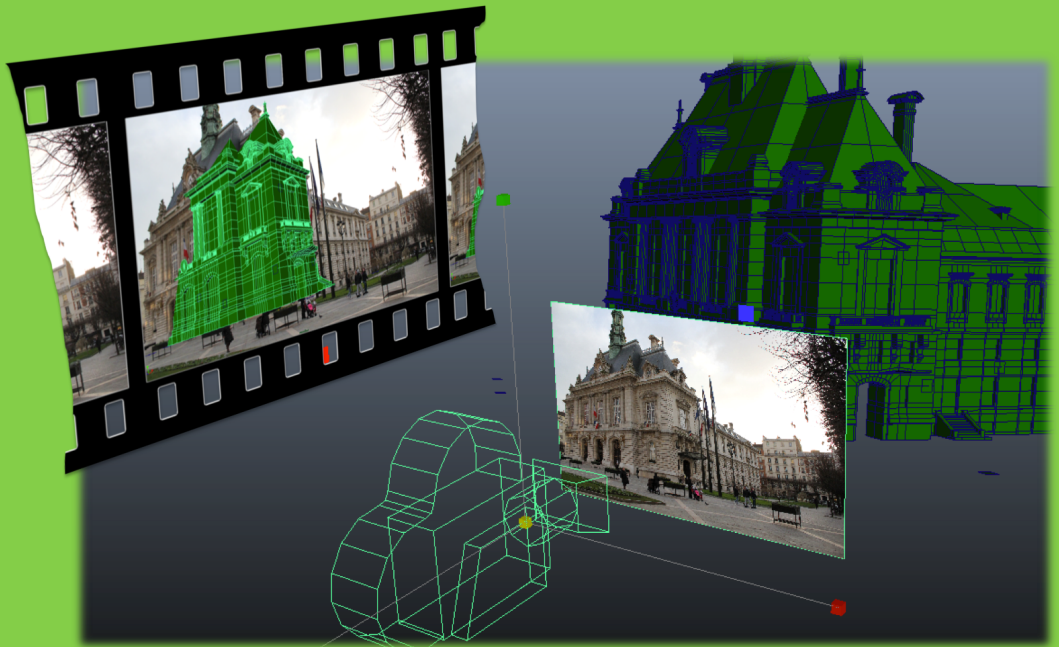


Année 2015

IRIT - ÉQUIPE VORTEX

- IMAGES ET 3D
- MULTIMÉDIA



Présentation

Le but de ce livret est de présenter aux visiteurs et partenaires potentiels les recherches menées durant l'année 2015 au sein de l'équipe VORTEX (Visual Objects from Reality to EXpression) de l'IRIT, ou plus précisément au sein de la partie de cette équipe localisée à l'ENSEEIH. Le style adopté est volontairement concis. Nous espérons cependant que le lecteur sera suffisamment intéressé pour proposer des collaborations qui, si elles s'avèrent fructueuses, feront à leur tour l'objet d'une rubrique dans l'édition 2016.

La première partie de ce livret regroupe les actions de recherche portant sur l'analyse d'images, la vision par ordinateur et la modélisation 3D en général. La deuxième partie décrit les actions de recherche relatives au multimédia, pour lesquelles le rôle de l'utilisateur est souvent central. Enfin, les projets industriels forment la troisième partie. Chaque rubrique est présentée sur deux pages : la page de gauche décrit une action de recherche ou un projet, celle de droite l'illustre par un ou plusieurs exemples.

Bonne lecture !

Liste des personnels de l'équipe VORTEX-ENSEEIH au 1^{er} décembre 2015

Permanents :

- Axel CARLIER, 05.34.32.22.00, axel.carlier@enseeiht.fr
- Sylvie CHAMBON, 05.34.32.21.78, schambon@enseeiht.fr
- Vincent CHARVILLAT, 05.34.32.21.71, charvi@enseeiht.fr
- Jean-Denis DUROU, 05.34.32.21.71, durou@irit.fr
- Emmanuel FAURE, 05.34.32.22.00, emmanuel.faure@irit.fr
- Simone GASPARINI, 05.34.32.21.72, sgaspari@enseeiht.fr
- Pierre GURDJOS, 05.34.32.21.77, pgurdjos@enseeiht.fr
- Zouhaier HAMROUNI, 05.34.32.21.75, hamrouni@enseeiht.fr
- Charlie KREY, 05.34.32.21.74, krey@enseeiht.fr
- Géraldine MORIN, 05.34.32.21.70, morin@enseeiht.fr

Non permanents :

- Vincent ANGLADON, vincent.angladon@etu.enseeiht.fr
- Clément AYMARD, blog.clm@gmail.com
- Chafik BAKEY, mohamedchafik.bakey@enseeiht.fr
- Marie-Anne BAUDA, marieanne.bauda@gmail.com
- Bastien DURIX, bastien.durix@enseeiht.fr
- Julien FAYER, julien.fayer@enseeiht.fr
- Thomas FORGIONE, thomas.forgione@irit.fr
- Damien MARIYANAYAGAM, damien.mariyanayagam@irit.fr
- Matthieu PIZENBERG, matthieu.pizenberg@enseeiht.fr
- Yvain QUÉAU, yvain.queau@enseeiht.fr
- Hatem RASHWAN, hatem.mahmoud@enseeiht.fr

Approche variationnelle différentielle de la stéréophotométrie

Mots-clés : méthodes variationnelles, EDP, parcimonie, reconstruction 3D, stéréophotométrie.

La résolution « classique » de la stéréophotométrie opère en deux étapes : estimation des normales, puis intégration des normales. Cela permet de décomposer le problème en deux sous-problèmes linéaires. Néanmoins, cette approche n'est pas vraiment adaptée à la prise en compte d'éclairages réalistes, et n'apporte aucune garantie sur la régularité de la surface estimée. Une façon d'éviter ce genre de difficultés numériques consiste à *estimer directement la profondeur*, en résolvant un système d'équations aux dérivées partielles (EDP) non linéaires issu du modèle photométrique.

Pour ce faire, nous suggérons une approche *variationnelle* de la stéréophotométrie différentielle, afin d'estimer une solution approchée du problème, optimale au sens d'un certain critère qui, s'il est judicieusement choisi, garantit la robustesse de l'estimation. Cette nouvelle méthode de résolution repousse certaines limites de la résolution classique en deux étapes : il n'est plus nécessaire d'imposer explicitement l'intégrabilité ; la robustesse de l'estimation est relativement facile à garantir ; les discontinuités de profondeur peuvent être retrouvées.

Nous introduisons, pour résoudre ce problème, un schéma numérique efficace fondé sur la combinaison d'un algorithme de point fixe et d'un algorithme ADMM. Ce schéma numérique s'étend facilement à des termes d'attache aux données et de régularisation parcimonieux (respectivement, la norme 1 et la variation totale), ce qui augmente encore la robustesse de l'estimation. Ceci nous permet finalement de proposer un cadre théorique et numérique original pour la résolution robuste de la stéréophotométrie.

Participants VORTEX : Jean-Denis DUROU, Yvain QUÉAU.

Partenaire : François LAUZE (DIKU, Copenhague, Danemark).

Publication associée :

Yvain QUÉAU, François LAUZE, Jean-Denis DUROU, *A L1-TV Algorithm for Robust Perspective Photometric Stereo with Spatially-Varying Lightings*, International Conference on Scale Space and Variational Methods in Computer Vision (SSVM 2015), Lège Cap Ferret, volume 9087 de Lecture Notes in Computer Science, pages 498-510, mai-juin 2015.



Exemple - Reconstruction 3D de l'avvers d'une pièce de 1 euro italienne. Première ligne : trois photographies (parmi cinq) de la pièce, prises sous le même angle, mais sous différents éclairages. Deuxième ligne : résultat de la reconstruction 3D par l'approche variationnelle différentielle de la stéréophotométrie.

Édition de surfaces par manipulation de leurs gradients

Mots-clés : méthodes variationnelles, intégration d'un champ de gradient, EDP, stéréophotométrie.

À partir d'une estimation du gradient d'une surface, il est possible de retrouver cette dernière en résolvant un problème d'*intégration*. Les méthodes variationnelles se prêtent particulièrement bien à la résolution de ce problème, car elles permettent de garantir une certaine robustesse au bruit et aux données aberrantes. Il est de plus relativement aisé d'adapter ces techniques à d'autres problèmes, notamment celui de l'édition de surfaces.

