



HAL
open science

Réalité augmentée en extérieur. État de l'art

Imane Zendjebil, Fakhr-Eddine Ababsa, Jean-Yves Didier, Emilie Lalagüe,
Fabrice Declé, Romuald Delmont, Luc Frauciel, Jacques Vairon

► To cite this version:

Imane Zendjebil, Fakhr-Eddine Ababsa, Jean-Yves Didier, Emilie Lalagüe, Fabrice Declé, et al.. Réalité augmentée en extérieur. État de l'art. Revue des Sciences et Technologies de l'Information - Série TSI: Technique et Science Informatiques, 2009, 28 (6-7), pp.857-890. 10.3166/tsi.28.857-890 . hal-00639331

HAL Id: hal-00639331

<https://hal.science/hal-00639331>

Submitted on 28 Feb 2024

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.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Réalité augmentée en extérieur

État de l'art

**Imane Maissa Zendjebil¹ — Fakhreddine Ababsa¹ Jean-Yves
Didier¹ — Emilie Lalagüe² — Fabrice Declé²
Romuald Delmont³ — Luc Frauciel⁴ — Jacques Vairon⁴**

1. Laboratoire IBISC - CNRS FRE 3190 40, rue du Pelvoux, F-91020Evry cedex
2. Projet IPARLA (INRIA - LaBRI) 351, cours de la Libération F-33405 Talence cedex
3. Archivideo 40, rue des Veyettes F-35051 Rennes
4. BRGM - Systèmes et Technologies de l'Information BP 36009 F-45060 Orléans cedex 2

RÉSUMÉ. La réalité augmentée en extérieur est une technologie émergente qui offre à son utilisateur des fonctionnalités de visualisation et d'interrogation de données (locales ou distantes) à l'aide de techniques d'interaction adaptées aux terminaux utilisés. Elle permet aussi de se localiser spatialement dans un environnement extérieur à l'aide d'un système multicapteur. Le but de cet article est de présenter un état de l'art sur les techniques et projets menés dans ce domaine.

ABSTRACT. Outdoor augmented reality is an emerging technology which allows to query and visualize (local or remote) data using interactions suited to the adapted terminals. It also allows to localize itself spatially using multi-sensors systems. The goal of this paper is to survey the various techniques and projects carried out in this field.

MOTS-CLÉS : réalité augmentée en extérieur, localisation 3D, visualisation 3D, interaction 3D.
KEYWORDS: outdoor augmented reality, 3D localization, 3D visualization, 3D interaction.

1. Introduction

La réalité augmentée (RA) est une technologie émergente suscitant un intérêt grandissant auprès des scientifiques et des industriels. Sa capacité à enrichir le monde réel par l'ajout d'entités virtuelles possède un grand potentiel et ouvre ainsi les portes à de nouvelles applications. Depuis son apparition, le potentiel de la réalité augmentée (voir figure 1) a été exploité, entre autres, dans des applications militaires, médicales et industrielles (maintenance ou téléopération) (Azuma *et al.*, 2001).

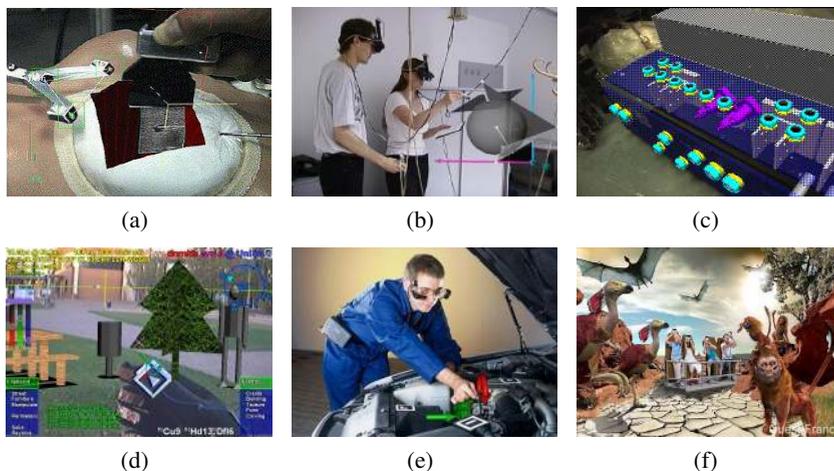


Figure 1. Exemples d'applications de réalité augmentée : (a) Biopsie en RA (State *et al.*, 1996) (b) Construct-3D (Kaufmann, 2002) (c) AMRA (Didier *et al.*, 2005) (d) Tinmith (Piekarski *et al.*, 2001) (e) MATRIS (Bleser *et al.*, 2006a) (f) Les Animaux du Futur (Futuroscope, 2008)

Cependant, dans leur grande majorité, ces systèmes sont dédiés à des environnements intérieurs préparés, restreints et contrôlables, limitant ainsi les problèmes souvent rencontrés en RA (variation de luminosité, gestion des occultations, alignement 3D, etc.). Toutefois, les avancées enregistrées dans les techniques de vision par ordinateur (robustesse, précision, rapidité, etc.), en informatique mobile (performances des terminaux mobiles) et dans les technologies des capteurs (miniaturisation, précision, etc.), ont rendu possible la mise en œuvre d'applications de RA de plus en plus complexes, destinées à des environnements extérieurs où les conditions de travail ne sont plus contrôlables. La réalité augmentée en milieu extérieur présente plusieurs problématiques scientifiques qui font toujours l'objet de recherches intenses, notamment la localisation 3D multicapteur, primordiale pour assurer la cohérence spatio-temporelle de la scène augmentée, la visualisation et l'interaction sur terminaux mobiles en adéquation avec les besoins des utilisateurs et l'exploitation de données géo-référencées distantes.

Cet article présente un tour d'horizon sur le déploiement de la réalité augmentée mobile en environnement extérieur. Dans la prochaine section, nous présentons les différents enjeux de la RA mobile en extérieur ainsi que les différentes problématiques qui en découlent. Par la suite, nous examinons les techniques et technologies spécifiques à ce type d'applications. Puis, nous présentons les principaux projets de recherche dans ce domaine avant de conclure.

2. RA en extérieur : enjeux et problématiques

2.1. Enjeux

Les applications de réalité augmentée ont été longtemps restreintes aux environnements intérieurs. La transposition de cette technologie dans des environnements extérieurs présente de véritables enjeux des points de vue technologique, méthodologique et industriel.

La mise en œuvre d'applications de RA mobile en extérieur offre la possibilité d'explorer la synergie entre différentes technologies, notamment entre les paradigmes de la réalité augmentée et de l'informatique mobile. D'une part, la réalité augmentée, par son principe, offre la possibilité de rompre les frontières entre le monde réel et le monde numérique. D'autre part, l'informatique mobile permet de migrer de l'espace de travail traditionnel (bureau) vers d'autres environnements (en extérieur, par exemple) où l'accès aux données distantes est contraignant.

La réalité augmentée en extérieur trouve son application dans divers domaines tels que le génie civil, les sciences et techniques de l'environnement (la géologie par exemple), la valorisation du patrimoine, le tourisme, etc. La RA mobile en extérieur offre à l'utilisateur un outil d'assistance qui lui permet d'accéder à des bases de données telles que les systèmes d'informations géographiques (SIG) et de représenter la connaissance sous forme visuelle. Ainsi, le géologue peut visualiser, sur le terrain, ses coupes géologiques superposées à la vue réelle du volcan. De même, le géomètre pourra visualiser *in situ* les réseaux d'assainissement modélisés en 3D. De plus, la RA mobile met en avant de nouvelles méthodes d'interaction afin de faciliter la manipulation des données par l'utilisateur final et d'enrichir les bases de connaissances (SIG, modèles CAO, etc.) en confrontant les données à la réalité du terrain et en les mettant à jour en temps réel si nécessaire.

Le déploiement des systèmes de RA en extérieur doit également faire face à des contraintes fortes imposées par le milieu extérieur, notamment les conditions climatiques, la technologie employée qui privilégie des solutions nomades à faible puissance de calcul et l'implantation de réseaux de télécommunication pour accéder à des données distantes et centralisées. La conception de systèmes mobiles de RA en extérieur englobe différentes techniques qui seront détaillées dans la section suivante.

2.2. Problématiques

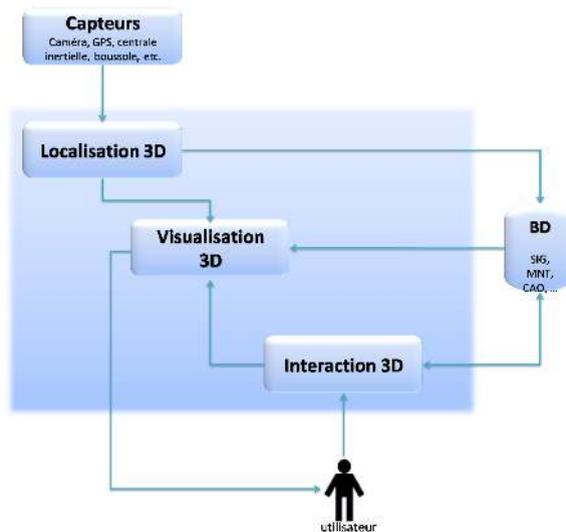


Figure 2. *Système de réalité augmentée mobile : architecture*

La mise en œuvre d'un système de réalité augmentée se focalise sur trois aspects complémentaires : la localisation, la visualisation et l'interaction (voir figure 2). La localisation permet de recalibrer le virtuel sur le réel et d'identifier, à partir des bases de données géo-référencées la donnée (modèle 3D, image, etc.) à visualiser. Or, la mobilité de l'opérateur, dans un environnement non restreint et non préparé, rend difficile le processus de localisation. Les systèmes de RA en extérieur combinent généralement une localisation absolue avec une localisation relative pour situer en permanence l'opérateur. La localisation absolue consiste à déterminer la position de l'opérateur dans l'environnement de travail, grâce à des balises de navigation ou un système GPS. La localisation relative utilise des capteurs de mouvement (accéléromètre, gyroscope, etc.) pour estimer le déplacement relatif de l'opérateur par rapport à une position de référence. La problématique de localisation réside dans la façon de combiner ces deux types de localisation. De plus, le processus de localisation doit faire face aux conditions de travail en extérieur, en particulier aux variations de luminosité, aux changements d'échelle et aux occultations.

La mobilité en milieu extérieur contraint à visualiser les données sur des terminaux, dits mobiles, tels que des tablettes-PC, des PDA ou des téléphones cellulaires. Toutefois, ces dispositifs offrent des capacités limitées de traitement et de stockage, d'où l'intérêt d'adapter les techniques de visualisation existantes ou de mettre au point de nouvelles méthodes pour visualiser en temps réel de grandes masses de données sur ce type de terminaux.

Aussi, les applications de RA doivent offrir la possibilité de manipuler et d'interagir avec les données 3D visualisées. Cette tâche est réalisée au moyen d'interfaces homme-machine (IHM). Dans les terminaux classiques (style PC), les outils d'interaction se basent sur des notions telles que les fenêtres, les icônes, les menus, etc. L'introduction de nouveaux dispositifs de visualisation ainsi que la contrainte de mobilité imposent le développement de nouvelles techniques d'interaction homme/machine mieux adaptées (cf. section 4.2).

