



HAL
open science

Une nouvelle interface avec le monde : Les guides virtuels.

Nehla Ghouaiel, Jean-Marc Cieutat, Jean Pierre Jessel

► To cite this version:

Nehla Ghouaiel, Jean-Marc Cieutat, Jean Pierre Jessel. Une nouvelle interface avec le monde : Les guides virtuels.. Journées de l'Association Française de Réalité Virtuelle, Augmentée et Mixte et d'Interaction 3D, France (2012), Oct 2012, Strasbourg, France. pp.1-5. hal-00739748

HAL Id: hal-00739748

<https://hal.science/hal-00739748>

Submitted on 8 Oct 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Une nouvelle interface avec le monde : Les guides virtuels.

Nehla Ghouaiel*
ESTIA Recherche, IRIT

Jean-Marc Cieutat†
ESTIA Recherche, IRIT

Jean-Pierre Jessel‡
IRIT

ABSTRACT

S'il est vrai que l'homme est devenu sédentaire au cours des âges, son aspiration à parcourir le monde est toujours aussi forte. Le tourisme est ainsi devenu la principale source de revenus de nombreux pays ; les revenus liés à ce secteur sont fortement excédentaires en France et à la base de l'économie de plusieurs de nos régions. À l'aube de l'informatique nomade, nous proposons une nouvelle interface avec le monde : l'homme découvre le monde accompagné de guides virtuels. Deux illustrations sont principalement proposées dans ce papier qui sont, d'une part, la découverte et l'observation de la terre et, d'autre part, la découverte l'héritage culturel et historique d'une ville.

Index Terms: Informatique mobile, Réalité augmentée, Réalité virtuelle, Réalité mixte, Humains virtuels, Recalage hybride.

1 INTRODUCTION

Nous perdons facilement nos repères en territoire inconnue, à l'étranger, à cause de la barrière de la langue bien sûr, mais en grande partie parce que la culture, les us et coutumes, les règlements nous sont moins familiers. Nous préparons alors notre séjour, en nous munissant des ouvrages touristiques les plus réputés, pour retenir les sites à visiter en fonction de nos centres d'intérêts, et choisir les itinéraires en conséquence. Malgré cela, nous demandons souvent de l'aide sur place, et nous finissons par privilégier les circuits animés par des guides touristiques parlant notre langue. Ce n'est qu'à la fin du séjour que nous prenons conscience de tout ce que nous aurions pu découvrir et apprendre. Il est aussi vrai que notre soif de découverte, de culture et de savoir n'a pas de limite.

Une idée originale porte sur la mise en scène de guides touristiques virtuels qui nous accompagnent dans notre découverte du monde. L'intérêt y est multiple. Le guide parle notre langue tout en connaissant les lieux. Il peut traduire les éléments d'une culture à une autre, présenter l'environnement tout en y situant les points historiques et géopolitiques, nous parler d'art et de créativité aussi bien que de dévoiler les traditions gastronomiques locales.

L'idée est intéressante à bien des égards. Le marché actuel le plus en expansion de l'informatique est celui de l'informatique mobile. Tous les jours nous pouvons télécharger de nouvelles applications logicielles. Nos smartphones et nos tablettes sont devenus des concentrés de technologie, et nous transportons avec nous nos ordinateurs d'antan démultipliant les possibilités grâce à l'informatique mobile. De nouveaux usages d'un smartphone apparaissent régulièrement sans surprendre pour autant.

Sur le plan de la communication, pour traduire les éléments d'une culture à une autre, faire appel à un personnage virtuel paraît être un choix judicieux comme la communication est autant une affaire verbale que non verbale. Un personnage virtuel peut communiquer par geste, posture et expression, jusqu'à traduire une

émotion. La lecture sur les lèvres facilite la compréhension, tout particulièrement pour les malentendants.