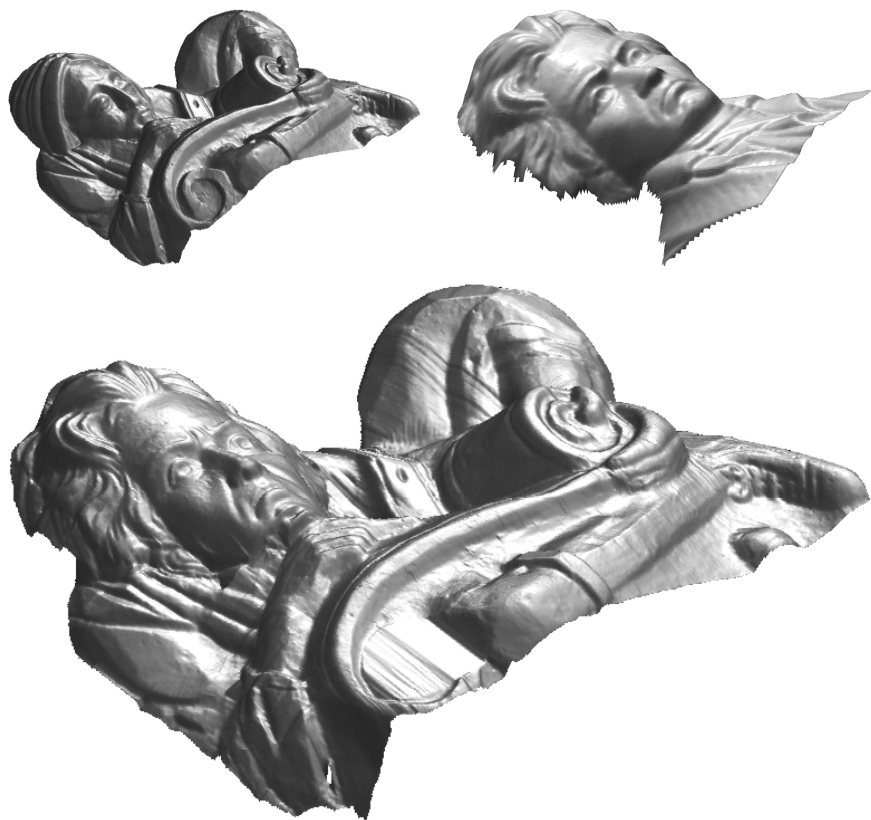
En effet, puisque le gradient d'une surface encode les variations les plus fines du relief, il est possible de modifier l'apparence d'un relief en changeant son gradient, avant de l'intégrer par les techniques mentionnées ci-dessus. On peut par exemple « gommer » certains détails en annulant le gradient sur ces détails, ou remplacer les valeurs du gradient, en un certain nombre de points, par de nouvelles valeurs. Dans les deux cas, le choix d'une méthode variationnelle appropriée permet de lisser les artéfacts aux frontières de la zone modifiée.

En adaptant des techniques bien connues en traitement d'images, notamment la minimisation en norme 1 ou en moindres carrés pondérés, nous avons montré qu'il était possible de garantir l'aspect lisse de la surface reconstruite, tout en préservant les éventuelles arêtes et discontinuités de profondeur.

Participants VORTEX : Jean-Denis DUROU, Yvain QUÉAU.

Publication associée :

Yvain QUÉAU, Jean-Denis DUROU, *Edge-Preserving Integration of a Normal Field: Weighted Least Squares and L1 Approaches*, International Conference on Scale Space and Variational Methods in Computer Vision (SSVM 2015), volume 9087 de Lecture Notes in Computer Science, pages 576-588, Lège Cap Ferret, mai-juin 2015.



Exemple - Première ligne : deux surfaces. Deuxième ligne : fusion variationnelle, fondée sur le gradient de profondeur, des deux surfaces. La tête de la première statue est remplacée par celle de la seconde, en préservant les détails du relief et sans que les raccords soient visibles.

Reconstruction 3D d'un objet par le squelette

Mots-clés : squelette, reconstruction 3D, topologie.

Nous proposons dans ce travail une approche de la reconstruction 3D d'un objet par le squelette. Un squelette est un modèle de représentation des formes qui correspond à une structure, curviligne ou surfacique, centrée à l'intérieur de celles-ci. En associant un rayon à chacun des points de cette structure, il est possible de retrouver complètement la forme. L'idée est donc de reconstruire le squelette 3D curviligne d'un objet à partir d'un ensemble d'images. Une fois le squelette 3D reconstruit, celui-ci peut être habillé d'un maillage pour visualiser la reconstruction 3D.

Afin de reconstruire le squelette 3D d'un objet, nous considérons un ensemble de points de vue sur celui-ci, calibrés de façons intrinsèque et extrinsèque. Nous estimons ensuite un squelette dit *projectif* pour chacune des images acquises, qui correspond à une projection du squelette 3D. Les branches de ces squelettes sont enfin associées entre les différentes images, pour permettre la triangulation branche par branche d'un squelette 3D.

Cette méthode présente l'avantage de générer une reconstruction complète donnant la topologie de l'objet, ainsi qu'un maillage exploitable. De plus, la connaissance du squelette 3D de l'objet permet de simplifier la recherche d'un squelette d'animation.

Participants VORTEX : Sylvie CHAMBON, Bastien DURIX, Géraldine MORIN.

Participants extérieurs : Lionel GARNIER, Céline ROUDET (Le2i, Université de Bourgogne).

Publication associée :

Bastien DURIX, Géraldine MORIN, Sylvie CHAMBON, Céline ROUDET, Lionel GARNIER, *Towards Skeleton based Reconstruction: From Projective Skeletonization to Canal Surface Estimation*, 3D Vision (3DV 2015), Lyon, octobre 2015.



Reconstruction 3D à partir d'un ensemble d'acquisitions - Première colonne : ensemble de cinq acquisitions calibrées d'une peluche. Deuxième colonne : estimation du squelette projectif de la peluche pour chacune des images. Troisième colonne : projection de l'estimation du maillage et du squelette 3D de la peluche sur les différentes images.

Évaluation d’algorithmes d’odométrie visuelle temps réel utilisant un capteur de profondeur

Mots-clés : odométrie visuelle, Smartphone, RGB-D, capteur de profondeur.

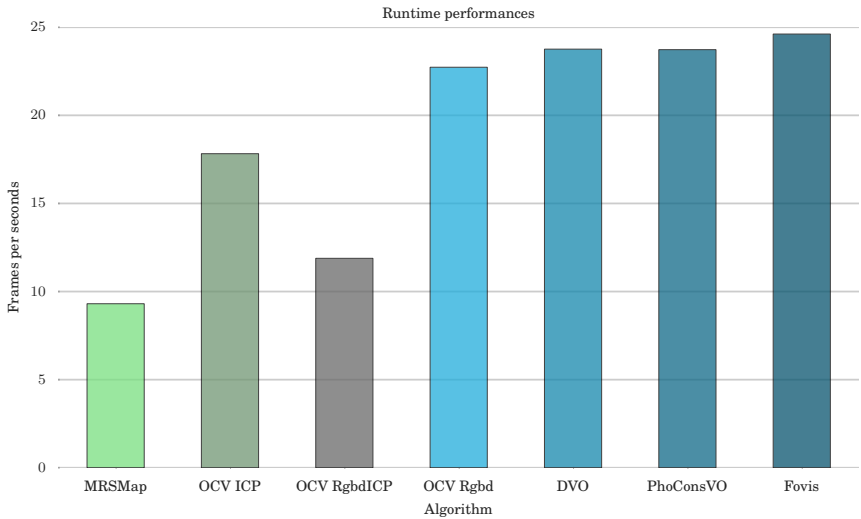
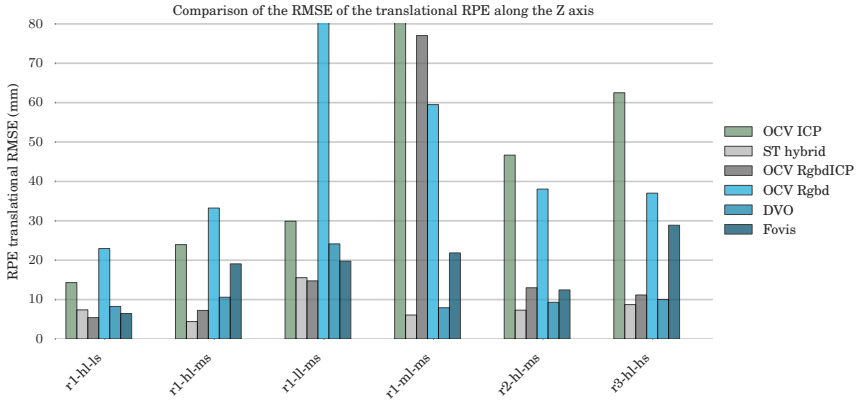
L’odométrie visuelle permet de connaître en temps réel la position et l’orientation de la caméra. C’est un composant essentiel pour toute application axée autour de la localisation ou de la perception de son environnement. Ce problème a été étudié sur PC par l’utilisation de capteurs de profondeur tels que la Kinect, ce qui évite l’ambiguïté sur le facteur d’échelle lors de la phase d’initialisation des méthodes monoculaires. Toutefois, très peu de solutions ont été proposées pour les Smartphones et les tablettes, qui commencent aujourd’hui à être équipés de tels capteurs.