3. Localisation 3D

Dans les applications de réalité augmentée, la connaissance de la position et de l'orientation du point de vue de l'utilisateur (ou de la caméra) est nécessaire, car elle permet d'assurer la cohérence spatio-temporelle de la scène augmentée. L'estimation des paramètres de localisation permet de modéliser une caméra virtuelle grâce à laquelle un rendu du monde virtuel est réalisé avec les mêmes caractéristiques que la caméra réelle, permettant ainsi d'aligner correctement les mondes réel et virtuel et de créer la vue dite augmentée.

3.1. Méthodes de localisation basées vision

Les méthodes basées vision sont classiquement utilisées en RA pour estimer la position et l'orientation de la caméra (estimation de la pose). Il existe deux grandes classes de méthodes d'estimation de la pose : les méthodes basées marqueurs et les méthodes sans marqueurs. Nous conseillons au lecteur de se référer à l'état de l'art de (Lepetit *et al.*, 2005) qui présente un panorama détaillé des méthodes basées vision.



Figure 3. Exemple d'utilisation de marqueurs artificiels en extérieur (Piekarski *et al.*, 2002)

Les méthodes basées marqueurs placent des cibles artificielles dans la scène réelle afin de faciliter le suivi visuel et l'estimation de la pose. Ces marqueurs sont particulièrement simples à détecter et contiennent des codes qui permettent de les distinguer.

Leurs positions dans le repère monde sont connues *a priori* et l'estimation de la pose suit le schéma suivant : détection des marqueurs, mise en correspondance 2D/3D, et enfin calcul effectif de la pose de la caméra. Cette classe de méthodes est assez précise. Néanmoins, son inconvénient majeur est qu'il faut disposer les marqueurs dans l'environnement de travail de façon à ce que la caméra puisse toujours en détecter au moins un, ce qui est souvent inadapté pour les environnements en extérieur. Toutefois, quelques applications de RA mobile en extérieur utilisent ces cibles codées. C'est le cas dans (Piekarski *et al.*, 2002) où les auteurs utilisent, pour le module de vision, des marqueurs de la bibliothèque logicielle ARToolKit (Kato *et al.*, 1999) d'une surface d'un mètre carré placés sur les bâtiments (voir figure 3). Ceci permet d'estimer la pose de la caméra lorsque l'utilisateur est proche des constructions.

Les méthodes sans marqueurs ou *markerless* représentent une alternative à l'utilisation des marqueurs artificiels en exploitant les caractéristiques naturelles de la scène réelle (coins, contours, segments de droites, etc.). Généralement, ces approches utilisent des modèles 3D qui constituent une connaissance *a priori* de l'environnement où évolue l'opérateur. Les données 2D extraites de l'image de la scène sont mises en correspondance avec les données 3D (du même type) extraites du modèle.



Figure 4. Exemple d'augmentation d'un environnement extérieur (Comport *et al.*, 2006)

Parmi les approches intéressantes pour des applications en extérieur, (Comport *et al.*, 2006) proposent un algorithme formulé en termes d'asservissement visuel virtuel pour un suivi local de contours. La loi de commande en boucle fermée minimise l'erreur entre la position courante et la position souhaitée de la caméra. L'utilisation d'un algorithme de type M-estimateurs permet de gérer les données aberrantes. Cette approche est robuste face aux occultations partielles et aux changements d'illumination. Elle converge rapidement pour de petits déplacements. L'estimation de la pose reste peu précise mais néanmoins satisfaisante pour la RA. En effet, l'estimation de la pose initiale et du déplacement fournissent des augmentations (voir figure 4), stables et sujettes à de faibles effets de glissement (*jitter*). Une initialisation manuelle est requise pour la première image. Afin d'automatiser cette procédure, les auteurs proposent d'utiliser des descripteurs géométriques locaux tels ceux calculés par la méthode SIFT (Lowe, 2003) associés à des images de références. À partir de l'appariement de plu-

sieurs ensembles de points, la transformation entre l'image courante et les images de références est estimée.

Le suivi planaire décrit dans (Simon *et al.*, 2002) est aussi exploitable en extérieur puisque les sites urbains (par exemple) sont principalement constitués de surfaces planaires. Cette méthode n'emploie pas explicitement un modèle 3D mais exploite la connaissance de la structure de la scène. La méthode permet d'estimer la meilleure homographie qui relie les points appartenant au même plan dans les images successives. Par son principe, la méthode est simple à mettre en œuvre, rapide et précise. Elle reste robuste aux occultations partielles et aux données aberrantes grâce à l'utilisation de l'algorithme RANSAC (Fischler *et al.*, 1981). Toutefois, elle présente un risque de dérive (*drift*) dû à l'accumulation des erreurs d'estimation. (Johansson *et al.*, 2002) exploitent les plans des façades des immeubles pour estimer la pose. Pour chaque construction, ils associent au plan de l'immeuble dominant une image contenant les contours saillants. La position et l'orientation de la caméra sont déduites à partir de l'homographie obtenue en alignant au mieux les images visibles avec les surfaces planaires dans l'image courante. L'approche proposée gère les occultations partielles.

Il est également intéressant d'utiliser des méthodes combinant différents types de primitives telles que les textures (points d'intérêt) et les contours afin de tirer profit de leurs avantages respectifs. En effet, les méthodes basées contours ont l'avantage d'être efficaces et robustes aux changements d'illumination mais cependant elles sont inadaptées aux scènes fortement texturées. En revanche, les méthodes basées textures sont robustes aux occultations mais restent sensibles aux changements d'illumination. Ces deux limitations ont fait émerger de nouvelles approches dites "hybrides" (Masson *et al.*, 2003; Vacchetti *et al.*, 2004; Pressigout *et al.*, 2007). Par exemple, dans (Vacchetti *et al.*, 2004), la pose est obtenue en minimisant simultanément l'erreur de re-projection des points d'intérêt et la distance entre les contours projetés et les contours image. La méthode proposée est plus stable et moins sensible aux dérives par rapport aux méthodes basées contours ou basées textures.

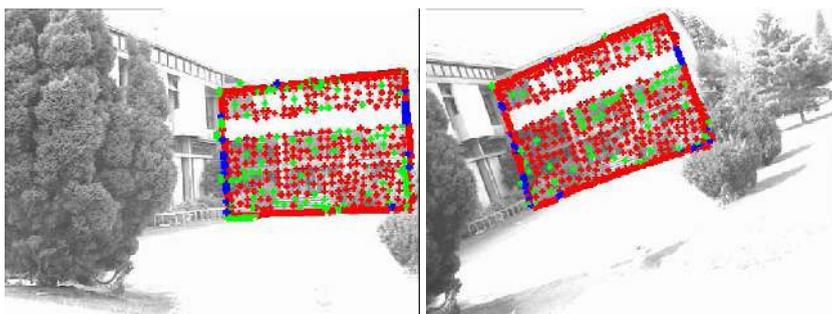


Figure 5. Suivi hybride (Pressigout et al., 2007) : résultats en milieu extérieur

Dans la même perspective, (Pressigout *et al.*, 2007) recouvrent la pose de la caméra en estimant une transformation obtenue à partir d'une minimisation itérative d'un critère qui combine les informations sur la texture et sur les contours de l'objet à suivre. Le critère à minimiser est divisé en deux : un critère de différence de luminosité pour la texture et un critère de distance entre les points des contours (points à fort gradient) extraits de l'image et les points des contours projetés. L'intégration d'une approche d'estimation de mouvement basée textures avec une méthode d'estimation de pose basée contours constitue une méthode de suivi robuste en temps réel. L'utilisation d'un M-estimateur permet d'éliminer les données aberrantes. De plus, le suivi est robuste aux occultations, aux ombres et aux spécularités. Des expérimentations ont été menées sous différentes conditions, y compris en milieu extérieur. La figure 5 présente le suivi d'une façade de bâtiment sous différentes perspectives.

L'utilisation de modèles 3D permet d'améliorer la robustesse et les performances des méthodes de localisation. Toutefois, la construction de modèles précis pour des environnements à grande échelle reste délicate. Pour cela, Stricker (Stricker *et al.*, 2001) propose d'utiliser, dans le cadre du projet ARCHEOGUIDE (Gleue *et al.*, 2001), une base d'images panoramiques de référence (voir figure 6) au lieu d'un modèle 3D. Lors de la phase de suivi, l'image courante est mise en correspondance avec les images de la base en utilisant une technique basée sur la transformée de Fourier. L'image de référence ayant le meilleur score est retenue. La transformation 2D entre cette image et l'image courante est ensuite estimée afin de déduire la pose de la caméra. L'approche est robuste face aux changements de luminosité. De plus, les différents traitements sont effectués avec un nombre fixe d'opération ce qui fait que la méthode s'exécute en temps réel.

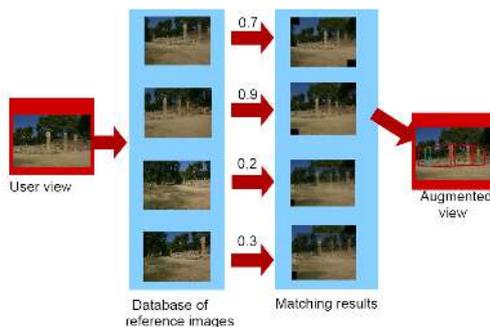


Figure 6. Suivi avec des images de références (Stricker *et al.*, 2001)

De même, les techniques SFM (*Structure From Motion*) (Nistér, 2003) et SLAM (*Simultaneous Localization And Mapping*) (Mouragnon *et al.*, 2006; Mei *et al.*, 2007) utilisées en robotique mobile sont intéressantes car elles permettent de s'affranchir (partiellement ou complètement) de l'utilisation des modèles 3D. Ces approches per-

mettent, simultanément, d'estimer la pose de la caméra et de reconstruire un modèle partiel de l'environnement. Par exemple, (Bleser *et al.*, 2006b) proposent une méthode robuste qui ne nécessite pas une connaissance *a priori* complète de la scène. L'approche proposée combine SFM, SLAM et un suivi basé modèle. Tant que le modèle 3D partiel est visible, la pose est obtenue en reprojetant les lignes du modèle et en estimant les paramètres en alignant au mieux les contours projetés avec les contours extraits de l'image. Lors de la phase de suivi, les mouvements de la caméra ne sont pas restreints à l'observation de la contrepartie réelle du modèle. En effet, la structure de la scène est estimée automatiquement lors du suivi, ce qui permet de compléter le modèle. L'approche utilise un suivi de points d'intérêt basé sur une mise en correspondance d'images. Pour gérer le changement d'échelle, les images sont obtenues à partir d'une pyramide d'images. À partir de la triangulation linéaire, l'estimation de la structure 3D est raffinée récursivement en utilisant un filtre de Kalman étendu. Une mesure de qualité basée sur la contrainte de colinéarité est incorporée au système pour gérer l'influence des caractéristiques sur l'estimation de pose. En effet, si la pose estimée est correcte, la structure est raffinée. La méthode permet de gérer les occultations lors de la phase de suivi.

Les méthodes basées vision fournissent une estimation précise et robuste de la localisation. Elles restent cependant sensibles à diverses contraintes dues à l'environnement extérieur. Afin d'assurer une localisation en continu, les travaux de RA en extérieur convergent vers la combinaison de capteurs de positionnement de natures hétérogènes.

