

**Année 2016**

# **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 2016 au sein de l'équipe VORTEX (Visual Objects from Reality To EXpression) de l'IRIT, ou du moins au sein de la partie de cette équipe localisée à l'ENSEEIH. Le style adopté est volontairement concis. Nous espérons 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 2017.

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 à vocation applicative 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<sup>er</sup> décembre 2016

## 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
- Chafik BAKEY, mohamedchafik.bakey@enseeiht.fr
- Thibault BLANC-BEYNE, thibault.blanc-beyne@irit.fr
- Bastien DURIX, bastien.durix@enseeiht.fr
- Julien FAYER, julien.fayer@enseeiht.fr
- Thomas FORGIONE, thomas.forgione@irit.fr
- Damien MARIYANAYAGAM, damien.mariyanayagam@irit.fr
- Sonia MEJBRI, sonia.mejbri@enseeiht.fr
- Jean MÉLOU, jeme@mikrosimage.eu
- Tsubasa MINEMATSU minematsu@limu.ait.kyushu-u.ac.jp
- Matthieu PIZENBERG, matthieu.pizenberg@enseeiht.fr
- Hatem RASHWAN, hatem.mahmoud@enseeiht.fr

# Reconstruction 3D multi-vue à partir du squelette

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

Le squelette est un modèle de représentation des formes correspondant à une structure, curviligne ou surfacique, centrée à l'intérieur de la forme. Un rayon est associé à chaque point de la structure, ce qui implique que la connaissance d'un squelette 2D ou 3D permet de retrouver complètement la forme associée. Des études ont montré que la perception des formes par l'être humain est liée à la notion de squelette, ce qui en fait un bon candidat pour une interaction intuitive avec un objet 3D.

Du fait que le squelette est un bon représentant d'une forme, nous proposons une méthode de reconstruction basée sur le squelette, qui utilise un ensemble d'acquisitions calibrées d'un objet. L'idée générale de la démarche est d'estimer le squelette 2D sur l'ensemble des images acquises, afin d'en déduire une information sur le squelette 3D de l'objet. Le résultat donné par l'algorithme est un squelette 3D, ainsi qu'un maillage de l'objet reconstruit.

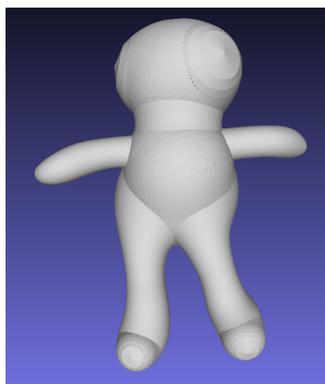
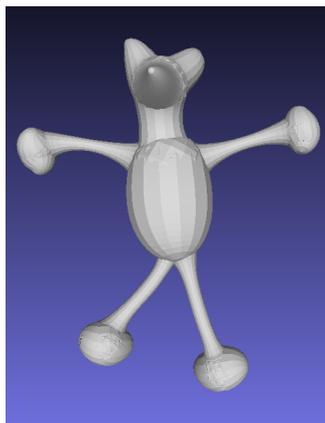
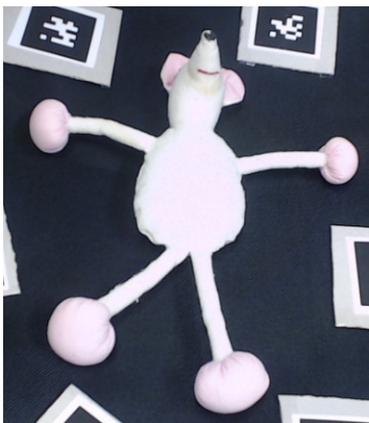
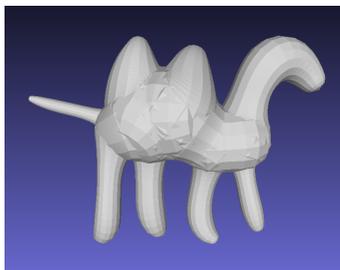
Cette méthode présente les avantages suivants : elle fournit une reconstruction qui restitue les informations sur la topologie de l'objet, ne nécessite que peu d'images (4 ou 5) pour être correctement effectuée, et peut être faite sans installation particulière. La connaissance du squelette permet en outre de faciliter 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, Dijon).

**Publication associée :**

Bastien DURIX, Géraldine MORIN, Sylvie CHAMBON, Céline ROUDET, Lionel GARNIER, *Skeleton-Based Multiview Reconstruction*, IEEE International Conference on Image Processing (ICIP 2016), Phoenix, États-Unis, septembre 2016.



**Exemples** - Reconstructions 3D obtenues avec notre méthode.

# Estimation de la réflectance à partir de données multi-vue

**Mots-clés :** vision 3D, réflectance, multi-vue, méthodes variationnelles.

L'acquisition du relief et de la réflectance d'un décor est fondamentale pour l'industrie du cinéma. Cela est nécessaire, en particulier, pour simuler correctement le ré-éclairage. Si plusieurs techniques bien rodées permettent d'obtenir une estimation précise et robuste du relief, il n'en va pas de même pour la réflectance. Faute de mieux, les images originales font généralement office de texture, c'est-à-dire qu'elles sont plaquées telles quelles sur le relief. Or, comme ces images combinent des caractéristiques de l'éclairage, de la géométrie et de la réflectance, les ré-éclairages obtenus sont alors très peu satisfaisants. Pour améliorer la procédure, il est indispensable d'estimer plus précisément la réflectance. Notre but est de séparer l'ombrage, qui dépend de l'éclairage et de la géométrie de la scène, de la réflectance. Nous utilisons plusieurs images de la surface acquises sous différents angles, sous l'hypothèse qu'une estimation préalable du relief, même grossière, est disponible.

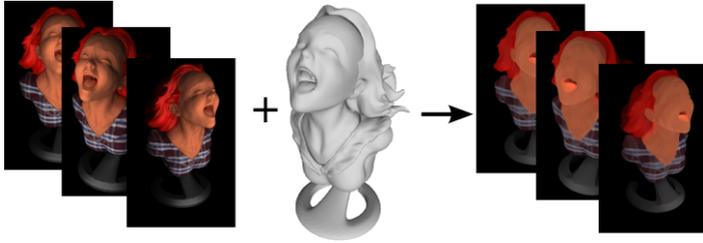
En supposant la surface lambertienne, ce qui signifie que sa réflectance est caractérisée par l'*albedo*, nous formulons ce problème sous forme variationnelle. Nous utilisons un terme d'attache aux données en norme  $L^1$ , afin de rendre le modèle robuste aux spéularités et aux alignements imparfaits, ainsi que deux autres termes concernant l'albédo : un terme d'a priori qui le force à être indépendant du point de vue, et un terme de régularisation de type « variation totale », qui vise à le rendre uniforme par morceaux.

**Participants VORTEX :** Jean-Denis DUROU, Jean MÉLOU.

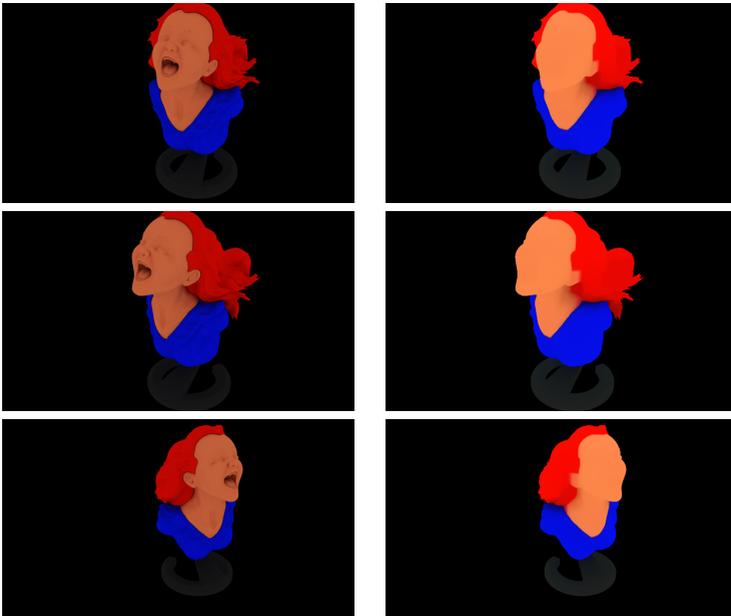
**Participants extérieurs :** Fabien CASTAN (société Mikros Image, Paris), Daniel CREMERS, Yvain QUÉAU (TUM, Munich, Allemagne).