Dans nos travaux, nous comparons la performance de ces algorithmes avec celui fourni par le SDK du capteur Structure Sensor d’Occipital sur des appareils mobiles et sur PC. Nous évaluons la précision RPE (*relative pose error*) des solutions proposées, le temps CPU et la charge CPU consommée, ainsi que la quantité de mémoire utilisée. Nos résultats présentent les algorithmes à privilégier en fonction du type de scène et de la précision souhaitée. Ces travaux confirment qu’il reste des progrès à faire en termes de recherche et d’optimisation logicielle, afin d’obtenir des solutions qui soient à la fois précises et temps réel pour un Smartphone.

Participants VORTEX : Vincent ANGLADON, Vincent CHARVILLAT, Simone GASPARINI.

Partenaire : société Telequid.

Publication en cours



Évaluation des performances - En haut : comparaison de la précision RPE des algorithmes étudiés sur des données issues du capteur Structure Sensor. Dans le cas présent, ST hybrid et OCV RgbdICP sont les plus précis. En bas : nombre d'images par seconde pouvant être traitées par chacune des solutions étudiées. Trois algorithmes n'atteignent pas le temps réel sur PC : MRSSMap, OCV ICP et OCV RgbdICP.

Décomposition non supervisée de formes géométriques 2D et recherche de similarités

Mots-clés : modèles d'éclairage, calibrage de caméra, reconstruction 3D, stéréophotométrie.

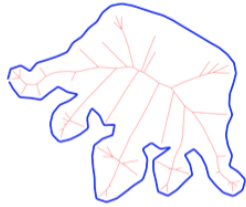
Nous avons développé un processus non supervisé pour l'analyse de la structure des formes 2D, qui utilise les propriétés intrinsèques de l'axe médian. Une telle analyse est importante pour de nombreuses applications telles que la comparaison de formes, l'édition ou la compression. Notre méthode est robuste et stable. Elle décompose une forme en plusieurs niveaux de hiérarchie et mesure la similarité entre parties en se fondant sur une distance à l'axe médian appelée WEDF (*weighed extended distance function*). Cette partition multi-résolution de la forme est invariante aux changements d'échelle et aux articulations. Une étude approfondie menée sur un ensemble d'utilisateurs avec une base de données de formes standard (*MPEG-7 shape 1*) montre que, d'un point de vue géométrique, nos résultats sont extrêmement cohérents avec la perception de l'utilisateur.

Participants VORTEX : Axel CARLIER, Géraldine MORIN.

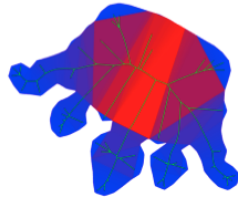
Participants extérieurs : Marie-Paule CANI, Stefanie HAHMANN (INRIA Rhône-Alpes), Lisa LARSSON (McGill University, Montréal, Canada), Kathryn LEONARD (California State University, Channel Islands, États-Unis).

Publication associée :

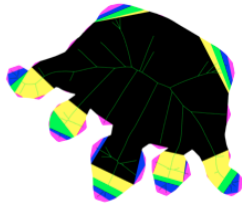
Lisa LARSSON, Géraldine MORIN, Antoine BEGAULT, Raphaëlle CHAINE, Jeannine ABIVA, Évelyne HUBERT, Monica HURDAL, Mao LI, Beatriz PANIAGUA, Giang TRAN, Marie-Paule CANI, *Identifying Perceptually Salient Features on 2D Shapes*, Research in Shape Modeling, pages 129-153, Springer International Publishing, 2015.



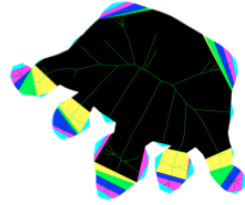
Forme 2D et axe médian



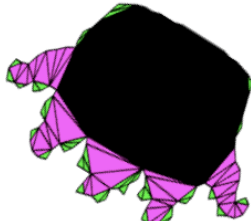
WEDF



Classification basée WEDF

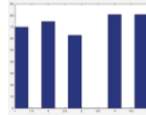


Hiérarchie

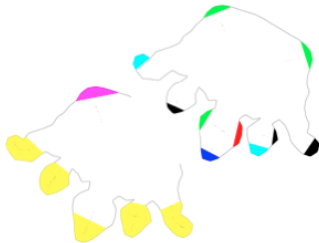


Annotation
de l'étude utilisateur

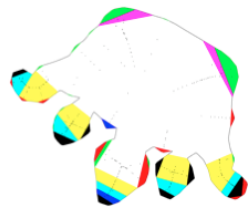
75 %



Score de similarité



Similarités
(niveaux 2 et 4 de la hiérarchie)



Similarités
(tous niveaux de hiérarchie)

Exemple - Segmentation et analyse de similarité d'une forme 2D.

Détection *a contrario* de lignes polygonales et d'arcs d'ellipses

Mots-clés : détection, lignes polygonales, arcs d'ellipses, approche *a contrario*, sélection de modèles.

Nous proposons une version continue de l'algorithme ELSD, dédié aux détections combinées de lignes polygonales et d'arcs elliptiques dans une image en niveaux de gris. Comme ELSD, le nouveau détecteur (ELSDc) contrôle le nombre de faux positifs et ne requiert aucun réglage de paramètre au regard d'images de natures très différentes. Il conserve le même *modus operandi* en trois étapes :

1. génération de régions candidates ;
2. validation de candidats en tant que supports de primitives discrétisées de type (i) ligne polygonale ou (ii) arc d'ellipse ;
3. sélection du meilleur modèle (i) ou (ii) pour les candidats validés.

Nous proposons un test statistique *a contrario*, utilisé à la fois pour la validation (étape 2) et pour la sélection de modèles (étape 3). Ce test calcule le nombre de fausses alarmes (NFA), qui est la probabilité pour que le degré de structuration (i) ou (ii) d'une région candidate, rapportée au nombre théorique maximal de candidats, ne soit pas dû au hasard. La spécificité de ELSDc est une formulation continue du NFA. Celle-ci permet d'obtenir de meilleures performances que la formulation discrète précédente du NFA, vis-à-vis des tâches de validation et de sélection, en comparaison des algorithmes de l'état de l'art [Etemadi, ICIP 1992] [Prasad et al., Pattern Recognition 2012] évalués à partir d'images synthétiques et d'images réelles.

Participants VORTEX : Pierre GURDJOS.

Participants extérieurs : Rafael GROMPONE VON GIOI (CMLA, ENS Cachan), Viorica PĂTRĂUCEAN (Université de Cambridge, Royaume-Uni).

Publication associée :

Viorica PĂTRĂUCEAN, Pierre GURDJOS, Rafael GROMPONE VON GIOI, *A Joint Ellipse and Line Detection Using A Contrario Theory* (soumis).