3.2. Méthodes de localisation hybrides

Les travaux récents de RA en extérieur couplent une approche basée vision avec un ou plusieurs autres capteurs. La vision est en effet le capteur privilégié de par sa précision mais elle reste sensible aux conditions qui règnent en extérieur (luminosité, occultations, etc.) et nécessite des temps de traitement relativement élevés. Ceci motive le besoin de l'hybrider avec d'autres capteurs pour gagner en robustesse, précision et rapidité lors de l'estimation de la pose. Cette hybridation permet également de compenser les faiblesses des différents capteurs lorsque ces derniers sont utilisés séparément. Ces systèmes de localisation couplent, généralement, une caméra avec une centrale inertielle, ou une caméra avec un GPS et une centrale inertielle. Le GPS est un système de positionnement par satellites d'une précision de l'ordre de 10 mètres pour les GPS grand public et du mètre pour les GPS professionnels. Toutefois, le GPS est généralement moins précis en environnement urbain à cause des réflexions multiples des ondes sur les façades des immeubles ou de l'occultation des satellites, ce qui le rend peu exploitable dans ces milieux. Les centrales inertielles sont composées d'accéléromètres et de gyroscopes. Les accéléromètres mesurent les accélérations linéaires et les gyroscopes estiment les vitesses angulaires. Les orientations sont obtenues par intégration. La position peut être déduite des accélérations par une double intégration mais ce processus d'intégration tend à accumuler les erreurs et produit des

dérives entre la position estimée et la position réelle. Les premiers systèmes de réalité augmentée en extérieur tels que MARS (Hollerer *et al.*, 1999) (*Mobile Augmented Reality System*) et BARS (Julier *et al.*, 2000) (*Battlefield Augmented Reality System*) utilisaient un GPS pour estimer la position absolue de l'utilisateur ainsi qu'un capteur inertielle couplé avec un compas électronique pour estimer l'orientation.

Il existe deux stratégies d'hybridation afin de localiser le système. La première consiste à fusionner les données des différents capteurs pour fournir une estimation de la position et de l'orientation. Par exemple, (You *et al.*, 1999) ont démontré la faisabilité d'un capteur hybride combinant un système de vision couplé à trois gyroscopes pour estimer l'orientation du point de vue dans un environnement extérieur. Les orientations 3D sont fusionnées avec le mouvement 2D de la caméra selon un modèle de prédiction/correction. Le système estime le mouvement 2D à partir des orientations fournies par les gyroscopes. Le suivi visuel corrige et affine cette estimation par la suite. L'approche proposée se restreint à recouvrir avec précision l'orientation. Afin d'estimer la position, un GPS et des accéléromètres sont aussi utilisés. (Ribo *et al.*, 2002) proposent de combiner un suivi basé points d'intérêt (coins) avec les données d'une centrale inertielle pour recouvrir la position et l'orientation. La fusion des données des deux capteurs est réalisée à l'aide d'un filtre de Kalman étendu. Les données du capteur inertielle permettent de prédire la pose, laquelle est ensuite corrigée par le suivi visuel. Le suivi permet d'estimer la pose en résolvant un problème PnP (Perspectives n-points) à partir d'un appariement 2D/3D entre les coins extraits des images et les points d'un modèle simplifié.

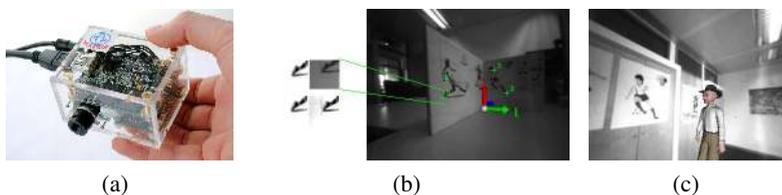


Figure 7. Capteur inertielle/vision pour la localisation (Bleser *et al.*, 2008) : (a) capteur hybride (b) à partir de rendu de l'imagette avec la prédiction, alignement avec l'image courante (c) résultat d'augmentation

Récemment, (Bleser *et al.*, 2008) ont proposé une méthode qui combine un suivi basé texture avec les données d'une centrale inertielle (voir figure 7). À l'aide d'un filtre de Kalman étendu, les auteurs prédisent la pose à partir des données fournies par l'accéléromètre. La pose prédite permet de réaliser un rendu du modèle CAO utilisé. Le modèle contient un ensemble de points 3D auxquels sont associées des imagettes de texture. Le rendu permet d'aligner itérativement ces imagettes dans l'image courante pour estimer le déplacement 2D. Ceci permet de mettre à jour l'estimation donnée par le filtre. Les résultats présentés montrent que la méthode peut fonctionner dans

un environnement à grande échelle. L'inconvénient majeur réside dans la phase de préparation du modèle (associer des imagerie à chaque point 3D).

La deuxième approche d'hybridation, consiste à pallier le capteur défaillant par un autre capteur. Ainsi, (Aron *et al.*, 2007) proposent d'utiliser le capteur inertiel seulement lorsque le suivi visuel échoue. En effet, les méthodes basées vision fournissent une bonne estimation de la pose mais échouent lors d'un mouvement brusque ou en présence d'occultations. Les orientations du capteur inertiel permettent alors de suivre les primitives visuelles en définissant une zone de recherche pour effectuer l'appariement. À partir des ensembles d'appariement, une homographie est estimée afin de déduire la pose de la caméra. Les erreurs du capteur inertiel sont prises en compte afin d'optimiser la zone de recherche des primitives. La méthode présente de bons résultats en milieu extérieur (voir figure 23). Cependant, le capteur inertiel n'est utilisé que pour estimer un mouvement de rotation. L'ajout d'un capteur pour estimer la position reste primordial. Nous retrouvons le même principe dans les travaux de (Maidi, 2007) qui estime la position et l'orientation de la caméra à partir des données d'une centrale inertielle lorsque les données visuelles sont incomplètes.



Figure 8. (a) Appariement lors d'un mouvement brusque (b) Résultat en milieu extérieur (Aron et al., 2007)

(Behringer, 1999) propose de recouvrir l'orientation de la caméra en exploitant la forme de la ligne d'horizon. À partir de la position fournie par un GPS, une ligne d'horizon à 360 degrés est calculée en utilisant le modèle numérique de terrain (MNT), puis elle est mise en correspondance avec les contours extraits de l'image. Cette méthode est intéressante pour les applications dans des sites panoramiques (désert, montagne, etc.). Pour estimer l'orientation, des mesures gyroscopiques peuvent être utilisées. Ceci facilite la mise en correspondance avec les contours extraits des images et corrige par la suite les données de positionnement.

Le système décrit par (Reitmayr *et al.*, 2006) combine un suivi visuel basé contours avec des mesures inertielles (vitesses angulaires, accélérations linéaires, champs magnétiques) (voir figure 9). Le suivi visuel est utilisé pour avoir une estimation précise de la localisation alors que le capteur inertiel permet de compenser les erreurs dues aux mouvements brusques et aux occultations. Les mesures de gravité et du champ

magnétique sont utilisées pour limiter le problème de dérive. Le gyroscope permet de réinitialiser automatiquement le processus de suivi. Les différentes mesures fournies par les deux capteurs sont fusionnées à l'aide d'un filtre de Kalman étendu en utilisant un modèle de vitesse constante. Cette combinaison améliore sensiblement la précision et la robustesse des approches de suivi visuel dans des environnements urbains. Concernant la position, (Reitmayr *et al.*, 2007) proposent d'utiliser un GPS pour fournir une approximation de la position 2D (la longitude et la latitude, l'altitude est égale à la taille moyenne de l'utilisateur). Ainsi, pour initialiser la procédure de suivi, les auteurs définissent une zone de recherche représentée par une ellipse d'axes égaux à 3 fois les variations de l'erreur autour de la position estimée par le GPS. La position de la caméra est estimée itérativement. Pour prédire l'erreur introduite par le GPS, cette dernière est modélisée par un processus gaussien construit à partir d'un ensemble de données d'apprentissage. Pour rendre le modèle dynamique, les auteurs proposent d'estimer en ligne la variance et la covariance des mesures.

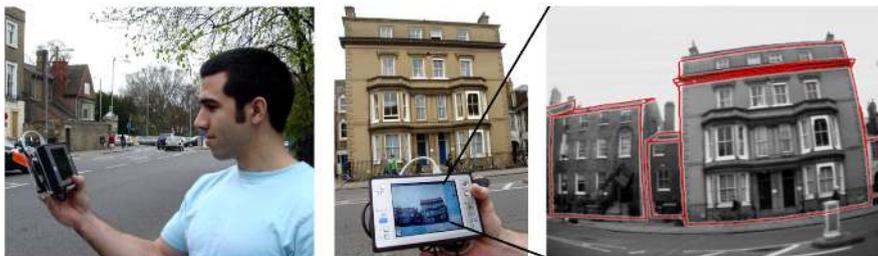


Figure 9. *Suivi de contours basé inertiel/vision (Reitmayr et al., 2006)*

L'hybridation de capteurs soulève une problématique essentielle qui concerne la calibration. En effet, chaque capteur fournit des mesures exprimées dans un système de coordonnées qui lui est propre. De plus, le processus de localisation fournit la position et l'orientation de la caméra par rapport au repère associé à la scène. De ce fait, nous avons besoin d'estimer les différentes transformations qui permettent d'aligner les différents repères au repère associé à la caméra. La procédure de calibration est importante et délicate. La précision de l'hybridation dépend de la précision de cette procédure.

Parmi les approches existantes, (Reitmayr *et al.*, 2006) subdivisent la transformation reliant la caméra à la centrale inertielle en deux rotations : une rotation entre le repère local de la centrale inertielle et le repère associé à la caméra et une rotation entre le repère monde et le système global de coordonnées de la centrale inertielle. Ces deux rotations sont estimées en utilisant la méthode décrite dans (Baillot *et al.*, 2003). La rotation entre le repère caméra et le repère associé à la centrale inertielle peut également être déduite à partir de la relation liant la vitesse angulaire des points de la scène exprimée dans le repère caméra avec celle exprimée dans le repère de la centrale inertielle (You *et al.*, 1999). Le mouvement angulaire par rapport au repère de la centrale

inertielle est fourni par le capteur inertielle lui-même. (Alves *et al.*, 2004) déterminent la rotation entre les deux référentiels en déplaçant le capteur hybride et en mesurant le champ de gravité (fourni par les accéléromètres) qui constitue la référence verticale. Le sens vertical du repère de la caméra est donné par un ensemble d'arêtes verticales parallèles. Ces deux séries de données permettent d'estimer la rotation entre les deux repères. Pour aligner les repères de la caméra et de la centrale inertielle, (Aron *et al.*, 2007) estiment l'orientation de la caméra à partir de l'orientation donnée par la centrale inertielle. Cette relation est exprimée sous la forme $AX = XB$. Aron et al. utilisent la méthode décrite par (Park *et al.*, 1994) pour résoudre ce système. Enfin, (Maidi *et al.*, 2005) montent un capteur hybride composé d'une caméra et une centrale inertielle sur un robot à 6 degrés de liberté (de type FANUC LR-Mate 200i) pour exploiter la précision du positionnement du robot dans le processus de calibration. La transformation qui fait correspondre le repère associé à la centrale inertielle avec le repère caméra est composée de trois rotations élémentaires. Le robot permet de définir deux orientations ce qui permet de déduire la troisième.

La phase de calibration est cruciale car la précision de la localisation hybride dépend de la précision de l'estimation des transformations. Or, les approches de calibration proposées sont souvent basées sur des hypothèses simplificatrices ce qui limite leur utilisation.

Les approches basées multicapteur représentent aujourd'hui la tendance majeure des travaux de RA en extérieur. Elles combinent les méthodes de vision qui sont précises mais moins robustes aux conditions de travail, avec des capteurs à la précision moindre et à la robustesse améliorée. Cette complémentarité permet de fournir une estimation continue de la position et de l'orientation. Actuellement, les approches de vision convergent vers des approches utilisant plusieurs informations visuelles et tendant à être indépendantes des informations *a priori* sur la scène ce qui donne plus de liberté de mouvement.