**Publication associée :**

Jean MÉLOU, Yvain QUÉAU, Jean-Denis DUROU, Fabien CASTAN, Daniel CREMERS, *Estimation de la réflectance à partir de données multi-vue* (soumis).



**Problématique** - À partir d'un ensemble d'images acquises sous différents angles, et d'une estimation grossière du relief de la scène obtenue par reconstruction 3D multi-vue, une carte de réflectance est estimée pour chaque vue.



**Résultats** - Première colonne : trois vues de synthèse (parmi treize) d'un objet lambertien, divisé en quatre parties d'albédo uniforme (cheveux, visage, chemise et socle), placé sous un éclairage de type « dôme éclairant ». Deuxième colonne : albedos colorés estimés par l'approche proposée.

# Analyse multi-résolution de formes 2D pour la segmentation et l'analyse de formes

**Mots-clés :** axe médian, mesure de l'aire, segmentation 2D, analyse multi-résolution.

Ce travail porte sur une analyse multi-résolution de formes 2D utilisée pour l'identification des similitudes entre les différentes parties d'une forme. Une telle analyse est importante pour de nombreuses applications telles que la comparaison de formes, l'édition ou encore la compression.

Nous proposons une méthode de décomposition robuste d'une forme en parties. Une topologie entre les différentes parties est donnée par construction, et une hiérarchie des différentes parties est calculée. Nous en déduisons une mesure de similarité entre ces parties. Notre méthode se fonde sur une fonction définie sur l'axe médian (WEDF, pour *Weighted Extended Distance Function*), qui permet de pallier l'instabilité de l'axe médian. La partition multi-résolution de la forme 2D obtenue est indépendante de l'échelle et des déformations par articulations (déformations semi-rigides).

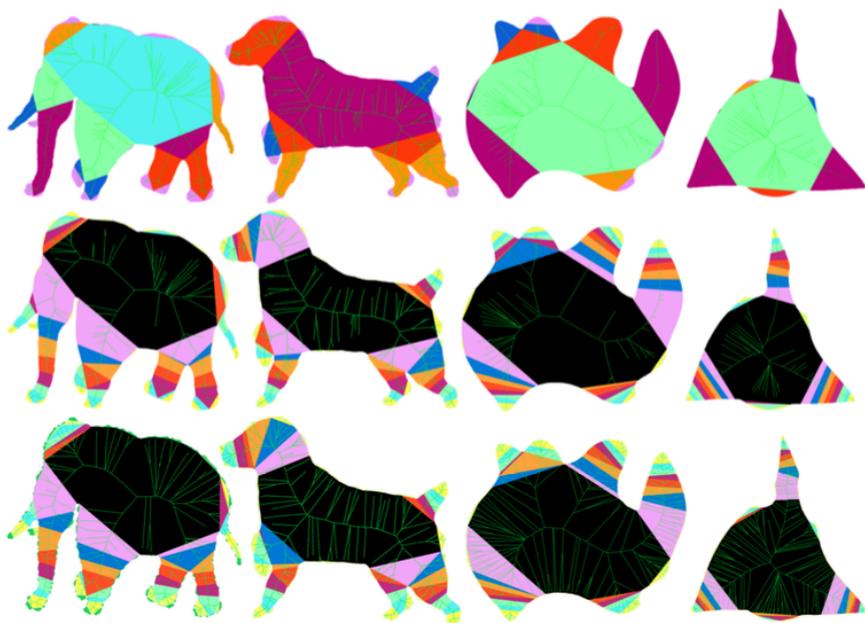
Finalement, la comparaison avec une étude utilisateur sur la base de données MPEG-7 démontre que nos résultats géométriques sont cohérents avec la perception humaine.

**Participants VORTEX :** Axel CARLIER, Géraldine MORIN.

**Participants extérieurs :** Stefanie HAHMANN (projet IMAGINE, INRIA, Grenoble), Kathryn LEONARD (California State University, Los Angeles, États-Unis).

**Publication associée :**

Kathryn LEONARD, Géraldine MORIN, Stefanie HAHMANN, Axel CARLIER, *A 2D Shape Structure for Decomposition and Part Similarity*, International Conference on Pattern Recognition (ICPR 2016), Cancún, Mexique, décembre 2016.



**Exemple** - Première ligne : niveaux de hiérarchie définis en fonction des valeurs de la fonction *WEDF*, pour quatre formes issues de la base MPEG-7 (deux animaux et deux formes abstraites). Chaque couleur caractérise un niveau de hiérarchie. Les niveaux sont déterminés indépendamment les uns des autres. Deuxième ligne : niveaux de la hiérarchie complète. Chaque niveau correspond à une couleur. La forme de base est noire. Ces niveaux étant déterminés selon un critère global sur la forme, les parties de même niveau sont comparables. Troisième ligne : hiérarchie complète sur les mêmes formes bruitées par un bruit blanc sur la frontière. Bien que l'axe médian soit très bruité, la hiérarchie obtenue reste stable.

# Segmentation hiérarchique de formes 3D basée sur le squelette

**Mots-clés :** segmentation, hiérarchisation, formes 3D, axe médian.

De nombreuses applications 3D, par exemple la modélisation, l'édition ou l'animation de formes 3D, impliquent le mélange et l'association de différents objets pour construire un ensemble cohérent. Ces méthodes supposent qu'une décomposition géométrique de la forme a été préalablement effectuée. Cependant, cette tâche n'est généralement pas triviale, et il est souvent demandé à l'utilisateur d'effectuer la segmentation à la main.

La segmentation que nous pratiquons se fonde sur le volume contenu dans la forme. Elle est robuste vis-à-vis des déformations non rigides de la forme. Pour décomposer une forme en parties hiérarchisées, nous utilisons son squelette curviligne. Ce travail généralise à la 3D des travaux existants en 2D : nous étendons des fonctions définies sur l'axe médian 2D à des fonctions sur le squelette curviligne 3D d'un objet.

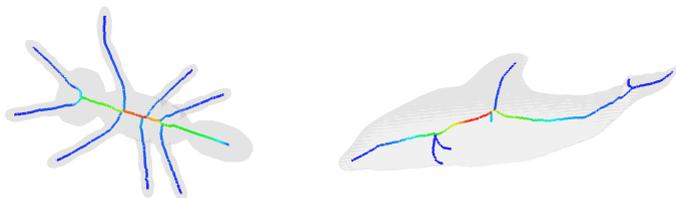
En particulier, nous proposons une fonction appelée *Weighted Extended Distance Function* (WEDF), qui sert à capturer le volume entourant les différentes branches du squelette curviligne, et qui généralise une fonction définie en 2D. Grâce à cette fonction, nous effectuons automatiquement la segmentation hiérarchique de la forme en parties cohérentes.

**Participants VORTEX :** Thibault BLANC-BEYNE, Axel CARLIER, Géraldine MORIN.

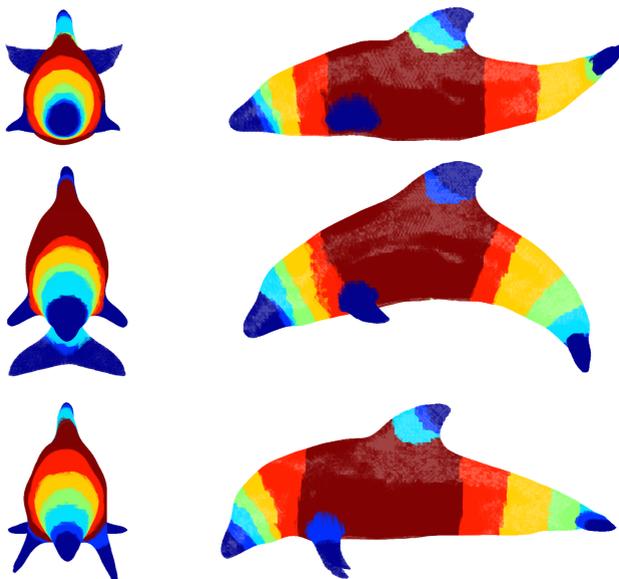
**Participants extérieurs :** Marie-Paule CANI, Stefanie HAHMANN (projet IMAGINE, INRIA, Grenoble), Kathryn LEONARD (California State University, Los Angeles, États-Unis).