La plupart des applications de réalité augmentée intègre des entités virtuelles statiques aux images réelles, qui ne varient pas d'une image à l'autre du flux vidéo, dans le but par exemple de documenter la réalité. Ici, au contraire, il s'agit d'intégrer des entités virtuelles dynamiques aux images réelles en y associant des animations. La synthèse des images réelles et des images artificielles, les techniques d'animation et le traitement d'image offrent de multiples possibilités que nous illustrons dans ce papier.

Proposer toutefois des animations convaincantes de personnages virtuels, qui captent l'attention de l'utilisateur et enrichissent l'état de ses connaissances, n'est pas une tâche aisée. Mais l'idée principale, mise en avant dans ce chapitre, est qu'il n'est pas nécessaire pour y parvenir de rechercher un réalisme absolu ; l'augmentation de la capacité des processeurs des systèmes mobiles n'y suffirait pas. C'est d'ailleurs la problématique générale de la Réalité Virtuelle de s'intéresser à la modélisation, aux modèles de simulation permettant de donner vie aux objets et aux êtres virtuels représentés à l'intérieur de ces mondes, avec un degré de crédibilité suffisant pour qu'un utilisateur puisse s'y sentir immergé. À ce propos, certains auteurs, comme Indira Mouttapa Thouvenin [13], opposent la crédibilité du perçu de l'interaction à la quête du réalisme absolu.

Ici, la crédibilité du perçu de l'interactivité avec un personnage virtuel repose sur sa connaissance de l'environnement et de celle de la langue et de la culture de l'utilisateur, pour traduire les éléments d'une culture à une autre. La mise en scène du personnage virtuel se trouve être néanmoins facilitée comme elle se rapporte à un élément visible de l'arrière-plan. Dès lors où notre smartphone est ouvert comme une lucarne sur un environnement devenu surligné, le simple fait d'établir un lien avec l'arrière-plan permet de simplifier le niveau des explications à apporter. On entre dans le domaine de la communication orale où un simple commentaire, une anecdote s'y rapportant, illustre une situation.

Nous développons dans ce papier deux cas d'applications qui sont, d'une part, la découverte et l'observation de la terre et, d'autre part, la découverte de l'héritage culturel et historique d'une ville.

Et de manière plus générale, nous finissons par constater que l'informatique nomade et l'informatique ubiquitaire favorisent l'interactivité entre un personnage virtuel et une scène réelle, comme c'est le cas entre un utilisateur et un personnage virtuel, ou encore entre un utilisateur et une scène réelle augmentée. Cela laisse entrevoir beaucoup de nouvelles perspectives auxquelles nous associons le terme de réalité augmentée interactive.

2 ETAT DE L'ART

2.1 Guide touristique virtuel autonome

La transition d'un guide touristique virtuel à un guide touristique virtuel autonome a été établie par Morgan Veyret [15] dans le cadre du projet ANR SIRENE [4] pour la visite guidée du musée d'Océanopolis de Brest. Le système est capable de construire dynamiquement une visite à partir de sa perception de l'environnement. L'architecture du système proposé est illustrée ci-dessous : Elle repose sur un module de perception visuelle de l'environnement, un module de décision et un module d'animation. C'est l'interaction de ces trois modules de manière asynchrone qui

*e-mail: n.ghouaiel@estia.fr

†e-mail: j.cieutat@estia.fr

‡e-mail: jean-pierre.jessel@irit.fr

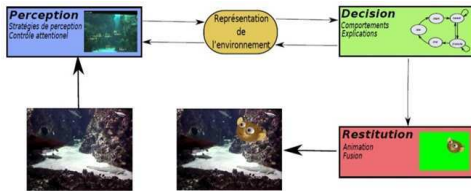


Figure 1: Processus pour la mise en scène d'un guide touristique virtuel autonome.

permet de mettre en scène une visite guidée.

2.2 Agents conversationnels

Les travaux portant sur les agents conversationnels sont remarquables et les illustrations très démonstratives. Parmi les agents conversationnels existants, Greta [1] se distingue tout particulièrement.