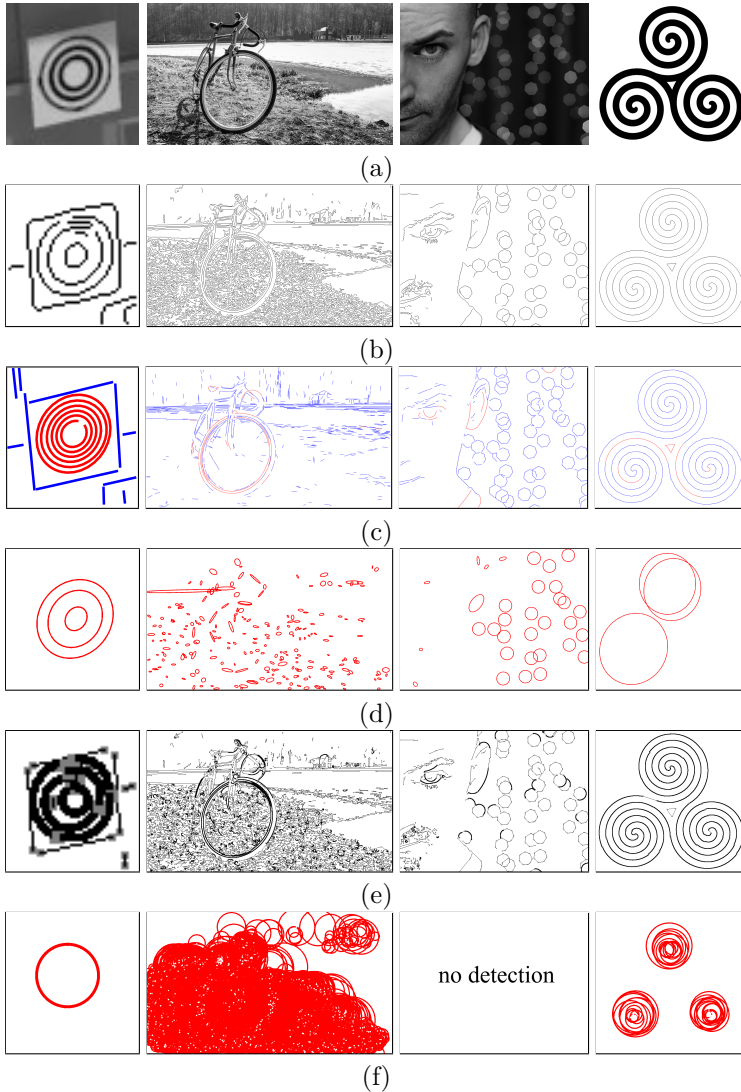


Illustration - (a) Images originales. (b) Images des points de contour. (c) Détecteur ELSDc (ellipses détectées en rouge, lignes polygonales détectées en bleu). (d) Détecteur de [Prasad 2012] (ellipses détectées en rouge). (e) Détecteur de [Etemadi 1992] (segments de droites détectés et, en gras, arcs de cercles détectés). (f) Détecteur de Hough (cercles détectés en rouge).

Détection et identification de codes à barres circulaires dans des conditions critiques

Mots-clés : marqueurs artificiels, codes à barres, détection, identification, suivi de caméra.

Dans le cadre du projet européen POPART, nous avons développé un système de marqueurs artificiels, appelés C^2 Tags, fondé sur l'utilisation de motifs plans en noir et blanc représentant des anneaux concentriques. Ces marqueurs sont conçus pour encoder « naturellement » trois points, à savoir le centre et les points cycliques, ce qui permet de calculer la pose et est utile dans toute application nécessitant le suivi de la caméra. Le système comprend des algorithmes précis de détection et d'identification des marqueurs, qui sont robustes dans des conditions d'acquisition critiques, par exemple : lorsque les marqueurs sont occultés ; lorsqu'ils sont situés sur des plans très inclinés par rapport à l'axe optique ou très éloignés de la caméra ; lorsque l'éclairage est faible ou irrégulier ; lorsqu'un déplacement rapide de la caméra provoque un flou de bougé.

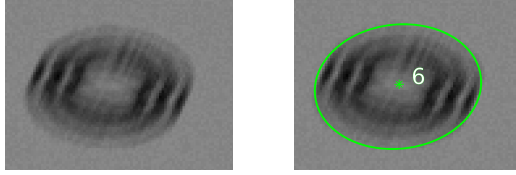
L'originalité de notre système consiste en une approche théorique de la détection, à partir d'une procédure de vote faisant intervenir les lignes de champ du gradient de l'image, et en une formulation du problème de rectification métrique de l'image restreinte à deux degrés de liberté (à savoir, les coordonnées du centre). Nous avons montré que ce système obtenait des performances au moins équivalentes aux systèmes existants, en particulier ARToolkitPlus [Prasad et al., WACV 2015] et RuneTag [Bergamasco et al., CVPR 2011], dans des conditions d'acquisition « favorables », voire nettement supérieures lorsque celles-ci deviennent critiques.

Participants VORTEX : Simone GASPARINI, Pierre GURDJOS.

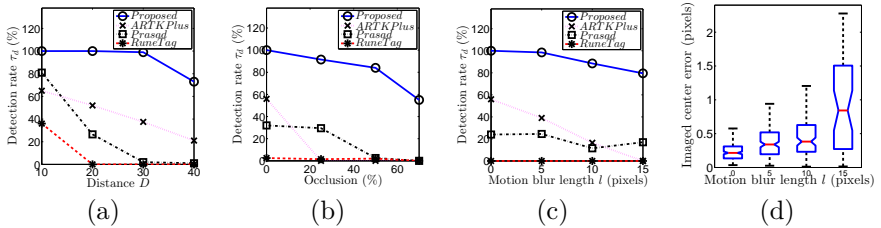
Participants extérieurs : Lilian CALVET, Carsten GRIWODZ (laboratoire Simula Research, Oslo, Norvège).

Publication associée :

Lilian CALVET, Pierre GURDJOS, Simone GASPARINI, Carsten GRIWODZ, *Detection and accurate localization of Circular Barcode under Highly Challenging conditions* (soumis).



Exemple - Détection réussie malgré un flou de bougé de 15 pixels.



Comparaison du système C²Tag aux systèmes existants - Taux de détection en fonction de : (a) la distance à la caméra ; (b) le pourcentage d'occultation ; (c) l'amplitude du flou de bougé. (d) Estimation de l'image du centre en fonction de l'amplitude du flou de bougé.



Mise en correspondance de marqueurs C²Tag et ARToolkitPlus - Seuls les marqueurs C²Tag sont correctement mis en correspondance.

Un réseau de distribution de copies numériques pour le cinéma

Mots-clés : distribution dématérialisée, pair à pair, VPN.

Les salles de cinéma sont passées à l'ère numérique. Pour la distribution des copies numériques de films, la DCI (*Digital Cinema Initiative*) a choisi d'encoder les films au format DCP (*Digital Cinema Package*). Notre activité est menée en collaboration avec les cinémas Utopia, dont la principale exigence est de fournir des logiciels libres et gratuits pour le cinéma numérique.

Nous avons mis en place un système de distribution de type CDN (*Content Delivery Network*) permettant de transférer les DCP des distributeurs vers les salles de cinéma. Il vise à proposer une alternative adaptée aux salles d'art et essai. Le système est opérationnel : il est actuellement déployé dans une trentaine de salles en France.

Ce CDN possède les caractéristiques d'un CDN classique (organisation hiérarchique des serveurs d'origine vers les salles de cinéma), mais aussi celles du fonctionnement en pair à pair (la distribution des contenus est réalisée par le protocole BitTorrent).

Participants VORTEX : Nicolas BERTRAND, Vincent CHARVILLAT, Jean-Denis DUROU.

Partenaires : cinémas Utopia, associations Tetaneutral.net, Cinémascop et FFDN.

Publication associée :

Nicolas BERTRAND, Jean-Denis DUROU, Vincent CHARVILLAT, *The DCP Bay: Toward an Art-House Content Delivery Network for Digital Cinema*, International Conference on Advances in Multimedia (MMEDIA 2015), Barcelone, Espagne, avril 2015.

The **DCP** Bay



Share a slice of the Internet !

Free Software, Net Neutrality, Interoperability
Peer to Peer, Virtual Private Network (VPN),
and a Wild Bunch of old school Nerds !



www.tdcpb.org

The DCP Bay - Principales vertus du CDN mis en place pour la distribution de copies numériques : logiciel libre, neutralité de l'Internet et interopérabilité.

Impact de signets sur la navigation dans un environnement 3D

Mots-clés : navigation 3D, *streaming*.

Dans ce travail, nous avons développé une interface permettant à des utilisateurs peu expérimentés de naviguer plus facilement dans une grande scène 3D, grâce à la présence de signets (qui matérialisent des points de vue recommandés). Ces signets sont affichés sous la forme de petits objets 3D qui se superposent à la scène, et qui proposent un déplacement à l'utilisateur. En cliquant sur un signet, la caméra de l'utilisateur va se déplacer jusqu'à un nouveau point de vue.

Nous avons montré que la présence de signets aidait les utilisateurs à accomplir une tâche de manière plus efficace : la recherche d'éléments cachés dans une scène est en moyenne deux fois plus rapide. La contrepartie de cette vitesse de navigation réside dans une moindre qualité de navigation, due à un *streaming* plus gourmand. En effet, un utilisateur qui navigue plus vite demandera plus de faces dans un intervalle de temps donné. En moyenne, l'utilisateur subira plus souvent l'absence de faces qu'il aurait dû voir, et donc observera une dégradation de qualité de l'expérience.

Cependant, nous avons aussi montré que la présence de signets conditionnait les utilisateurs, de sorte que le fait de savoir que l'utilisateur est susceptible de cliquer un signet permet d'améliorer le système de trois manières :

- Lorsqu'un utilisateur navigue dans une scène, il est susceptible d'avoir un comportement similaire aux utilisateurs qui ont navigué avant lui, et notamment de cliquer sur les mêmes signets dans le même ordre.
- Lorsqu'un utilisateur clique sur un signet, la caméra va se déplacer de manière fluide jusqu'à la position recommandée et le temps de ce mouvement peut être utilisé pour précharger la zone d'arrivée.
- S'il n'est pas envisageable de faire du rendu côté serveur pour savoir quelles sont exactement les faces visibles (impossibilité de passer à l'échelle), il est en revanche possible de précalculer les rendus sur les positions des signets, de façon à optimiser le préchargement de celles-ci.