**Publication associée :**

Thibault BLANC-BEYNE, Géraldine MORIN, Kathryn LEONARD, Stefanie HAHMANN, Marie-Paule CANI, Axel CARLIER, *A 3D shape hierarchy based on the skeleton*, Journées Françaises d'Informatique Graphique (JFIG), Grenoble, novembre-décembre 2016.



**Fonction WEDF** - Calcul de la fonction WEDF sur le squelette curviligne de deux formes (fourmi et dauphin). L'évolution de cette fonction le long du squelette est codée en couleur, du bleu pour les valeurs faibles au rouge pour les valeurs élevées. L'évolution de la fonction WEDF le long des branches est proportionnelle au volume qui les entoure.



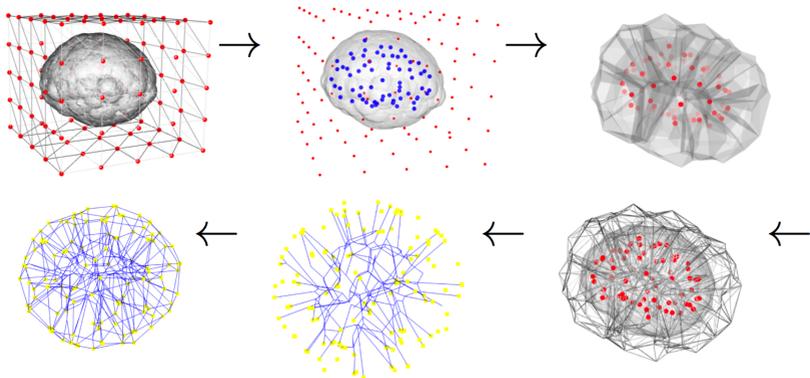
**Résultats** - Décompositions hiérarchiques d'un dauphin dans trois postures. La décomposition est appliquée au squelette curviligne, puis la forme est décomposée par mise en correspondance entre le squelette et la surface. Cette décomposition est robuste vis-à-vis de la déformation de la forme.

# Compression des données 3D+t de la morphogenèse animale

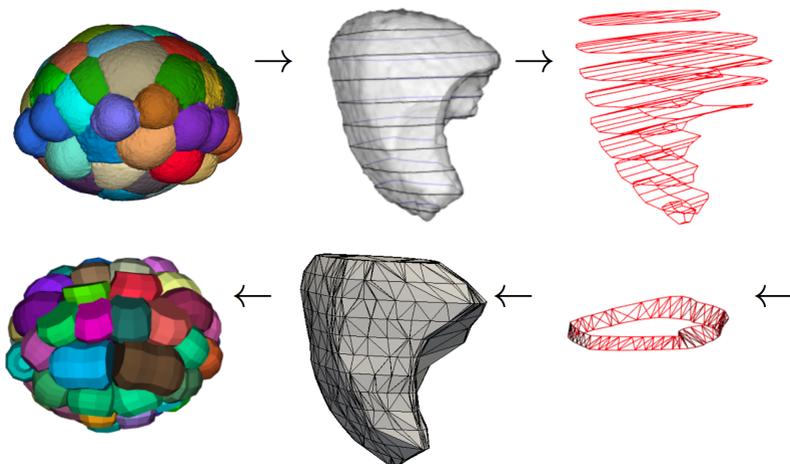
**Mots-clés :** maillage 3D, Voronoï, biologie du développement.

Le développement embryonnaire représente l'histoire permettant de passer, en quelques heures, d'une cellule unique, l'œuf tout juste fécondé, à un organisme complexe constitué de cellules différenciées soigneusement organisées en tissus et organes. Des travaux récents ont permis, pour la première fois, à l'aide d'une imagerie rapide et de haute résolution, de segmenter les cellules et de reconstruire l'histoire de chaque cellule d'un organisme marin : *l'ascidie*. À partir des segmentations 3D sous forme volumique des données temporelles, ce projet a permis d'évaluer différentes méthodes de représentation adéquates pour obtenir un modèle de maillage de l'ensemble des cellules, adapté au *streaming* de données sur le web, c'est-à-dire léger, progressif et possédant d'excellentes qualités de rendu visuel.

**Participants VORTEX :** Axel CARLIER, Emmanuel FAURE, Sonia MEJBRI, Géraldine MORIN.



**Pipeline** - Construction du diagramme de Voronoï 3D fermé.



**Illustration** - Étapes de construction de la représentation des cellules à la fois par leurs contours, et en utilisant les points situés aux extrémités du squelette 2D de chaque coupe.

# Stéréophotométrie différentielle

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

La reconstruction 3D par stéréophotométrie permet de retrouver les structures géométriques très fines d'une surface. Cependant, cette technique est peu adaptée aux images RVB, car la méthode de résolution classique consiste à estimer séquentiellement l'albédo relatif aux trois canaux de couleur, les normales à la surface et enfin la fonction de profondeur. La précision de la reconstruction 3D finale dépend donc de celle de chaque sous-tâche. Or, cette approche séquentielle induit nécessairement un biais.

Afin d'éviter ce biais, nous proposons une nouvelle formulation de la stéréophotométrie adaptée aux images RVB, fondée sur l'utilisation de quotients entre paires de niveaux de couleur, ce qui permet d'éliminer l'albédo des inconnues. En outre, en adoptant une approche différentielle, le problème est reformulé comme un système d'équations aux dérivées partielles (EDPs) ayant pour unique inconnue la fonction de profondeur. Ce système d'EDPs présente l'énorme avantage d'être linéaire, ce qui bien sûr simplifie sa résolution, mais il est généralement incompatible à cause du bruit.

Pour surmonter ce problème, nous introduisons une approche variationnelle permettant de résoudre le système d'EDPs au sens des moindres carrés, ce qui garantit une certaine robustesse au bruit de mesure. La pertinence de notre approche est démontrée sur données réelles par la reconstruction 3D de structures macroscopiques et microscopiques.

**Participants VORTEX :** Jean-Denis DUROU, Yvain QUÉAU (ex-doctorant VORTEX).

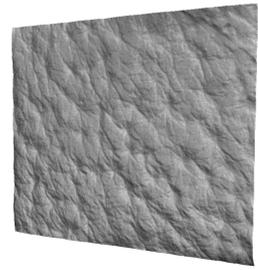
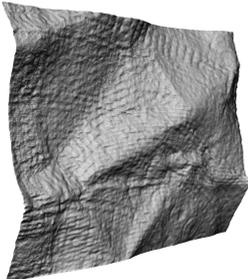
**Participants extérieurs :** Roberto MECCA (University of Cambridge, Royaume-Uni).

**Publication associée :**

Yvain QUÉAU, Roberto MECCA, Jean-Denis DUROU, *Unbiased Photometric Stereo for Colored Surfaces : A Variational Approach*, IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016), Las Vegas, États-Unis, pages 3707-3716, juin-juillet 2016.



**Exemple 1** - Reconstruction 3D d'une scène colorée à l'échelle macroscopique. À gauche : une des huit photographies d'une scène prises depuis le même point de vue, mais sous différents éclairages. À droite : résultat de la reconstruction 3D par l'approche différentielle proposée.



**Exemple 2** - Reconstruction 3D de structures microscopiques. Première ligne : Une photographie (parmi quinze) de différents objets (un billet de 10 euros, une pièce de 50 centimes espagnole et un échantillon de peau synthétique). Deuxième ligne : résultat de la reconstruction 3D par l'approche différentielle proposée.

# Stéréophotométrie à deux images

**Mots-clés :** reconstruction 3D, stéréophotométrie, *graph-cut*.

La stéréophotométrie à deux images constitue le cas intermédiaire entre la stéréophotométrie conventionnelle, qui utilise au moins trois images et est un problème bien posé, et le *shape-from-shading*, qui est mal posé.

Nous proposons d'abord une analyse théorique des ambiguïtés apparaissant dans ce cas intermédiaire. À partir de cette analyse théorique, nous montrons que, lorsque l'albédo de la surface observée est connu, la résolution des ambiguïtés peut être reformulée comme un problème d'étiquetage binaire, en utilisant la contrainte d'intégrabilité et un modèle d'Ising non stationnaire.