Greta est une plate-forme qui permet d'animer des agents conversationnels 3D conformément au type standard d'animation MPEG-4. Avec Greta, un humain virtuel peut communiquer en utilisant une palette riche de comportements verbaux et non verbaux. La communication s'effectue aussi aux moyens d'expressions du visage, de gestes et de mouvements de la tête.

Afin de spécifier les gestes effectuées avec les bras et les ajuster en fonction du contexte de communication, Greta utilise un langage de script se basant sur des contraintes temporelles obtenues à partir d'un synthétiseur de parole. Un geste est caractérisé dans ce langage par sa classe, sa durée en secondes, l'animation de l'un ou des deux bras, la position des bras ou des doigts dans chaque frame.

3 MISE EN SCÈNE D'UN GUIDE TOURISTIQUE VIRTUEL

La Réalité Cliquable est largement commentée dans la littérature. Elle repose sur la métaphore d'interaction de nos smartphones ouverts sur un environnement devenu cliquable. Le champ d'applications de la Réalité Cliquable étant vaste, nous nous situons dans un contexte où l'utilisateur peut sélectionner une augmentation qui lui est proposée et choisir le mode de l'animation qui lui est associé, allant d'un simple commentaire audio à une animation multimédia traditionnelle, jusqu'à une animation en 3D. C'est ce dernier cas de figure qui nous intéresse tout particulièrement pour une visite accompagnée d'un guide touristique virtuel.

À noter, tout d'abord, que le choix des centres d'intérêts et des itinéraires, celui d'un personnage, le contenu des commentaires doivent pouvoir être définis par les professionnels du tourisme, tels que les offices de tourisme, les musées, ... Il y a déjà, dans un premier temps, à concevoir un logiciel adapté à la création du contenu des augmentations par les professionnels du tourisme. Dans un deuxième temps, il s'agit de mettre en scène le personnage virtuel qui va jouer la scène à partir du contenu numérique défini précédemment. Un dossier de projet européen impliquant les professionnels du tourisme a été déposé sur ce point.

Le comportement d'un guide virtuel semble être préprogrammé. Tel un vrai guide, il intervient le long de l'itinéraire, aux abords d'un site touristique, pour en commenter les points remarquables. Mais, pour gagner en crédibilité, nos travaux de recherche portent tout naturellement sur les moyens de procurer une autonomie apparente à un guide virtuel et d'augmenter d'autre part l'interactivité avec la scène.

Pour gagner en autonomie, un guide virtuel doit pouvoir répondre à une question se rapportant à un site touristique et l'illustrer éventuellement par rapport à une référence culturelle compréhensible par l'utilisateur. Un guide virtuel en somme doit passer avec succès le test d'Alan Turing [14]. C'est le premier

point. Nous nous appuyons, pour cela, sur les avancées remarquables liées aux agents conversationnels.

Concernant l'interactivité avec la scène, elle est de divers ordres, comme entre l'utilisateur et le guide virtuel, nous venons d'ailleurs d'en parler, mais également entre l'utilisateur et les augmentations. Pour ce dernier point, il s'agit de s'adapter aux besoins de l'utilisateur et à son niveau de compréhension. Pour cela, on peut être amené à cacher des augmentations jugées inutiles, ajouter de nouvelles augmentations, augmenter ou diminuer le niveau des informations numériques à apporter,

Un mode d'interactivité nouveau est celui à développer entre le guide virtuel et la scène réelle. S'appuyant sur un mode de communication multimodale, un guide virtuel peut joindre une expression à la parole, mais également un geste à la parole [6]. Pour mieux le comprendre, nous pouvons d'ailleurs illustrer nos propos par un exemple. La gare du midi de Biarritz (figure 2), aujourd'hui transformée en une salle de spectacle, était la seule gare de France où les trains arrivaient par le toit de la gare ; les passagers descendaient ensuite au rez-de-chaussée pour sortir de la gare par le hall d'entrée. Pour le mettre en scène, de l'emplacement où se trouve l'utilisateur (figure 2), nous mettons alors en scène un personnage virtuel qui se tourne vers la gare, lève le doigt pour désigner où se situait le terminus des trains, tout expliquant à l'utilisateur la singularité de cette ancienne gare. Pour ce faire, il faut créer un script, par exemple en langage BML de la plate-forme Greta, qui associe une communication non verbale à appropriée la communication verbale du personnage virtuel.



