



### **Science Arts & Métiers (SAM)**

is an open access repository that collects the work of Arts et Métiers Institute of Technology researchers and makes it freely available over the web where possible.

This is an author-deposited version published in: <https://sam.ensam.eu>  
Handle ID: <http://hdl.handle.net/10985/9997>

#### **To cite this version :**

Emmanuel DURAND, Julien ROGER, Christian PÈRE, Frédéric MERIENNE - Une implémentation de la réalité augmentée pour le patrimoine - In: Confere 2012, Italie, 2012-07-05 - Confere 2012 - 2012

Any correspondence concerning this service should be sent to the repository

Administrator : [scienceouverte@ensam.eu](mailto:scienceouverte@ensam.eu)



# UNE IMPLEMENTATION DE LA REALITE AUGMENTEE POUR LE PATRIMOINE

**Emmanuel DURAND, Frédéric MERIENNE, Julien ROGER, Christian PERE**

emmanueldurand@gmail.com, frederic.merienne@ensam.eu, julien.roger@on-situ.com,  
christian.pere@ensam.eu  
on-situ / Institut Image Chalon sur Saône, Institut Image Chalon sur Saône, on-situ, Institut Image  
Chalon sur Saône

## RESUME

Ce papier décrit la démarche et les choix mis en œuvre pour proposer un dispositif de réalité augmentée adapté à un circuit de visite d'un lieu historique. L'accent a été mis sur la production d'images réalistes et sur une intégration fluide entre objets virtuels et réels.

*Réalité virtuelle, design industriel, valorisation de l'innovation*

## 1 INTRODUCTION

Sur les quelques dernières années, et particulièrement du fait du fort taux de pénétration des Smartphones dans la société, la réalité augmentée (ou RA) s'est démocratisée et est l'objet de nombreuses expérimentations auprès du public. De tous les champs d'application de la RA, la mise en valeur du patrimoine est particulièrement intéressante puisqu'il devient possible d'apporter in-situ à l'utilisateur les connaissances et informations dont il a besoin pour appréhender le lieu qu'il parcourt. La difficulté lors du développement d'une application de réalité augmentée pour le patrimoine est de garder à l'esprit que le sujet doit rester le lieu visité, l'application de réalité augmentée ne devant en aucun cas concentrer l'attention. De plus puisque le but est culturel, il est important de mettre l'accent sur l'intérêt du contenu plutôt que l'aspect divertissement qui peut émerger d'une telle application.

## 2 ANTECEDENTS

Les applications de réalités augmentées peuvent être regroupées en deux familles, selon qu'elles soient mobiles ou fixes. Les dispositifs fixes sont moins représentés, les implémentations mobiles s'inscrivant plus dans la continuité des audioguides proposés par la plupart des sites culturels. On peut malgré tout citer le Virtual Museum Guide de l'institut Fraunhofer [Fraunhofer 2010] qui propose d'intégrer des éléments contextuels (image ou texte) sur un tirage très grand format appliqué sur un mur, le dispositif prenant la forme d'un écran rotatif derrière lequel est placée une caméra ; ou le projet de « télescope de réalité augmentée » du CITU [CITU 2008], dont le prototype a permis durant une semaine d'avoir accès à des informations textuelles ou graphique sur la ville de Paris à partir d'un point de vue situé sur l'Arc de Triomphe. Enfin, citons les implémentations de Total Immersion [Total Immersion] qui, contrairement aux applications précédentes n'intègrent pas des objets virtuels dans un environnement réel contextuellement (par exemple en plaçant un édifice détruit à sa position d'époque), mais permettent à l'utilisateur de manipuler une maquette virtuelle : elles n'ont donc pas vocation à être disposées à un emplacement spécifique sur le site.

Du côté des applications mobiles, le projet précurseur Archeoguide [Vassilios et al. 2001] se proposait de faire découvrir les édifices grecs antiques pour la plupart détruits, en représentant ceux-ci en 3D à leur emplacement historique. Le dispositif mis en place reposait alors sur une paire de lunettes de type see-through et un GPS, les images étant calculées par un ordinateur portable placé dans un sac à dos porté par le visiteur. Le projet iTACITUS [Zoellner et al. 2007], plus récemment, a profité de l'évolution technologique pour prendre la forme d'un UMPC (Ultra mobile PC). Le positionnement est alors assuré par GPS en extérieur, associé à une détection de points clés dans l'image capturée par la caméra. Le contenu diffusé par ce dispositif est basé sur des gravures et images d'époques.

