Ecully et Bron.

Les activités du laboratoire sont regroupées dans deux départements thématiques : "Image" et "Données, Connaissances, Services".

Le département Image est structuré en cinq équipes réunies autour de trois pôles de compétences :

- coopération, analyse et modélisation géométrique pour les objets 3D (multirésolution et graphes ; géométrie discrète ; géométrie algorithmique),
- simulation, rendu réaliste et réalité augmentée,
- extraction d'information et reconnaissance d'objets dans les images, vidéos, documents numériques.

Le département Données, Connaissances, Services est constitué de cinq équipes réunies en deux pôles de compétences :

- découverte de connaissances (data mining, modélisation des systèmes complexes, ingénierie des connaissances),
- ingénierie des données et des services (sécurité et confidentialité, modélisation, intégration et interrogation, composition de services).

Le LIRIS réalise une activité de recherche de fond sur ces deux thèmes scientifiques, tout en développant un savoir-faire dans des domaines applicatifs pour la société :

- les loisirs numériques (jeux vidéos, cinéma d'animations, traitement de données multimédia...),
- la biologie et la santé (fouille de données, analyse et modélisation de systèmes complexes, e-santé...),
- la culture et le patrimoine (bibliothèque numérique, édition critique, numérisation de documents anciens, archivage, musée virtuel 3D...),
- l'intelligence ambiante (systèmes pervasifs, réseaux de capteurs, vidéosurveillance intelligente, objets communicants sécurisés...)

Le centre de lutte contre le cancer Léon Bérard (CLB) est, comme vingt autres centres en France, un établissement de santé privé à but non lucratif, participant au service public hospitalier. A ce titre, ces centres sont financés par l'assurance maladie et contrôlés par le Ministère de la Santé dans les mêmes conditions que les hôpitaux publics. Leur statut comporte explicitement une vocation de recherche fondamentale et appliquée.

Depuis plusieurs années, le laboratoire travaille en collaboration avec ETOILE et aussi avec le personnel du CLB : Christian Carrie, Sébastien Clippe (radiothérapeute et chef de clinique assistant) et l'équipe des médecins, dont Chantal Ginestet.

## 1.2 L'hadronthérapie et le Centre ETOILE

L'hadronthérapie est, comme la radiothérapie, une technique de traitement du cancer visant à détruire les cellules cancéreuses en les irradiant avec un faisceau de particules. La radiothérapie conventionnelle utilise essentiellement des rayons X, l'hadronthérapie d'autres types de particules : des hadrons, notamment des protons (protonthérapie) et des particules de carbone (carbonothérapie).

Avant le traitement, une technique d'imagerie médicale est réalisée afin de déterminer la taille, la forme et la localisation de la tumeur. Ensuite, les images sont transférées sur le système de planification de traitement (SPT), où le radiophysicien (ou le médecin) identifie la tumeur et les zones saines à éviter. Ensuite, un plan de traitement est généré par le radiophysicien en indiquant la balistique et la dose désirée au niveau de la tumeur en épargnant « au mieux » les tissus sains. Durant le traitement, le patient est positionné sur une table dans la même position qu'il avait pendant le procédé d'imagerie médicale précédemment fait. Ensuite, un grand nombre de faisceaux sont administrés au patient. Avant chaque faisceau, le radiophysicien peut réaliser une imagerie X, pour repositionner le patient en fonction de la direction des faisceaux et en corrigeant les mouvements éventuels du patient et de la tumeur pour assurer un « ciblage » précis.

Le centre ETOILE (Espace de Traitement Oncologique par Ions Légers dans le cadre Européen) est le résultat d'un projet (appelé également ETOILE) de construction d'un centre médical dédié au traitement de certaines

tumeurs cancéreuses par faisceaux d'ions carbone issus d'un accélérateur. Ce sera un centre de soins et de recherche clinique. Les premiers résultats cliniques obtenus depuis 1994 au NIRS (Japon) et depuis 1997 au GSI (Allemagne) ont démontré la faisabilité du traitement « expérimental » par ions carbone avec des résultats prometteurs, en particulier pour les cancers inopérables et résistants aux rayonnements conventionnels.

L'Université Claude Bernard Lyon 1 (UCBL) a pris l'initiative en 1999 de confier à des médecins et des physiciens la rédaction d'un cahier des charges pour la création d'un centre d'hadronthérapie par ions légers. Cette étude a reçu le soutien financier de l'ANVAR, de la Région Rhône-Alpes, de la Communauté Urbaine de Lyon et du Ministère de la Recherche.

L'avant-projet qui a suivi s'est concrétisé en 2002. Il est le fruit d'une collaboration étroite entre médecins (oncologues, radiothérapeutes des Hospices Civils de Lyon, du Centre Hospitalier Universitaire de Grenoble, du Centre Léon Bérard de Lyon), scientifiques (physiciens, informaticiens, économistes des Universités de Lyon et Grenoble, radiophysiciens de Lyon, Grenoble, ainsi que de Nice et Orsay) et ingénieurs spécialistes des accélérateurs (CEA et CNRS dans le cadre d'une convention avec l'UCBL). Il a donné lieu à une première coopération européenne avec le Centre GSI à Darmstadt, le CERN à Genève et la fondation TERA à Milan, puis à la constitution d'un réseau européen (ENLIGHT++ : European Network for LIGHT ion Hadron Therapy) dans le cadre du 5ème Programme Cadre de Recherche et Développement (PCRD).

ETOILE, qui est maintenant devenu un « GCS » (Groupe de Coopération Sanitaire), situé au BioParc de Lyon, est dans sa phase de réalisation d'un projet de Partenariat Public Privé (PPP) avec le premier patient programmé pour le début 2013. Le centre ETOILE prévoit de traiter à terme entre 1000 et 2000 patients par an avec un recrutement national, voire européen.

### 1.3 DOSIsoft et AS2I, les industriels intéressés par le projet

DOSIsoft est une spin-off de l'Institut Gustave Roussy et de l'Institut Curie, deux centres majeurs impliqués dans la lutte contre le cancer, créée en 2002. Elle est spécialisée aujourd'hui dans le développement, la commercialisation et la maintenance de systèmes de planification en radiothérapie destinés aux hôpitaux. DOSIsoft propose des plateformes logicielles d'imagerie et de contournage (IMAgO<sup>TM</sup>), de simulation virtuelle (SIMAgO<sup>TM</sup>) et de dosimétrie (ISOgray<sup>TM</sup>) qui sont le fruit du partenariat avec les plus grands établissements et instituts de lutte contre le cancer et centres de recherche en informatique et analyse d'images.

Leader français, DOSIsoft applique une stratégie d'innovation continue et est impliquée dans de nombreux projets de recherche français et européens visant à optimiser les outils de planification en radiothérapie et à en faire bénéficier le plus grand nombre de patients. Parallèlement, dans le cadre d'une politique de développement internationale, DOSIsoft est en train de créer un réseau de distributeurs dans les pays occidentaux.

AS2I est une entreprise française spécialisée dans le conseil, l'ingénierie et l'étude de la réalisation de machines et de systèmes robotiques et complexes. AS2I est actuellement le coordinateur du projet POROS (POsitionnement RObotisé du patient)<sup>2</sup>, dont DOSIsoft fait également partie du consortium.

### 1.4 Le stage

Des accidents sont recensés tous les ans dans les centres de radiothérapie. Souvent à cause des mouvements imprévus du patient<sup>3</sup>, ou des mauvaises manipulations de l'équipement par le personnel. Toute erreur significative dans le positionnement du patient ou des équipements peut avoir des conséquences graves, du fait de l'utilisation d'un faisceau de rayonnement directement focalisé sur le patient ou bien des sources radioactives placées au contact des tissus<sup>4</sup>. Aujourd'hui, on s'oriente vers l'automatisation des traitements contre le cancer pour pallier ces problèmes. Les progrès en mécanique, en physique et en robotique permettent d'aborder de nouvelles solutions, et de nombreux travaux ont déjà été présentés. Cependant, ils proposent soit des systèmes de contention ou basés sur la spirométrie qui sont très gênants pour le patient, soit des systèmes à base des rayons X ou d'infrarouges, très coûteux [Mazeron et al., 2007]. Tous ces travaux se centrent sur le patient, et n'utilisent pas les informations sur l'environnement de la salle.

L'automatisation des traitements permettrait principalement de réduire le nombre d'interventions d'ordre humain et donc la probabilité d'erreurs. Ceci implique une diminution du temps de traitement assez importante, et donc un système moins gênant pour le patient qui pourra tenir plus facilement dans une position donnée sans bouger. Ce stage se focalise sur le besoin, exprimé par les médecins et les radiophysiciens, d'avoir une vue globale

---

2. pour plus d'information : [www.dosisoft.com/dosisoft-imagerie/projets-partenariats/poros.htm](http://www.dosisoft.com/dosisoft-imagerie/projets-partenariats/poros.htm)

3. pour plus d'information : [www.asn.fr](http://www.asn.fr)

4. source : Institut de Radioprotection et de Sécurité Nucléaire

du système du traitement pour pouvoir suivre le patient et la totalité des équipements mobiles ou non, dans le but ultime de sécuriser la manipulation de l'ensemble, améliorer la précision du traitement et faire du contrôle automatique.

Imaginons un salle de traitement par radiothérapie comme une salle carrée ou rectangulaire avec une seule porte. Dans cet espace fermé, on retrouve des objets propres à la radiothérapie, essentiellement le bras d'irradiation et une table où le patient sera allongé. Certains de ces objets ne sont pas toujours présents dans la salle. On dispose des caméras fixes permettant de visualiser tout l'intérieur. A cette salle est associée une salle de contrôle située juste à coté, avec la possibilité d'avoir une fenêtre, toujours fermée, vers la salle de traitement. Nous nous intéressons au moment de la réalisation du traitement, où le patient sera le seul individu à être dans la salle et les objets seront presque statiques. Comme le milieu radioactif l'impose, la porte de cette salle restera fermée pendant toute la durée du traitement.

Dans cette configuration, on veut pouvoir suivre le patient et les objets dans la salle à l'aide des caméras installées. On impose l'utilisation d'un repère global pour le suivi de la position du patient et des objets. Ce qui permet de localiser le patient en fonction des équipements pour assister son positionnement. On veut visualiser pour superviser, contrôler et finalement corriger. Nous songeons, avec la suite de ce travail, d'améliorer la précision et la robustesse du traitement. Il est envisagé également d'utiliser des méthodes de reconnaissance de scénario.

Le cancer est une des premières causes de mortalité au niveau mondial, à l'origine de 7,9 millions de décès en 2007<sup>5</sup>. De plus, le nombre de nouveaux cas de cancers est en croissance constante en raison du vieillissement de la population. D'où l'importance de l'automatisation du traitement de cette maladie. Le système étudié ici pourrait aussi être utilisé dans d'autres domaines nécessitant du positionnement.

Les difficultés sont nombreuses dans le domaine de la vision par ordinateur (calibrage, lumière, sémantique, ...). Il y a aussi le coté humain. Le patient est déjà mal à l'aise par le simple fait d'être malade et de se retrouver seul dans une salle avec des équipements parfois effrayants. Nous devons être attentifs aux choix faits pour ne pas gêner d'avantage le patient.

Certains travaux se focalisent sur la possibilité de simuler le mouvement des organes internes (principalement au niveau du thorax dans le cadre du cancer du poumon [Didier et al., 2007]), mais ils ont besoin des données de externes pour être fonctionnels. Notre système viendrait compléter ces travaux, notamment ceux développés au sein de SAARA [Viguiet et al., 2007].