Figure 2: Photo de la gare du midi de Biarritz.

4 DÉCOUVERTE ET OBSERVATION DE LA NATURE

Le compagnon virtuel nous amène en randonnée pour découvrir la nature. L'urbanisation de ces dernières décennies et l'intérêt croissant pour les nouvelles technologies ont conduit l'homme à s'éloigner de la nature et à ne plus connaître son environnement. Couplé par exemple aux résultats du projet ReVeS [3] "REconnaissance de VEgétaux pour des interfaces Smartphones", financé par l'Agence Nationale de la Recherche dans le cadre de l'appel à projet CONTINT 2010, un guide virtuel peut décrire toutes les formes d'arbres et d'arbustes à partir de la simple analyse d'une de leurs feuilles. De nombreux scénarii de ce type sont envisageables pour nous aider à redécouvrir la nature et à comprendre l'importance de la biodiversité.

Dans Jean-Marc Cieutat al. [8], nous avons même exposé l'idée de joindre l'utile à l'agréable, et de profiter des temps de loisirs pour partager nos données environnementales pour l'observation de la terre. Cela permettrait, par exemple, de surveiller l'érosion du littoral, le niveau d'eau de nos lacs, la fonte des glaciers ou encore la déforestation. Une idée similaire a été également proposée par [5] pour surveiller la propagation des mauvaises herbes dans les paysages naturels.

Toute observation objective, toute utilisation à caractère scientifique de l'information, présuppose que les données employées soient précises. Cela est désormais rendu possible grâce au géoréférencement de nos photos. À cet effet, un géotag [12], un marqueur à caractère géographique, est inséré directement dans le

format de l'image. Cette balise de géolocalisation peut être constituée des coordonnées GPS (latitude, longitude, hauteur), de la direction de la vue en coordonnées polaires, de la date et d'une saisie [7]. Citons, par exemple, le format Exif (Exchangeable image file format) établi par le JEIDA (Japan Electronic Industry Development) qui permet d'intégrer automatiquement une balise dans un format image. En l'absence de traitement automatique, il reste toujours également possible de tatouer la balise de géolocalisation directement dans l'image. Cette technique utilise généralement le bit de poids faibles, le moins significatif, des premiers octets de l'image pour insérer directement l'information dans l'image.

Une des illustrations de l'article [8] porte sur un algorithme de morphing intelligent qui rend compte de l'érosion de nos côtes. Les contours entre la mer, la côte et le ciel sont extraits de l'image source I_s et de l'image destination I_d dans un premier temps (figure 3) ; puis, les images intermédiaires I_α de l'animation de morphing sont ensuite calculées par interpolation à partir des coordonnées des contours en appliquant l'équation 1 suivante :

$$I_\alpha = I_s * \alpha + I_d * (1 - \alpha) \quad (1)$$



Figure 3: Exemple de morphing intelligent

De cette manière, l'illusion d'un rendu, qui retranscrit fidèlement le recul de la côte au cours des âges sous l'effet de la mer, est alors parfaite. Un personnage virtuel, aux allures d'un célèbre physicien, vient expliquer la constitution des roches pour justifier ici les deux jumeaux de la falaise d'Hendaye aient si peu changé au cours des âges alors que sur les arrières de la falaise recule (figure 4).



Figure 4: Explication du phénomène naturel qu'est l'érosion par un personnage scientifique.