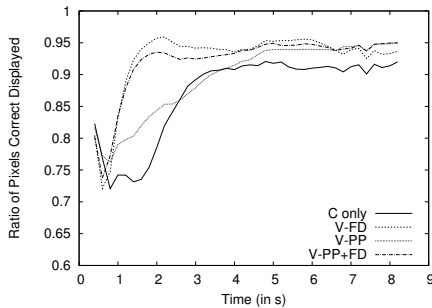
Participants VORTEX : Axel CARLIER, Vincent CHARVILLAT, Thomas FORGIONE, Géraldine MORIN.

Participant extérieur : Wei Tsang OOI (National University of Singapore, Singapour).

Publication en cours



Illustration - Exemple de scène avec des signets.



Résultats - Courbes représentant le nombre moyen de pixels rendus correctement pendant les quelques secondes qui suivent un clic sur un signet, pour différentes politiques de préchargement :

- *C only* : sans tenir compte des signets.
- *V-FD* : en préchargeant uniquement la destination lors d'un clic sur un signet.
- *V-PP* : en tentant de prédire le prochain signet cliqué et de le précharger.
- *V-PP+FD* : en tentant de prédire le prochain signet cliqué et de le précharger, et en chargeant la destination lors d'un clic sur un signet.

Préchargement de signets pour une interface de vidéo à la demande

Mots-clés : interface vidéo, préchargement, signets, *video browsing*.

Avec l'augmentation constante du nombre de vidéos mises en ligne sur Internet (cinq nouvelles heures de vidéo apparaissent sur Youtube chaque seconde) et l'avènement des réseaux sociaux, le comportement des utilisateurs évolue vers une consommation toujours plus rapide des contenus. Une étude de Chen et al. de 2014 qualifie ce nouveau comportement de *video browsing* et montre que les utilisateurs ne regardent pas dans leur totalité 80% des vidéos, et que 60% des sessions comportent au moins un accès non linéaire.

Face à ce constat, les politiques de transmission de contenu deviennent obsolètes. Sur Youtube par exemple, le téléchargement d'une vidéo est linéaire et adapté à un faible niveau d'interactions de la part des utilisateurs.

Dans ce travail, nous proposons de modifier la transmission des contenus vidéo en tirant parti de la présence de signets sur la barre de navigation temporelle de l'interface. Ces signets, qui peuvent être déterminés automatiquement (par exemple, par la détection d'événements spécifiques) ou inférés à partir des interactions des utilisateurs, indiquent des moments-clés d'une vidéo. Ils sont représentés par des points blancs sur la figure. La littérature montre que les signets concentrent naturellement une proportion importante des accès non linéaires des utilisateurs.

Nous montrons que cette proportion d'accès non linéaires aux signets peut être encore plus importante en modifiant la politique de préchargement des vidéos, afin de télécharger en priorité les premières secondes de chaque signet. Le niveau de préchargement est indiqué en gris clair sur la figure (à la suite des signets). L'effet de ces modifications se traduit par une meilleure qualité de service. En particulier, la latence subie lors d'accès non linéaires est en moyenne beaucoup plus faible avec notre méthode.

Participants VORTEX : Axel CARLIER, Vincent CHARVILLAT.

Participant extérieur : Wei Tsang OOI (National University of Singapore, Singapour).

Publication associée :

Axel CARLIER, Vincent CHARVILLAT, Wei Tsang OOI, *A Video Timeline with Bookmarks and Prefetch State for Faster Video Browsing*, ACM Multimedia (MM 2015), pages 967-970, Brisbane, Australie, octobre 2015.



Capture d'écran de l'interface vidéo - Les signets sont représentés par des points blancs, et leur niveau de téléchargement apparaît en gris clair. Un survol des signets laisse apparaître une prévisualisation de la vidéo à cet instant.

Crowdsourcing et ludification en segmentation interactive

Mots-clés : ludification, *crowdsourcing*, segmentation interactive.

La segmentation interactive est généralement conduite par un utilisateur expert (cf. figure (a)), mais nous avons étudié dans des travaux antérieurs comment l'utilisation de techniques dites de *crowdsourcing* pouvaient apporter des résultats d'une qualité presque comparable aux experts. Nous avons également présenté un jeu, *Ask'nSeek*, qui permet de collecter des données pouvant ensuite être utilisées pour la segmentation.

Dans ce travail, nous comparons les performances de ces deux approches : *crowdsourcing* d'une part, « ludification » d'autre part. L'utilisation du *crowdsourcing* permet de disposer rapidement, et à moindre coût, d'une grande quantité de données. En contrepartie, une part non négligeable de ces données sont erronées (cf. figure (b)) et doivent donc être traitées avec précaution. Nous montrons également que certains utilisateurs recrutés sur des plateformes de *crowdsourcing*, et donc non experts, ont un niveau de performance potentiellement supérieur à celui des experts (cf. figure (c)).

Nous discutons également des avantages et des inconvénients de la ludification. Un de ses avantages est que les données collectées sont très peu bruitées : en effet, pour pouvoir gagner une partie, les utilisateurs sont obligés d'apporter des informations correctes pendant la partie. En revanche, le mécanisme du jeu encourage les utilisateurs à fournir des données très redondantes, qui sont insuffisantes pour obtenir une bonne segmentation (cf. figure (d)). Un des enjeux majeurs de la ludification est donc d'inclure un mécanisme de jeu qui favorise la diversité des données collectées.

Participants VORTEX : Ferran CABEZAS (ex-stagiaire VORTEX), Axel CARLIER, Vincent CHARVILLAT.

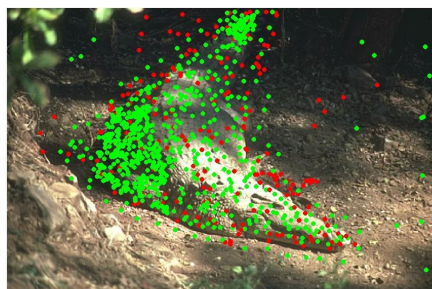
Participants extérieurs : Xavier GIRO-I-NIETO, Amaia SALVADOR (Universitat Politècnica Catalunya, Barcelone, Espagne), Oge MARQUES (Florida Atlantic University, Boca Raton, États-Unis).

Publication associée :

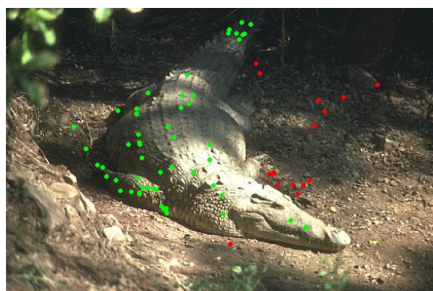
Axel CARLIER, Amaia SALVADOR, Ferran CABEZAS, Xavier GIRO-I-NIETO, Vincent CHARVILLAT, Oge MARQUES, *Assessment of crowdsourcing and gamification loss in user-assisted object segmentation*, Multimedia Tools and Applications, septembre 2015.



(a)



(b)



(c)



(d)

Exemple de données de segmentation interactive collectées - Les clics sur l'objet à segmenter sont indiqués en vert, ceux sur l'arrière-plan sont indiqués en rouge. Ils ont été effectués par : (a) des experts; (b) des utilisateurs non experts (*crowdsourcing*); (c) les meilleurs utilisateurs non experts; (d) des joueurs.

Une représentation basée image efficace pour les environnements 3D virtuels en ligne

Mots-clés : représentation basée image, 3D *image-warping*, environnements 3D virtuels en ligne.

Nous proposons une nouvelle représentation basée image efficace pour les environnements 3D virtuels en ligne, appelée *sprite tree*. Cet arbre (de type *octree*) stocke des *sprites* (imassettes pouvant être déplacées par rapport au fond de l'écran) issus de plusieurs images de référence, et permet de calculer le rendu de scènes virtuelles complexes plus efficacement qu'un rendu classique. La création d'un *sprite tree* est effectuée à partir d'un ensemble d'images RGBD obtenues par des utilisateurs effectuant le rendu d'une scène 3D, et d'un algorithme de rendu utilisant ces données. Les résultats montrent que, d'une part, le *sprite tree* code efficacement la scène 3D originale à partir de centaines d'images de référence différentes et, d'autre part, que le rendu obtenu est plus efficace qu'un rendu classique, tout en garantissant une bonne qualité visuelle. Nous proposons, en outre, une mesure de similarité des *sprites* qui permet de considérablement réduire : les artéfacts d'éclairage dans les images rendues ; la redondance, et donc la taille de l'arbre, avec une perte non significative de la qualité visuelle.