4. Visualisation et interaction 3D

4.1. Visualisation 3D

Dans cette section, nous allons expliquer comment la génération des éléments additionnels est réalisée. Le défi ici est d'effectuer le rendu graphique des éléments virtuels de manière à ce qu'ils apparaissent comme appartenant au monde réel. Pour parvenir à ce mélange idéal, il faut diriger son attention sur la cohérence au niveau des profondeurs (l'objet le plus proche doit masquer un objet plus éloigné) ainsi que la cohérence au niveau des éclaircissements. Pour les systèmes de réalité augmentée à base de vidéo, il est également important de prendre en compte la cohérence photométrique. Nous décrivons ici comment ces trois cohérences peuvent être assurées dans une application de réalité augmentée, puis, dans la dernière sous-section, comment ces techniques sont réalisées pour fonctionner sur des terminaux mobiles dont les capacités sont souvent inférieures à celles des ordinateurs de bureau.

4.1.1. *Cohérence des profondeurs*

Lorsque l'on porte son attention sur la cohérence des profondeurs, on va désirer gérer les occultations. Ce phénomène est réglé depuis longtemps en réalité virtuelle où les mondes numériques sont entièrement connus (la méthode la plus répandue est celle du z-buffer). En réalité augmentée, la tâche est plus difficile car les profondeurs des objets réels ne sont pas connues *a priori*.

La technique la plus largement utilisée est celle décrite dans (Breen *et al.*, 1996) : les éléments virtuels cachés par des éléments réels sont rendus en utilisant la couleur noir, permettant une fusion de la vidéo du monde réel et des éléments virtuels plus facile. Deux solutions existent pour connaître quelles sont les parties à rendre en noir.

La première consiste à aligner un modèle numérique 3D de l'environnement réel sur l'environnement lui-même, et de calculer les occultations après cet alignement. Cette partie peut être effectuée en même temps que la localisation puisque cette dernière peut employer elle aussi, selon les techniques qui ont été exposées plus tôt dans le document, le modèle 3D de l'environnement.

La seconde technique utilise des algorithmes de vision par ordinateur : les profondeurs des objets réels sont calculées en utilisant les flux vidéos provenant d'une caméra (technique décrite dans (Breen *et al.*, 1996)) ou deux (cas de la stéréovision exposé dans (Kanbara *et al.*, 2000)).

(Fischer *et al.*, 2003) proposent une méthode permettant de gérer l'occlusion des objets virtuels par des éléments réels dynamiques. Cette méthode utilise un modèle géométrique de la scène réelle qu'il faut acquérir avant d'utiliser l'application. L'algorithme détecte les différences entre le modèle acquis et l'image : ces dernières mettent en évidence des objets dynamiques qui ne doivent pas être occultés par les objets virtuels. Cet algorithme fonctionne en temps réel mais peut introduire des incohérences car il ne tient pas compte de la profondeur des objets réels dynamiques.

4.1.2. *Cohérence de l'éclairément*

Le but de cette section est d'assurer les interactions lumineuses entre les objets réels et virtuels de la manière la plus réaliste possible. L'éclairément étant une tâche complexe en réalité virtuelle où tous les paramètres sont connus, elle est d'autant plus compliquée à réaliser en réalité augmentée où les informations nécessaires ne sont que partiellement disponibles. Le problème peut être scindé en deux parties : l'éclairément des objets virtuels par des lumières réelles et l'éclairément des objets réels par des lumières réelles.

4.1.2.1. L'éclairément des objets virtuels par des lumières réelles

Ce problème peut être résolu selon deux principales méthodes. La première méthode consiste à retrouver les paramètres des lumières réelles grâce à diverses techniques. La méthode de (Nakamae *et al.*, 1986) ne marche que pour des scènes extérieures pour lesquelles, grâce à l'heure, la date, et la position et l'orientation de la

caméra, la position du soleil est déterminée. Puis l'intensité lumineuse du soleil est déterminée, ainsi que l'intensité d'une lumière ambiante. Enfin, les éléments virtuels sont rajoutés à la scène.

(Sato *et al.*, 1999) proposent d'estimer les sources de lumière entrant en jeu dans une scène en analysant les ombres projetées par un objet connu, sur une surface connue, puis d'utiliser cette estimation pour calculer l'éclairage des objets virtuels. (Kanbara *et al.*, 2004) ont conçu une méthode qui estime en temps réel l'éclairage d'une scène en utilisant une sphère réfléchissante devant se situer dans le champ de la caméra en permanence : l'éclairage de la scène est estimé grâce à la réflexion dans la sphère. (Jensen *et al.*, 2006) proposent une méthode pour mettre à jour en temps réel les interactions lumineuses en extérieur. La première étape consiste à calculer les paramètres de l'éclairage. Les auteurs font l'hypothèse que le soleil est la principale source de lumière. Ils peuvent estimer les paramètres de l'éclairage (lumière provenant du soleil ou réfléchi par les surfaces de la scène) à partir d'un modèle 3D de la scène, de surfaces définies comme étant diffuses et de la vidéo de la scène. On peut alors rendre l'objet de synthèse en utilisant les paramètres calculés précédemment et le modèle de (Phong, 1975). Il faut cependant prendre en compte les ombres dessinées par l'objet, et celles dessinées sur l'objet. Pour cela, les auteurs utilisent l'algorithme des *Shadow Volumes* de (Crow, 1977) modifié pour que les objets virtuels ne produisent une ombre qu'aux endroits où il n'y en a pas déjà et que les objets réels ne produisent une ombre que sur les objets virtuels, puisque celles créées sur les objets réels est déjà présente. Pour améliorer le rendu, on se sert d'un *Irradiance Volume* (Greger *et al.*, 1998) obtenu à partir d'un *environment map* HDR. Les auteurs se servent des accélérations graphiques et obtiennent un taux de rafraîchissement de 10 images par seconde. Un résultat de l'algorithme est visible sur la figure 10.



(a) Situation initiale (la statue est virtuelle) (b) Sans mise à jour des paramètres (c) Avec mise à jour dynamique des paramètres

Figure 10. Résultat de l'algorithme de Jensen

Une autre méthode permettant d'obtenir l'éclairage des objets virtuels est de les placer dans une texture d'environnement (*environment map*) extraite du monde réel. (Debevec, 1998) propose une méthode permettant de mesurer les radiances en tant qu'images HDR (*High Dynamic Range*) et d'ajouter les objets virtuels à une scène avec le bon éclairage en utilisant une *environment map* faite à partir des images HDR.

(Heymann *et al.*, 2005) proposent une méthode de capture de l'éclairage par vidéo pour le réappliquer sur des éléments virtuels. On place dans une scène réelle un marqueur et un miroir sphérique. On calcule en prétraitement une *cube map* pour chaque sommet du maillage de l'objet à ajouter, et dans laquelle on stocke les facteurs de forme (auto-occultations des objets). Puis pendant le rendu, une première étape consiste à analyser dans la vidéo la sphère pour en retirer une *cube map* correspondant à l'éclairage ; enfin on multiplie les 2 *cube maps* pour trouver l'éclairage diffus à appliquer en chaque sommet. Cette méthode permet, comme on peut le voir sur la figure 11, d'obtenir des ombres douces et l'éclairage indirect. Comme les auteurs se servent des capacités de la carte graphique, ils obtiennent un résultat en temps réel pour des objets constitués de dizaines de milliers de sommets.



Figure 11. Résultat de Heyman *et al.*

(Pomi *et al.*, 2004) proposent une méthode étendant *OpenRT* (qui permet de faire un rendu à partir de tracé de rayon) afin de pouvoir rendre les ombres et les réflexions. Ils se servent de tracé de rayon distribué, pour que le rendu soit interactif. Leur adaptation à la réalité mixte consiste à faire une méthode de rendu différentiel pour appliquer à la vidéo d'origine les modifications dues à la présence d'un objet virtuel.

4.1.2.2. L'éclairage des objets réels par des lumières virtuelles

Dans cette étape, la cohérence de l'éclairage doit être assurée lorsque des sources de lumière virtuelles sont ajoutées à l'environnement. (Loscos, 1999) propose dans sa thèse deux méthodes pour réaliser le ré-éclairage et la remodelisation de scènes réelles dans un contexte de vision par vidéo. Une première étape, commune aux deux méthodes, consiste à obtenir une modélisation de la scène réelle renfermant la géométrie d'une part et les propriétés de réflectance des matériaux, d'autre part. La première méthode se base sur un seul éclairage connu. Un premier traitement est effectué pour obtenir les textures sans qu'elles dépendent de l'éclairage. Puis un modèle de réflectance est déterminé pour chaque objet. La seconde méthode utilise plusieurs photographies prises avec des éclairages et des temps d'exposition différents pour mieux estimer les radiances, pixel par pixel.

(Hughes *et al.*, 2004) proposent une méthode permettant de calculer l'éclairage réel sur les objets virtuels et *vice versa*, ainsi que les ombres projetées par les objets virtuels sur les objets réels et *vice versa*. Les auteurs font l'hypothèse que le modèle

du monde réel est connu (car nécessaire au calcul d'occultation et d'inter-réflexion) ainsi que la position et l'orientation de la caméra. L'éclairage est obtenu en faisant une moyenne par pixel entre la couleur de l'image d'origine et le résultat de l'application d'un modèle d'éclairage avec la couleur de l'image d'origine. Pour rendre les ombres, les auteurs se servent d'un algorithme des *Shadow Volumes* (Crow, 1977) modifié. En effet, pour pouvoir projeter une ombre sur un objet réel, les auteurs ont fait le choix d'assombrir l'image, et de ré-éclairer les endroits censés être illuminés. On peut voir un résultat sur la figure 12.



Figure 12. Résultat de l'algorithme de Hughes et al.

4.1.3. Cohérence photométrique

Dans les systèmes de RA à base de vidéo, les éléments virtuels peuvent être distincts de la réalité parce qu'ils ne sont pas soumis aux effets de la caméra contrairement aux objets réels : il peut s'agir du bruit dû à l'acquisition ou bien du flou cinétique lorsqu'un objet ou la caméra bouge rapidement. On prend également en compte ici les problèmes de mise au point ; en effet, dans le système de vision humaine tout comme dans les caméras, un seul plan peut être net à la fois.

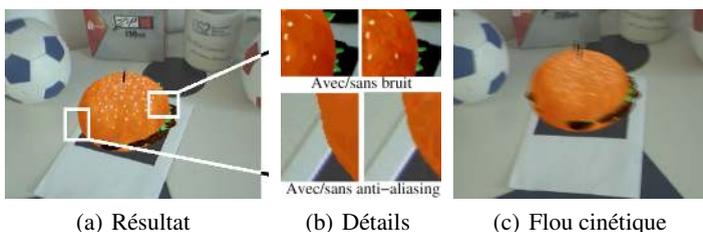


Figure 13. Résultat de Fischer

(Fischer et al., 2006) proposent une approche pour simuler le bruit (en imitant celui de la caméra qui a été calibrée) et le flou cinétique (en analysant le mouvement

courant de la caméra et en recopiant dans la direction de la caméra les objets virtuels) sur les objets virtuels. Les auteurs proposent également de faire de l'anti-aliasing aux frontières entre les objets réels et virtuels. Cette méthode utilise les capacités de la carte graphique et permet d'atteindre des taux de rafraîchissement interactifs. Un résultat est visible figure 13.