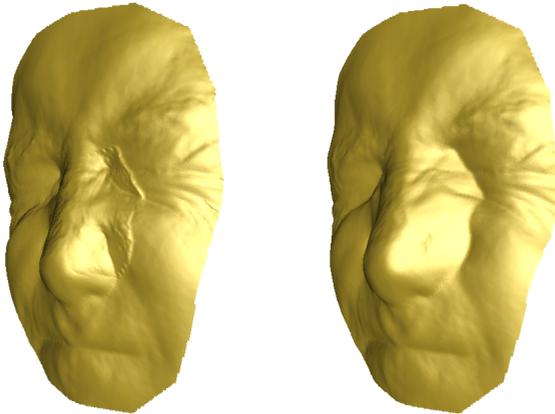
La reconstruction 3D est alors reformulée comme un problème d'optimisation discrète, qui peut être résolu de façon efficace par coupure de graphe (*graph-cut*). Ces contributions théoriques et numériques sont enfin validées sur données réelles, par une application au problème de la stéréophotométrie à trois images en présence d'ombres.

**Participants VORTEX :** Jean-Denis DUROU, Yvain QUÉAU (ex-doctorant VORTEX).

**Participants extérieurs :** Xavier DESCOMBES (projet Morpheme, INRIA, Sophia Antipolis), Roberto MECCA (University of Cambridge, Royaume-Uni).

**Publication associée :**

Yvain QUÉAU, Roberto MECCA, Jean-Denis DUROU, Xavier DESCOMBES, *Photometric Stereo with Only Two Images: A Theoretical Study and Numerical Resolution*, Image and Vision Computing, volume 57, pages 175-191, janvier 2017.



**Exemple -** En haut : photographie RVB prise en éclairant simultanément la scène par trois sources colorées. Au milieu : canaux R, V et B correspondants. Ces canaux peuvent être utilisés pour la stéréophotométrie à trois images, mais certaines données sont inexploitable à cause des ombres. En bas : reconstructions 3D par stéréophotométrie à trois images, et par l'approche proposée à deux images, où les données inexploitable sont éliminées.

# Habillage par *quads* d'un squelette

**Mots-clés :** axe médian, maillages par *quads*, habillage.

Un squelette permet de décrire simplement et intuitivement une forme, tandis que sa « surface limite » permet de visualiser la forme. Le but de ce travail est de construire un maillage quadrilatère multiple, aussi simple que possible, autour d'un squelette unidimensionnel. Être capable de trouver un maillage simple autour d'un squelette est une première étape nécessaire pour fournir deux représentations différentes d'un même objet : un squelette, et une représentation surfacique (ici, le maillage).

Nous avons montré que, en supposant que chaque section du maillage est un polygone ayant un nombre constant de sommets, notre solution, c'est-à-dire un maillage à section quadrangulaire, est optimale. Ce maillage peut servir de support pour une représentation surfacique : il fournit la topologie de base pour un maillage qui peut être affiné, mais aussi un domaine pour une représentation paramétrique (un maillage par *quads* est idéal pour des produits tensoriels de *splines*) ou, avec le squelette, un support de grille pour lancer des rayons à travers une surface implicite et ainsi la visualiser.

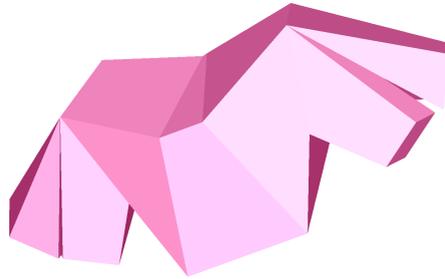
Nous fournissons un algorithme constructif pour dériver un tel maillage par *quads*. Les exemples ci-contre montrent que le maillage obtenu correspond tout à fait à l'habillage attendu autour d'un squelette.

**Participants VORTEX :** Géraldine MORIN.

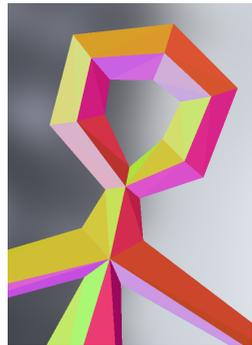
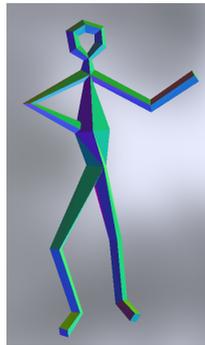
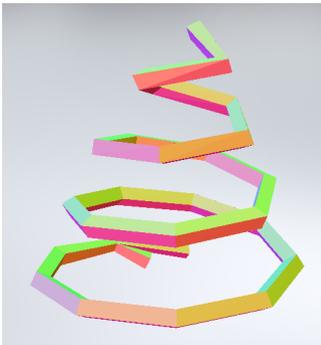
**Participants extérieurs :** Athina PANOTOPOULOU (Dartmouth College, Hanover, États-Unis), Kathrin WELKER (Trier University, Allemagne), Elissa ROSS (MESH Consultants Inc., Fields Institute for Research in the Mathematical Sciences, Canada), Evelyne HUBERT (projet GALAAD, INRIA, Sophia Antipolis).

**Publication associée :**

Athina PANOTOPOULOU, Kathrin WELKER, Elissa ROSS, Evelyne HUBERT, Géraldine MORIN, *Scaffolding a Skeleton* (soumis).



**Exemple 1** - Habillage d'un même squelette avec des rayons peu élevés (en haut) et plus importants (en bas).



**Exemple 2** - À gauche : spirale illustrant la rotation du maillage constituant le squelette. Au milieu : maillage entier. À droite : détail du cou.

# Rectification métrique de l'image d'un marqueur à contour circulaire

**Mots-clés :** réalité augmentée, marqueur fiduciaire, rectification métrique.

Nous nous intéressons à la rectification métrique de marqueurs circulaires plans, afin de reconnaître des objets rigides dans un environnement industriel, et de déterminer leurs poses relativement à la caméra. Ces marqueurs doivent être de taille réduite, mais doivent pouvoir être scannés à une distance relativement importante (cf. figure du haut). Parmi les marqueurs fiduciaires existants, nous nous intéressons au RuneTag et au C2Tag, qui sont circulaires, ce qui leur offre des propriétés géométriques intéressantes, par exemple la robustesse à l'occultation ou au flou de bougé.

Lorsque la caméra n'est pas calibrée, il existe une famille de méthodes de rectification métrique à partir d'un cercle, dépendant de deux paramètres inconnus qui peuvent être interprétés géométriquement comme les coordonnées de l'image du centre du cercle. Dans le cas calibré, il est important de connaître la précision de la détection, de l'ajustement de l'ellipse et du calibrage de la caméra. Malheureusement, lorsque le plan supportant le marqueur est très éloigné de la caméra, il est impossible d'effectuer l'étape de calibrage, ni de garantir ce niveau de précision pour l'ellipse. Dès lors, comment trouver une solution précise, à partir d'une détection d'ellipse qui ne l'est pas, et sans connaissance spécifique sur la caméra ?

Notre approche se fonde sur des tendances nouvelles, rendues possibles par les capacités de calcul des ordinateurs récents. À partir d'une connaissance incertaine sur les paramètres de l'ellipse détectée et d'un calibrage « par défaut », des candidats sont proposés pour les deux coordonnées de l'image du centre du cercle. Le meilleur candidat est conservé, au regard d'un critère photométrique de rectification. La figure du bas permet de comparer notre méthode aux approches classiques.

**Participants VORTEX :** Sylvie CHAMBON, Vincent CHARVILLAT, Pierre GURDJOS, Damien MARIYANAYAGAM.

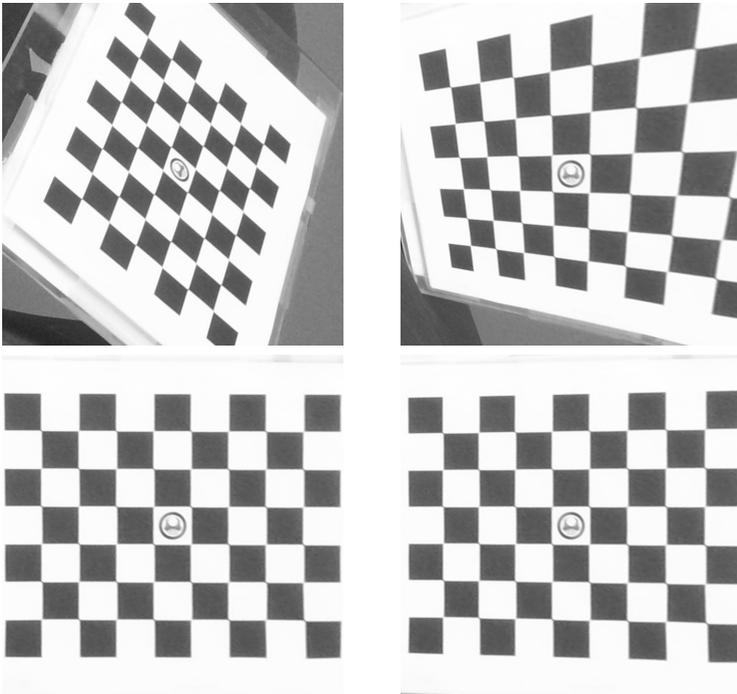
**Participants extérieurs :** Florent BRUNET (société UBLEAM, Toulouse).