Participants VORTEX : Vincent CHARVILLAT, Géraldine MORIN.

Participants extérieurs : Wei Tsang OOI, Minhui ZHU (Université Nationale de Singapour, Singapour).

Publication associée :

Minhui ZHU, Géraldine MORIN, Vincent CHARVILLAT, Wei Tsang OOI, *Sprite Tree: An Efficient Image-based Representation for Networked Virtual Environments* (soumis).



(a) Image de fond.



(b) Premier plan.



(c)



(d)



(e)



(f)

Création de *sprites* à partir d'une image RGBD - Image de référence composée (a) d'une image de fond et (b) d'un premier plan. (c-f) Quatre des *sprites* créés pour modéliser le premier plan, insérés dans l'arbre de *sprites*.



(g)



(h)

Comparaison des résultats - (g) Rendu à partir de l'arbre de *sprites*. (h) Vérité terrain rendue à partir du modèle 3D original. Le temps de rendu moyen est de 70 *ms* pour le rendu classique, et de 40 *ms* pour le rendu à partir de l'arbre de *sprites*.

Données pour l'évaluation de la résistance des algorithmes de tatouage vidéo au *camcording*

Mots-clés : tatouage d'images, *camcording*.

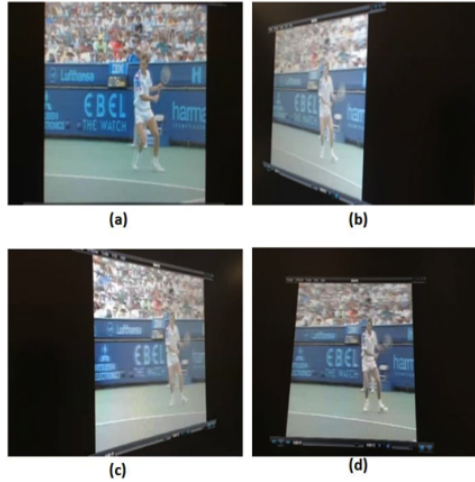
Avec le développement des techniques de traitement numérique, il est nécessaire d'inventer de nouvelles techniques permettant de sécuriser les médias. Le tatouage vidéo, qui consiste à insérer une marque invisible, robuste aux attaques, va dans ce sens. Or, si la robustesse des algorithmes existants de tatouage vidéo est souvent testée avec des attaques simples telles que la rotation, la translation ou l'ajout de bruit, il semble nécessaire de la tester avec des attaques plus sophistiquées. La plus connue de ces attaques est le *camcording*, qui consiste à filmer la vidéo à l'aide d'un Smartphone ou d'une caméra, puis à la diffuser après application de certaines transformations visant à détruire les marques.

En étudiant le comportement de vrais pirates suite à la mise en ligne du film *Jurassic World*, nous avons constaté qu'après *camcording*, le pirate applique souvent des traitements à la vidéo avant de la partager (*cropping*, compression, ajout de couleurs, rotation, ...). Nous avons donc eu l'idée de proposer un nouveau protocole d'évaluation de la robustesse des méthodes de tatouage vidéo face à des attaques de type *camcording*. Pour ce faire, nous avons filmé des vidéos préalablement tatouées, projetées dans une salle de cinéma mise à notre disposition. Nous avons utilisé cinq appareils de capture (un caméscope Panasonic, un caméscope Canon, un Smartphone GalaxyNote2, un Smartphone Iphone 4 et une webcam Logitech) et quatre angles de vue (depuis la cabine de projection; depuis la salle de cinéma, à droite, à gauche ou en face de l'écran). Les vidéos projetées sont deux dessins animés au format « Full HD » (*Big Buck Bunny* et *Sintel*), la vidéo *Stéphan*, un clip vidéo et une publicité. Ces vidéos ont été tatouées par l'algorithme multi-fréquentiel DWT-SVD-DCT, par un algorithme DWT-SVD et par un algorithme de tatouage spatial. Nous disposons donc dorénavant d'une base de données permettant d'évaluer la résistance des algorithmes de tatouage vidéo au *camcording*.

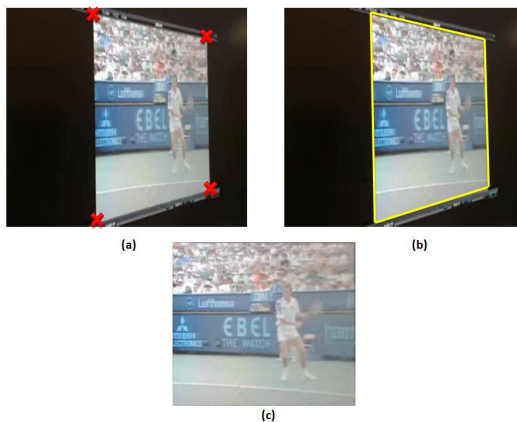
Participant VORTEX : Vincent CHARVILLAT.

Participants extérieurs : Asma KERBICHE, Ezzedine ZAGROUBA (Université Virtuelle de Tunis, Tunisie).

Publication en cours



Données brutes - Exemple d'image tatouée, projetée sur un écran de cinéma, filmée sous quatre angles : (a) depuis la salle de projection ; depuis la salle de cinéma, (b) à droite, (c) à gauche ou (d) en face de l'écran.



Donnée « camcordée » - Image obtenue en filmant une image tatouée, projetée sur écran, depuis la partie gauche de la salle de cinéma. Après rectification homographique (c), la marque est-elle toujours visible ?

Projet MobVille

Mots-clés : application mobile, reconnaissance d'objets 3D, points d'intérêt pour l'appariement 2D-3D.

Le projet MobVille du programme régional Lapérouse, mené en collaboration avec la société Imajing, vise à créer des applications mobiles géolocalisées pour la collecte et la gestion visuelles d'alertes émises par les citoyens. La première fonctionnalité consiste à centraliser les alertes, à les localiser géographiquement et à les rendre visuellement accessibles, au travers d'un service Internet, par différents acteurs : citoyens, élus, services opérationnels des collectivités. La seconde fonctionnalité consiste à mettre à la disposition des citoyens une application mobile pour déclencher une alerte via une simple photographie prise pour identifier le problème.

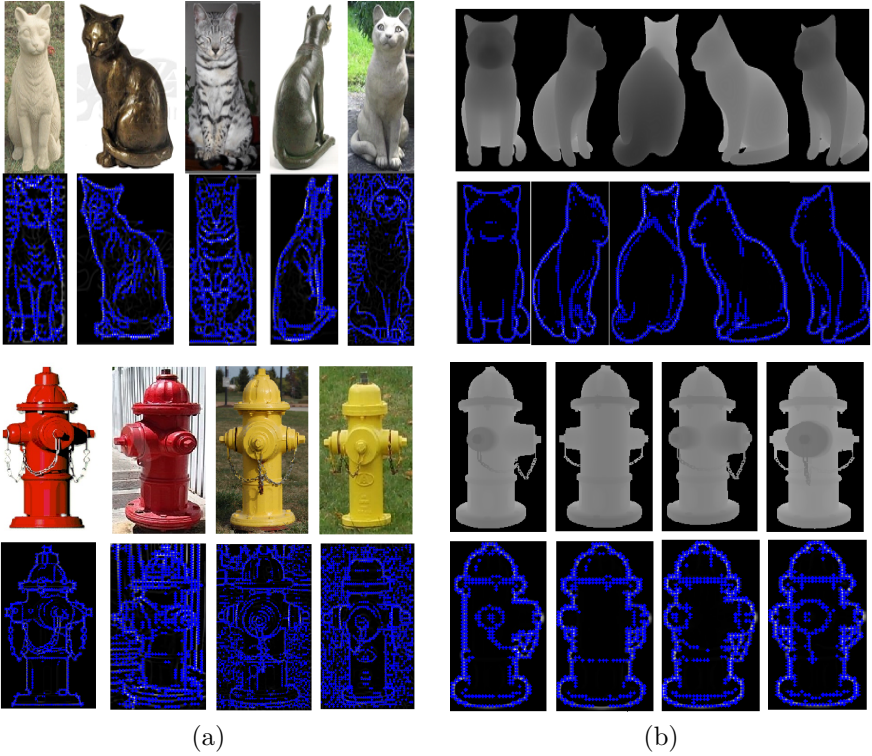
Les travaux développés dans l'équipe VORTEX portent, d'une part, sur l'évaluation de l'alerte visuelle afin de déterminer sa pertinence, et d'autre part, sur la classification de cette alerte permettant de faciliter son traitement par le gestionnaire. Cette qualification visuelle s'appuie sur l'appariement entre des images 2D et des modèles 3D issus d'un inventaire.