(Okumura *et al.*, 2006) proposent une méthode pour appliquer le flou cinétique et le flou lié à la mise au point sur les objets virtuels. Pour cela, il part d'un flux vidéo contenant un marqueur permettant de suivre la position de l'objet virtuel. Si les arêtes du dessin du marqueur sont floues, cela se traduit par une intensité décroissante sur l'image : on peut alors analyser cette décroissance pour en extraire les paramètres du flou et les appliquer à l'objet virtuel. Cette technique permet également d'améliorer l'alignement en utilisant les paramètres du flou pour estimer la position et l'orientation de la caméra. Les auteurs annoncent un résultat de 10 images par seconde.

4.1.4. Mise en œuvre

Les techniques que nous venons de décrire ont été développées pour des systèmes de réalité augmentée, sans prendre en compte les applications de réalité augmentée en extérieur. Comme nous l'avons établi, la réalité augmentée en extérieur se fait par le biais de terminaux mobiles. Ces derniers ont des puissances de calcul, capacités graphiques et mémoires limitées, malgré leurs augmentations constantes. Aussi, des adaptations sont nécessaires.

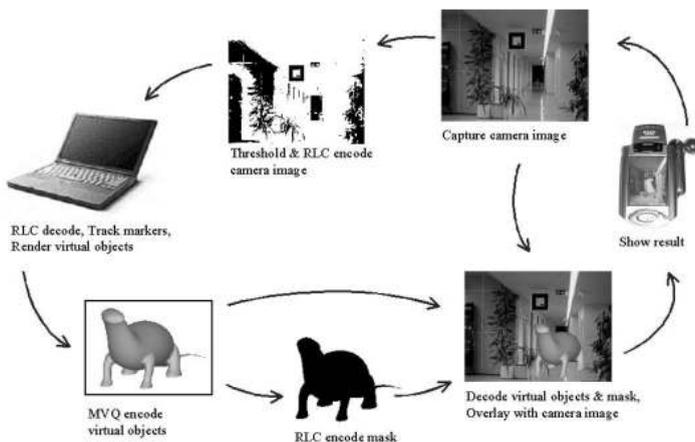


Figure 14. Schéma de rendu à distance de Pasman *et al*

Le rendu/visualisation 3D interactive à distance (en anglais *remote rendering*) consiste en un schéma client/serveur dans lequel le client sert à saisir les interactions utilisateurs et à afficher les images qui sont calculées, à distance, par un serveur dédié.

Parmi les travaux qui ont exploité ce concept, nous citons la bibliothèque Distributed IRIS GL de SGI (*Integrated Raster Imaging System Graphics Library* ancêtre de la bibliothèque OpenGL). Les commandes graphiques GL sont interceptées par la bibliothèque cliente, envoyées par le réseau au serveur et transmises à la carte graphique du serveur pour y être exécutées. Le produit OpenGL Vizserver de SGI permet de créer une session de rendu à distance entre un ordinateur de bureau et des serveurs graphiques de type SGI Onyx. Différentes implémentations de ces techniques ont été proposées plus récemment (Pouderoux, 2007) dans le cadre d'une utilisation sur des terminaux mobiles du type assistants personnels (PDAs). Dans le contexte de la réalité augmentée, cette technique a été employée par (Pasman *et al.*, 2004) qui réalise le suivi à partir de vidéos et le rendu à distance (voir le schéma figure 14). Ces techniques donnent de bons résultats mais nécessitent une excellente connectivité entre le serveur et le client, ce qui est souvent difficile à garantir dans un cadre de mobilité réel.

Les techniques de rendu adaptatif consistent en des algorithmes pouvant adapter en temps réel la quantité de données à afficher en fonction des ressources disponibles sur l'ordinateur client. L'avantage de ces techniques sur les techniques classiques est qu'elles peuvent fonctionner sur une classe de machines hétérogènes ce qui est le cas des terminaux mobiles en particulier. Par exemple, Mage11an (Marvie, 2004) est une plate-forme de visualisation mise au point au sein de l'IRISA qui a pour but de faciliter le développement de nouvelles solutions de visualisation de scènes 3D interactives distribuées sur des machines hétérogènes. Elle offre un ensemble de classes systèmes encapsulant des appels systèmes (*sockets, threads, etc.*) afin d'assurer la portabilité. Un avantage important est l'aspect *cross-platform* (Windows/Linux/SunOS) du logiciel. L'équipe Iparla a procédé au portage de la plate-forme Mage11an sur les terminaux mobiles basés sur le système d'exploitation Windows Mobile (cas des PocketPCs et Smartphones en particulier). La nouvelle version de la plate-forme, baptisée Elkano, inclut un module de rendu distribué, la capacité pour le serveur de construire un arbre des scènes, etc.

4.1.5. De la faisabilité de la RA sur les terminaux mobiles

Les quatre classes de terminaux mobiles de nouvelle génération sont les tablettes-PC, les UMPCs, les PDAs et les téléphones portables. Si les deux premières catégories possèdent une capacité de calcul suffisante pour calculer la localisation et le rendu graphique de manière confortable pour une application de RA, il n'en est pas de même pour les deux autres catégories.

Il est toutefois possible de faire fonctionner des applications de RA mobile sur des téléphones cellulaires. En 2004, l'Université de Bauhaus (Mohring *et al.*, 2004) a développé une application simple utilisant un téléphone mobile issu de la production de série : le Nokia 7650. Cette application permet d'afficher des objets 3D sur le flux vidéo provenant de la caméra, tout en utilisant des algorithmes de localisation simples utilisant des cibles codées. L'application développée sur des PDA (basés sur une architecture Intel XScale cadencée à 400 MHz) par l'Université de Graz (Wagner *et al.*, 2005) en 2005 met en place un jeu où deux utilisateurs employant chacun leurs

dispositifs peuvent interagir avec un train virtuel en modifiant sa vitesse ou les aiguillages de la piste sur laquelle le train évolue. Ce démonstrateur, qui a été présenté dans le cadre de plusieurs expositions, met en évidence que ces dispositifs possèdent un excellent taux d'acceptation de la part du public (par opposition aux systèmes employant des casques de réalité virtuelle). En revanche quelques points ont été soulevés concernant l'autonomie de tels dispositifs. Le problème principal est la capacité de calcul exigée par les applications de réalité augmentée. A l'heure actuelle, ces dispositifs utilisent les techniques les plus simples de localisation et de visualisation disponibles en RA.

4.2. Interaction 3D

L'interaction entre un utilisateur et une application se fait au travers d'interfaces. Dans une configuration classique de bureau, ces interfaces sont principalement basées sur le modèle WIMP (*Windows, Icons, Menus, Pointing device.*). Les contraintes des terminaux mobiles étant différentes de celles des PCs de bureau, les méthodes d'interaction homme/machine ont dû évoluer.

Concernant l'interaction 3D, une mutation similaire des interfaces est nécessaire. En effet, les méthodes d'interaction qui ont été développées dans un contexte de bureau ne sont plus forcément adaptées au contexte de mobilité. Interagir avec des environnements 3D virtuels est un processus complexe. On distingue généralement deux grandes familles de tâches d'interaction : la *manipulation* d'entités (objets, parties d'objets) et la *navigation* dans la scène 3D.

- Manipulation : tout déplacement d'un objet dans l'espace peut être décomposé suivant 6 transformations géométriques indépendantes, qui sont les translations le long des 3 axes x , y , z et les rotations autour de ces mêmes axes. Ainsi, si l'on souhaite permettre à l'utilisateur de placer librement un objet, il est nécessaire de lui permettre d'agir sur les 6 degrés de liberté (DDL).

- Navigation : elle peut être séparée en deux composantes : la *composante motrice*, qui permet le déplacement effectif de la vue d'un point à un autre, et la *composante cognitive*, aussi appelée *recherche d'itinéraire*, qui correspond aux processus intellectuels mis en jeu pour se repérer, appréhender l'environnement 3D, ou trouver une trajectoire à suivre pour atteindre un point précis (par exemple : exploration, recherche, manœuvre, mouvement prédéfini).

Il existe plusieurs dispositifs pour l'interaction 3D sur terminaux mobiles, à savoir : les écrans tactiles, les touches, les caméras, et les capteurs de positions et orientations.

4.2.1. Écrans tactiles

Les appareils mobiles d'aujourd'hui possèdent de plus en plus fréquemment un écran tactile. Celui-ci peut être utilisé à l'aide d'un stylet (comme sur la plupart des TabletPCs et assistants personnels) ou bien au doigt (comme sur l'iPhone d'Apple par exemple).



Figure 15. *Navidget sur un TabletPC (Hachet et al., 2008)*

L'utilisation d'un stylet permet une plus grande précision, et ainsi d'utiliser certaines techniques classiques qui ont été développées dans un contexte de bureau. Par exemple, dans le cadre de la navigation, il est possible de désigner le point d'intérêt directement au travers d'un processus de pointage (Mackinlay *et al.*, 1990). Cette technique est très largement utilisée dans les applications 3D sous le nom de "Go to". La vitesse de déplacement est fonction de la distance entre l'observateur et le point d'arrivée. De la même manière, Navidget (Hachet *et al.*, 2008) permet de sélectionner un point d'intérêt et de définir précisément la position et l'orientation de la caméra afin d'observer au mieux la zone visée, à l'aide de gestes simples (voir figure 15).

D'autres techniques, comme la "Sphère virtuelle" (Chen *et al.*, 1988) ou l'ensemble de techniques de Ware, *Eye in Hand*, *Scene in Hand*, *Flying Vehicle* (Ware *et al.*, 1990) peuvent aussi être adaptées. HoverCam (Khan *et al.*, 2005) permet d'inspecter facilement des objets 3D, et de les adapter à l'utilisation sur écran tactile.

Dans le cas de la manipulation d'objets, des techniques utilisées depuis longtemps sur PC sont aussi applicables, comme les contrôleurs de type "Jack" (Bier, 1986) ou le "Skitter" (Nielson *et al.*, 1987). Il faut cependant noter que le contrôle de ces interfaces est plus délicat avec un stylet qu'avec une souris.



Figure 16. *Scriboli (Hinckley et al., 2005) permet de contrôler une application à l'aide de gestes*

En revanche, certaines techniques classiques en utilisation de bureau deviennent peu adaptées sur de petits terminaux mobiles, à cause des contraintes de taille et des différences d'ergonomie de ces appareils. Par exemple, pour le contrôle de l'appli-

cation, l'utilisation de menus sur un écran de petite taille est problématique de par l'occultation quasi totale de la fenêtre de visualisation. Des méthodes alternatives ont alors été proposées, comme Scriboli (Hinckley *et al.*, 2005). Ces techniques se basent sur l'interprétation du tracé de l'utilisateur pour le contrôle de l'application (figure 16).

D'autres approches se basent sur le tracé de l'utilisateur pour la réalisation de tâches d'interaction 3D. Parmi ces techniques, citons l'Unicam (Zelevnik *et al.*, 1999) ou la manipulation 3D par gestes (Schmidt *et al.*, 2008). Ces techniques se basent sur la reconnaissance de gestes, et la prédiction de ce que souhaite faire l'utilisateur.

L'apparition d'écrans tactiles à plusieurs points sur terminaux mobiles permet d'envisager l'utilisation de techniques telles que celles de (Hancock *et al.*, 2007) ou (Zelevnik *et al.*, 1997). Ces techniques offrent une alternative intéressante pour la manipulation d'objets 3D. L'utilisation de plusieurs points permet de définir des opérations géométriques de manière intuitive (translations, rotations, facteur d'échelle).