#### 1.4.1 Le sujet et les objectifs

Le sujet du stage est : développer un module capable, à partir des images acquises avec des caméras et des données enregistrées dans des fichiers descripteurs, de localiser et de suivre les mouvements d'un patient et des objets dans une salle de traitement par radiothérapie, et ceci en temps réel. Le module doit pouvoir fournir des informations sur l'emplacement du patient par rapport aux machines. Il doit donc être capable de construire une modélisation 3D de la salle en temps réel.

Les objectifs sont :

1. Localisation du patient et des équipements dans la salle de traitement, à partir des informations contenues dans des fichiers descriptifs (fichiers xml).
2. Mise en relation des éléments retrouvés dans la salle, à partir d'un unique repère global.
3. Modélisation de la salle et des éléments retrouvés.
4. Suivi des positions du patient et des équipements.

Ce stage, qui a été effectué dans le laboratoire LIRIS avec le financement du projet ETOILE, a été conçu dans l'espoir d'intéresser les deux industriels, DOSIsoft et AS2I, dans le cadre d'une bourse CIFRE.

Du fait de la complexité du sujet, il a été décidé de réaliser un prototype pour démontrer la faisabilité d'un tel système, qui pourra être amélioré dans la suite par un travail de recherche.

Ces informations doivent permettre de connaître la position, voire les mouvements, du patient en temps réel, et en fonction des équipements présents dans la salle. Le système doit aussi gêner le moins possible le patient.

Dans un premier temps, notre système doit être capable des signaler aux opérateurs si un problème se présente. Dans un futur proche, on souhaite que le système puisse apporter des corrections à la position du patient et des machines.

---

5. source : Organisation Mondiale de la Santé

## 2 État de l'art

Nous cherchons à réduire les erreurs dues à des variations de positionnement de patients ou aux mouvements physiologiques qui peuvent être la cause d'échec d'un traitement par radiothérapie. Pour ceci, nous souhaitons utiliser un système de suivi externe, à partir de caméras. Dans ce chapitre, nous exposons le résultat de notre recherche bibliographique dans les domaines de la vision par ordinateur et du positionnement en radiothérapie.

Deux axes ont été importants dans ces recherches :

- les systèmes médicaux,
- les systèmes de suivi des individus et des objets.

### 2.1 Les systèmes médicaux

#### 2.1.1 Systèmes de positionnement utilisés actuellement

**Novalis**<sup>6</sup> : Le système Novalis est un système de radiochirurgie conçu par les sociétés BrainLAB et Varian. Dans ses premières versions, il a été entièrement dédié à la Radio-Neurochirurgie Stéréotaxique<sup>7</sup> et plus de 70 hôpitaux dans le monde ont adopté ce système.

Aujourd'hui, le système Novalis inclut un système de positionnement par repérage et suivi. Le système positionne automatiquement le patient avec un système robotique à six degrés de liberté pour aligner avec précision la tumeur avec le faisceau. Un masque personnalisé sert à l'immobilisation du patient, et non au positionnement, comme c'est le cas des autres systèmes. Un système de surveillance par infrarouge suit en permanence le patient pendant toute la durée du traitement. Deux unités d'imagerie par rayon X placés dans la fosse et deux détecteurs plats en silicium permettent de numériser les images de l'intérieur du patient et les marqueurs implantés avant et pendant le traitement. Le radiothérapeute contrôle le système depuis la salle de contrôle, et réalise des images X selon les besoin. Après la prise d'une image, un processus automatique vérifie la position du patient et calcule la correction nécessaire à appliquer. Ensuite, le système affiche et réalise les mouvements nécessaires pour aligner les tumeurs et le faisceau. Le processus de mise en place prend moins de deux minutes et permet une précision de 1 mm. Plus de 65 hôpitaux à travers le monde s'appuient sur ce système de positionnement.

**Cyberknife**<sup>8</sup> : Le système CyberKnife est le seul système de radiochirurgie qui utilise la robotique intelligente (dite robotique de seconde génération) pour traiter des tumeurs dans tout le corps avec une précision inférieure à 1 millimètre. Il utilise une technique où de multiples faisceaux de rayonnement convergent avec une grande précision vers la tumeur tout en minimisant l'impact sur les tissus environnants sains. Grâce à sa précision, aucun cadre de stéréotaxie n'est nécessaire.

Composition du Système :

- Accélérateur linéaire compact : Cette section est montée sur le bras du robot manipulateur, car elle plus petite et moins lourde comparée aux accélérateurs linéaires standard.
- Bras robotisé : Avec 6 axes de mouvement, il est conçu pour déplacer et diriger la partie accélératrice. Le temps de réponse du bras de robot est rapide, et permettrait même de cibler des tumeurs qui bougent avec la respiration en temps réel.
- Table de traitement RoboCouch : Également robotisée avec 6 degrés de liberté.
- Système d'imagerie guidé permettant le positionnement en continu : Il est principalement composé de deux sources de rayons X et deux détecteurs au sol.

Le Cyberknife intègre les scans et des autres données d'imagerie médicale sur sa station de planification de traitement. Ensuite, le radiophysicien (ou le médecin) identifie les caractéristiques de la tumeur et les zones saines à éviter, et le système de planification détermine automatiquement les manipulations des faisceaux à réaliser. Durant le traitement le patient est positionné sur la table qui est automatiquement replacée dans la position de traitement. Le traitement dure de 30 à 90 minutes, durant lesquelles 100 à 200 faisceaux sont administrés dans différentes directions toutes les 10 à 15 secondes. Avant chaque faisceau, le système peut réaliser à la demande une imagerie X et la comparer au CT original, pour guider les faisceaux et corriger les mouvements éventuels du patient et assurer un ciblage précis.

Le système Cyberknife dispose aussi d'un système de suivi qui synchronise le faisceau délivré avec les mouve-

---

6. pour plus d'information : [www.varian.com](http://www.varian.com)

7. La stéréotaxie est une technique utilisée en neurochirurgie pour atteindre des zones du cerveau de manière précise. Avec un système de contention sur le crâne du patient et des moyens d'imagerie médicale, elle assure une meilleure précision tout en étant moins invasive qu'une intervention classique.

8. pour plus d'information : [www accuray.com](http://www accuray.com)

ments éventuels de la tumeur lors de la respiration. Ce système est composé d'un gilet porté par le patient sur lequel sont disposés 3 émetteurs IR (Infra Rouges) et d'un capteur mural fixe permettant de suivre les mouvements respiratoires.



FIGURE 1 – Les systèmes Novalis (à gauche) et Cyberknife (à droite)

Les systèmes Novalis et Cyberknife offrent des fonctionnalités qui répondent à la problématique du positionnement automatisé, mais à un coût très élevé. La composition de leur système de positionnement commence à devenir un standard : un bras et une table robotisés avec six degrés de liberté, plus un système de suivi externe. Leur efficacité a été démontrée, mais ces systèmes ne répondent pas aux besoins que nous cherchons à satisfaire. Ils manquent des informations sur l'environnement, tout le système est centré sur le patient. De plus, leurs fonctionnalités ont été développées spécifiquement pour leur système de traitement, et ne peuvent pas être exportées vers d'autres. Ceci nous a permis d'avoir un premier aperçu de ce qu'on arrive à faire aujourd'hui en matière de positionnement en radiothérapie.

### 2.1.2 La Réalité Augmentée dans les systèmes médicaux

La modélisation 3D ou 4D (c'est-à-dire 3D + le temps) du patient à partir des images TDM (TomoDensitoMétrie) ou IRM (Imagerie par Résonance Magnétique) permettent aujourd'hui de mieux planifier l'acte chirurgical, grâce à une meilleure connaissance du patient avant l'opération. Une copie précise et fiable du patient offre la possibilité de s'entraîner aux gestes opératoires par une simulation personnalisée de l'opération, et donc de réduire sensiblement les erreurs médicales. Une autre utilisation du patient virtuel est le suivi temporel permettant de visualiser plus facilement l'efficacité d'une thérapie. La réalité augmentée propose d'offrir un guidage plus efficace du geste opératoire superposant au patient les informations pré-opératoires, ce qui n'est pas exploité par les simulateurs actuels. La réalité augmentée donne au chirurgien une vue en transparence de son patient qui permettra demain d'obtenir une automatisation des gestes les plus complexes.

[Soler et al., 2004] introduisent ce principe dans leur article pour nous présenter ces travaux appliqués aux opérations chirurgicales de l'appareil digestif. Les rares travaux réalisés sur la région abdominale n'offrent qu'une précision relativement faible dans les informations fournies, du fait des mouvements possibles des organes en fonction de la respiration. Afin de dépasser les limites de l'analyse 3D des images médicales, de la simulation pré-opératoire, et de la réalité augmentée sur les organes et pathologies de l'appareil digestif, les auteurs présentent un ensemble d'outils qui utilisent la bibliothèque ARToolKit (voir en 2.2), dont l'objectif sera de fournir aux chirurgiens de l'appareil digestif une aide aussi bien avant que pendant l'acte chirurgical.

Également, [Krupa et al., 2003] présentent dans un article un système robotique basé vision qui gère et positionne automatiquement des instruments chirurgicaux au cours des interventions chirurgicales robotisées par laparoscopie<sup>9</sup>. Un instrument est installé dans l'extrémité du bras d'un robot chirurgical qui est contrôlée par asservissement visuel. Un système doté de diodes et de fibres optiques, joint à l'extrémité de l'instrument, est utilisé pour projeter des points laser sur la surface des organes. Ces marqueurs optiques sont détectés dans l'image endoscopique et permettent de localiser l'instrument par rapport à la scène. L'instrument est déplacé et centré dans le plan d'image par le biais d'algorithmes d'asservissement visuel en utilisant des points caractéristiques dans l'image. Avec ce système, le chirurgien peut spécifier une position relative souhaitée entre l'instrument et l'organe

9. La laparoscopie est une technique chirurgicale mini-invasive de diagnostic et d'intervention sur la cavité abdominale, de plus en plus utilisée sur l'appareil digestif, en gynécologie, et en urologie. Elle fait partie des techniques d'endoscopie chirurgicale.

visé. Cette approche a été validée avec succès dans un environnement chirurgical réel en réalisant des expériences sur des tissus vivants dans la salle de formation chirurgicale de l'Institut de Recherche sur les Cancers de l'Appareil Digestif (IRCAD) de Strasbourg.

Le développement continu des simulateurs dans un cadre médical apporte de nouvelles idées sur l'utilisation de la Réalité Augmentée dans ce domaine. Le fait de simuler des actions avant d'être réellement exécutées offre un amélioration importante de performance. Mais pour ceci, on doit pourvoir créer un modélisation de la scène à traiter la plus fidèle possible.

## 2.2 Outils existants de suivi par vision

La capture de mouvements (*motion capture* ou mocap) est le procédé qui capte les mouvements d'un élément réel afin d'être enregistrés ou transmis en temps réel vers d'autres systèmes. Elle a initialement été développée dans un cadre médical pour faire correspondre des mouvements d'un être vivant à ceux d'un « avatar » ou d'une représentation conceptuelle. Cette technique est aujourd'hui utilisée pour des films et des émissions de télé avec des décors virtuels, pour certains jeux vidéo et pour l'élaboration d'études scientifiques, ainsi que dans les domaines sportif et artistique.

On distingue trois grands systèmes de capture de mouvements : les systèmes optiques, mécaniques et magnétiques. Nous nous intéressons aux systèmes optiques, et nous présentons ici certain d'entre eux. Pour des questions d'implémentation, nous avons donné la préférence aux algorithmes implémentés dans la bibliothèque OpenCV. Voici donc une sélection de systèmes optiques intéressants pour notre projet.