Le même personnage nous explique encore les phénomènes naturels que sont les marées et les courants, pour nous montrer à quoi ressembleraient nos côtes si, par exemple à la même latitude, elles étaient refroidies par le courant du Labrador au lieu d'être réchauffées par le courant du Gulf Stream. En hiver, les paysages seraient alors recouverts de neige et les lacs et les rivières seraient gelés.

La détection des contours, la segmentation des images en régions sont calculées par les méthodes traditionnelles, comme l'algorithme de Canny pour les contours ou encore comme la classification en régions au moyen de l'analyse spectrale. Dans ce dernier cas, une région peut être constituée de l'union de plusieurs classes de couleurs spatio-colorimétrique (bleu et gris pour le ciel, ...).

5 DÉCOUVERTE DE L'HÉRITAGE CULTUREL ET HISTORIQUE D'UNE VILLE

5.1 Visite guidée interactive

Durant une visite en environnement urbain, le guide virtuel accompagne le visiteur dans sa découverte de l'héritage culturel et

historique d'une ville [11]. Faisant preuve d'autonomie, le guide présente au visiteur les centres d'intérêts les plus proches du lieu où il se trouve, comme les bâtiments historiques ou encore les centres culturels. Il anime une visite en associant le verbal et le non-verbal, pour agrémenter la narration et rendre la visite plus attractive.

La structure de données utilisée pour implémenter la mise en scène d'une visite urbaine par un guide virtuel est un automate. L'automate de la figure 5 illustre une visite composée de trois centres d'intérêts remarquables que l'on appelle des POI (Points d'Intérêt) (POIs). Les états de l'automate représentent les POIs tandis que les arcs représentent les transitions possibles entre les états. Nous avons de plus opté pour l'utilisation de la plate-forme de développements d'agents conversationnels Greta pour animer un guide virtuel.

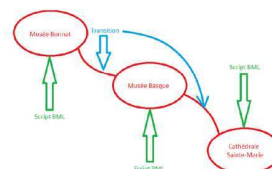


Figure 5: Visite guidée dans le centre-ville de Bayonne.

Dans Greta, la communication non verbale se traduit par les expressions du visage de l'agent conversationnel, ses gestes et les mouvements de son corps. Le corps est un modèle 3D, dont le modèle géométrique s'articule autour d'un squelette. Aux extrémités de chaque partie du squelette, on trouve des nœuds qui sont à la base du mouvement. Pour définir un mouvement, on commence par associer un déplacement à un nœud suivant les trois dimensions de l'espace (en x, en y et en z). C'est le déplacement simultané de plusieurs nœuds du squelette qui déclenche le mouvement d'un personnage virtuel. La coordination du mouvement est assurée grâce à la description de tous les mouvements possibles des nœuds du squelette dans un fichier associé.

Concernant la communication verbale, à chacune des syllabes d'un mot à prononcer est associé un mouvement particulier des lèvres. C'est un synthétiseur vocal de type text to speech qui permet de traduire les mots en paroles. Le texte du discours à prononcer est préalablement mémorisé et une visite complète est composée d'une succession de discours. Dans la plate-forme Greta, à chaque discours correspond un script BML. Un script BML est en fait un script XML qui associe le langage verbal et le langage non-verbal.

La structure générale d'un script BML est la suivante :

```
< Bml >
< Speech >< /speech >
< Mouvements >< /mouvements >
< /Bml >
```

Dans un script, les balises Bml marquent le début et la fin du script. Une balise Speech contient le texte que doit prononcer un guide virtuel. On trouve ensuite les balises de mouvements.

Comme si l'illustration d'un lieu touristique ne suffisait pas, le guide virtuel doit pouvoir également proposer un début de réponse à une question posée par le visiteur. Pour cela, comme le montre la figure 6, on associe à chacun des états de l'automate des temps de visite interactive prévus à cette intention. De la même manière, on pourrait ajouter un temps interactif à une transition d'un état à un autre de l'automate pour justifier le choix d'une visite d'un site touristique plutôt qu'un autre.