4.2.2. Touches

Dans le cas des interactions 3D, l'utilisation du clavier pose un problème dû au signal binaire qu'il fournit. Par exemple, dans le cas de la navigation, il est difficile de fixer la vitesse de déplacement lors de l'appui sur une touche. Néanmoins plusieurs stratégies ont été développées. Par exemple, la sélection d'un point à l'aide de touches peut se faire en déplaçant un curseur, mais ce déplacement peut être long. Afin d'améliorer la phase de sélection, quelques études ont été menées. Pour sélectionner un point particulier à l'écran, (Hachet *et al.*, 2007) proposent le *Jump and Refine*, qui divise la tâche en deux phases : la sélection grossière dans une grille, qui se fait rapidement, puis la sélection précise à l'intérieur de cette zone. Le *ZoneZoom* (Robbins *et al.*, 2004) a une approche similaire où l'on agrandit la zone de l'écran intéressante. Dans le cas de la navigation 3D, le *Z-Goto* (Hachet *et al.*, 2006) se propose comme une extension discrète du « goto », adaptée à la sélection d'un point à l'aide uniquement des touches de l'appareil. Le principe est de diviser la vue courante en un faible nombre de « sections », dont la taille croît avec la distance. L'utilisateur sélectionne ainsi la section qui l'intéresse, puis la position du point d'intérêt au sein de cette section.

Il est à noter que les appareils mobiles actuels délaissent de plus en plus les touches au profit des écrans tactiles.

4.2.3. Caméras

Les caméras embarquées sur les terminaux mobiles ont beaucoup été utilisées dans des applications de réalité augmentée, comme la *NaviCam* (Rekimoto *et al.*, 1995). Ces approches sont basées sur la détection de marqueurs devant être déposés précédemment dans l'environnement réel. *Tangimap* (Hachet *et al.*, 2005) est une interface bimanuelle qui a été développée pour interagir avec des scènes 3D. En déplaçant une cible devant la caméra d'un terminal mobile, l'utilisateur peut contrôler simultanément

3 DDL. Enfin, de nombreuses interfaces d'entrée sont basées sur les mouvements du terminal mobile tenu dans la main. Le *ScrollPad* (Fallmana *et al.*, 2004) est un PDA équipé d'une souris optique sur la face arrière et permettant la visualisation de grands documents en le déplaçant sur une table. La technique *Peephole* (Yee, 2003) est une interface comparable où l'utilisateur déplace dans l'espace le PDA relié à la mécanique d'une souris par un système de câbles et de poulies.

4.2.4. Capteurs

Les appareils mobiles disponibles actuellement possèdent pour certains différents capteurs : GPS pour la géolocalisation, accéléromètres pour détecter l'orientation de l'appareil par exemple. Des travaux ont étudié l'utilisation de ces capteurs dans le cas de l'interaction avec des environnements 3D. Par exemple, (Kela *et al.*, 2006) utilise des accéléromètres pour détecter les mouvements de l'utilisateur, et ainsi contrôler l'application. Cependant, l'angle de vision des écrans des portables est parfois limité, et devoir tourner l'appareil pour interagir peut influencer sur la visibilité.

5. Projets et travaux

Depuis quelques années, nous assistons à la multiplication de projets de RA en extérieur. Ces derniers touchent des domaines d'applications aussi divers que le domaine militaire, le génie civil, l'héritage culturel, etc. Dans ce qui suit, nous allons effectuer un tour d'horizon des applications de RA en extérieur. Il montre l'intérêt de la RA mobile en extérieur et cerne l'enjeu industriel.



Figure 17. La plate-forme MARS et son interface de navigation (Hollerer *et al.*, 1999)

L'équipe de Feiner (Computer Graphics and User Interfaces Lab, université de Colombia, 1996) est pionnière dans ce domaine avec le système MARS "*Mobile Augmented Reality System*" (Hollerer *et al.*, 1999). Le principal objectif était d'explorer la synergie entre la réalité augmentée et l'informatique mobile et de démontrer la faisabilité d'un système en extérieur pour fournir des informations sur l'environnement, permettant ainsi de créer "une machine de guide personnalisée" (voir figure 17). À partir de la position obtenue avec un DGPS (Differential GPS) et un capteur inertiel, l'information est retournée à l'utilisateur *via* un écran tactile (l'unité de traitement étant un ordinateur portable porté sur le dos) utilisé uniquement pour la visualisation et l'accès au moteur de recherche.

Ce système a servi de modèle pour la mise en œuvre du système BARS "*Battlefield Augmented Reality System*" (Livingston *et al.*, 2006) (Naval Research Laboratory, Washington, 2000). Ce projet avait pour objectif de proposer aux militaires un système de navigation et de localisation qui leur permettait de visualiser, en plein champ de bataille, des informations pertinentes telles que la position des tireurs embusqués (voir 18). À partir de la localisation fournie par un GPS couplé à une centrale inertielle, le soldat peut visualiser les données à travers un casque semi-transparent "see-through HMD" et interagir avec celles-ci à l'aide d'une souris ou d'une commande vocale.



Figure 18. Le prototype BARS testé par un militaire (Julier *et al.*, 2000) (a) Plateforme matérielle (b) Exemple d'augmentation

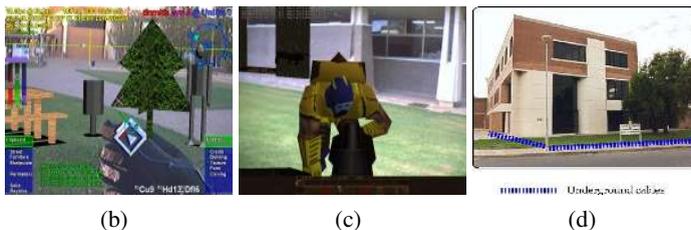
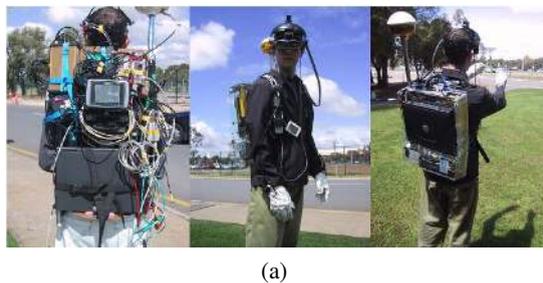


Figure 19. Plat-forme Tinnith (Piekariski *et al.*, 2001) (a) La plate-forme matérielle (b) Exemple de visualisation (c) ARQuake (Piekariski *et al.*, 2002) (d) Exemple d'utilisation de la plate-forme pour l'architecture (Thomas *et al.*, 1999)

La plate-forme Tinnith (Piekariski *et al.*, 2001) (Wearable Computer Lab, université d'Australie, 1998) a pour objectif de mettre au point une architecture (matérielle

et logicielle) pour un système de navigation mobile utilisant la RA. Le système développé a connu plusieurs déclinaisons tel le jeu *ARQuake* (Piekarski *et al.*, 2002), ou la visualisation de plans architecturaux (Thomas *et al.*, 1999). De plus, la plate-forme *Tinmith* a été utilisée dans la conception du système *ARVino* (King *et al.*, 2005) pour créer une nouvelle façon de visualiser des données géolocalisées en relation avec la viticulture. Ainsi, la RA fournit une assistance sur le terrain, en permettant de superposer à la parcelle de viticulture les données la concernant.

Autre exemple, la plate-forme collaborative du projet *Studierstube* (Reitmayr *et al.*, 2003) (Graz University of Technology, Université de Vienne, 2004) permet à un ou plusieurs utilisateurs d'interagir lors d'une application de navigation, de recherche et d'affichage d'informations (voir figure 20).



Figure 20. La plate-forme *Studierstube* (Reitmayr *et al.*, 2003) (a) La plate-forme matérielle (b) Exemple d'augmentation

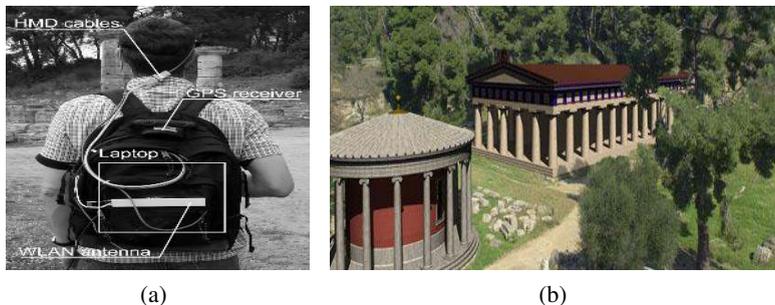


Figure 21. Architecture développée dans le projet *ARCHEOGUIDE* (Gleue *et al.*, 2001) (a) La plate-forme (b) Exemple de visualisation : Un temple grec reconstruit en 3D

En 2000, le projet *ARCHEOGUIDE* "Augmented Reality-based Cultural Heritage On-site GUIDE" (Gleue *et al.*, 2001) est lancé. Son principal objectif était de concevoir de nouvelles méthodes interactives pour l'accès aux informations sur l'héritage culturel. Ce projet englobait tous les aspects d'un système de RA : la visualisation 3D, l'informatique mobile et les méthodes d'interaction multimodales. En fonction de la

position du visiteur sur le site, *ARCHEOGUIDE* filtre les informations et n'affiche que celles qui correspondent aux monuments vus par l'utilisateur. De plus, l'interface développée permet de choisir entre plusieurs thèmes (en différentes langues) et supports multimédias. *ARCHEOGUIDE* offre également la possibilité de visualiser en 3D les parties endommagées des vestiges.

Lancé en avril 2006, le projet *Vidente* (Schall *et al.*, 2007) a pour but de développer une solution de réalité augmentée mobile pour la visualisation 3D des caractéristiques du sous-sol. Suivant la position et l'orientation de l'utilisateur, ce dernier visualise, en temps réel, le réseau de câbles et de tuyaux enfouis sous terre sous forme de graphiques 3D (voir figure 22). L'objectif du projet est de fournir un outil supplémentaire aux solutions SIG mobiles existantes, et de contribuer ainsi à un traitement optimisé pour l'accès aux données.

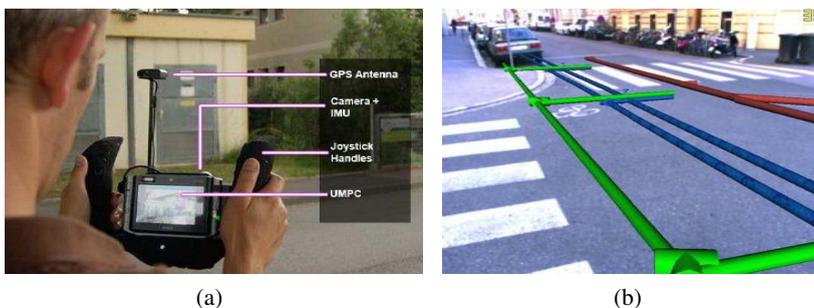


Figure 22. (a) Prototype de *Vidente* (b) Visualisation des tuyaux de canalisation (Schall *et al.*, 2007)



Figure 23. *IpCity* (Markus *et al.*, 2007)

IPCity "Integrated Project on Interaction and Presence in Urban Environments" (Markus *et al.*, 2007) est un projet européen lancé dans la même année et dont l'objectif est de développer des techniques d'interactions qui permettent d'exporter la réalité mixte (MR) dans des environnements urbains. Ceci permet de faciliter la participation et la communication entre les différents acteurs d'un même projet (par exemple des

élus, des citoyens ou des riverains, dans le cas d'un projet urbain) au moyen d'interfaces mobiles et légères.