**Publication associée :**

Damien MARIYANAYAGAM, Pierre GURDJOS, Sylvie CHAMBON, Vincent CHARVILLAT, Florent BRUNET, *Rectification métrique de l'image d'un marqueur à contour circulaire à partir d'une caméra calibrée par défaut* (soumis).



**Exemple** - Lecture dans une usine d'un marqueur *bleam*, proposé par la société Ubleam en collaboration avec l'équipe VORTEX.



**Résultats** - Première ligne : image originale et rectification basée sur la détection de deux ellipses. Deuxième ligne : rectification avec un cercle (cas calibré) et rectification « par défaut » avec sélection du meilleur candidat.

# Détection et localisation de marqueurs circulaires en conditions critiques

**Mots-clés :** marqueurs artificiels, cercles concentriques, détection, flou de bougé, suivi de caméra.

Les systèmes de marqueurs « artificiels » sont utilisés dans une large gamme d'applications, en particulier lorsque des indices visuels de référence font défaut dans la scène, par exemple lors du suivi de la caméra dans un environnement non texturé. Plus généralement, les systèmes conçus pour de telles applications doivent être robustes aux occultations partielles des marqueurs, aux variations des distances d'acquisition et des angles de prise de vues, aux conditions d'éclairage lorsque celui-ci est faible ou irrégulier, ainsi qu'aux mouvements rapides de la caméra, lorsque ceux-ci provoquent un flou de bougé.

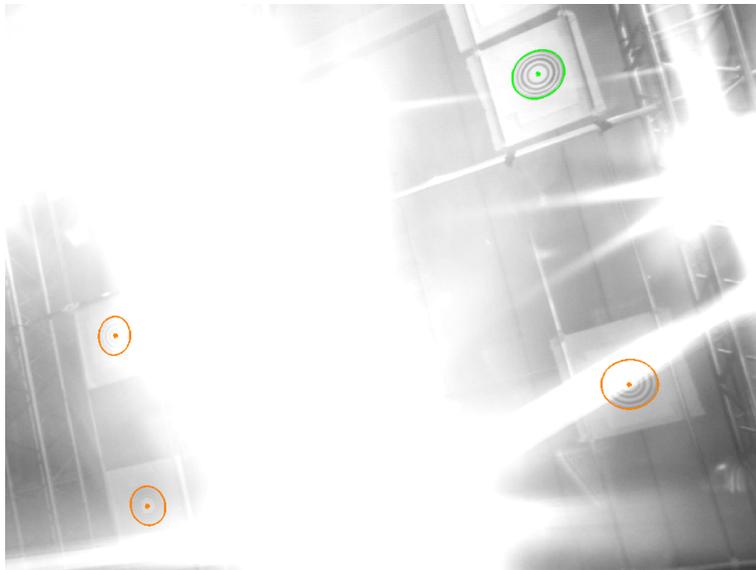
Nous proposons un système de marqueurs circulaires utilisant des motifs plans en noir et blanc représentant des anneaux concentriques. À partir d'algorithmes fondés sur des résultats issus de la géométrie projective, le système détecte les marqueurs dans les images, tout en étant robuste aux conditions critiques mentionnées plus haut, et localise de façon extrêmement précise la position de l'image du centre des cercles. L'originalité de notre approche, fondée sur la conservation du flux d'un champ de vecteurs, consiste à utiliser une procédure de vote faisant intervenir les lignes de champ du gradient de l'image. Nous avons montré que le système atteignait des performances au moins équivalentes à celles des systèmes existants (en particulier ARToolkitPlus et RuneTag) lorsque les conditions d'acquisition sont « favorables », voire nettement supérieures lorsque celles-ci deviennent critiques.

**Participants VORTEX :** Simone GASPARINI, Pierre GURDJOS.

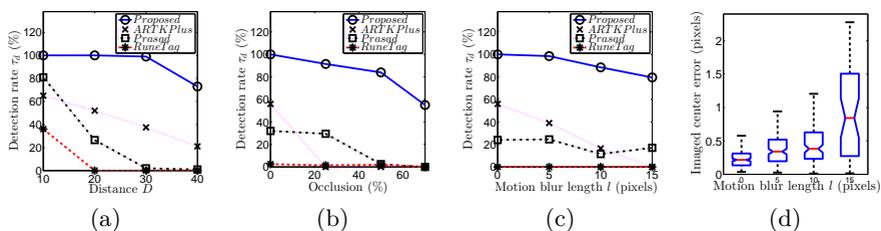
**Participants extérieurs :** Lilian CALVET, Carsten GRIWODZ (laboratoire Simula Research, Oslo, Norvège).

**Publication associée :**

Lilian CALVET, Pierre GURDJOS, Carsten GRIWODZ, Simone GASPARINI, *Detection and Accurate Localization of Circular Fiducials Under Highly Challenging Conditions*, International Conference on Computer Vision and Pattern Recognition (CVPR 2016), Las Vegas, États-Unis, pages 562-570, juin-juillet 2016.



**Exemple** - Les images des cercles détectés sont surimprimées en vert. Le système proposé obtient des performances bien supérieures à celles des systèmes existants, lorsque les conditions d'acquisition deviennent critiques.



**Évaluation** - Comparaison du système proposé avec les systèmes ARToolkitPlus [Prasad et al., WACV 2015] et RuneTag [Bergamasco et al., CVPR 2011]. Taux de détection en fonction de : (a) la distance à la caméra, (b) le pourcentage d'occultations et (c) l'amplitude du flou directionnel. (d) Estimation de l'image du centre en fonction de l'amplitude du flou directionnel.

# Co-segmentation interactive par entourage

**Mots-clés :** co-segmentation interactive.

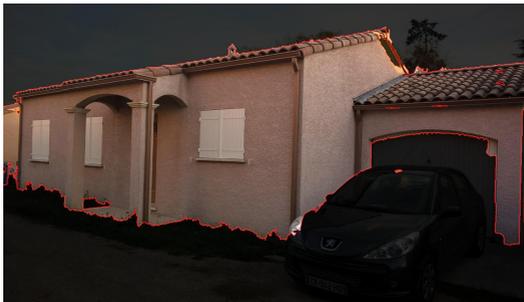
La co-segmentation consiste à segmenter un objet vu dans plusieurs images sous différents angles. Malgré la redondance induite par la présence de l'objet dans différentes images, ce problème reste ardu en raison des occultations, des conditions d'éclairage, des ombres portées ou des différents angles de vue. Pour améliorer les résultats, nous mettons à contribution un utilisateur humain dans une approche de « co-segmentation interactive ».

Dans la plupart des travaux sur la co-segmentation interactive, les interactions sont effectuées par des utilisateurs initiés au traitement d'images. Nous proposons, au contraire, une interaction facile pour des utilisateurs n'ayant aucune connaissance en traitement d'images. Nous demandons aux utilisateurs d'entourer l'objet d'intérêt dans les images, ce qui présente l'avantage d'être naturel et facile à adapter aux appareils mobiles tactiles.

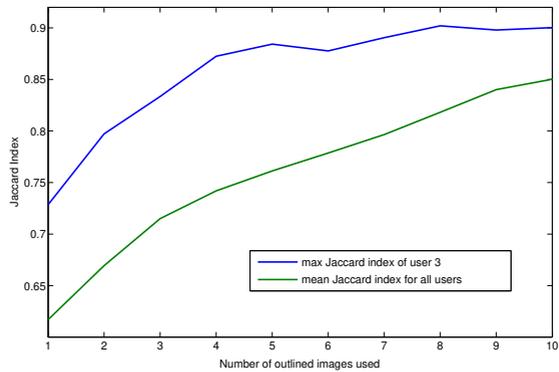
Pour réaliser la co-segmentation de l'objet d'intérêt dans la collection d'images, nous réutilisons un algorithme (cf. l'article *Interactive Cosegmentation Using Global and Local Energy Optimization*) qui construit un modèle d'apparence à partir de gribouillis (*scribbles*) décrivant des parties de l'objet (*foreground*) et de l'arrière-plan (*background*). Nous transformons donc les entourages pour générer deux ensembles de gribouillis servant de données d'entrée à l'algorithme de co-segmentation. Sur la figure en haut à gauche, l'entourage d'une maison par un utilisateur est indiqué en rouge, et les gribouillis générés automatiquement par un algorithme de squelettisation sont indiqués en vert. Les gribouillis rouges servent à définir un modèle d'apparence de l'arrière-plan. Quant aux gribouillis verts, ils servent de points de départ du modèle d'apparence de l'objet étudié (ici, la maison).