L'interaction verbale avec un compagnon virtuel se fait en plusieurs étapes, qui comportent la reconnaissance vocale, le choix et la formulation de la réponse. La reconnaissance vocale est la technologie informatique permettant de dialoguer avec un ordinateur. Elle s'appuie sur différents domaines comme le traitement du

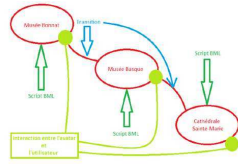


Figure 6: Répondre aux interrogations de l'utilisateur.

signal et l'intelligence artificielle. Après quelques recherches, nous avons opté pour la librairie Voce qui est une librairie de reconnaissance multiplateforme.

Suite à la reconnaissance vocale avec la librairie Voce, une question posée par l'utilisateur est convertie au format BML pour qu'un guide virtuel puisse formuler la réponse. Trouver la réponse à la question posée nécessite de parcourir préalablement un fichier de type AIML. Le format de ce fichier est en fait un dérivé du XML, qui permet de gérer les connaissances associées à un site touristique. Ainsi, dans ce fichier, nous mémorisons les questions susceptibles d'être posées et les réponses s'y rapportant. Bien sûr il est impossible de connaître à l'avance toutes les questions que poseront les utilisateurs. Aussi ce fichier, de taille indéterminée, peut contenir parfois uniquement des mots clés et les quelques associations qui s'y rapportent.

Le guide virtuel, intégré à la scène réelle, joue le rôle d'interface entre l'utilisateur et la scène. En effet, l'utilisateur regardant la scène réelle à travers son smartphone ouvert comme une lucarne sur le monde, voit apparaître sous ses yeux l'humain virtuel. L'impression de coexistence entre la scène réelle et le guide virtuel repose sur plusieurs points. Plusieurs mises en scène sont en effet possibles. Dès lors où le smartphone scrute la scène, le personnage peut être dans une posture narrative : sa position est alors relative par rapport à l'écran, pas nécessairement par rapport à la scène. Si l'on cherche à renforcer le sentiment de présence du guide virtuel dans la scène, tout particulièrement lors d'une interaction avec la scène, il faut alors maintenir la cohérence spatio-temporelle entre le guide virtuel et la scène. L'univers créé devient un univers de réalité mixte. Sur la figure 6, notre guide touristique semble marcher sur les pavés qui conduisent à l'église Sainte-Eugénie de Biarritz tout en décrivant l'architecture toute particulière de cette église. Maintenir la cohérence spatio-temporelle lors d'une posture interactive du guide avec la scène, présuppose que l'utilisateur n'effectue pas de mouvements brusques avec son smartphone. Nous avons fait le choix d'une méthode hybride basée centrale inertielle et basée vision pour effectuer le recalage des augmentations avec la scène. La méthode est décrite dans la section 5.2.



Figure 7: Histoire de l'église Sainte-Eugénie Biarritz.

5.2 Méthode de Recalage

5.2.1 Modèle de contours 3D d'un bâtiment

Le recalage des augmentations s'appuie sur la reconnaissance et le suivi des contours 3D des bâtiments. On parle souvent dans la littérature de méthode de suivi 3D ; il s'agit plus exactement d'une méthode de suivi 2D/3D. Pour cela, il faut préalablement connaître les modèles 3D des contours des bâtiments. La figure ci-dessous montre le modèle de contours 3D du musée basque (figure 8). Notre méthode de reconstruction utilise plusieurs captures d'un bâtiment, prises depuis plusieurs points de vue. Le logiciel Autodesk ImageModeler est utilisé pour construire le modèle 3D à partir de l'appariement de points caractéristiques présents dans les différentes captures.

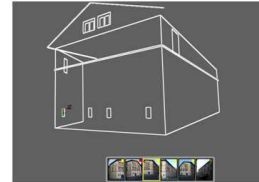


Figure 8: Modèle 3D de contours de musée basque.