**Optical Flow** : Le Flux Optique (*Optical Flow*) est le modèle des mouvements apparents des objets, des surfaces, et les arêtes d'une scène visuelle provoquée par le mouvement relatif entre un observateur (un œil ou une caméra) et la scène. Depuis que l'on sait que le mouvement et la production de cartes mentales de la structure de notre environnement sont des composantes essentielles de la vision animale (et humaine), la conversion de cette capacité de l'humain à l'ordinateur est aussi cruciale dans le domaine de la vision par ordinateur. De ce modèle dérive donc une technique de détection des mouvements, appelé également « Flux Optique ». L'application de ce modèle inclut le problème de déduire non seulement le mouvement de l'observateur et des objets dans la scène, mais aussi la structure des objets et de l'environnement.

Trois méthodes sont utilisées dans OpenCV pour l'estimation du flux optique : l'algorithme de [Horn and Schunck, 1980], l'algorithme de [Lucas and Kanade, 1981], et l'estimation par appariement des blocs. L'approche proposée par Horn & Schunck est basée sur un modèle d'observation reliant la luminance dans l'image au vecteur vitesse recherché et sur une régularisation spatiale du champ des vitesses. D'autre part, l'algorithme d'estimation de flux optique de Lucas & Kanade s'intéresse à la différence entre deux images d'une séquence vidéo, sur l'hypothèse de la conservation de l'intensité (ou couleur) d'un pixel ou voxel. Cette intensité ne doit pas changer de façon significative entre deux images, ce qui permet d'établir des équations de contraintes sur les images. Une version itérative de cette méthode, présenté par [Bouguet, 2000], est également implémentée dans la bibliothèque OpenCV. Elle calcule les coordonnées des points caractéristiques (ou points d'intérêt) sur l'image en fonction de ses coordonnées dans l'image précédente. La troisième méthode utilise le principe suivant : pour chaque bloc de l'image précédente, avec une taille de bloc donnée, la méthode cherche un bloc similaire dans un voisinage donné ou dans une position déduite grâce au vecteur déplacement calculé précédemment.

**Camshitf** [Bradski<sup>10</sup>, 1998] : C'est un algorithme de suivi basé couleur développé et utilisé pour le suivi des visages humains. Cet algorithme utilise une technique robuste non paramétrique qui explore les gradients de densité pour trouver le mode<sup>10</sup> (maximums locaux) des probabilités de distributions, appelé algorithme de « Mean Shift ». Dans notre cas, nous voulons trouver le mode d'une distribution de couleurs dans une scène vidéo. Par conséquent, l'algorithme de « Mean Shift » est modifié pour faire face à l'évolution dynamique des probabilités des distributions de couleurs provenant des images d'une séquence vidéo, et il est appelé l'algorithme de « Continuously Adaptive Mean Shift (CAMSHIFT) ». La tolérance de la méthode aux bruits, ces contraintes et les performances ont été étudiés dans l'article.

---

<sup>10</sup>. On appelle mode, une valeur correspondant à la densité de probabilité maximum. Parfois, il peut arriver que la courbe de densité de probabilité présente plusieurs « bosses ». On dit alors qu'elle est multimodale, sinon elle est unimodale

**ARToolKit**<sup>11</sup> : ARToolkit est une bibliothèque qui permet le développement d'applications de RA (Réalité Augmentée). Elle a été conçue pour connaître le point de vue de l'utilisateur. Une application de RA nécessite des informations sur la façon dont l'utilisateur regarde le monde réel. Les fonctions de cette bibliothèque permettent de calculer la position réelle de la caméra et son orientation, par rapport à des marqueurs, en temps réel. ARToolKit inclut certaines caractéristiques :

- Suivi de la position et de l'orientation avec une seule caméra,
- Des algorithmes de suivi qui utilisent des carrés noirs simples comme marqueurs,
- La capacité d'utiliser n'importe quel marqueur carré avec motif,
- Le calibrage de la caméra est simple grâce aux fonctions données,
- Assez rapide pour des applications AR en temps réel,
- Coût d'utilisation très bas,
- Une bibliothèque multiplateforme (Windows, Linux, Mac OS X, SGI),
- OpenSource avec licence GPL pour un usage non commercial.



FIGURE 2 – ARToolKit : des exemples d'application (à gauche) et de motif (à droite)

**Blob Tracking** : Un « blob » représente un amas de pixels segmentés par une méthode quelconque dans une image. En aucun cas il ne s'agit d'un objet, qui dans ce cas est clairement identifié (un visage, une boîte, etc.). Le *blob tracking* est la méthode qui consiste à suivre les déplacements d'un blob. Il s'agit alors d'associer des blobs aux objets en mouvement qu'on cherche à suivre, et de mettre en relation les blobs entre deux images qui se suivent dans une séquence vidéo. La segmentation des blobs dans une image est faite par différentes approches, et elle retourne toujours la position, la taille et le centre des gravité des blobs. Cependant, elle ne garantit pas l'ordre dans lequel les blobs sont indexés. C'est donc à nous d'assurer la cohérence entre les blobs segmentés à l'instant  $t$  et ceux de l'instant  $t - 1$ .

Il y a essentiellement 3 cas à prévoir pour pouvoir reconstruire correctement des mouvements et établir des successions de positions cohérentes pour chaque blob :

- Suivi standard : À partir de la position d'un blob à l'instant  $t - 1$  et du mouvement qu'il a effectué, on calcule une zone où il devrait se trouver au temps  $t$  s'il avait continué son mouvement. Il faut ensuite chercher la meilleure solution en fonction de l'ensemble des zones calculées, car plusieurs nouveaux blobs peuvent se trouver dans une des zones calculées, ou l'inverse, un blob obtenu à l'instant  $t$  peut être présent dans des différents zones calculées.
- Fusion de blobs : On se retrouve dans le cas où un blob à l'instant  $t - 1$  correspond à plusieurs blobs à l'instant  $t$ , ou vice-versa. Ceci arrive quand deux objets (ou plus) que l'on suit s'approchent ou s'éloignent, de manière que la différence des distances à l'instant  $t - 1$  et  $t$  permettent de faire « un saut » dans la segmentation. Le problème majeur est qu'on ne peut pas poser des contraintes sur le nombre des blobs à chercher lors la segmentation, ni la taille des blobs, pour pouvoir associer les blobs de deux images qui se suivent. La solution se trouve dans la continuité des mouvements. En tenant compte de cela, on peut « prédire » la position où devrait se trouver le blob à l'instant suivant. Ensuite, quand les blobs se détachent, c'est la logique et la continuité du mouvement qui prime pour décider auquel objet on associe les blobs.
- Disparition de blob : Un dernier cas à gérer est quand l'objet quitte momentanément le champ de vision de la caméra. La parade est là aussi très simple et intuitive. Il s'agit d'attendre que le blob réapparaisse par le

11. pour plus d'information : [artoolkit.sourceforge.net](http://artoolkit.sourceforge.net)



même côté de l'image où il a disparu avant de continuer à traiter l'objet auquel on a associé ce blob.

Pour que cette méthode soit efficace, la fréquence de capture de la caméra doit être suffisamment élevée. Car ce faisant, le nombre de centres de gravité des blobs dont on dispose pour créer les chemins sera plus important, et le mouvement sera plus facile à suivre. Si ce n'est pas le cas, alors les changements brutaux de direction, provoqueront des erreurs dans la prédiction du mouvement, et tout le travail qui suit en sera fortement perturbé.

**Background Subtraction** : Une approche commune pour l'identification des objets en mouvement est la soustraction de fond (*Background Subtraction*). Cette méthode utilise la différence entre une image d'une séquence vidéo, et un modèle de base de la même séquence vidéo. Il y a de nombreux défis dans le développement des algorithmes de soustraction de fond. Tout d'abord, l'opération doit être robuste contre les changements d'illumination. Deuxièmement, il convient d'éviter la détection dans des scènes avec beaucoup d'éléments en mouvement. Enfin, le modèle du fond doit réagir rapidement aux changements à l'arrière-plan, pour les objets qu'on ne cherche pas à suivre et que l'on considère comme presque stationnaires. Dans la littérature, on retrouve différentes approches avec des techniques simples comme la différence de valeurs entre les images et le filtrage par moyenne, ou des techniques plus sophistiquées utilisant des outils probabilistes.

**Systèmes Infrarouges** : Le rayonnement infrarouge (IR) est un rayonnement électromagnétique d'une longueur d'onde supérieure à celle de la lumière visible. L'infrarouge peut être de la lumière visible ou sous la forme de micro-ondes. Dans sa forme visible, l'infrarouge peut être considéré comme le rouge ou violacé. L'infrarouge est associé à la chaleur car, à température ambiante ordinaire, les objets émettent spontanément des radiations dans le domaine infrarouge. Parce que le rayonnement infrarouge peut être mesuré et observé avec un équipement spécial, il existe aujourd'hui de nombreuses façons d'afficher les objets, dans l'obscurité ou dans des conditions normales. Chaque objet émet généralement un rayonnement infrarouge différent, ce qui en fait des signatures faciles pour repérer des éléments d'une scène.

L'infrarouge est un des systèmes les plus utilisés pour la capture de mouvements optiques. La technique est basée sur la prise de vue synchronisée de plusieurs caméras et l'utilisation de marqueurs avec des propriétés caractéristiques sur l'infrarouge. Le signal est capté directement par les cellules photosensibles des caméras infrarouges et filtré, en même temps, à travers une paroi plate (pas de problème de distorsion d'image) traitée à cet effet. Le regroupement des informations de chaque caméra (2 caméras minimum donc) permet de déterminer la position du marqueur dans l'espace virtuel. La synthèse des coordonnées  $(x, y)$  d'un même objet vu sous différents angle, permet de déduire ses coordonnées spatiales  $(x, y, z)$ . Les informations reçues par les caméras subissent donc de nombreuses manipulations. En revanche, un calibrage correct de tous ces paramètres permet une grande précision des données récoltées.

On retrouve principalement deux types de marqueurs : actif et passif. Pour les marqueurs actifs (micro-LED), c'est le principe « d'émetteur-récepteur » qui est utilisé. Les marqueurs émettent directement un signal infrarouge captés par les cellules photosensibles des caméras. Les caméras utilisées ne sont sensibles qu'à un type de longueur d'onde et visualisent les marqueurs sous forme des taches vidéo blanches (ou en niveau de gris pour les caméras les plus récentes). Ces marqueurs sont donc identifiés automatiquement en temps réel comme étant des points dynamiques avec des coordonnées 3D dans l'espace virtuel, grâce à une séquence d'émission multicomplexe. De l'autre côté, le principe de fonctionnement des marqueurs passifs est analogue à celui des radars : les caméras émettent un rayonnement infrarouge, réfléchi par les marqueurs (dont la surface est composée d'une matière ultra réfléchissante) puis renvoyé à ces mêmes caméras. Comme auparavant, celles-ci ne sont sensibles qu'à un type de longueur d'onde. Ce type de marqueur est le plus utilisé.

### 2.2.1 Comparatif des ces méthodes

Ici, nous présentons un tableau comparatif des différents algorithmes de suivi. Ce tableau est basé sur nos besoins et a été réalisé dans le but de choisir une des méthodes pour l'implémenter dans notre système. Comme l'aspect médical l'impose, il est très important de considérer la précision et la robustesse comme des critères cruciaux pour cette évaluation. Le patient et les machines doivent être localisés à tout moment et avec le moins d'erreur possible, pour permettre un positionnement optimal.

Un autre aspect important à prendre en compte est la complexité. La méthode choisie doit être peu gourmande en ressources, car le système a des autres tâches à exécuter, comme l'acquisition des images, la mise en commun des informations récupérées, le calcul des erreurs, etc. Les algorithmes de suivi par ordinateur sont souvent destinés

à faire partie d'une interface perceptuelle, et doivent être rapides et efficaces du fait de cette interaction avec l'utilisateur. De plus, il y a d'autres processus propres au traitement par radiothérapie qui s'exécutent en parallèle (spécialement pour les systèmes les plus récents), comme la gestion de la dose d'irradiation calculée et la mise en place des faisceaux, mais également le processus robotique de positionnement qui donne aux machines les commandes de déplacement à effectuer pour placer le patient dans la bonne position.