Notre étude a montré que cette approche permettait d'obtenir une bonne qualité de segmentation (la figure du bas montre que l'indice de Jaccard obtenu par l'utilisateur 3 est égal à 0.9), en plus d'être naturelle. L'étude a également montré qu'il n'était pas nécessaire d'interagir sur toutes les images. Ici, la qualité de la co-segmentation atteinte avec cinq images choisies « intelligemment » est presque identique à celle qui utilise la totalité des dix images de la collection.

**Participants VORTEX :** Axel CARLIER, Vincent CHARVILLAT, Matthieu PIZENBERG.



**Résultats** - En haut : entourage effectué par l'utilisateur (en rouge) et squelette calculé (en vert). En bas : résultat de la co-segmentation.



**Validation** - Évolution de l'indice de Jaccard de la co-segmentation de dix images, en fonction du nombre d'images dont on utilise les gribouillis.

# Photomaton 3D

**Mots-clés :** stéréophotométrie, reconstruction 3D, LEDs.

La stéréophotométrie est une technique de reconstruction 3D qui permet d'estimer le relief d'un objet à partir d'un ensemble d'images prises sous différents éclairages, mais sous le même point de vue. Cette méthode est efficace pour estimer les détails fins du relief. Elle permet également d'estimer l'*albédo* de l'objet, c'est-à-dire sa couleur, ce qui permet de simuler le ré-éclairage de l'objet.

Pour respecter au mieux les contraintes de la stéréophotométrie, nous avons aménagé une salle dédiée, dans le cadre d'un projet de transfert technologique piloté par la société Toulouse Tech Transfer. Le but du transfert était l'utilisation de la stéréophotométrie pour la reconstruction 3D de visages humains. Les murs de la salle de démonstration ont été peints en noir, afin d'éviter les réflexions parasites de la lumière. Le dispositif en lui-même est composé d'un appareil photographique et de colonnes sur lesquelles sont fixées des diodes électro-luminescentes (LEDs, pour *light-emitting diodes*). Ce dispositif est contrôlé par un ordinateur qui synchronise les prises de vue avec l'allumage et l'extinction des LEDs.

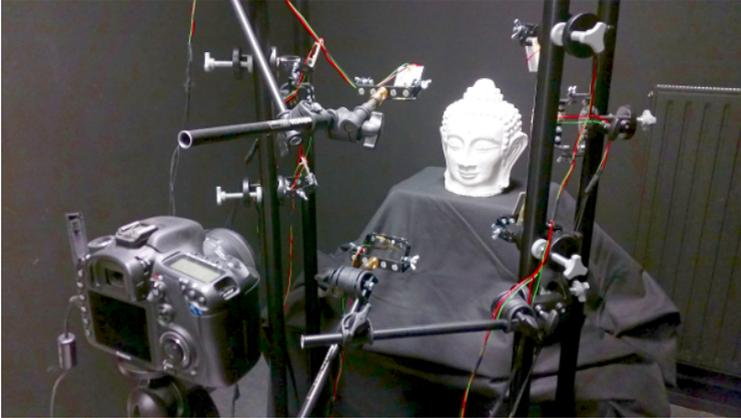
Nous proposons à chacun de nos visiteurs d'effectuer la reconstruction 3D de son visage dans cette salle, qui constitue donc un « photomaton 3D ».

**Participants VORTEX :** Bastien DURIX, Jean-Denis DUROU, Yvain QUÉAU (ex-doctorant VORTEX).

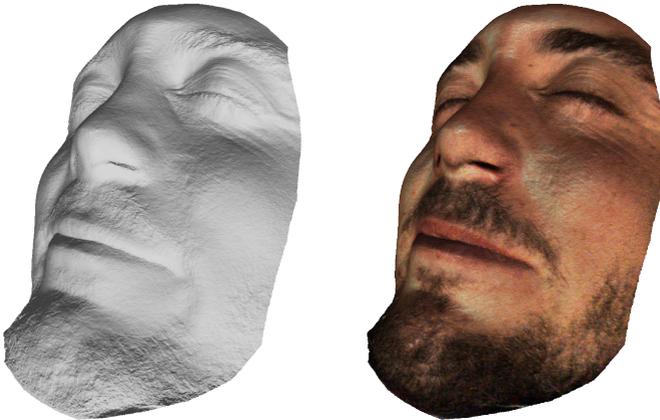
**Participants extérieurs :** François LAUZE (University of Copenhagen, Danemark).

**Publication associée :**

Yvain QUÉAU, Bastien DURIX, Jean-Denis DUROU, François LAUZE, *An Operational 3D-reconstruction Device using Photometric Stereo and LEDs – Modeling, Calibration and Resolution* (soumis).



**Salle noire** - Photographie de la salle d'acquisitions stéréophotométriques. On voit ici l'ensemble des LEDs éclairant le relief à estimer, ainsi que l'appareil photographique permettant d'acquérir les images.



**Exemple de reconstruction 3D** - Cet exemple illustre les deux avantages de la stéréophotométrie : les détails fins, comme le grain de la peau ou les rides, sont bien restitués ; la « vraie » couleur est estimée, ce qui permet de simuler le ré-éclairage du visage.

# Sélection de régions d'intérêt pour la transmission intelligente de contenus 3D

**Mots-clés :** *streaming* 3D, régions d'intérêt.

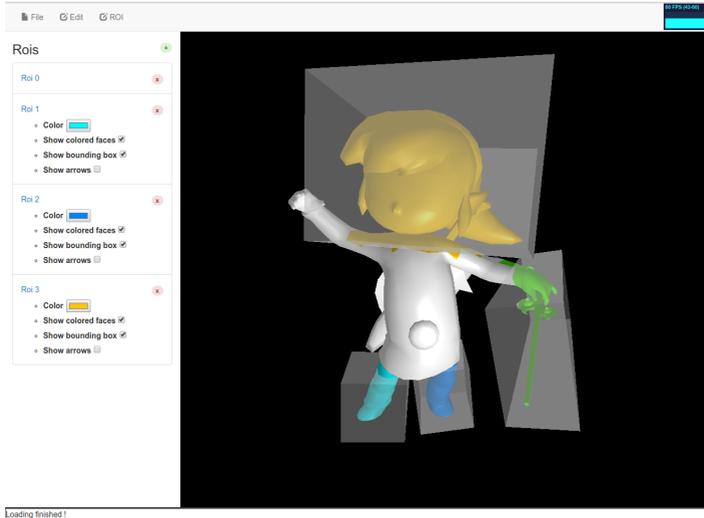
Avec la démocratisation des dispositifs d'acquisition de contenus 3D (*Kinect*, *Google Tango*, etc.) et des technologies d'accès distant à ces contenus (*WebGL*, *Three.js*, etc.), la problématique de la transmission (*streaming*) de contenus 3D se pose.

La plupart des travaux actuels en *streaming* 3D se focalisent sur les approches progressives, qui consistent à envoyer d'abord une version de basse résolution, qui représente la globalité du modèle, puis à affiner le modèle. Ces techniques sont très performantes pour des modèles représentant des objets ou des personnages, mais ne sont pas du tout adaptées aux cas de grandes scènes 3D navigables. L'approche que nous proposons s'articule en deux temps.

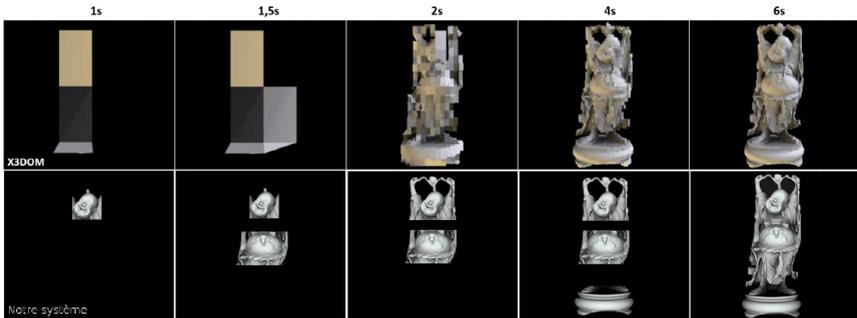
Dans un premier temps, nous proposons une interface permettant d'importer un modèle 3D, puis de le segmenter en différentes régions d'intérêt. L'auteur du modèle peut désigner plusieurs sous-parties du modèle, ainsi que leurs intérêts respectifs. L'interface transmet ensuite une version adaptée du modèle 3D, région d'intérêt par région d'intérêt, selon l'importance et la requête de l'utilisateur. On parle alors de « *streaming* adaptatif par région d'intérêt ».