Les applications de réalité augmentée mobiles sont maintenant principalement destinées aux possesseurs de Smartphones. Le matériel n'étant alors pas la propriété du lieu culturel, il est envisageable de mettre en place des surfaces de visite très importante, de la taille d'une ville entière dans le cas de l'application StreetMuseum du Museum of London [Museum of London 2010]. Ce secteur est en plein expansion actuellement, et est soutenu par l'apparition de kits de développement qui réduisent considérablement le temps de développement d'applications de RA comme par exemple les produits de Layar (Anon.).

### 3 UNE IMPLEMENTATION ORIENTEE PHOTOREALISME

Une limitation technique forte des exemples donnés dans la section précédente réside dans les possibilités d'intégration des objets virtuels dans la scène réelle, sur un plan qualitatif. Cette intégration peut être caractérisée par deux éléments : la précision du positionnement des objets virtuels, et la cohérence visuelle entre l'image réelle et le rendu des objets virtuels.

La précision du positionnement des objets virtuels est dépendante des méthodes de capture de la position à notre disposition. Il peut s'agir d'un positionnement par GPS [Reitmay et Drummond 2006], ou un équivalent local lorsque celui-ci n'est pas disponible (voir les technologies de positionnement par triangulation Wifi ou équivalent). Ce n'est cependant pas suffisant pour avoir un calage correct des objets virtuels sur l'image réelle. Les solutions les plus avancées proposent donc d'y ajouter un suivi visuel de points clés sur l'image capturée, associé parfois à une centrale inertielle [Chandaria, Thomas, et Stricker 2007]. Ces méthodes donnent de bons résultats bien qu'étant limités à des environnements réduits.

Le second élément, la cohérence visuel, peut être abordé de deux manières différentes. On peut d'un côté opter pour un rendu non réaliste de l'image finale, en filtrant la capture de la scène réelle pour obtenir un style aisément reproductible pour le rendu de nos objets virtuels. Fischer et al. [Fischer, Bartz, et Strasser 2005] et Chen et al. [Chen, Turk, et MacIntyre 2008] ont choisi cette approche en reproduisant un style s'approchant du cel-shading (avec un rendu cartoon) consistant principalement à réduire la palette de couleurs.

L'autre choix est de simuler un rendu réaliste pour les objets virtuels, si possible en reproduisant les interactions entre ceux-ci et l'environnement virtuel dans lequel ils sont destinés à être placés. Debevec et al. [Callieri et al. 2005] et Pessoa et al. [S. A. Pessoa et al. 2003; S. Pessoa et al. 2010] ont mis en place des méthodes de reproduction de l'éclairage à partir d'images (ou IBR pour Image Based Rendering) associées à la simulation d'aberrations visuelles liées à la caméra telles que le flou lumineux.

C'est cette dernière direction que nous avons décidé de prendre pour développer notre dispositif de réalité augmentée. Ce choix a été largement justifié par le site pilote : l'église abbatiale de Cluny, en Bourgogne. Celle-ci a été en majeure partie détruite après la Révolution française, et il n'en subsiste qu'à peine huit pourcents. Un tel lieu justifie que l'on prenne soin de reproduire les jeux de lumière et les méthodes de type IBR sont actuellement celles offrant les meilleurs résultats.

### 4 SUPPORT MATERIEL DE NOTRE IMPLEMENTATION

Les critères en terme de photoréalisme induits par le contexte applicatifs nous ont poussé à passer outre les limites liées à la puissance de calcul à notre disposition. De plus les méthodes de reproduction de l'éclairage à partir d'images impliquent le stockage et le traitement d'une quantité de données importante, augmentant avec le nombre de degrés de liberté accordés à l'utilisateur pour ses déplacements. Aussi avons-nous opté pour un dispositif fixe nous permettant d'utiliser un ordinateur standard, sans contrainte spécifique.