Au niveau du développement, on considère aussi les aspects difficulté d'implémentation et coût des équipements nécessaires. Avec une augmentation croissante du nombre de cas de cancer, la rapidité du développement du système jouera un rôle important. On considère aussi que le prototype que l'on souhaite réaliser sera développé par une seule personne, avec l'aide de l'équipe des chercheurs du LIRIS. D'autre part, les équipements pour le traitement par radiothérapie sont déjà assez coûteux, ce qui rend le traitement moins accessible. Un système à bas coût permettra d'offrir un traitement moins cher, ce qui permettra de traiter un plus grand nombre des patients. En fin, le système sera plus facilement adaptable à d'autres domaines qui nécessitent du positionnement.

On a pris en compte les résultats exposés et des tests qui nous avons réalisés dans des conditions de lumière et d'espace semblables à une salle de radiothérapie. La plupart de ces algorithmes sont implémentés dans la bibliothèque OpenCV, ce qui nous a permis de passer assez rapidement à la phase de test.

Finalement, on a classé toutes les méthodes selon chaque critère. Dans le tableau, nous représentons cette classification par le nombre de signes « + » présents. Plus ce nombre est grand, plus la méthode est intéressante dans ce critère par rapport aux autres méthodes, et plus elle répond à nos besoins. Par exemple, pour le critère « Complexité », la méthode « Background Subtraction » a été classée comme étant la méthode la plus rapide pour fournir une réponse, et la moins gourmande en ressources, et donc la plus intéressante pour notre système.

Méthode	Précision	Robustesse	Complexité	Implémentation	Coût
Optical Flow	+	+	++++	+++++	++++
Camshift	++	+++	++	++	++++
ARToolKit	+++++	++++	+++	+++	+++++
Blob Tracking	++++	+++++	+	+	++++
Background Subtraction	+++	++	+++++	+++++	++++

Le principale avantage de *ARToolKit*, est qu'il permet de récupérer des informations 3D à partir d'une seule caméra. En effet, on peut obtenir non seulement les positions des marqueurs par rapport à la caméra, mais aussi leur orientation. Les autres méthodes proposent un suivi 2D par image, ce qui implique l'utilisation d'au minimum deux caméras pour pouvoir obtenir des informations 3D. Le seul inconvénient de *ARToolKit* sur les autres méthodes est d'être la seule à utiliser des marqueurs. L'utilisation des marqueurs implique la perte du suivi quand le marqueur n'est pas (ou très peu) visible, même si l'objet associé est toujours visible par la caméra. Elle reste cependant la méthode la moins coûteuse du fait du nombre des caméras à utiliser.

Les méthodes de *Camshift* et de *Background Subtraction* ont besoin d'un processus d'apprentissage. Ce qui les rend sensibles aux forts changements de décor et aux bruits, et peu adaptées au suivi de plusieurs objet ou très complexes si on applique la méthode à chaque image autant de fois que le nombre d'objets que l'on suit. L'avantage de la méthode *Optical Flow* est la simplicité de l'algorithme qu'elle utilise.

Les systèmes infrarouges offrent une grande précision et rapidité en comparaison aux méthodes précédentes. Malheureusement, ce type de système est de coût très élevé et nous n'avons pas pu le tester. Dans un milieu radioactif ce matériel peut être facilement endommagé, et son remplacement peut devenir inabordable. C'est pour ceci que nous avons donné la préférence aux webcams, et aux méthodes qui peuvent travailler avec.

## 2.3 Outils en développement liés à la problématique

**Reconstruction d'humanoïde à partir de plusieurs vues** [Michoud et al., 2005] : En utilisant des caméras calibrées et synchronisées, ce travail propose une méthode capable d'acquérir automatiquement un personnage ainsi que sa posture à partir de plusieurs vues et sans marqueurs, en utilisant le principe du « Shape from Silhouette ». L'avantage de cette méthode est qu'elle n'impose aucune contrainte sauf l'équilibre stable de la personne filmée. Il n'y a pas donc des poses « type ». La méthode utilise un modèle à base de squelette articulé (squelette d'animation), et se base sur une approche prédiction/correction. Des mesures anthropométriques sont utilisées pour déterminer les articulations du squelette, suivant un ordre donné qui permet de faire des corrections sur les articulations précédemment retrouvées. Après l'estimation de posture de la personne filmée, la méthode propose une estimation de son enveloppe volumique, représentant les volumes qui décrivent les parties rigides. Ces volumes sont définis par un ensemble de cônes tronqués.

**A provably robust and efficient algorithm for reconstruction using structured light** [Viguier et al., 2007] : Dans le domaine médical, plus spécifiquement dans la radiothérapie, les besoins du suivi des mouvements deviennent de plus en plus grands. Ce rapport de recherche propose une solution robuste à cette demande. Un grand nombre des méthodes existent pour le suivi des individus, mais la plupart sont inadaptées au domaine de la médecine. Ce travail propose de construire la modélisation 3D d'une scène en projetant des points de lumière dans la scène filmée par 2 caméras. Ensuite, on cherche la correspondance entre les points projetés et les points vus par chaque caméra à l'aide d'un graphe d'appariement. La géométrie épipolaire est utilisée pour les calculs. Aucune hypothèse n'est faite sur la continuité ou l'homogénéité des surface observés, ou les occultations éventuelles.

**Using an external registration system for daily patient repositioning in protontherapy** [Pinault et al., 2007] : Cet article propose une méthode itérative pour le positionnement d'un patient en protonthérapie en utilisant des robots industriels et un système de localisation temps réel infrarouge. La principale difficulté dans ce type de manipulation est la connaissance de la position du volume traité au cours de la mise en place du patient. Ces travaux ont permis de réduire le temps du traitement par l'automatisation de la procédure d'étalonnage du système de positionnement et la mise en place du principe multi capteurs. Le processus de positionnement est réalisé par un robot industriel avec 6 degrés de liberté et le système infrarouge NDI Polaris. Les auteurs expliquent comment l'erreur du positionnement du volume traité peut être réduit au minimum grâce à un système de suivi externe.

**Real-time markerless tracking for augmented reality : The virtual visual servoing framework** [Pressigout et al., 2006] : Les principales exigences des systèmes de suivi sont une haute précision et un court temps de latence à un coût raisonnable. Afin de répondre à ces demandes, cet article propose un algorithme de suivi en temps réel basé modèle, pour un système de vision monoculaire par vidéo. Le suivi des objets dans la scène filmée se base sur le calcul de la matrice de transformation entre la caméra et les objets. L'estimation non linéaire de cette matrice est formulée par un approche d'asservissement visuel virtuel. Dans ce contexte, les matrices d'interaction sont indiquées pour les différentes primitives géométriques 3D, y compris des lignes droites, des cercles, des cylindres et sphères. Un suivi des mouvements locaux des bords est utilisé pour fournir le suivi en temps réel de points aux contours de l'objet. La robustesse est obtenue par l'ajout d'un M-estimateur sur la loi de contrôle visuel à travers une pondération itérative, mise en œuvre par moindres carrés. Cette approche est alors étendue au problème de la réalité augmentée sans modèle 3D. La méthode présentée dans cet article a été validée sur plusieurs séquences d'images complexes y compris les espaces extérieurs. Les résultats montrent que la méthode est robuste à l'occultation, les changements d'illumination, et le suivi incomplet ou erroné.

**Real-time tracking of multiple skin-colored objects with a possibly moving camera** [Argyros and Lourakis, 2004] : Cet article présente une méthode pour suivre plusieurs objets par leur couleur dans des images acquises par une caméra, qui elle-même peut être en mouvement. Les objets d'intérêt dans l'article sont des objets de couleur chair, comme les mains et le visage. La méthode proposée englobe un ensemble de techniques qui permettent la modélisation et la détection de ces objets ainsi que leur association temporelle dans les séquences d'images. Les objets sont détectés grâce à un classificateur Bayésien qui est initialisé par apprentissage. Ensuite, un apprentissage interactif hors-ligne est utilisé pour affiner le classificateur en utilisant des images supplémentaires d'apprentissage. Pour faire face aux changements d'illumination, le classificateur est adapté en-ligne en suivant les probabilités de la couleur chair dans les images. Le suivi dans le temps est réalisé grâce à une nouvelle technique qui peut gérer plusieurs objets. Ces objets peuvent se déplacer selon des trajectoires complexes et en cacher ou être caché par d'autres, et aussi dans l'idée des caméras mobiles. Entre autres, le nombre d'objets à suivre peut varier dans le temps. Des résultats expérimentaux représentatifs d'un prototype sur des séquences d'images sont également présentés dans cet article.

**Breathing thorax simulation based on pleura physiology and rib kinematics** [Didier et al., 2007] : Toujours dans la lutte contre le cancer et en répondant au besoin de suivi des tumeurs mobiles, ce travail propose un modèle biomécanique du poumon fondé sur la méthode du « finite helical axis ». Il est basé sur la physiologie

de la plèvre<sup>12</sup>, et utilise les données fournies par des experts médicaux. Le modèle est centré sur la cinématique de la cage thoracique, qui joue un rôle important sur les mouvements de la surface de la plèvre extérieure, et donc sur les mouvements du poumon.

## 2.4 Les méthodes retenues

Le suivi précis du patient ne nous intéresse qu'à partir du moment où ce dernier se trouve sur la table du système de traitement. Nous avons donc décidé d'implémenter une méthode différente pour le suivi global et pour le suivi pendant le traitement. On ne s'intéresse pas à la façon dont le patient bouge ses bras ou ses jambes si le traitement n'a pas encore commencé. On a juste besoin de connaître sa position par rapport aux machines, dans le but de savoir si celles-ci peuvent effectuer des déplacements sécurisés dans une zone donnée.

Les objets seront toujours suivis grâce à l'utilisation de la bibliothèque ARToolKit. Cette bibliothèque a présenté des excellents résultats dans son utilisation dans des systèmes de simulation médicale, par leur précision et simplicité d'utilisation. Pour éviter les problèmes liés à l'occultation des motifs, chaque objet en disposera plusieurs. Ceci permettra aussi d'améliorer les performances du suivi.

Le patient sera suivi dans la partie globale du système par une implémentation de la méthode de BlobTracking. Cette méthode est adaptée à notre système du fait que pendant l'utilisation de la partie globale (en dehors de la durée du traitement), seul le patient est en déplacement. Les objets dans la salle sont quasi-stationnaires en dehors de la durée du traitement. En plus, dans un futur, cette méthode nous permettra de diviser la salle de traitement en fonction des probabilités de présence du patient, du fait que cette méthode enregistre un chemin de déplacement de l'élément suivi. Avec cette information, on pourra restreindre l'espace de recherche du patient dans la salle. Pour la vue de près, nous proposons l'utilisation des travaux introduits par [Viguier et al., 2007], pour son adaptation aux systèmes médicaux.

Nous avons décidé finalement d'inclure la notion de « fichier descripteur » à notre système. Ces fichiers contiendront des informations sur l'objet, comme sa taille, sa forme, sa couleur, etc. Des informations sur les contraintes de l'objet seront aussi incluses dans ces fichiers, comme le nombre de degrés de liberté de déplacement, des contraintes sur sa position (exemples : l'objet est toujours posé sur le sol, il est toujours à une distance donnée d'un autre objet, etc.). Finalement, ces fichiers incluront des informations sur la façon dont l'objet est vu et suivi par le système, comme le nombre de motifs qui lui sont associés et leur position relative à l'objet, la probabilité que l'objet soit entièrement (ou partiellement) visible, etc. En apportant des informations au système par le biais de ces fichiers descripteurs, on peut appliquer des corrections et améliorer la précision du système.