Il existe de très nombreux travaux sur la caractérisation et l'appariement de primitives 2D. Par ailleurs, de nombreux outils ont été développés pour tenter de détecter des points d'intérêt sur un modèle 3D. En revanche, les travaux de recherche sur la façon de qualifier et d'apparier des points d'intérêt entre une prise de vue 2D et un modèle 3D sont beaucoup plus récents et moins nombreux. Dans le contexte du projet, nous souhaitons qualifier l'état d'un objet à partir d'un modèle connu de cet objet, comme par exemple un abribus ou un élément du mobilier urbain, sachant que l'utilisateur fournit une image 2D, et que nous avons accès à un modèle 3D de cet objet.

L'approche retenue consiste à projeter le modèle 3D suivant différents points de vue et à rechercher la projection la plus proche de l'image 2D acquise. Nous utilisons la courbure dans les images pour détecter les points d'intérêt, à la fois dans les images 2D et dans les différentes projections de l'objet 3D.

Participants VORTEX : Sylvie CHAMBON, Vincent CHARVILLAT, Pierre GURDJOS, Hatem RASHWAN.

Participant extérieur : Étienne LAMORT DE GAIL (société Imajing).



Approche proposée - Extraction des points : (a) sur les images réelles ;
 (b) sur les cartes de profondeur.

Projet POPART

Mots-clés : localisation par images, effets spéciaux.

Le projet européen POPART (*Previz for On-set Production - Adaptive Realtime Tracking*) a démarré en janvier 2015 et se terminera en juin 2016. Ce projet concerne le suivi de caméra pour la prévisualisation d'effets spéciaux en temps réel sur un tournage de film. En post-production cinématographique, les algorithmes de vision par ordinateur sont utilisés pour trouver les mouvements de caméra, afin d'incruster des éléments synthétiques dans des scènes naturelles. Néanmoins, ces traitements sont le plus souvent effectués plusieurs semaines après le tournage. Le projet POPART vise à combler le fossé entre la production et le processus de post-production des films. L'idée principale est de développer un système qui puisse permettre au réalisateur du film d'avoir un aperçu sommaire des effets numériques (rendu 3D) qui seront ajoutés en post-production.

L'approche innovante proposée par POPART est la construction d'une base de données (visuelle) au cours d'une étape de préparation du tournage, dans laquelle les points d'intérêt utilisés pour le suivi de caméra (que ce soient des marqueurs ou des points d'intérêt de la scène) sont stockés avec leurs descripteurs SIFT et leurs points 3D associés. Cela permet de créer une base de connaissances contenant un modèle de la scène, qui peut ensuite être interrogée au cours des tournages suivants afin de récupérer la position de la caméra dans la scène (*image-based localization*).

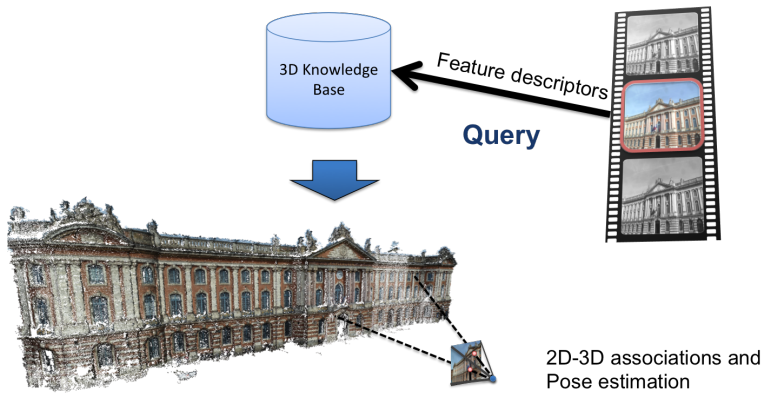
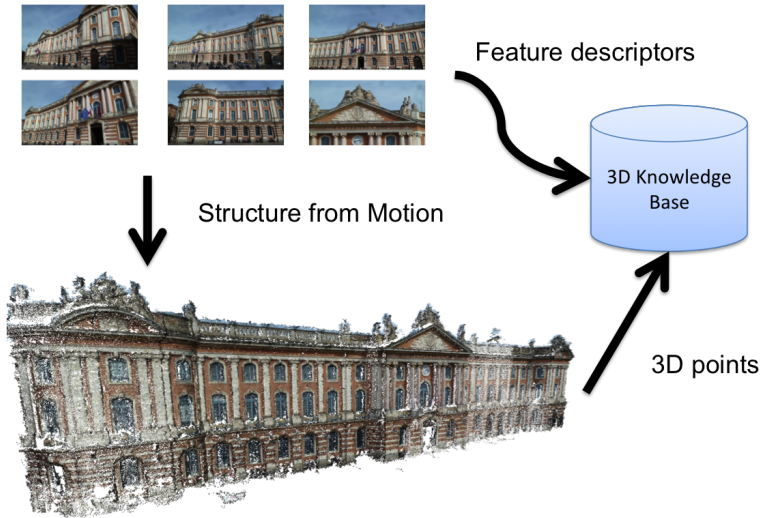
Tous les logiciels produits seront librement accessibles (*open-source*).

Participants VORTEX : Clément AYMARD, Simone GASPARINI, Pierre GURDJOS, Thierry MALON (ex-stagiaire VORTEX).

Partenaires : laboratoire Simula Research (Oslo, Norvège), sociétés Mikros Image (Paris), Labo Mixed Reality (Oslo, Norvège) et Band Pro (Munich, Allemagne).

Site web : <http://www.popartproject.eu/>

Illustration vidéo : <https://youtu.be/I6o1tk9iVUU>



Illustrations - En haut : construction de la base de données visuelle avec des techniques de *structure-from-motion*. En bas : localisation de l'image courante par interrogation de la base de données.

Reconstruction 3D du relief cutané

Mots-clés : reconstruction 3D microscopique, stéréophotométrie.

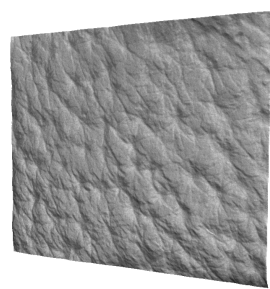
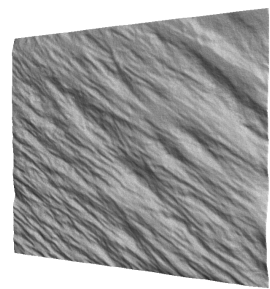
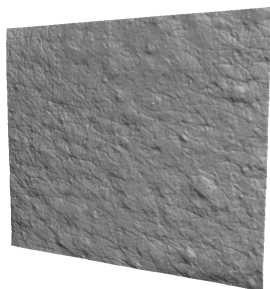
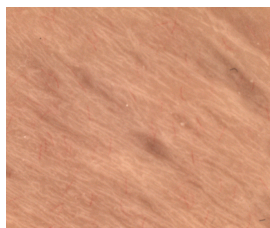
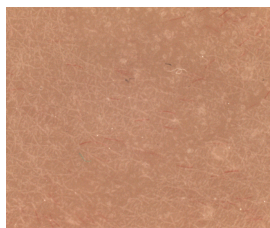
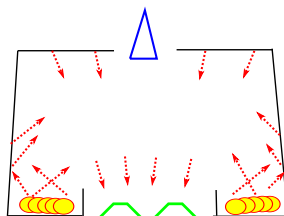
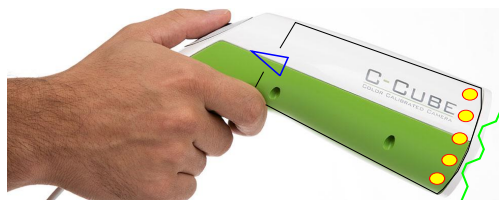
Les structures les plus fines d'un relief (de l'ordre de quelques micromètres) sont souvent difficiles à retrouver par les systèmes de mesure 3D « classiques ». Or, ces structures sont facilement mises en évidence en provoquant des variations d'éclairage. La technique de la stéréophotométrie, qui vise à reconstruire le relief à partir de plusieurs photographies prises sous le même angle, mais sous différents éclairages, est donc particulièrement adaptée à la reconstruction 3D de telles structures.

Dans le cadre d'un projet de transfert technologique entre l'IRIT et la société Pixience, financé et piloté par la SATT de Toulouse Midi-Pyrénées (société Toulouse Tech Transfer), nous avons pu transformer un dermoscope numérique (dispositif permettant d'observer le relief cutané à très haute résolution) en un scanner 3D permettant d'effectuer la reconstruction 3D des détails de la peau des patients. Ces détails vont du mélanome (quelques millimètres) aux rides les plus fines (quelques microns).