Enfin, le projet *RAXENV* (*Réalité augmentée en extérieur appliquée aux sciences de l'environnement*) (Vairon et al., 2007) a pour objectif de démontrer la faisabilité d'un système de réalité augmentée en extérieur dans le domaine des sciences et techniques de l'environnement (voir figure 24), que ce soit en termes de technologie mais aussi d'adoption par les utilisateurs finaux. Le projet se démarque des autres projets car il explore la mise en œuvre d'un système englobant la localisation multicapteur, la visualisation sur terminaux mobiles et l'interaction avec des données (SIG, modèles 3D, coupes géologiques, etc.) en situation de mobilité.

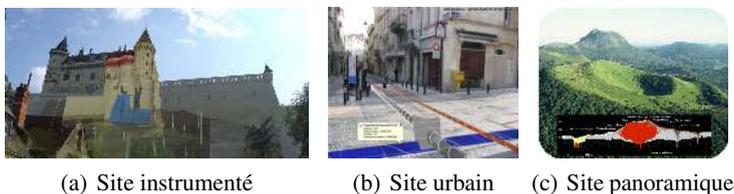


Figure 24. Les trois scénarios envisagés dans *RAXENV* (Vairon et al., 2007)

6. Conclusion

Cet article comprend un panorama des techniques et technologies pour la réalité augmentée en extérieur. Dans ce tour d'horizon, nous présentons les différentes problématiques à prendre en compte pour la conception d'un système de RA mobile. Cela comprend les aspects complémentaires de localisation, de visualisation et d'interaction. La localisation permet d'assurer la cohérence spatio-temporelle de la scène augmentée et d'indexer les données à visualiser et à mettre à disposition de l'utilisateur. La tendance des techniques de localisation converge vers des méthodes multicapteur qui fusionnent différents types de données complémentaires pour fournir un positionnement précis, fiable et robuste aux conditions de travail en environnement extérieur. L'utilisation de capteurs hybrides doit prendre en compte plusieurs aspects importants, à savoir : exprimer les différentes données fournies dans un même référentiel, quantifier la qualité des mesures fournies par les différents capteurs, et opter pour la meilleure stratégie de fusion. La visualisation des augmentations pour les applications de RA en extérieur se fait *via* des terminaux mobiles qui ont des puissances de calcul, capacités graphiques et mémoires limitées. De ce fait, les techniques de rendu proposées essaient d'adapter l'existant conformément aux caractéristiques de ces nouveaux types de terminaux pour ainsi offrir des vues augmentées réalistes. La visualisation de données dans une application de réalité augmentée en extérieur doit davantage tenir compte des conditions extérieures (luminosité accrue par exemple). En revanche, l'ergonomie des appareils mobiles diffère de celle des

stations de travail. De nouveaux concepts d'interaction ont dû être développés, se basant sur l'utilisation des touches ou du stylet. Cependant, les appareils mobiles actuels possèdent pour beaucoup un écran tactile, voire multitactile. Ceci engendre de nouvelles façons d'interagir, et, si quelques recherches ont été effectuées sur l'ergonomie de ces appareils pour des applications 2D, utiliser de tels dispositifs pour interagir avec un environnement 3D reste un problème ouvert.

Remerciements

Cet article a été rédigé dans le cadre du projet RAXENV cofinancé par l'Agence Nationale de la Recherche - Programme Technologies Logicielles 2006.

7. Bibliographie

- Alves J., Lobo J., Diaz J., « Camera-inertial sensor modeling and alignment for visual navigation », *Journal of Robotic Systems*, vol. 21, p. 6-12, 2004.
- Aron M., Simon G., Berger M., « Use of inertial sensors to support video tracking : Research Articles », *Comput. Animat. Virtual Worlds*, vol. 18, n° 1, p. 57-68, 2007.
- Azuma R., Baillo Y., Behringer R., Feiner S., Julier S., MacIntyre B., « Recent Advances in Augmented Reality », *IEEE Computer Graphics and Applications*, vol. 21, n° 6, p. 34-47, 2001.
- Baillo Y., Julier S., Brown D., Livingston M., « A Tracker Alignment Framework for Augmented Reality », *Proceedings of the The 2nd IEEE and ACM International Symposium on Mixed and Augmented Reality*, p. 142, 2003.
- Behringer R., « Registration for Outdoor Augmented Reality Applications Using Computer Vision Techniques and Hybrid Sensors », *VR '99 : Proceedings of the IEEE Virtual Reality*, IEEE Computer Society, Washington, DC, USA, p. 244, 1999.
- Bier E., « Skitters and Jacks : Interactive 3D Positioning Tools », *Workshop on Interactive 3D Graphics*, ACM, p. 183-196, 1986.
- Bleser G., Stricker D., « Advanced tracking through efficient image processing and visual-inertial sensor fusion », *VR*, p. 137-144, 2008.
- Bleser G., Stricker D., Wohlleber C., Felsberg M., Gustafsson F., Hol J., Schön T., Skoglund J., Slycke P., Smeitz S., Chandaria J., Thomas G., Bartczak B., Koester K., Koch R., Becker M., « Real-Time Camera Tracking in the MATRIS project », *Proceedings of IBC*, Amsterdam, NL, 7-11 September, 2006a.
- Bleser G., Wuest H., Stricker D., « Online camera pose estimation in partially known and dynamic scenes », *Fifth IEEE and ACM International Symposium on Mixed and Augmented Reality*, Santa Barbara, USA, p. 56-65, 22-25 Oct, 2006b.
- Breen D., Whitaker R., Rose E., Tuceryan M., « Interactive Occlusion and Automatic Object Placement for Augmented Reality », *Computer Graphics Forum*, vol. 15, n° 3, p. 11-22, 1996.
- Chen M., Mountford S. J., Sellen A., « A study in interactive 3-D rotation using 2-D control devices », *SIGGRAPH '88 : Proceedings of the 15th annual conference on Computer graphics and interactive techniques*, ACM Press, New York, NY, USA, p. 121-129, 1988.

- Comport A. I., Pressigout M., Marchand E., Chaumette F., « Real-Time Markerless Tracking for Augmented Reality : The Virtual Visual Servoing Framework », *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, n° 4, p. 615-628, 2006.
- Crow F. C., « Shadow Algorithms for Computer Graphics », *ACM SIGGRAPH Computer Graphics*, vol. 11, n° 2, p. 242-248, 1977.
- Debevec P., « Rendering Synthetic Objects into Real Scenes : Bridging Traditional and Image-Based Graphics with Global Illumination and High Dynamic Range Photography », *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques*, vol. 32, p. 189-198, 1998.
- Didier J.-Y., Roussel D., Mallem M., Otmame S., Naudet S., Pham Q.-C., Bourgeois S., Mégard C., Leroux C., Hocquard A., « AMRA : Augmented Reality assistance in train maintenance tasks », *Workshop on Industrial Augmented Reality (ISMAR'05)*, Vienna (Austria), October 5th, 2005.
- Fallmana D., Lund A., Wiberg M., « ScrollPad : Tangible Scrolling with Mobile Devices », *HICSS '04 : Proceedings of the Proceedings of the 37th Annual Hawaii International Conference on System Sciences (HICSS'04) - Track 9*, IEEE Computer Society, Washington, DC, USA, p. 90286.3, 2004.
- Fischer J., Bartz D., Handling Photographic Imperfections and Aliasing in Augmented Reality, Technical report, Wilhelm Schickard Institute for Computer Science, Graphical-Interactive Systems (WSI/GRIS), University of Tubingen, 2006.
- Fischer J., Regenbrecht H., Baratoff G., « Detecting Dynamic Occlusion in Front of Static Backgrounds for AR Scenes », *Proceedings of the Workshop on Virtual Environments 2003*, ACM Press New York, NY, USA, p. 153-161, 2003.
- Fischler M. A., Bolles R. C., « Random sample consensus : a paradigm for model fitting with applications to image analysis and automated cartography », *Commun. ACM*, vol. 24, n° 6, p. 381-395, 1981.
- Gleue T., Daehne P., « Design and Implementation of a Mobile Device for Outdoor Augmented Reality in the ARCHEOGUIDE Project », *Virtual Reality, Archaeology, and Cultural Heritage International Symposium*, Glyfada, Nr Athens, Greece, 28-30, 2001.
- Greger G., Shirley P., Hubbard P., Greenberg D., « The Irradiance Volume », *IEEE Computer Graphics and Applications*, vol. 18, n° 2, p. 32-43, 1998.
- Hachet M., Declé F., Guitton P., « Z-Goto for efficient navigation in 3D environments from discrete inputs », *Proceedings of the ACM symposium on Virtual reality software and technology*, 2006.
- Hachet M., Declé F., Knoedel S., Guitton P., « Navidget for Easy 3D Camera Positioning from 2D Inputs », *Proceedings of Symposium on 3D User Interfaces (3DUI) 2008*, 2008.
- Hachet M., Pouderoux J., Guitton P., « A Camera-Based Interface for Interaction with Mobile Handheld Computers », *Proceedings of I3D'05 - ACM SIGGRAPH 2005 Symposium on Interactive 3D Graphics and Games*, ACM Press, p. 65-71, 2005.
- Hachet M., Pouderoux J., Tyndiuk F., Guitton P., « "Jump and refine" for rapid pointing on mobile phones », *CHI '07 : Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM Press, New York, NY, USA, p. 167-170, 2007.
- Hancock M., Carpendale S., Cockburn A., « Shallow Depth 3D Interaction : Design and Evaluation of One-, Two-, and Three-Touch Techniques », *Proceedings of ACM CHI'2007 Conference on Human Factors in Computing Systems. San Jose, CA*, 2007.

- Heymann S., Smolic A., Müller K., Fröhlich B., « Real-Time Video Capture for Illumination Reconstruction in Augmented Reality Applications », *I3D 2005, ACM SIGGRAPH 2005 Symposium on Interactive 3D Graphics and Games*, 2005.
- Hinckley K., Baudisch P., Ramos G., Guimbretiere F., « Design and analysis of delimiters for selection-action pen gesture phrases in scriboli », *CHI '05 : Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM Press, New York, NY, USA, p. 451-460, 2005.
- Hollerer T., Feiner S., Terauchi T., Rashid G., Hallaway D., « Exploring MARS : Developing Indoor and Outdoor User Interfaces to a Mobile Augmented Reality System », *Computers and Graphics, Elsevier Publishers*, vol. 23, n° 6, p. 779-785, December, 1999.
- Hughes C., Kontinen J., Pattanaik S., « The Future of Mixed Reality : Issues in Illumination and Shadows », *The Interservice/Industry Training, Simulation & Education Conference (IITSEC)*, NTSA, 2004.
- Jensen T., Andersen M., Madsen C., « Real-Time Image Based Lighting for Outdoor Augmented Reality under Dynamically Changing Illumination Conditions », *Proceedings of the International Conference on Graphics Theory and Applications*, p. 364-371, 2006.
- Johansson B., Cipolla R., « A system for automatic pose-estimation from a single image in a city scene », *IASTED Int. Conf. Signal Processing, Pattern Recognition and Applications*, 2002.
- Julier S., Baillot Y., Lanzagorta M., Brown D., Rosenblum L., « BARS : Battlefield Augmented Reality system », *NATO Symposium on Information Processing Techniques for Military Systems*, Istanbul, Turkey, 9-11 October, 2000.
- Kanbara M., Okuma T., Takemura H., Yokoya N., « A Stereoscopic Video See-Through Augmented Reality System Based on Real-Time Vision-Based Registration », *Proceeding of IEEE Virtual Reality Conference*, p. 255-262, 2000.
- Kanbara M., Yokoya N., « Real-time Estimation of Light Source Environment for Photorealistic Augmented Reality », *Proceedings of the 17th International Conference on Pattern Recognition*, vol. 2, p. 911-914, 2004.
- Kato H., Billinghurst M., « Marker Tracking and HMD Calibration for a Video-Based Augmented Reality Conferencing System », *IWAR '99 : Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality*, IEEE Computer Society, Washington, DC, USA, p. 85, October, 1999.
- Kaufmann H., *Construct3D : An Augmented Reality Application for Mathematics and Geometry Education*, Technical report, Institute of Software Technology and Interactive Systems, 2002.
- Kela J., Korpiää P., Mäntyjärvi J., Kallio S., Savino G., Jozzo L., Marca D., « Accelerometer-based gesture control for a design environment », *Personal Ubiquitous Comput.*, vol. 10, n° 5, p. 285-299, 2006.
- Khan A., Komalo B., Stam J., Fitzmaurice G., Kurtenbach G., « HoverCam : interactive 3D navigation for proximal object inspection », *I3D '05 : Proceedings of the 2005 symposium on Interactive 3D graphics and games*, ACM Press, New York, NY, USA, p. 73-80, 2005.
- King G., Piekarski W., Thomas B., « ARvino - outdoor augmented reality visualisation of viticulture GIS data », *Proceedings of the Fourth IEEE and ACM International Symposium on Mixed and Augmented Reality*, oct, 2005.