## 3 Développement de la solution

### 3.1 La solution proposée

À partir des objectifs clairs et avec les méthodes retenues dans la recherche bibliographique, nous sommes capables de proposer une solution à la problématique. Nous présentons donc un système composé de 2 parties : une vue globale et une vue de près.

La vue globale permettra de voir la salle tout entière et de suivre globalement le patient et les objets présents. Nous utilisons la méthode « Blobtracking » pour suivre le patient. Pour le suivi des objets, nous proposons l'utilisation des marqueurs en forme de motifs et des fonctions disponibles dans la bibliothèque ARToolKit pour leur détection. Cette vue globale est réalisée par au moins une caméra fixée au plafond. La vue de près permet de suivre avec plus de précision les objets et le patient dans la zone de la salle où se réalise le traitement. Pour le suivi du patient dans la vue de près, nous proposons un système de projection de lumière structurés, composé de deux caméras et d'un vidéo-projecteur. Le système se sert des fichiers au format XML avec des informations décrivant la salle et des données nécessaires pour reconnaître les objets.

### 3.2 Le cahier des charges

Dans le but de démontrer la faisabilité de la solution proposée, nous avons travaillé sur le développement d'un prototype. On a commencé par poser des objectifs et donner des priorités aux différentes phases du développement.

---

12. La plèvre est une membrane séreuse (qui sécrète une liquide organique, analogue à la lymphe) qui revêt les parois du thorax et enveloppe les poumons.

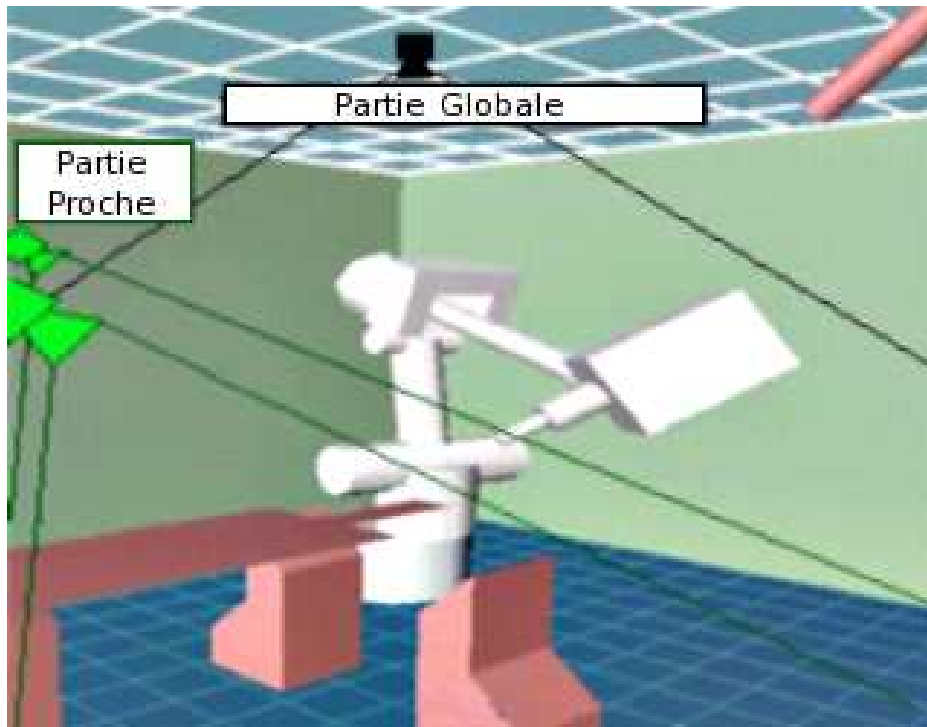


FIGURE 3 – Modelisation possible du système

### 3.2.1 Les besoins pour démarrer

Les besoins pour le développement ont été matériels et logiciels. Dans le laboratoire on dispose d'un grand nombre de ressources, dont nous avons fait un inventaire de ce qui nous était utile :

- matériel : une salle aménagée pour l'acquisition des vidéos, des ordinateurs connectés en réseau, des webcams, des vidéo-projecteurs.
- logiciel : les bibliothèques OpenGL, OpenCV et ARToolKit, des logiciels de réglage de caméras et de capture des images et des vidéos.

On a encore besoin :

- matériel : des marqueurs en forme des motifs.
- logiciel : d'un modèle objet générique, des fonctions permettant la lecture de nos fichiers descripteurs au format xml, d'une interface nous permettant la visualisation des résultats.

À ces derniers, on ajoute les possibles besoins de renouvellement du matériel déjà existant, ou d'acquérir du matériel plus adapté à nos besoins. Mais même avec cette considération, le prix du matériel reste abordable.

### 3.2.2 Les contraintes de l'environnement d'exécution du système

Principalement, il faut prendre en compte le fait que le système doit traiter des personnes malades que l'on souhaite guérir. On se retrouve dans un cadre médical et on a donc les mêmes contraintes, spécialement l'exigence de précision et de robustesse. Il est aussi nécessaire d'avoir un système résistant au milieu radioactif ou, le cas échéant, un système peu coûteux qui soit facilement renouvelable. L'exposition continue peut endommager le système et affecter les résultats.

L'éclairage est un point important dans un traitement à base d'acquisition vidéo. Les salles de radiothérapie ont un éclairage artificiel, du fait qu'elles sont totalement isolées pendant le traitement. Elles sont souvent éclairées par des néons, qui représentent de la lumière non continue. Actuellement en radiothérapie, on dispose de machines volumineuses qui peuvent en cacher facilement d'autres. De même le patient peut cacher, ou être caché par ces machines. On souhaite pouvoir détecter et suivre le patient ou un objet à tout moment, même quand il est temporairement et/ou partiellement caché.

### 3.2.3 Elaboration du planning

La recherche bibliographique a duré un mois. Il nous restait 5 mois pour le développement du prototype, mais une version avancée devait être prête pour une présentation après deux mois et demi de développement. Le temps de préparation de cette présentation, ainsi que la soutenance du stage, et le temps de rédaction du rapport de stage doivent aussi être pris en compte. Il a été prévu de faire des séances de discussion une fois par semaine si possible, pour suivre le développement et traiter les possibles problèmes retrouvés en cours. Des séminaires avec le groupe de recherche du projet ETOILE ont lieu deux fois par mois les vendredi après-midi. Des outils de visualisation des résultats devaient également être développés. Le temps utilisé pour réaliser des tests et du débogage doivent aussi être ajoutés au calcul prévisionnel du développement.

Prenant en compte tous les points antérieurement présentés, le planning prévisionnel suivant a été réalisé. Le temps prévu pour les différentes activités est calculé par semaine « effective », c'est-à-dire, on considère des semaines de cinq jours, avec cinq heures de travail effectif par jour.

Travaux		Temps Prévu
Outils XML		1/2 semaine
Outils Visualisation		1 semaine et 1/2
Le modèle Objet		2 semaines et 1/2
Partie Globale (Objets)	Implémentation ARToolkit	2 semaines et 1/2
	Analyse des résultats	1 semaine
	Tests et débogage	1 semaine
Partie Globale (Patient)	Implémentation Blobtracking	3 semaines
	Analyse des résultats	1 semaine
	Tests et débogage	1 semaine
Préparation de la présentation à l'industriel		1 semaine
Rapport de Stage		3 semaine
Préparation de la soutenance de Stage		1 semaine
Partie Proche (Patient)	Intégration du module [Viguiet et al., 2007]	Temps restant
	Analyse des résultats	
	Tests et débogage	

Ce planning prévisionnel, dont nous présenterons les résultats dans le chapitre 4, couvre largement les cinq mois restant de la durée du stage. Les parties « Analyse des résultats » réfèrent à la conversion des données récupérés par les algorithmes de suivi, à des données exploitables pour notre modèle générique objet. Le temps prévu pour les tests et le débogage peut fortement changer en fonction des problèmes rencontrés. La partie proche ne sera ajoutée au prototype que si le temps le permet, c'est pour ça que nous n'avons pas donné une durée plus précise. Nous signalons finalement que l'ordre donné dans ce planning n'est pas forcément l'ordre chronologique à suivre, et que le temps indiqué est bien le temps total.

## 3.3 Les outils choisis

### 3.3.1 Environnement logiciel

Le langage que nous avons pris pour développer le prototype est le C++, du fait que c'est un langage orienté objet et que les bibliothèques que l'on souhaite utiliser ont été écrites en langage C. De plus, les connaissances de notre équipe sont centrées sur ce langage.

Nous utilisons la bibliothèque OpenCV, qui nous permet principalement l'acquisition des images à partir des caméras, à travers des fonctions très simples à mettre en place. De même, cette bibliothèque propose un grand nombre d'implémentations des algorithmes de traitement et analyses d'images qui sont très utiles. Elle propose spécialement un ensemble de fonctions permettant la mise en place d'un système de suivi par la méthode de BlobTracking.

La bibliothèque OpenGL est utile pour la visualisation des résultats. Cette librairie propose des outils puissants de modélisation et de création des environnements virtuels. Le suivi des marqueurs en forme de motifs sera fait à l'aide de la bibliothèque ARToolkit.

Nous avons finalement développé des outils permettant la fusion de ces bibliothèques, spécialement la conversion des données entre les différents types de variable « Image » utilisées par chaque bibliothèque.

### 3.3.2 Système d'acquisition vidéo

Nos premières sources vidéo sont des webcams conventionnelles. Les bibliothèques OpenCV et ARToolkit sont suffisantes pour l'interaction entre les caméras et l'application. Des fonctions de calibrage sont également disponibles dans ces bibliothèques.

Il reste à mettre en place un système de synchronisation pour l'acquisition d'images à partir de plusieurs caméras, du fait du suivi temps réel.

### 3.3.3 Les méthodes

Nous proposons une méthode de suivi des objets présents dans la salle de traitement, qui consiste à suivre les motifs visibles par les caméras indépendamment de l'objet auquel ils sont associés. Aucun ordre n'est donc donné pour la recherche des motifs. Une fonction de la bibliothèque ARToolKit permet de faire ceci avec un seul appel, et renvoie comme résultat un tableau contenant des informations sur les motifs cherchés dans l'image passée en paramètre. L'idée est ensuite de vérifier pour chaque objet le statut de ces motifs (s'ils sont vus par la caméra ou non, et leur localisation), et de retrouver la position de l'objet dans le repère global en fonction des leur positions.

Pour que cela soit possible l'application doit connaître, au moment de l'initialisation, tous les motifs qui sont susceptibles d'être présents dans la salle, leur objet associé, et les données nécessaires pour modéliser avec précision l'objet à partir de ses motifs. Un des buts des fichiers descripteurs est de contenir toutes ces informations. Chaque motif sera donc une variable à mettre à jour à chaque fois qu'une image est acquise, et à consulter lors de la localisation des objets. Ce principe a été mis en place lors de la création du modèle d'objet générique.

L'avantage de cette méthode est que l'objet pourra être suivi tant qu'au moins un de ses motifs est visible de manière très précise, même s'il est partiellement caché. Le résultat sera toujours plus précis si on dispose de plus d'information, c'est-à-dire, si plusieurs motifs de l'objet sont détectés.

Le suivi globale du patient est fait par la méthode de Blocktracking, et il sert principalement à activer le suivi de près lorsque le patient se trouve dans la zone désirée. La méthode du Blocktracking n'offre pas des informations 3D, mais cela ne pose pas problème. A partir des informations 2D données par la méthode et des informations contenues dans un fichier descripteur du patient, on peut déduire un volume englobant du patient nous permettant de décider si le patient est vu et peut être suivi par la partie de près du système, et donc de manière plus précise.