5.2.2 Méthode de Recalage

L'algorithme de suivi visuel utilisé, s'inspire des travaux de Georg Klein et de Tom Drummond [10], eux-mêmes inspirés par les travaux de 1990 de Chris Harris et de Carl Stennett et de leur algorithme RAPiD [9]. Dans l'algorithme RAPiD, la pose de la caméra est mise à jour en minimisant l'erreur entre l'objet suivi, qui est un objet de la scène réelle, et la projection du modèle 3D de ce dernier calculée à partir de la pose courante de la caméra dans la scène réelle. L'erreur au niveau de l'arête projetée i est obtenue en calculant la distance perpendiculaire l entre cette dernière et sa correspondante dans l'image :

$$l = d \sin(\beta - \alpha) \quad [9] \quad (2)$$

α : l'angle entre l'arête du modèle et l'axe X.

β : un angle qui peut être $\frac{\pi}{2}, \frac{\pi}{4}, 0$.

d : l'unité de longueur de focale.

La méthode, que nous employons pour le recalage, est une méthode hybride basée centrale inertielle et vision, expliquée précédemment pour la méthode de suivi visuel, afin de combler les imperfections du GPS et les occultations visuelles. Le GPS et les capteurs inertiels permettent d'initialiser la pose de la caméra. Ils se substituent aussi au système de suivi visuel en cas de défaillance de ce dernier. En plus, le suivi inertielle peut remplacer le suivi visuel dans le cas d'une défaillance due à un flou de mouvement. Dans ce cas, la vitesse de mouvement de la caméra dépasse un seuil limite, ce qui rend la détection de contours impossible et par la suite le suivi visuel non fiable, voire erroné. Comme solution à ce problème, on utilise traditionnellement le filtre de kalman ou le filtre de particule pour prédire le mouvement de la caméra.

Notre idée est d'utiliser principalement le gyroscope du smartphone pour déterminer la rotation de la caméra. Déterminer la rotation revient à estimer les trois composantes du vecteur du mouvement $u(u_1, u_2, u_3)$ référencé dans [10].

L'intégration des données fournies par le gyroscope introduit un biais que l'on peut corriger au moyen du filtre complémentaire qui tire bénéfice de l'ensemble des capteurs embarqués dans un smartphone ; en position presque statique, les données fournies par les accéléromètres sont précises à leur tour. Le vecteur de translation T est approximé en utilisant le système GPS [2]. Calculer la translation T revient à calculer la distance entre la position GPS de

l'utilisateur et la position GPS du bâtiment enregistré dans la base de données des POIs.

L'objectif de nos travaux de recherche en recalage hybride est de trouver une méthode de recalage permettant de traiter les éléments de grande taille. Nous supposons que l'utilisateur ne peut regarder qu'une partie de la scène à travers son smartphone, étant donnée la taille des bâtiments. Ainsi, il est indispensable que la méthode de recalage employée puisse gérer le cas de scènes partiellement visibles, contrairement à l'algorithme référencé dans [10] efficace seulement pour des scènes totalement visibles.

6 CONCLUSION

Dans cet article nous explorons les applications de la technologie de la réalité augmentée dans le secteur du tourisme. Nous proposons deux types d'applications qui sont, d'une part, la découverte et l'observation de la terre et, d'autre part, la découverte l'héritage culturel et historique d'une ville.

L'originalité des travaux porte sur la mise en scène de guides touristiques virtuels, qui jouent le rôle d'interface avec le monde réel augmenté, ce qui a pour effet de créer des environnements de réalité mixte qui offrent de nouvelles possibilités.

Sur le plan de la communication, l'idée est intéressante. Cela permet déjà d'associer communication verbale et non verbale pour apporter une meilleure compréhension ; le guide virtuel désigne par exemple un élément de la scène en même temps qu'il explique. Mais cela dépasse ce cadre. Nous proposons un nouveau mode de communication pour traduire des connaissances qui, ici, sont de nature scientifique et culturelle.