Le dermoscope est constitué d'une caméra HD et d'un ensemble de LED qui éclairent la scène de façon diffuse. En allumant différentes combinaisons de LED situées à l'intérieur du dermoscope, les conditions d'utilisation de la stéréophotométrie sont réunies : il est alors possible de reconstruire en 3D les détails du relief, ce qui peut être très utile pour le suivi médical de l'évolution de certaines maladies, ou pour visualiser l'effet de produits cosmétiques sur la peau.

Participants VORTEX : Vincent CHARVILLAT, Jean-Denis DUROU, Matthieu PIZENBERG, Yvain QUÉAU.

Partenaires : sociétés Pixience et Toulouse Tech Transfer.



Reconstruction 3D à l'échelle micrométrique - En haut à gauche : dermoscope numérique, sur lequel l'appareil photographique est mis en évidence en bleu, les LED en jaune, et la surface observée en vert. En haut à droite : représentation schématique de l'intérieur du dermoscope. En utilisant différentes combinaisons de LED, différentes configurations d'éclairage sont obtenues, ce qui permet d'utiliser la stéréophotométrie. Au milieu : photographies obtenues sous une configuration d'éclairage particulière, pour trois échantillons de fausse peau en silicone présentant des « rugosités » différentes. En bas : reconstructions 3D correspondantes.

Publications 2015

Articles dans des revues internationales :

Axel CARLIER, Amaia SALVADOR, Ferran CABEZAS, Xavier GIRO I NIETO, Vincent CHARVILLAT, Oge MARQUES, *Assessment of crowdsourcing and gamification loss in user-assisted object segmentation*, Multimedia Tools and Applications, septembre 2015.

Yvain QUÉAU, François LAUZE, Jean-Denis DUROU, *Solving Uncalibrated Photometric Stereo using Total Variation*, Journal of Mathematical Imaging and Vision, volume 52, numéro 1, pages 87-107, mai 2015.

Conférences et workshops internationaux avec actes édités et comités de lecture :

Bastien DURIX, Géraldine MORIN, Sylvie CHAMBON, Céline ROUDET, Lionel GARNIER, *Towards Skeleton based Reconstruction: From Projective Skeletonization to Canal Surface Estimation*, 3D Vision (3DV 2015), Lyon, octobre 2015.

Axel CARLIER, Vincent CHARVILLAT, Wei Tsang OOI, *A Video Timeline with Bookmarks and Prefetch State for Faster Video Browsing*, ACM Multimedia (MM 2015), pages 967-970, Brisbane, Australie, octobre 2015.

Ferran CABEZAS, Axel CARLIER, Amaia SALVADOR, Xavier GIRO I NIETO, Vincent CHARVILLAT, *Quality Control in Crowdsourced Object Segmentation*, International Conference on Image Processing (ICIP 2015), Québec, Canada, pages 4243-4247, septembre 2015.

Yvain QUÉAU, Jean-Denis DUROU, *Some Illumination Models for Industrial Applications of Photometric Stereo*, International Conference on Quality Control by Artificial Vision (QCAV 2015), volume 9534 de Proceedings of SPIE, Le Creusot, juin 2015.

Yvain QUÉAU, Jean-Denis DUROU, *Edge-Preserving Integration of a Normal Field: Weighted Least Squares and L1 Approaches*, International Conference on Scale Space and Variational Methods in Computer Vision (SSVM 2015), volume 9087 de Lecture Notes in Computer Science, pages 576-588, Lège Cap Ferret, mai-juin 2015.

Yvain QUÉAU, François LAUZE, Jean-Denis DUROU, *A L1-TV Algorithm for Robust Perspective Photometric Stereo with Spatially-Varying Lightings*, International Conference on Scale Space and Variational Methods in Computer Vision (SSVM 2015), Lège Cap Ferret, volume 9087 de Lecture Notes in Computer Science, pages 498-510, mai-juin 2015.

Nicolas BERTRAND, Jean-Denis DUROU, Carsten GRIWODZ, Vincent CHARVILLAT, *Toward Free and Open Source Projection for Digital Cinema*, EUROMEDIA, Lisbonne, Portugal, avril 2015.

Nicolas BERTRAND, Jean-Denis DUROU, Vincent CHARVILLAT, *The DCP Bay: Toward an Art-House Content Delivery Network for Digital Cinema*, International Conference on Advances in Multimedia (MMEDIA 2015), Barcelone, Espagne, avril 2015.

Vincent ANGLADON, Simone GASPARINI, Vincent CHARVILLAT, *The Toulouse vanishing points dataset*, Multimedia Systems Conference (MMSys 2015), Portland, États-Unis, mars 2015.

Marie-Anne BAUDA, Sylvie CHAMBON, Pierre GURDJOS, Vincent CHARVILLAT, *Geometry-Based Superpixel Segmentation Introduction of Planar Hypothesis for Superpixel Construction*, International Conference on Computer Vision Theory and Applications (VISAPP 2015), Berlin, Allemagne, mars 2015.

Marie-Anne BAUDA, Sylvie CHAMBON, Pierre GURDJOS, Vincent CHARVILLAT, *Image Quality Assessment for Photo-consistency Evaluation on Planar Classification in Urban Scenes*, International Conference on Pattern Recognition Applications and Methods, Lisbonne, Portugal, janvier 2015.

Conférences et workshops nationaux avec actes édités et comités de lecture :

Bastien DURIX, Géraldine MORIN, Sylvie CHAMBON, Céline ROUDET, Lionel GARNIER, *Caractérisation de la projection du squelette d'une surface canal 3D : Application à la reconstruction 3D à partir de deux images*, Journées du Groupe de Travail en Modélisation Géométrique (GTMG 2015), Poitiers, avril 2015.

Vincent ANGLADON, Simone GASPARINI, Vincent CHARVILLAT, *TVPD : un jeu de données pour évaluer les algorithmes d'estimation de points de fuite*, ORASIS, Amiens, janvier 2015.

Yvain QUÉAU, Jean-Denis DUROU, Xavier DESCOMBES, *Que peut-on apprendre d'une scène vue par une webcam à partir d'images prises au cours d'une journée ensoleillée ?*, ORASIS, Amiens, janvier 2015.

Contributions à des ouvrages de synthèse :

Lisa LARSSON, Géraldine MORIN, Antoine BEGAULT, Raphaëlle CHAINE, Jeannine ABIVA, Évelyne HUBERT, Monica HURDAL, Mao LI, Beatriz PANIAGUA, Giang TRAN, Marie-Paule CANI, *Identifying Perceptually Salient Features on 2D Shapes*, Research in Shape Modeling, pages 129-153, Springer International Publishing, 2015.

Thèses :

Nicolas BERTRAND, *Transition vers le cinéma numérique : contributions pour le cinéma d'art et essai*, thèse de l'Université de Toulouse, décembre 2015.

Yvain QUÉAU, *Reconstruction tridimensionnelle par stéréophotométrie*, thèse de l'Université de Toulouse, novembre 2015.

Rabih AMHAZ, *Détection automatique de fissures dans des images de chaussée par sélection de chemins minimaux*, thèse de l'Université de Nantes Angers Le Mans, mars 2015.

Sommaire

Images et 3D

Approche variationnelle différentielle de la stéréophotométrie	3
Édition de surfaces par manipulation de leurs gradients	5
Reconstruction 3D d'un objet par le squelette	7
Évaluation d'algorithmes d'odométrie visuelle temps réel	9
Décomposition non supervisée de formes géométriques 2D . . .	11
Détection <i>a contrario</i> de lignes polygonales et d'arcs d'ellipses	13
Détection et identification de codes à barres circulaires	15

Multimédia

Un réseau de distribution de copies numériques pour le cinéma	17
Impact de signets sur la navigation dans un environnement 3D	19
Préchargement de signets pour une interface de VOD	21
<i>Crowdsourcing</i> et ludification en segmentation interactive . . .	23
Représentation basée image pour les environnements 3D virtuels	25
Résistance des algorithmes de tatouage vidéo au <i>camcording</i> . .	27

Projets

Projet MobVille	29
Projet POPART	31
Reconstruction 3D du relief cutané	33

Publications 2015

35

**IRIT (UMR CNRS 5505)
ÉQUIPE VORTEX
(Visual Objects from Reality To EXpression)**

ENSEEIHT
**2, rue Camichel, BP 7122
31071 TOULOUSE Cedex 7
05.34.32.21.71**

Illustration de la page de garde :

Système de suivi de caméra par localisation / Projet européen POPART (programme Horizon 2020), regroupant trois partenaires industriels (Mikros Image - France, LABO Mixed Realities AS - Norvège et Band Pro - Allemagne) et deux laboratoires (IRIT - France et Simula - Norvège)

www.popartproject.eu