Le dispositif prend la forme d'un écran de 32'' en format portrait orientable autour de deux axes, et est destiné à être installé sur site. L'orientation est modifiée par le visiteur simplement en guidant manuellement le dispositif. Ce mode d'utilisation impose au visiteur d'être face à l'écran, une main de part et d'autre de celui-ci. Nous avons alors une valeur approximative de sa distance à l'écran et donc du champ de vision couvert par l'écran. Cette donnée est utilisée pour reproduire un champ de vision équivalent dans notre image finale. En orientant précisément l'image à l'écran, les objets virtuels se retrouvent dans le prolongement des objets réels environnants et le dispositif devient une fenêtre vers une réalité alternative.



**Figure 1 : Le dispositif ray-on**

Outre par le contenu virtuel qui est, dans le cas de l'église abbatiale de Cluny, le fruit de recherches archéologique et historique, et d'un travail de reconstruction virtuelle détaillée, l'aspect réalité virtuelle est apporté de deux manières : premièrement en intégrant visuellement l'église (virtuelle) dans le Cluny actuel (capturé par une caméra) ; deuxièmement en reproduisant l'éclairage réel sur nos objets virtuels. A noter que selon la définition proposée par Milgram [Milgram et Jr 1999] ce dernier point se rapprocherait plutôt de la virtualité augmentée.



**Figure 2 : un dispositif pensé comme une fenêtre vers une réalité alternative**

#### **4.1 Mise en œuvre de la capture de la scène réelle**

Un intérêt supplémentaire de l'aspect fixe de notre dispositif tient dans la réduction de la complexité de mise en œuvre du mélange entre l'image réelle capturée par la caméra et l'image virtuelle simulée : la position de la caméra est fixe dans la scène et nous pouvons donc créer un masque permanent qui nous permettra ce mélange. Cette approche permet une précision du positionnement des objets virtuels qui n'est pas accessible par les méthodes de positionnement disponibles sur les appareils mobiles, et elle s'accorde avec notre but d'offrir une réalité augmentée photoréaliste.

La caméra utilisée pour la capture de la scène réelle est une caméra industrielle haute définition (Prosilica GE1910C), équipée d'une optique grand angle offrant un champ de vision de 95°. Les déformations optiques, particulièrement sensibles sur ce type d'optique grand angle, sont corrigées immédiatement après la capture. Ceci permet de faciliter par la suite l'intégration d'objets virtuels dans la vue réelle, ceux-ci étant rendus sans déformation.

Enfin, la connectique utilisée pour ce type de caméra est de type gigabit, c'est à dire qu'elle dispose d'un débit approximatif de 100Mo/sec pour transmettre les captures. Ce débit est tout juste suffisant pour avoir une fréquence de rafraichissement correcte en RGB 8bits, puisque nous atteignons alors environ 16 images par secondes. De plus la dynamique des luminances représentables par un tel format est très limitant dans le cas d'une scène extérieure puisqu'elle sera dans le meilleur des cas de 8 stops.

La solution adoptée est de récupérer la capture brute, avant conversion par la transformation de Bayer en une image en couleur. Cela procure deux avantages : tout d'abord la taille des images à transférer est bien moindre puisque en noir et blanc, soit un seul canal contre trois pour une image RGB ; ensuite notre caméra permet dans ce mode de capture d'accéder à non pas 8 mais 12 bits d'information par pixel et par canal. La transformation de Bayer est effectuée sur l'ordinateur, ce qui nous donne accès à la dynamique complète de la caméra, tout en réduisant considérablement le débit nécessaire pour chaque image puisque nous passons de 5.9Mo à 3.75Mo par image. La fréquence de rafraichissement est alors de 26 images par seconde.

#### 4.2 Capture de l'éclairage réel

La reproduction de l'éclairage réel passe tout d'abord par sa capture. Comme nous nous concentrons sur l'éclairage d'un édifice complet, une approximation valable est de supposer que la majeure partie de cet éclairage provient du ciel (durant la journée du moins). Nous avons donc opté pour une capture de l'éclairage à l'aide d'une caméra équipée d'un objectif de type fish-eye qui offre un champ de vision de 180°. Celle-ci est pointée en direction du ciel, et est située dans un endroit dégagé afin qu'aucun bâtiment ne fasse obstruction à la capture du ciel.