- Lepetit V., Fua P., « Monocular model-based 3D tracking of rigid objects », *Found. Trends. Comput. Graph. Vis.*, vol. 1, n° 1, p. 1-89, 2005.
- Livingston M. A., Brown D., Julier S. J., Schmidt G. S., « Military Applications of Augmented Reality », , NATO Human Factors and Medicine Panel Workshop on Virtual Media for Military Applications, 13–15 June, 2006.
- Loscos C., Ré-éclairage et Remodélisation Interactifs des Scènes réelles pour la Réalité Augmentée, PhD thesis, Ph. D. Thesis, Université Grenoble I, Grenoble, 1999.
- Lowe D., « Distinctive image features from scale-invariant keypoints », *International Journal of Computer Vision*, vol. 20, p. 91-110, 2003.
- Mackinlay J., Card S., Robertson G., « Rapid controlled movement through a virtual 3D workspace », *Proceedings of the 17th annual conference on Computer graphics and interactive techniques*, ACM Press New York, NY, USA, p. 171-176, 1990.
- Maidi M., Suivi Hybride en présence d'Occultations pour la Réalité Augmentée, PhD thesis, Université d'Évry Val d'Essonne, Évry, Novembre, 2007.
- Maidi M., Ababsa F., Mallem M., « Vision-Inertial System Calibration for Tracking in Augmented Reality », *2nd International Conference on Informatics in Control, Automation and Robotics*, p. 156-162, 2005.
- Markus S., Dieter S., « Urban Sketcher : Mixed Reality on Site for Urban Planning and Architecture », *6th International Symposium on Mixed and Augmented Reality*, IEEE, ACM, Nara, Japan, p. 27-30, November, 2007.
- Marvie J.-E., Visualisation Interactive d'Environnements Virtuels Complexes à travers des Réseaux et sur des Machines à Performances Variables, PhD thesis, INSA, de Rennes, France, 2004.
- Masson L., Dhome M., Jurie F., « Contour/texture approach for visual tracking », *Scandinavian Conference on Image Analysis*, 2003.
- Mei C., Rives P., « Cartographie et localisation simultanée avec un capteur de vision », *Journées Nationales de la Recherche en Robotique*, October, 2007.
- Mohring M., Lessig C., Bimber O., « Video See-Through AR on Consumer Cell-Phones », *ISMAR '04 : Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality*, IEEE Computer Society, Washington, DC, USA, p. 252-253, 2004.
- Mouragnon E., Lhuillier M., Dhome M., Dekeyser F., Sayd P., « Monocular Vision Based SLAM for Mobile Robots », *ICPR '06 : Proceedings of the 18th International Conference on Pattern Recognition*, IEEE Computer Society, Washington, DC, USA, p. 1027-1031, 2006.
- Nakamae E., Harada K., Ishizaki T., Nishita T., « A Montage Method : the Overlaying of the Computer Generated Images onto a Background Photograph », *ACM SIGGRAPH Computer Graphics*, vol. 20, n° 4, p. 207-214, 1986.
- Nielson G., Jr D. O., « Direct manipulation techniques for 3D objects using 2D locator devices », *SI3D '86 : Proceedings of the 1986 workshop on Interactive 3D graphics*, ACM Press, New York, NY, USA, p. 175-182, 1987.
- Nistér D., « Preemptive RANSAC for Live Structure and Motion Estimation », *ICCV '03 : Proceedings of the Ninth IEEE International Conference on Computer Vision*, IEEE Computer Society, Washington, DC, USA, p. 199, 2003.

- Okumura B., Kanbara M., Yokoya N., « Augmented Reality Based on Estimation of Defocusing and Motion Blurring from Captured Images », *Proceedings of ISMAR 2006 the IEEE/ACM International Symposium on Mixed and Augmented Reality*, p. 219-225, 2006.
- Park F., Martin B., « Robot sensor calibration : solving $AX=XB$ on the euclidean group », *IEEE Transactions on Robotics and Automation*, vol. 5, p. 717-721, 1994.
- Pasman W., Woodward C., Hakkarainen M., Honkamaa P., Hyvääkkä J., « Augmented Reality With Large 3D Models on a PDA : Implementation, Performance and Use Experiences », *Proceedings of the 2004 ACM SIGGRAPH International Conference on Virtual Reality Continuum and its Applications in Industry*, ACM Press New York, NY, USA, p. 344-351, 2004.
- Phong B. T., « Illumination for Computer Generated Images », *Publication of the ACM*, 1975.
- Piekarski W., Thomas B., « Tinmith-evo5 - An Architecture for Supporting Mobile Augmented Reality Environments », *2nd Int'l Symposium on Augmented Reality*, New York, p. 177-178, Octobre, 2001.
- Piekarski W., Thomas B., « ARQuake : the outdoor augmented reality gaming system », *Communications of the ACM*, vol. 45, n° 1, p. 36-38, 2002.
- Pomi A., Slusallek P., « Interactive Mixed Reality Rendering in a Distributed Ray Tracing Framework », *Proc. of IEEE and ACM International Symposium on Mixed and Augmented Reality04, Student Colloquium*, 2004.
- Pouderoux J., Création semi-automatique de modèles numériques de terrain - Visualisation et interaction du terminaux mobiles communicants, PhD thesis, Université Bordeaux 1, France, 2007.
- Pressigout M., Marchand E., « Real-Time Hybrid Tracking using Edge and Texture Information », *Int. Journal of Robotics Research, IJRR*, vol. 26, n° 7, p. 689-713, July, 2007.
- Reitmayr G., Drummond T., « Going Out : Robust Model-based Tracking for Outdoor Augmented Reality », *IEEE ISMAR*, Santa Barbara, California, USA, 2006.
- Reitmayr G., Drummond T., « Initialisation for Visual Tracking in Urban Environments », *IEEE ISMAR*, Nara, Japan, 2007.
- Reitmayr G., Schmalstieg D., Collaborative Augmented Reality for Outdoor Navigation and Information Browsing, Technical Report n° TR-188-2-2003-28, Interactive Media Systems Group, Vienna University of Technology, 2003.
- Rekimoto J., Nagao K., « The World through the Computer : Computer Augmented Interaction with Real World Environments », *symposium User Interface Software and Technology*, 1995.
- Ribo M., Lang P., Ganster H., Brandner M., Stock C., Pinz A., « Hybrid Tracking for Outdoor Augmented Reality Applications », *IEEE Comput. Graph. Appl.*, vol. 22, n° 6, p. 54-63, 2002.
- Robbins D., Cutrell E., Sarin R., Horvitz E., « ZoneZoom : map navigation for smartphones with recursive view segmentation », *AVI '04 : Proceedings of the working conference on Advanced visual interfaces*, ACM Press, New York, NY, USA, p. 231-234, 2004.
- Sato I., Sato Y., Ikeuchi K., « Illumination Distribution from Shadows », *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, p. 306-312, 1999.

- Schall G., Mendez E., Reitinger B., Junghanns S., Schmalstieg D., « Handheld Geospatial Augmented Reality using Urban 3D Models », , MSI Workshop, CHI, 2007.
- Schmidt R., Singh K., Balakrishnan R., « Sketching and Composing Widgets for 3D Manipulation », *Eurographics 2008 (To Appear)*, 2008.
- Simon G., Berger M.-O., « Pose Estimation for Planar Structures », *IEEE Comput. Graph. Appl.*, vol. 22, n° 6, p. 46-53, 2002.
- State A., Livingston M., Hirota G., Garrett W., Whitton M., Fuchs H., Pisano E., « Techniques for Augmented-Reality Systems : Realizing Ultrasound-Guided Needle Biopsies », *Proceedings of SIGGRAPH*, New Orleans, p. 439-446, 4-9 August, 1996.
- Stricker D., Kettenbach T., « Real-Time and Markerless Vision-Based Tracking for Outdoor Augmented Reality Applications », *ISAR '01 : Proceedings of the IEEE and ACM International Symposium on Augmented Reality (ISAR'01)*, IEEE Computer Society, Washington, DC, USA, p. 189, 2001.
- Thomas B., Piekarski W., Gunther B., « Using Augmented Reality to Visualize Architecture Designs in an Outdoor Environment », *In Design Computing on the Net*, Sydney, NSW, Nov, 1999.
- Vacchetti L., Lepetit V., Fua P., « Combining Edge and Texture Information for Real-Time Accurate 3D Camera Tracking », *ISMAR '04 : Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'04)*, IEEE Computer Society, Washington, DC, USA, p. 48-57, 2004.
- Vairon J., Frauciel L., Ababsa F., Didier J., Zendjebil I., Delmont R., Guitton P., Hachet M., Roudiere B., « RAXenv : a new AR research project to serve outdoor environmental activities », , 4e conférence du réseau Intuition, 4-5 octobre, 2007.
- Wagner D., Pintaric T., Ledermann F., Schmalstieg D., *Pervasive Computing*, vol. 3468/2005 of *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, chapter Towards Massively Multi-user Augmented Reality on Handheld Devices, p. 208-219, 2005.
- Ware C., Osborne S., « Exploration and virtual camera control in virtual three dimensional environments », *Proceedings of the 1990 symposium on Interactive 3D graphics*, ACM Press New York, NY, USA, p. 175-183, 1990.
- Yee K., « Peephole displays : pen interaction on spatially aware handheld computers », *CHI '03 : Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM Press, New York, NY, USA, p. 1-8, 2003.
- You S., Neumann U., Azuma R., « Orientation Tracking for Outdoor Augmented Reality Registration », *IEEE Computer Graphics and Applications*, vol. 19, n° 6, p. 36-42, 1999.
- Zeleznik R., Forsberg A., « UniCam - 2D gestural camera controls for 3D environments », *Proceedings of Symposium on Interactive 3D Graphics*, p. 169-173, 1999.
- Zeleznik R., Forsberg A., Strauss P., « Two pointer input for 3D interaction », *Proceedings of the 1997 symposium on Interactive 3D graphics*, 1997.