Dans un deuxième temps, nous nous en servons pour envoyer le contenu de manière intelligente : l'utilisateur pourra choisir la région d'intérêt qu'il désire télécharger en priorité, ou bien le système pourra déterminer quelle région d'intérêt doit être transmise, en fonction par exemple de la position de l'utilisateur, des régions voisines, et des interactions d'autres utilisateurs.

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



**Capture d'écran** - Interface de sélection des régions d'intérêt.



**Illustration** - Évolution du chargement des modèles en fonction du temps pour deux systèmes différents, et la façon dont le chargement devrait évoluer sur notre système. En haut : la version progressive de X3DOM [Max Limper, 2014]. En bas : notre système, qui reçoit en priorité les régions d'intérêt définies par l'éditeur du modèle.

# Base de données de formes 2D segmentées par des utilisateurs

**Mots-clés :** segmentation 2D, analyse multi-résolution, étude utilisateurs, *crowdsourcing*.

Dans le cadre de notre travail sur l'analyse multi-résolution de formes 2D, nous avons mis au point une interface de segmentation interactive de formes et avons réalisé une étude utilisateurs afin de recueillir des segmentations correspondant à une interprétation humaine de ces formes.

L'interface présentée ci-contre permet aux utilisateurs de décomposer une forme en une partie principale (en noir), en parties secondaires (en magenta et en rouge) et en détails (en vert).

Nous avons ensuite réalisé une étude impliquant 2861 utilisateurs, chacun annotant 20 formes parmi les quelques 1250 formes extraites de la base de données MPEG-7. Au total, ce sont plus de 40000 annotations de formes que nous avons récoltées au cours de cette étude. Les utilisateurs ont été recrutés par le biais d'une chasse au trésor en ligne organisée par un des co-auteurs. L'ensemble des données récoltées est consultable à l'adresse <http://2dshapesstructure.github.io/index.html>.

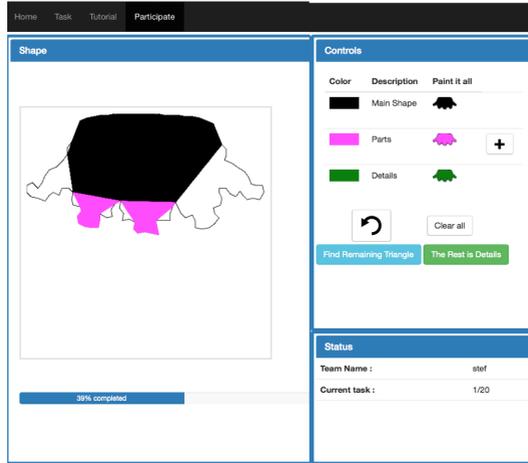
Les données récoltées (un exemple est donné ci-contre) soulèvent de nombreuses remarques et questions. Tout d'abord, il est clair, au vu des résultats, qu'il existe plusieurs interprétations possibles d'une même forme, et donc plusieurs segmentations également valides. Nous décrivons, dans la publication associée à cette base de données, comment prendre en compte cette variabilité. Nous y expliquons également comment nous avons traité les données aberrantes inhérentes à la collecte de données par *crowdsourcing*.

**Participants VORTEX :** Axel CARLIER, Géraldine MORIN.

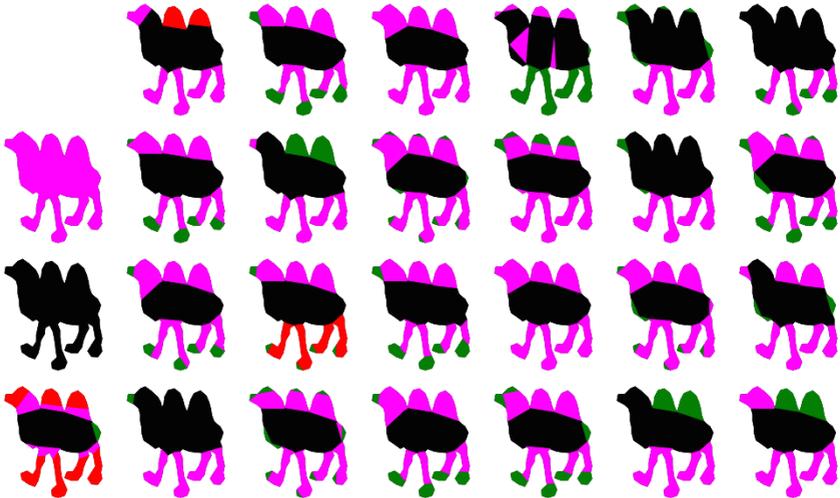
**Participants extérieurs :** Stefanie HAHMANN (projet IMAGINE, INRIA, Grenoble), Kathryn LEONARD (California State University, Los Angeles, États-Unis).

**Publication associée :**

Axel CARLIER, Kathryn LEONARD, Stefanie HAHMANN, Géraldine MORIN, Misha COLINS, *The 2D shape structure dataset : A user annotated open access database*, Computer & Graphics, volume 58, pages 23-30, août 2016.



**Exemple** - Capture d'écran de l'interface de segmentation de formes 2D.



**Exemple** - Exemple de données récoltées pour une forme de chameau.

## Métre d'une salle sur tablette

**Mots-clés :** perception 3D, scène d'intérieur, *smartphone*, capteur de profondeur, RVB-D.

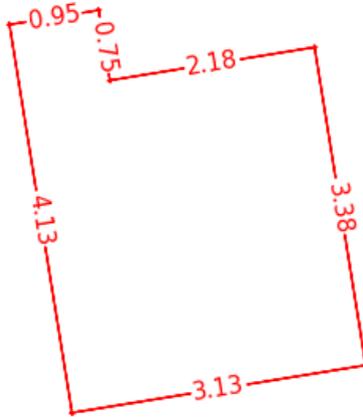
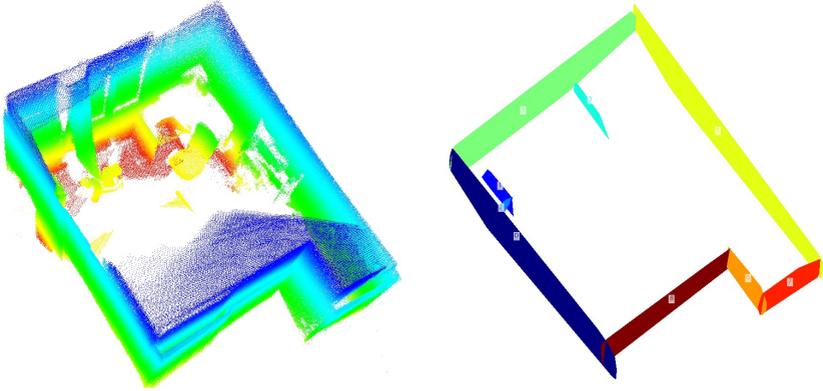
L'objectif de notre projet est de générer le plan d'une salle à partir d'une tablette équipée d'un capteur de profondeur tel que *Google Tango*. Des solutions commerciales permettent déjà d'effectuer des métrés à partir de données 3D, en envoyant ces données sur un service réalisant « hors ligne » un traitement automatique ou manuel du nuage de points 3D final. L'originalité de notre approche consiste à réaliser tous les calculs « en ligne » (au fil de l'arrivée des données sur la tablette), en mettant l'utilisateur à contribution.

Après avoir travaillé sur les problèmes de dérives de la localisation de la tablette, nous nous concentrons maintenant sur la perception des murs de la pièce. Cette tâche est complexifiée par le mobilier, qui occulte des portions de murs et des ouvertures. En outre, certaines surfaces sont incorrectement mesurées par le capteur, par exemple les miroirs qui réfléchissent les rayons infrarouges et le verre qui les absorbe. Enfin, les données de profondeur peuvent être fortement bruitées, à cause de la distance ou de l'inclinaison du capteur relativement aux surfaces perçues.