Si la dynamique (rapport de luminosité entre le point le plus sombre et le plus le plus lumineux) d'une scène extérieure est élevée, la dynamique du ciel l'est encore plus et dépasse souvent 20 stops (rapport de luminosité de  $2^{20}$ ). Ainsi, une caméra ou un appareil photo standard n'est pas capable de capturer correctement une telle image. La solution mise en place consiste à produire une image haute dynamique (HDRI, pour High Dynamic Range Image) en associant les informations tirées de plusieurs images pour lesquelles la valeur d'exposition varie de l'une à l'autre.

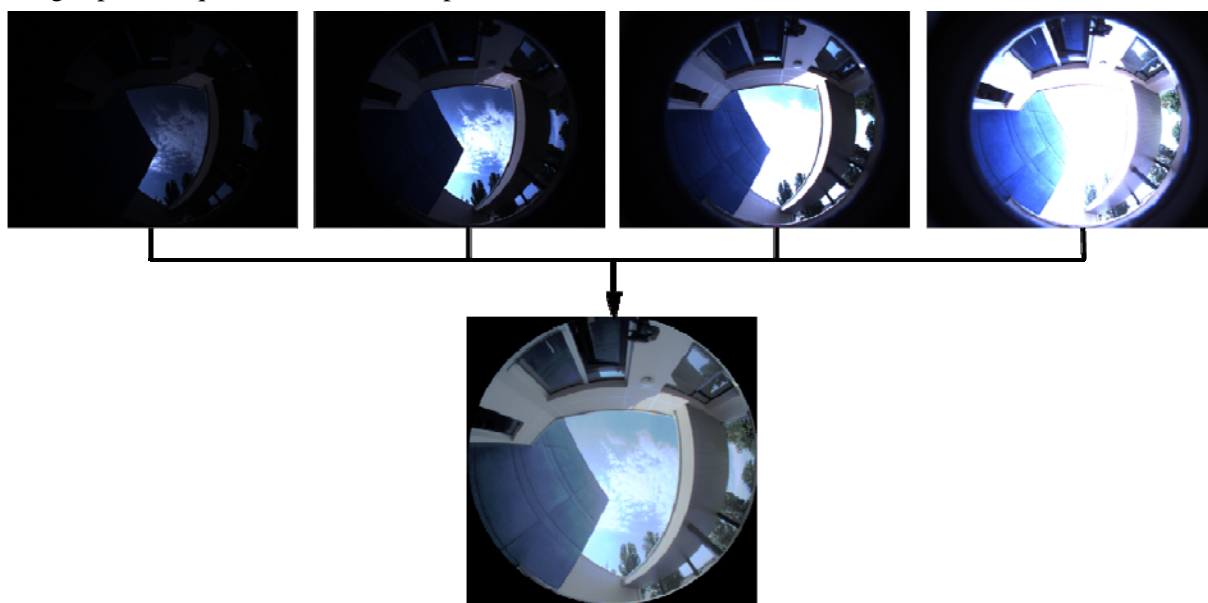


Figure 3 : création d'une image HDR à partir d'images à faible dynamique

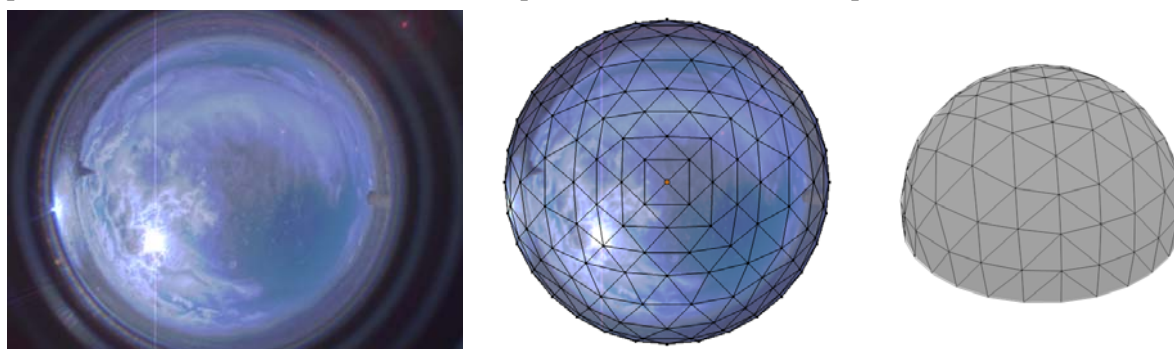
## 5 REPRODUCTION DE L'ECLAIRAGE ET MISE EN OEUVRE DE LA RA

Comme nous l'avons vu plus haut, nous avons choisi de nous orienter vers une méthode de reproduction précise de l'éclairage réel sur les objets virtuels. Notre méthode est de la famille des méthodes de rendu à base d'image (IBR). Celles-ci imposent de précalculer la contribution de différentes sources de lumière pour l'éclairage de la scène virtuelle, qui seront accumulées en fonction des luminances mesurées sur la capture de l'éclairage réel. En contrepartie, il devient possible d'utiliser toutes les possibilités des moteurs de rendu dédiés à la production d'images photoréalistes, qui ne sont normalement pas adaptés aux applications interactives.

## 5.1 Reproduction de l'éclairage réel

La reproduction de l'éclairage se déroule en deux phases : discrétisation de la capture de l'éclairage, selon la précision et la rapidité de rendu désirées ; puis accumulation des images précalculées selon les informations extraites de la capture de l'éclairage.

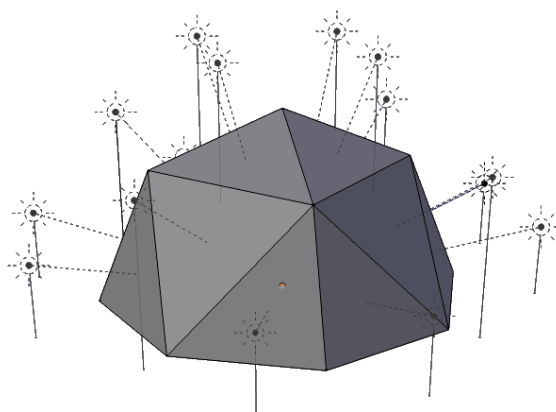