Dans la conception du stage, il était prévu d'utiliser la méthode présentée par [Viguier et al., 2007] pour le suivi du près. Pour son utilisation dans notre système, nous proposons la méthode suivante. Une fois le patient est dans le champ de vision des caméras de près, de la lumière structurée est projetée sur lui. Ensuite, la méthode est appliquée pour construire la modélisation 3D du patient, que nous complétons en utilisant les informations contenues dans le fichier descripteur du patient.

## 3.4 Création d'un modèle objet générique

Dans l'idée de l'utilisation des fichiers descripteurs, il était nécessaire de définir un modèle ayant la faculté de décrire un objet quelconque. Différents types de description doivent être possibles, car on n'utilise pas les mêmes informations pour suivre un objet et pour le modéliser. Nous avons développé plusieurs versions, et nous présentons la version finale.

Nous avons opté pour une structure en arbre, du fait de son adaptabilité à la représentation des objets divisibles. Ce type de structure permet facilement la sauvegarde des informations en utilisant un langage par balises, comme le langage XML. Nous définissons deux types d'objets : simples et complexes. Les objets simples sont ceux qui peuvent être définis par une seule forme géométrique de base (cube, sphère, etc.), et les objets complexes sont d'objets composés par des objets simples et/ou par des objets complexes. On se rapproche du principe de la géométrie de construction de solides (CSG), car nous utilisons également des opérateurs pour associer les objets.

Pour décrire un objet, nous avons donné une définition par composition des éléments suivant : un repère, des paramètres, des opérateurs et facultativement des motifs. Tout objet a un repère. Il est soit global, si c'est un objet à suivre, soit local s'il fait partie d'un autre objet (et donc défini en fonction du repère de l'objet qui le contient). Les paramètres permettent de définir des propriétés aux objets, comme la couleur, les dimensions, la forme, les statistiques, etc. Les opérateurs ont une double utilité. Ils peuvent combiner plusieurs objets à l'aide d'opérateurs géométriques (union, intersection, etc.), ou associer plusieurs objets par une propriété (ex : distance constante donnée, notion de « posé sur », alignement, etc.). Finalement, selon la configuration d'un objet, des motifs sont associés à un objet (qui peuvent être divisés entre les objets qui le composent). Les motifs ont également un repère local, définit en fonction du repère de l'objet.

Nous avons créé un diagramme UML du modèle contenant les différents types d'éléments précédemment présentés (figure 4).

Nous avons implémenté le modèle par des classes en langage C++. Une description plus détaillée est disponible

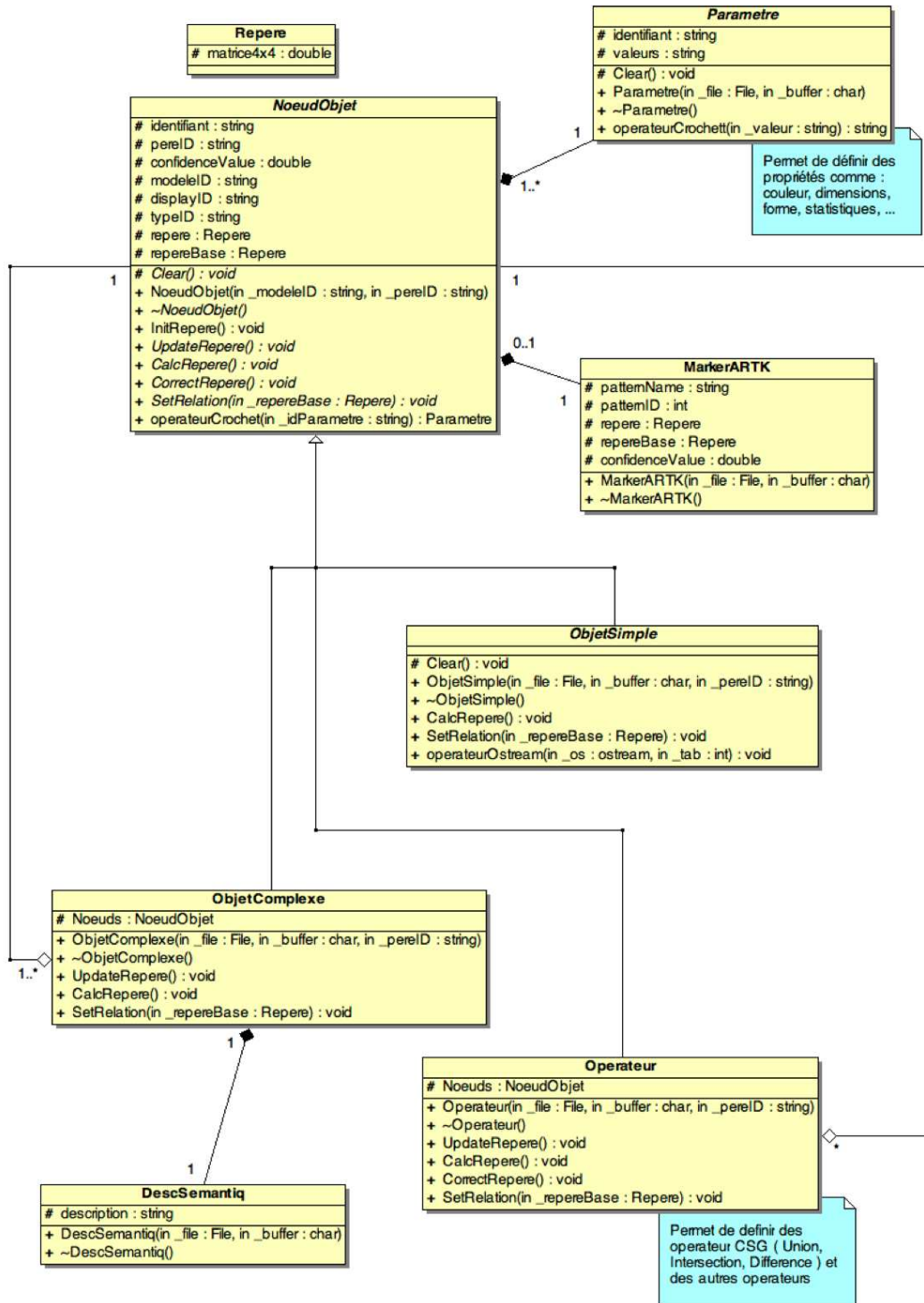


FIGURE 4 – Diagramme UML du modèle

dans la documentation du code faite à l'aide de Doxygen<sup>13</sup>.

Ce modèle a été également conçu de manière à permettre, dans le futur, la création d'une bibliothèque d'objets de base avec laquelle l'utilisateur de l'application sera capable de définir n'importe quelle salle, en ajoutant bien

13. pour plus d'information : [fr.wikipedia.org/wiki/Doxygen](http://fr.wikipedia.org/wiki/Doxygen)



sûr des informations personnalisées.

### 3.5 Les bases de l'application

On définit l'utilisateur de cette application comme étant le radiophysicien (ou le médecin) ayant la formation pour pouvoir pratiquer des traitements par radiothérapie. Aucune notion de programmation, ni de traitement et analyse d'image, ni de manipulation de matériel video ne lui est demandée. L'installation et la configuration du système seront faites par un technicien. L'utilisateur pourra ensuite se servir du système sans avoir besoin de faire appel au technicien à chaque utilisation. Le technicien n'interviendra que pour la maintenance ou en cas de panne ou dysfonctionnement.

#### 3.5.1 Le schéma de l'application

Notre système a été divisé en plusieurs parties pour le développement. Elles peuvent être divisées en trois groupes : « Mise en exécution », « Boucle d'exécution » et « Arrêt ». Dans le premier groupe on retrouve deux parties : « démarrage » et « initialisation ». La partie « démarrage » inclut la mise en place de tout le matériel nécessaire. Les caméras et leur fichiers de configuration, les fichiers descripteurs, des machines et leur motifs associés, et tout autre matériel nécessaire doit être à un emplacement donné et correctement configuré. Ce sont des manipulations que le technicien ou l'utilisateur doivent faire pour pouvoir mettre en marche l'application, et passer à la partie « initialisation ». Dans cette deuxième étape, les fichiers descripteurs et les fichiers de configuration sont lus et interprétés. Ceci permet principalement d'initialiser les variables globales de l'application.

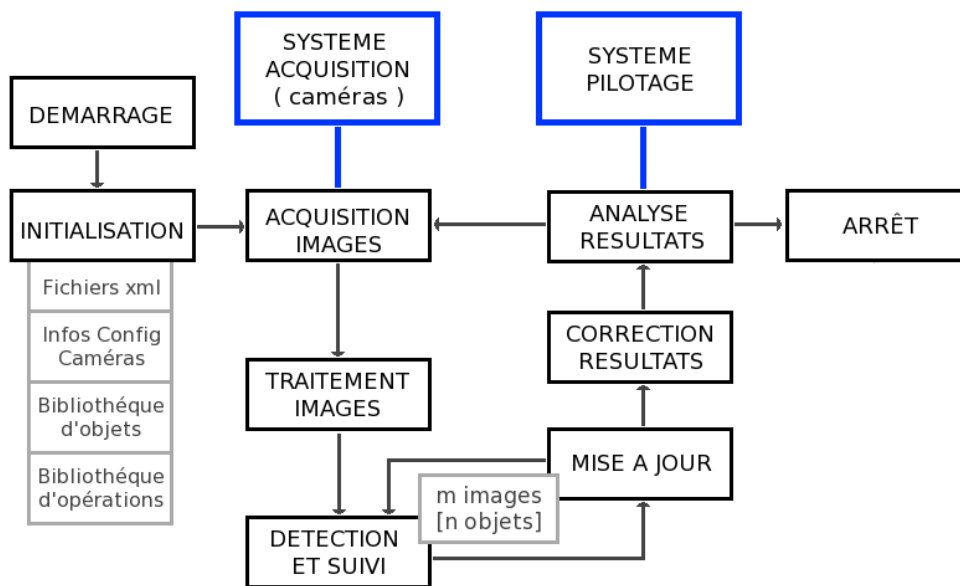


FIGURE 5 – Schéma fonctionnel

Une fois les deux premières parties du système exécutées et validées, on peut exécuter les parties contenues dans le groupe « Boucle d'exécution ». Ce deuxième groupe peut être encore divisé en trois sous-groupes, chacun composé de deux parties : « Vision », « Suivi » et « Contrôle ». Le sous-groupe « Vision » inclut les parties « acquisition d'images » et « traitement d'images ». La partie « acquisition d'images » fait le lien entre le système d'acquisition par vidéo et l'application. Elle est chargée de récupérer les images captées par les caméras, pour ensuite être traitées. La partie « traitement d'images » applique les opérations nécessaires aux images pour qu'elles soient directement exploitables pour les parties suivantes dans la boucle.

Les parties du sous-groupe « Suivi » ont la tâche d'appliquer les algorithmes de suivi sur les images acquises et traitées dans le sous-groupe « Vision ». Une boucle entre ces parties est effectuée, et cette boucle se termine quand tous les objets ont été cherchés dans toutes images. La partie « détection et suivi » applique les algorithmes de suivi sur les images et la partie « mise à jour » se charge de récupérer et mettre en commun les informations obtenues par les algorithmes de suivi.

Le dernier sous-groupe « Contrôle » est en communication avec le système de pilotage de la partie robotique du système de positionnement, et contient les parties « Correction des résultats » et « Analyse des résultats ». La partie « Correction des résultats » utilise les informations et les contraintes interprétées dans les fichiers descripteurs pour détecter des possibles erreurs dans les informations obtenues du sous-groupe « Suivi », et y apporter des corrections. La partie « Analyse des résultats » récupère les informations au dernier stade, et en communication en temps réel avec le système de pilotage, apporte des informations sur l'emplacement actuel des objets et du patient et l'emplacement connu par le système de pilotage. Elle communique également avec la partie « acquisition d'images » pour autoriser un nouveau passage dans la boucle si il n'y a pas eu de demande d'arrêt du système.