L'interactivité du nouveau système proposé a été abordée à plusieurs reprises, au point que nous avons avancé le terme réalité augmentée interactive. Nous avons, à ce sujet, abordé l'interactivité du guide virtuel avec la scène, celle de l'utilisateur avec les augmentations, et enfin celle de l'utilisateur avec le guide virtuel.

Pour renforcer le sentiment de présence du guide virtuel dans la scène, nous avons parlé de posture interactive. Le compagnon virtuel est alors intégré à la scène. Maintenir l'illusion de coexistence de la scène réelle et du guide virtuel présuppose de gérer la cohérence spatio-temporelle de l'ensemble des informations. Pour y parvenir, nous avons proposé une méthode de suivi hybride que nous avons décrit dans la section 5.

7 PERSPECTIVES

L'algorithme de Georg Klein et de Tom Drummond est efficace seulement pour des objets entièrement visibles, de petite taille pour être toujours présents dans le champ de vision. Nous continuons à étendre les résultats prometteurs de notre méthode hybride pour gérer le suivi 3D de bâtiments de très grande taille, partiellement visibles dès que l'on s'en rapproche.

Concernant le dernier mode d'interactivité que nous avons évoqué, l'humain virtuel se doit de répondre aux questions que se pose le visiteur. Pour ce faire, l'application consulte un fichier de type AIML dont le format, nous l'avons vu, est en fait un dérivé du XML. Pour doter l'humain virtuel de plus d'autonomie apparente, l'une des perspectives de notre travail est de construire une base de connaissances qui soit consultable à distance.

REFERENCES

- [1] <http://www.chatbots.org/chatbot/greta>.
- [2] http://fr.wikipedia.org/wiki/transverse_universelle_d'emercator.
- [3] <http://iris.cnrs.fr/reves/>.
- [4] <http://www.anr-sirene.com/>.
- [5] A. S. Australia, P. Ghadirian, and I. D. Bishop. Composition of augmented reality and gis to visualize environmental changes, 2002.
- [6] R. A. Bolt. put-that-there: Voice and gesture at the graphics interface. *SIG-GRAPH Comput. Graph.*, 14(3):262–270, 1980.
- [7] F. Evennou. *Techniques et technologies de localisation avancées pour terminaux mobiles dans les environnements indoor*. PhD thesis, Institut de Microélectronique, Electromagnétisme et Photonique (IMEP), 2007.
- [8] N. Ghouaiel, J.-M. Cieutat, and J.-P. Jessel. Partage de données environnementales pour la découverte et l'observation de la terre. *Actes des 7ème journées francophones Mobilité et Ubiquité*, Juin 2011.
- [9] C. Harris and C. Stennet. Rapid a video rate object tracker. In *In British Machine Vision Conference*, pages 73–77, september 1990.
- [10] G. Klein and T. Drummond. Robust visual tracking for non-instrumented augmented reality. In *Proc. Second IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'03)*, pages 113–122, October 2003.
- [11] J. Park. Evaluation of interaction tools for augmented reality based digital storytelling. In *ICHIT (1)*, pages 263–270, 2012.
- [12] R. C. Pasley, P. D. Clough, and M. Sanderson. Geo-tagging for imprecise regions of different sizes. In *GIR '07: Proceedings of the 4th ACM workshop on Geographical information retrieval*, pages 77–82. ACM, 2007.
- [13] I. M. Thouvenin. *Interaction et connaissance : construction d'une expérience dans le monde virtuel*. PhD thesis, Université de Technologie de Compiègne, Mai 2009.
- [14] A. M. Turing. On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, 42:230–265, 1936.
- [15] M. Veyret. *Un guide virtuel autonome pour la description d'un environnement réel dynamique. – Interaction entre la perception et la prise de décision*. PhD thesis, Université de Bretagne Occidentale, March 2009.