Nous avons utilisé la proposition de Robart et al [Robart 1999] pour notre discrétisation. Elle se base sur une pyramide à base carrée où chaque arête est subdivisée pour former sur chaque face triangulaire quatre nouveaux triangles. A chaque étape les sommets sont projetés sur la demi-sphère dans laquelle la pyramide de base s'inscrit. L'intérêt de cette méthode est de produire des facettes dont la surface reste sensiblement identique. Un exemple d'une telle subdivision est illustré sur la figure [figure]. A chaque nouvelle étape de subdivision, on se rapproche de l'image d'origine mais en contrepartie le nombre d'images à précalculer et à accumuler augmente : chaque facette correspond en effet à une direction d'éclairage qu'il faudra simuler. Durant nos essais, nous avons pu observer qu'une subdivision de niveau 6 (c'est à dire en subdivisant cinq fois la pyramide d'origine, donnant finalement 4096 facettes) produit un résultat visuellement très proche de l'utilisation de la capture de l'éclairage complète dans une résolution de 480x480 pixels. Nous passons donc d'une représentation de l'éclairage par environ 181000 directions (l'ensemble des pixels de la capture) à une représentation par 4096 directions, soit une division de la quantité de données à traiter par un facteur de 44.



**Figure 4 : de gauche à droite : la capture HDRI ; la discrétisation ; une vue de 3/4 de la demi-sphère discrétisée**

La base de données d'image est constituée d'autant d'images qu'il y a de facettes à notre discrétisation. Chaque image représente la contribution d'une source de lumière unitaire directionnelle, c'est à dire semblant venir de l'infini dans la direction définie par la facette associée. Le choix d'une source directionnelle est une approximation valable puisque nous ne nous occupons que du ciel en tant que source de lumière, qui peut être considéré comme étant à l'infini.