Finalement, le groupe « Arrêt », qui ne contient qu'une partie appelée également « Arrêt », a pour but de surveiller le bon arrêt du système. Ceci implique différentes tâches, comme la sauvegarde des dernières configurations du système et de la salle.

### 3.5.2 L'interface Homme-Machine

L'utilisateur communique avec le système à travers les fichiers descripteurs et des périphériques d'entrées standard (clavier et souris). Dans un premier temps, l'utilisateur doit être capable de pouvoir créer des fichiers descripteurs. Ils doivent être donc simples et facilement lisibles et modifiables. Des touches permettront d'activer ou de désactiver des fonctionnalités de l'application.

La visualisation est faite par le biais d'une fenêtre représentant de modélisation 3D de la scène. Deux vues sont possibles : une vue caméra fixe et une vue caméra mobile. La vue caméra fixe permet de visualiser la scène depuis le point de vue d'une des caméras réelles. Elle inclut le principe de la Réalité Augmentée, car on peut regarder les objets virtuels, modélisés « incrustés » dans la vue réelle (image directement acquise depuis la caméra). La vue caméra mobile permet de se déplacer dans la modélisation virtuelle 3D de la scène, et de la visualiser depuis n'importe quelle position et direction.

### 3.5.3 Les données d'initialisation

**Fichiers descripteurs XML :** Ces fichiers contiennent des informations sur les objets (ou sur le patient) respectant la syntaxe XML. Des balises sont prédéfinies pour permettre à l'utilisateur une lecture plus simple et la possibilité d'apporter des modifications sur les fichiers. Les principales balises représentent les différentes classes du modèle objet présenté précédemment, et ces attributs.

**Fichiers de calibrage des caméras :** Les paramètres des calibrage des caméras sont enregistrés dans des fichiers pour éviter de réaliser cette manipulation à chaque utilisation. Ces fichiers sont dans un format défini par la bibliothèque ARTollKit, qui dispose de fonctions pour la lecture et l'utilisation. Cette bibliothèque dispose également de fonctions permettant de créer ce type de fichier, avec un procédé de calibrage.

**Fichiers de définition des motifs :** La bibliothèque ARToolKit offre un vaste choix de motifs prédéfinis, avec des fichiers définissant leur structure. Ces fichiers sont indispensables pour l'utilisation des fonctions de la bibliothèque ARToolKit, car ils contiennent des informations précises permettant aux fonctions de chercher les motifs dans les images. ARToolKit propose également des outils pour définir d'autres motifs, avec des fonctions d'apprentissage qui génèrent le fichier associé au nouveau motif.

## 4 Résultats

Nous avons mis en place la méthode de suivi d'objets présentée en 3.3.3. Plusieurs problèmes, que nous présentons dans la suite, ont ralenti le temps de développement que nous avons prévu pour cette partie du système. La partie du suivi du patient n'a pas pu être donc développée. Nous travaillons actuellement sur l'implémentation du principe du Blob Tracking pour le suivi global du patient.

Les tests ont été réalisés sur un ordinateur Mac portable<sup>14</sup> personnel. L'interface graphique développée a permis de valider de façon visuelle les résultats, et un affichage sur console a permis de réaliser des comparaisons plus détaillées entre les résultats des différents tests. Dans nos tests, nous avons surtout joué avec les facteurs « lumière » et « visibilité » (des objets et motifs).

---

14. Apple MacBook Pro 15" - processeur : Intel Core 2 Duo T7500 2.2 GHz, 2 Go de mémoire vive, webcam Isight intégré.

```

<xml version="1.0" encoding="UTF-8">
<objet_complexe identifiant="Planche" display="union">
  <desc_semantiq>
    Ceci est un exemple de la description d une planche !!!
  </desc_semantiq>
  <opérateur identifiant="UnionTable" display="union">
    <objet_simple identifiant="Planche" display="box">
      <parametre identifiant="couleurRGBA">
        <valeur identifiant="R" vlr="255"/>
        <valeur identifiant="G" vlr="255"/>
        <valeur identifiant="B" vlr="255"/>
        <valeur identifiant="A" vlr="1.0"/>
      </parametre>
      <parametre identifiant="dimension">
        <valeur identifiant="longueur" vlr="670"/>
        <valeur identifiant="largeur" vlr="1000"/>
        <valeur identifiant="hauteur" vlr="10"/>
      </parametre>
      <pattern id="0" name="Data/patt.hiro" width="80">
        <repere>
          1.0 0.0 0.0 185
          0.0 1.0 0.0 350.0
          0.0 0.0 1.0 0.0
          0.0 0.0 0.0 1.0
        </repere>
      </pattern>
      <pattern id="1" name="Data/patt.kanji" width="80">
        <repere>
          1.0 0.0 0.0 -185
          0.0 1.0 0.0 350.0
          0.0 0.0 1.0 0.0
          0.0 0.0 0.0 1.0
        </repere>
      </pattern>
    </objet_simple>
  </opérateur>
</objet_complexe>

```

FIGURE 6 – Exemple de fichier descripteur.

## 4.1 Les premiers résultats

**Le modèle Objet et les fichiers descripteurs** : Notre modèle a démontré par la pratique qu’il peut décrire un objet ou une scène complexe. De même, cette description peut être sauvegardée dans un fichier descripteur au format XML en utilisant la syntaxe donnée par notre modèle. Il reste à valider le modèle pour n’importe quel objet ou scène.

**Localisation des motifs dans le repère global** : On a obtenu des résultats satisfaisants pour la localisation des motifs dans la scène filmée. Les informations données par la bibliothèque ARToolkit sont définies dans le repère de la caméra, et elles sont ensuite transformées au repère global (avec une perte négligeable de l’ordre de 0,001 millimètre) grâce à des matrices de transformation. Ces résultats sont plus liés aux performances de la bibliothèque ARToolkit, notre travail ayant consisté dans leur interprétation et conversion dans un repère global.

**Localisation d’un objet à l’aide des motifs** : Les informations contenues dans les fichiers descripteurs ont effectivement permis de localiser un objet en fonction des motifs associés. Les tests ont montré que la robustesse de cette localisation est proportionnelle au nombre de motifs associés à cet objet. La précision est cependant proportionnelle au nombre de motifs visibles par la caméra.

**Correction des résultats obtenus** : L’imprécision de la localisation a été nettement améliorée par les contraintes données dans la description d’un objet. Des contraintes de type « degrés de liberté » et « distance à un autre objet » ont donné de bons résultats, spécialement pour les objets qui sont toujours posés par terre.

### 4.1.1 Les problèmes rencontrés et leur solutions

Le premier problème rencontré est en relation avec la fusion des bibliothèques OpenCV, ARToolkit et OpenCV. En effet, la mise en commun des structures de données entre les bibliothèques a demandé une étude plus détaillée que nous ne l’avions prévu. La documentation n’était pas assez précise sur la structure interne des variables utilisées par les bibliothèques, et la différence des formats nous a posé encore plus de problèmes. Nous avons appliqué une méthode simple pour déduire leur structure interne. Nous avons fait des tests en modifiant les valeurs des variables afin de visualiser leurs conséquences. Nous avons finalement réussi à faire ces conversions, et pour faciliter la tâche, nous avons opté pour la réalisation d’un module « Image », dans le but d’éloigner ces traitements du principe de l’application.

La qualité et la quantité des informations acquises a été un autre problème à traiter. La bibliothèque ARToolKit évalue, par une valeur entre 0 et 1, la précision de la localisation d'un motif (valeur rendue à chaque appel de la fonction permettant la recherche des motifs - la valeur 0 est donnée quand le motif n'est pas retrouvé dans l'image). Plusieurs cas de figure se présentent, chacun avec un défi à relever.

Le premier cas de figure est quand un seul motif de l'objet est trouvé. La localisation du motifs est précise, mais cependant, elle est n'est pas stable. On peut voir donc « vibrer » la modélisation d'un objet qui ne se déplace pas. Pour remédier ceci, nous avons opté pour la définition des intervalles dans lequel on peut décider si un objet se déplace ou non. Actuellement, nous prenons trois intervalles fixes (un pour chaque coordonnée XYZ) avec des valeurs moyennes des résultats. Cependant, d'après le tests, cette instabilité est accentuée par l'éloignement de l'objet par rapport à la caméra. Les intervalles fixe ne sont plus valables à partir d'une distance donnée par rapport à la caméra. Nous travaillons actuellement pour surpasser ce problème, que nous avons partiellement résolu à l'aide des corrections apportées par les informations des fichiers descripteurs.

Ensuite, on a le cas où plusieurs motifs sont localisés. Du fait des erreurs présentes lors de la localisation, les motifs ne sont pas retrouvés exactement dans les emplacements donnés dans les fichiers descripteurs. De ce fait, il est plus difficile de retrouver la position de l'objet car les motifs ne concordent avec la description donnée. Il nous a fallu trouver un compromis entre les informations fournies par chaque motifs, leur poids de précision de localisation (valeur entre 0 et 1) et la description de l'objet. Nous avons décidé d'utiliser une approche par interpolation, en cherchant d'abord le centre du repère de l'objet dans le repère et ensuite son orientation. Les tests ont démontré qu'une description plus détaillée de l'objet dans son fichier descripteur permet une localisation plus précise de l'objet.

## 4.2 Présentation à l'industriel

Le 8 juillet 2008 nous nous sommes déplacés sur le site du Centre de Protonthérapie de l'Institut Curie à Orsay (CPO), pour présenter une première version de notre prototype à des industriels. Étaient présents : Hanna KAFROUNI, docteur en Physique Médicale et directeur de DOSIsoft, Samuel PINAULT, ingénieur informaticien à l'AS2I, et Régis FERRAND, Physicien et Directeur de projet au CPO.

### 4.2.1 Résultats obtenus

Nous avons commencé cette présentation par l'introduction de l'équipe de travail, et par une série des diapositives rappelant la problématique et expliquant de façon peu détaillée la solution que nous proposons. On a finalisé avec une démonstration du prototype, qui a consisté dans le suivi d'une chaise de la salle de présentation.

Nous avons préparé à l'avance un fichier descripteur contenant uniquement le sol de la scène et une chaise. Avant la présentation, nous avons vérifié que le fichier s'adaptait bien à la salle, et nous avons ajouté les informations manquantes sur la hauteur de la caméra par rapport au sol (l'ordinateur portable était posé sur une table) et les dimensions de la chaise. Nous avons ensuite mis en route l'application et elle s'a exécuté comme prévu : elle suivait bien la chaise. Nous avons présenté les différents affichages de l'application, et la vue en Réalité Augmentée a marché comme voulu.

De plus, cette présentation nous a permis de valider la portabilité du système.

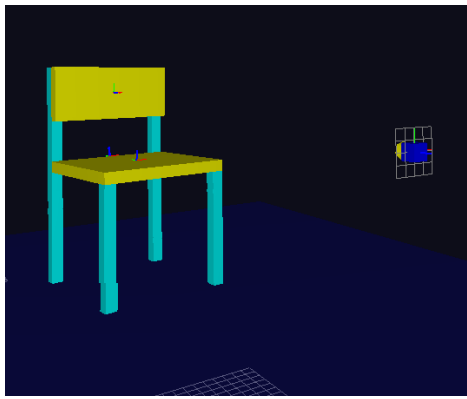
### 4.2.2 La discussion

Une partie de la discussion a commencé au milieu de la démonstration, mais la plus grande partie s'est déroulée sur la fin. Elle a touché à plusieurs sujets, principalement d'ordre technique et académique. Nous avons discuté sur les méthodes et les outils utilisés, ainsi que les problèmes rencontrés. Des questions sur les résultats obtenus au laboratoire ont été également posées. J'étais préparé pour répondre à ces questions, mais malheureusement je n'ai pas su répondre aux questions concernant la quantification exacte du calcul de précision.