Les images sont accumulées selon un coefficient qui correspond à l'irradiance (en  $W.m^{-2}$ ) de chacune



**Figure 5 : ensemble des directions de lumière pour une subdivision de niveau 2**

des facettes. L'irradiance de chaque facette est calculée en multipliant la moyenne des luminances des pixels contenus dans chacune d'elles par leur angle solide. L'irradiance totale sur la demi-sphère est alors calculée comme suit :

$$E_{\Omega} = \sum_{i \in \Omega} \frac{S_i}{P_i} \sum_{p \in i} L_p$$

$E_{\Omega}$  étant l'énergie sur la demi-sphère,  $S_i$  et  $P_i$  l'angle solide et le nombre de pixel de la facette  $i$ , et  $L_p$  la luminance du pixel  $p$ .

Finalement nous obtenons une simulation de l'éclairage des objets virtuels par la capture de l'éclairage dans une qualité dépendant exclusivement de la précision du travail de modélisation et du rendu des images de notre base de données. Si celles-ci ont été calculées avec un objectif de photoréalisme, nous avons donc la possibilité d'appliquer l'éclairage réel sur une scène virtuelle avec une qualité photoréaliste.

## 5.2 Mise en œuvre sur site, à Cluny (Bourgogne)

Le site de l'église abbatiale de Cluny, qui a été avant celle de St Pierre de Rome la plus grande de la chrétienté, nous a permis d'expérimenter plusieurs scénarii de notre chaîne de réalité augmentée. Les dimensions de l'édifice, et le fait que seule une faible portion subsiste, justifient l'installation de plusieurs dispositifs sur le parcours de visite, en intérieur comme en extérieur. Deux approches ont été mises en œuvre.

La première consiste à intégrer l'église à un point de vue global de l'emplacement historique de celle-ci. L'emplacement choisi dans notre cas est en surplomb de la ville ce qui nous permet d'avoir une vue réellement globale sur l'édifice, tout en limitant le nombre d'objets mobiles susceptibles de passer devant, ce qui facilite l'intégration du rendu virtuel de l'église dans la capture de la caméra. Dans cette configuration le point de vue est fixe, toujours dans un souci d'obtenir finalement la meilleure intégration possible entre réel et virtuel.

Outre l'éclairage qui est reproduit sur la maquette virtuelle à l'identique de la scène réelle, nous avons également simulé les projections d'ombres entre les bâtiments réels et virtuels. Ceci prend la forme d'une étape supplémentaire dans notre chaîne de réalité augmentée puisque les bâtiments proches de l'église ont été modélisés (succinctement) puis un masque en niveau de gris décrivant la projection d'ombre de l'église aux alentours a été calculé pour chaque direction de lumière. Le résultat est une nouvelle base d'images qui est elle aussi accumulée en fonction de la capture de l'éclairage réel. Le résultat est ensuite utilisé pour obscurcir les zones virtuellement ombragées par l'édifice détruit.



**Figure 6 : Intégration de l'église abbatiale à une vue réelle de Cluny**

La seconde approche est moins conventionnelle en temps qu'application de réalité augmentée puisqu'elle ne comporte plus d'élément réel visible en tant que tel. Cette approche, utilisée pour deux points de vue intérieurs, permet d'offrir une vue sur ce que serait l'intérieur de l'église abbatiale si celle-ci était complète, à l'instant où le visiteur manipule le dispositif. L'interaction entre réel et virtuel se fait uniquement par la reproduction de l'éclairage dans une scène totalement virtuelle.

En revanche cette approche profite pleinement des possibilités d'orientation offertes par notre dispositif. Le visiteur a la possibilité d'orienter son regard dans toutes les directions, ce qui lui permet d'apprécier les phénomènes lumineux à l'intérieur de l'édifice. La reproduction de l'éclairage se fait dans ce cas sur un panorama complet centré sur la position du dispositif, la portion de panorama affichée à l'écran dépendant de l'orientation de celui-ci.



Figure 7 : panorama d'un dispositif en intérieur (voir Figure 1 : Le dispositif ray-on), éclairé par la capture en bas à gauche

## 6 CONCLUSION

Nous proposons dans ce papier une implémentation de réalité augmentée qui se concentre sur la production d'images photoréalistes. L'interactivité a été conservée dans la limite de ce qui est possible en conservant cet impératif de produire et diffuser des images réalistes sur notre dispositif. Une méthode laissant accès aux meilleurs moteurs de rendu d'image a donc été privilégiée.

Ainsi sur le site pilote de Cluny, le visiteur a ainsi face à lui une simulation réaliste de ce que pourrait être de nos jours l'église abbatiale détruite. Cette vue uchronique est renforcée par la reproduction réaliste des conditions lumineuses réelles, et la simulation des interactions lumineuses entre les édifices réels et virtuels. L'accueil pour les quatre dispositifs installés le long du parcours de visite de l'église, le premier ayant été installé durant l'été 2007, a été très bon aussi bien du côté des guides qui les utilisent pour illustrer leur discours, que des visiteurs pour qui la compréhension du lieu est facilitée.

## REFERENCES

- CITU (2008) « Telescope AR du CITU ». <http://www.citu.info/>. <http://www.citu.info/>.
- Fraunhofer Institute (2010) « Virtual Museum Guide ». <http://www.fraunhofer.de/en/press/research-news/2010/02/virtual-museum-guide.jsp>. <http://www.fraunhofer.de/en/press/research-news/2010/02/virtual-museum-guide.jsp>.
- Museum of London (2010) « StreetMuseum ». <http://www.museumoflondon.org.uk/streetmuseum.htm>. <http://www.museumoflondon.org.uk/streetmuseum.htm>.
- Total Immersion, <http://www.t-immersion.com/>. <http://www.t-immersion.com/>.
- Layar, <http://www.layar.com/>. <http://www.layar.com/>.
- Callieri, M., Paul E. Debevec, J. Pair, et R. Scopigno. (2005) « A realtime immersive application with realistic lighting: The Parthenon ».
- Chandaria, Jigna, Graham A. Thomas, et Didier Stricker. (2007) « The MATRIS project: real-time markerless camera tracking for Augmented Reality and broadcast applications ». *Journal of Real-Time Image Processing* 2: 69-79.



- Chen, Jiajian, Greg Turk, et Blair MacIntyre. (2008) « Watercolor Inspired Non-Photorealistic Rendering for Augmented Reality ». *VRST*: 231-234.
- Fischer, Jan, Dirk Bartz, et Wolfgang Strasser. (2005) « Stylized Augmented Reality for Improved Immersion ». *Virtual Reality (mars)*: 195-202.
- Milgram, Paul, et Herman Colquhoun Jr. (1999) « A Taxonomy of Real and Virtual World Display Integration ». *International Symposium on Mixed Reality*.
- Pessoa, S., G. Moura, J. Lima, V. Teichried, et J. Kelner. (2010) « Photorealistic rendering for Augmented Reality: A global illumination and BRDF solution ». In *Virtual Reality Conference*, 3-10.
- Pessoa, Saulo A., Eduardo L. Apolinario, Guilherme de S. Moura, et João Paulo S. do M. Lima. (2008) « Illumination Techniques for Photorealistic Rendering in Augmented Reality ». *ISMAR*: 223-232.
- Reitmay, Gerhard, et Tom W. Drummond. (2006) « Going out: Robust Model-based Tracking for Outdoor Augmented Reality ». *International Symposium on Mixed and Augmented Reality*: 109-118.
- Robart, Mathieu. (1999) « Simulation des interactions lumière - matière pour la modélisation de la réflectance par ondelettes en synthèse d'images réalistes ». Université de Toulouse 3. <http://cat.inist.fr/?aModele=afficheN&cpsidt=201292>.
- Vassilios, Vlahakis, Karigiannis John, Tsotros Manolis, Gounaris Michael, Almeida Luis, Stricker Didier, Gleue Tim, Christou Ioannis T., Carlucci Renzo, et Ioannidis Nikos. (2001) « ARCHEOGUIDE: First results of an Augmented Reality, Mobile Computing System in Cultural Heritage Sites ». *VAST*.
- Zoellner, Michael, Didier Stricker, Gabi Bleser, et Yulian Pastarmov. (2007) « iTacitus - Novel Interaction and Tracking Paradigms for Mobile AR ». *VAST*.