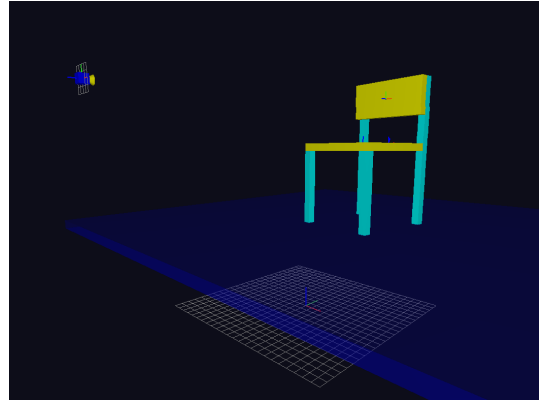
## 5 Perspectives

### 5.1 Améliorations possibles du système

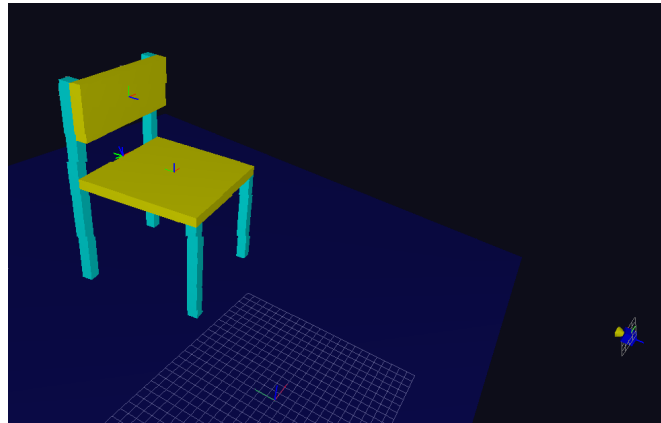
Pas seulement parce qu'il nous a été demandé, mais parce que c'est un outil indispensable pour la validité du système, nous devons inclure le calcul de la précision. Ce calcul se fera sur deux points cruciaux : les résultats



(a) Environnement virtuel - Vue 1



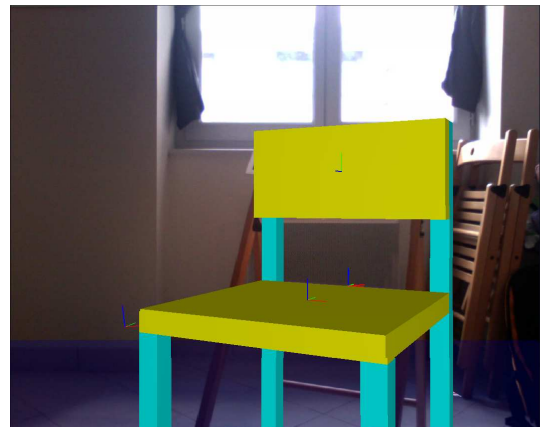
(b) Environnement virtuel - Vue 2



(c) Environnement virtuel - Vue 3



(d) Réalité Augmentée - Vue 1



(e) Réalité Augmentée - Vue 2

FIGURE 7 – Résultats visuels de l'application

obtenus avant leur correction versus les informations contenues dans les fichiers descripteurs, et les résultats obtenus après leur correction (grâce aux fichiers descripteurs) versus les données réelles. Pour le deuxième point, nous devons trouver un moyen de transmettre des données réelles à l'application pour que ce calcul puisse être fait en temps réel.

Il sera également souhaitable d'inclure le principe d'« Inter-séances ». Depuis les débuts de la radiothérapie conformationnelle, le plan de traitement est personnalisé au patient. Les informations référentes au positionnement du patient dans les séances déjà effectuées peuvent être exploitables dans les séances suivantes. De même, une connaissance sur la localisation des objets lors des séances précédentes (pas forcément du même patient) ne peuvent qu'apporter des avantages à l'application. Ceci suggère la mise en place d'un module d'apprentissage des

caractéristiques du patient (poids, mobilité, etc.) et des objets (trajectoires, zones plus fréquentées, etc.).

Finalement, il est évidant que l'utilisation d'une méthode de suivi sans marqueurs (comme la méthode présentée par [Pressigout et al., 2006]) serait un plus dans notre application. Cependant, ce type de méthode n'a pas encore montré de meilleurs résultats que la méthode avec marqueur. Nous avons réfléchi à fusionner un méthode sans marqueur avec notre méthode actuelle, dans le but d'améliorer la robustesse de l'application (pouvoir suivre l'objet même quand tous ses motifs sont cachés).

## 5.2 La suite du travail

Après la présentation face aux industriels, ceux-ci nous ont fait part de leur intérêt pour le développement de notre système. Les modules de suivi externe représentent un grand besoin pour le positionnement en radiothérapie, et le module que nous présentons propose un approche qui n'avait pas encore été traitée, celle du suivi global du système du traitement.

# 6 Conclusion

## 6.1 La solution proposée

Dans ce rapport, nous avons présenté un module de suivi de mouvements et localisation d'un patient par rapport à un ensemble d'équipements dans une salle de traitement par radiothérapie. La particularité de ce type d'application est qu'elle touche à d'autres sujets qui présentent encore des défis pour les chercheurs, comme le suivi précis des mouvements et la modélisation fidèle des humanoïdes.

Du fait des capacités des machines récentes (processeurs multicœurs et cartes graphiques) et du matériel d'acquisition vidéo, les avancées en informatique graphique font penser que prochainement on disposera de méthodes de surveillance vidéo plus précises. Ceci nous encourage à nous investir dans le développement des applications utilisant des systèmes de suivi optique.

Bien que le prototype de l'application n'ait atteint qu'un niveau primitif, elle a permis de valider la faisabilité de la solution proposée. Notre modèle objet est suffisamment générique pour s'adapter aux différentes configurations que l'on peut trouver dans une salle de radiothérapie, et à ses changements éventuels.

## 6.2 Le stage

Après avoir regardé avec plus de détail mon travail, je me sens plus fier de la formation que j'ai choisi et des différents choix que j'ai du faire. Cette expérience professionnelle dans le domaine de l'informatique, qui n'était pas la première (un stage de 3 mois en 3<sup>ème</sup> année, et un en 4<sup>ème</sup> année de 3 semaines), a été très instructif et m'a permis de m'améliorer encore dans des différents aspects.

C'est sûrement par le fait d'être entouré des chercheurs au sein du laboratoire LIRIS et du groupe de recherche d'ETOILE, que ma curiosité et mon « appétit » des nouvelles connaissances ont été considérablement augmentés. Je n'en ai pas seulement acquis des notions dans le domaine de l'informatique, mais aussi dans la physique et le médical, qui ont été très utiles dans le développement de l'application. Bien sûr, mes connaissances avant le stage ont été également d'une grande utilité.

J'ai dû être plus organisé que d'habitude dans mon travail, car plus j'apprenais, plus j'avais des idées et des capacités pour les développer. Au début, j'ai eu du mal à suivre les conseil de mon maître de stage à ce sujet, car je n'avais jamais participé à l'organisation et au développement d'un travail si complexe. Ça ne ressemblait pas aux projets lors de ma formation, où on se contente de suivre les indications du sujet. Ici, les conditions et les besoins changent en fonction des choix et des effort fournis. Lors de la préparation de la présentation du prototype aux industriels, j'ai pu mieux comprendre ces conseils, car j'avais déjà quelques mois d'expérience acquis.

Au LIRIS, beaucoup des permanentes n'ont pas hésité à m'apporter leur aide et j'ai pu également échanger avec des autres stagiaires sur nos connaissances et nos expériences. J'ai ainsi pu observer le fonctionnement d'un laboratoire de recherche, et me rendre compte que même si les apports personnels sont importants, il y a des choses propres à un métier dont on ne peut pas se passer. C'est l'expérience et le contact avec des experts qui nous permettent de les apprendre.

Je suis extrêmement satisfait de ma décision d'étudier en France car elle m'a apporté, non seulement les meilleurs enseignements et un gratifiant échange culturel, mais aussi une facilité d'intégration et d'adaptabilité aux groupes,

et une force d'esprit qui sont des atouts majeurs dans le monde de l'entreprise.

## Références

- [Argyros and Lourakis, 2004] Argyros, A. and Lourakis, M. (2004). Real-time tracking of multiple skin-colored objects with a possibly moving camera. In *ECCV04*, pages Vol III : 368–379.
- [Bouguet, 2000] Bouguet, J.-Y. (2000). Pyramidal implementation of the Lucas Kanade feature tracker.
- [Bradski, 1998] Bradski, G. R. (1998). Computer vision face tracking for use in a perceptual user interface. *Intel Technology Journal*, (Q2) :15.
- [Didier et al., 2007] Didier, A.-L., Villard, P.-F., Bayle, J.-Y., Beuve, M., and Shariat, B. (2007). Breathing thorax simulation based on pleura physiology and rib kinematics. In *MEDIVIS '07 : Proceedings of the International Conference on Medical Information Visualisation - BioMedical Visualisation*, pages 35–42. IEEE Computer Society.
- [Horn and Schunck, 1980] Horn, B. K. and Schunck, B. G. (1980). Determining optical flow. Technical report, Cambridge, MA, USA.
- [Krupa et al., 2003] Krupa, A., Gangloff, J., Doignon, C., de Mathelin, M., Morel, G., Leroy, J., Soler, L., and Marescaux, J. (2003). Autonomous 3-d positioning of surgical instruments in robotized laparoscopic surgery using visual servoing. *IEEE Trans. on Robotics and Automation*, 19(5) :842–853.
- [Lucas and Kanade, 1981] Lucas, B. D. and Kanade, T. (1981). An iterative image registration technique with an application to stereo vision. In *Proceedings of the 1981 DARPA Image Understanding Workshop*, pages 121–130.
- [Mazeron et al., 2007] Mazeron, J.-J., Lang, P., Simon, J.-M., Toubiana, T., Feuvret, L., and Boisserie, G. (2007). Les nouveaux équipements en radiothérapie [new devices in radiation oncology]. In *Bulletin du Cancer*, volume 94, pages 109–113.
- [Michoud et al., 2005] Michoud, B., Guillou, E., and Bouakaz, S. (2005). Reconstruction d'humanoïde à partir de plusieurs vues. In *Journées AFIG 2005*, pages 171–181.
- [Pinault et al., 2007] Pinault, S., Morel, G., Ferrand, R., Auger, M., and Mabit, C. (2007). Using an external registration system for daily patient repositioning in protontherapy.
- [Pressigout et al., 2006] Pressigout, M., Comport, A. I., Marchand, E., and Chaumette, F. (2006). Real-time markerless tracking for augmented reality : The virtual visual servoing framework. *IEEE Transactions on Visualization and Computer Graphics*, 12(4) :615–628.
- [Soler et al., 2004] Soler, L., Ayache, N., Nicolau, S., Pennec, X., Forest, C., Delingette, H., and Marescaux, J. (2004). Traitement d'images médicales pour la planification, la simulation et l'aide intra-opératoire des actes chirurgicaux. *La Revue de l'Electricité et de l'Electronique*, pages 64–71.
- [Viguiet et al., 2007] Viguiet, B., Desserée, E., and Moreau, J.-M. (2007). A provably robust and efficient algorithm for reconstruction using structured light. Technical Report RR-LIRIS-2007-032, LIRIS UMR 5205 CNRS/INSA de Lyon/Université Claude Bernard Lyon 1/Université Lumière Lyon 2/Ecole Centrale de Lyon.