À chaque nouvelle acquisition de données 3D, les primitives planaires sont détectées en temps réel et fusionnées avec les primitives précédemment détectées, ce qui construit incrémentalement un modèle par plans de la pièce. Nous analysons les occultations entre régions planaires et la cohérence de la visibilité des points 3D pour compléter les données manquantes de ce modèle. Ces primitives planaires, éventuellement complétées, sont ensuite étiquetées comme étant un mur ou pas, puis affichées à l'utilisateur en réalité augmentée. Une vue de dessus permet de visualiser la génération du plan et les parties de la scène qui n'ont pas encore été observées, ce qui incite l'utilisateur à scanner entièrement la pièce concernée. Ce dernier peut à tout moment proposer une région planaire manquante ou corriger l'étiquette proposée en « cliquant » sur un des murs. Enfin, une étape de post-traitement permet de corriger les petits défauts d'orthogonalité et de retoucher manuellement le plan proposé.

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

**Partenaires :** société Telequid.



**Illustrations** - En haut à gauche : concaténation des données 3D acquises lors du métré d'une pièce avec la tablette *Google Tango*. On notera la présence de données manquantes à cause du mobilier, dans le coin en haut et sur les murs adjacents. En haut à droite : les primitives planaires (verticales) détectées. Les primitives 1, 2 et 3 correspondent à du mobilier. En bas : le plan généré avec les dimensions des murs, après post-traitement.

# Protocole et logiciel pour la capture et l'édition simplifiée de vidéos pédagogiques

**Mots-clés :** *e-learning*, MOOC, édition de vidéos, acquisition vidéo, *micro-teaching*, SCORM, EDL.

Ce projet vise la mise au point de protocoles simplifiés d'acquisition vidéo en mini-studio, grâce à un outillage logiciel original, de telle sorte qu'un enseignant sans expertise technique, voire sans affinité particulière avec l'informatique, puisse réussir en autonomie à produire des contenus à vocation pédagogique. Nous proposons ainsi des interfaces logicielles de pilotage et de supervision de la capture vidéo, de manière à réduire la frontière entre acquisition et montage.

L'objectif de l'étape d'acquisition vidéo est de créer des interactions minimales de validation des plans, pour que seul le plan validé soit exploité par la suite. L'enseignant doit donc pouvoir, de façon autonome, créer et valider des plans avec des outils intuitifs, qui le guident et minimisent les interactions requises et les autres actions ou manipulations techniques.

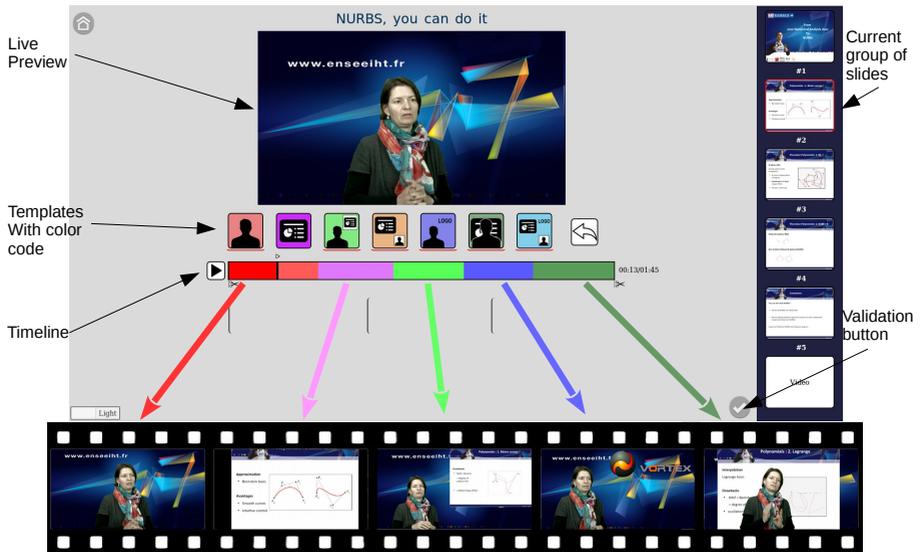
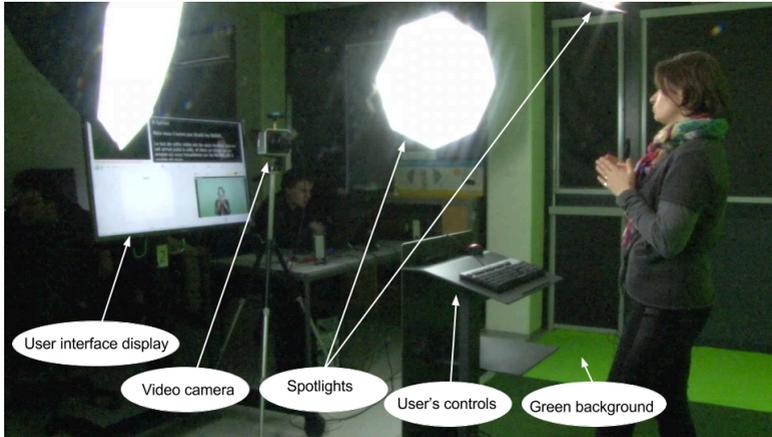
L'étape d'édition de vidéos concerne une simplification du montage final, qui est également semi-supervisé, au travers d'interactions minimales durant la lecture des plans validés. Par « interaction minimale », nous entendons, pour commencer, le basculement simple entre la capture de l'enseignant ou des supports qu'il commente ou annote à l'aide d'icônes simples, comme illustré sur la figure ci-contre. Cette approche est enrichie par des mécanismes de recommandation de types de plans.

**Participants VORTEX :** Chafik BAKEY, Axel CARLIER, Vincent CHARVILLAT, Géraldine MORIN, Matthieu PIZENBERG.

**Participants extérieurs :** Eithne BALAVOINE (Université Paul Sabatier, Toulouse).

**Publication associée :**

Chafik BAKEY, Matthieu PIZENBERG, Axel CARLIER, Eithne BALAVOINE, Géraldine MORIN, Vincent CHARVILLAT, *Protocols and Software for Simplified Educational Video Capture and Editing (soumis)*.



**Illustration** - En haut : studio « à fond vert » pour l'acquisition des vidéos.  
 En bas : interface permettant l'édition des vidéos.

# Détection de changements dans une vidéo

**Mots-clés :** détection d'objets en mouvement, apprentissage profond.

La quantité de flux vidéo produits par les caméras de vidéo-surveillance est telle que son traitement nécessite un très grand nombre d'opérateurs humains. Afin de réduire les coûts, il serait très utile de détecter automatiquement les situations « irrégulières » telles que des accidents de circulation ou des activités criminelles.

Pour détecter ces situations, il convient dans un premier temps de définir ce que sont les situations « régulières ». Dans notre cas, il s'agit d'une scène ne comportant pas d'objet mobile. Nous utilisons dans notre approche l'apprentissage profond : un réseau de neurones convolutif est entraîné à l'aide de situations régulières (première ligne de la figure). Lorsqu'une image correspondant à une situation irrégulière est fournie en entrée du réseau, celui-ci produit en sortie une image dans laquelle les objets mobiles apparaissent altérés, alors que le reste de la scène, qui ressemble plus ou moins à la situation régulière, n'est pas modifié. Ainsi, en comparant les images d'entrée et de sortie, il est possible de détecter les objets mobiles, et donc de prédire la régularité ou l'irrégularité d'une scène.

L'intérêt majeur de notre approche est qu'elle peut être appliquée à tous types de caméras, par exemple aux caméras PTZ (*Pan-Tilt-Zoom*), ou bien aux caméras embarquées sur des voitures ou sur des drones.

**Participants VORTEX :** Vincent CHARVILLAT, Tsubasa MINEMATSU.

**Participants extérieurs :** Rin-ichiro TANIGUCHI (Kyushu University, Fukuoka, Japon).



**Exemple** - Première ligne : exemple de situation régulière (source [Goyette et al., CVPRW 2012]). Deuxième et troisième lignes : les images de gauche sont des exemples de situations irrégulières, comportant des objets mobiles (ici, des piétons). Les images de droite ont été reconstruites par notre réseau de neurones, en altérant (c'est-à-dire en effaçant) les objets mobiles.

# Politiques optimales de préchargement de contenus multimédia

**Mots-clés :** *streaming* adaptatif, préchargement, processus décisionnels de Markov.

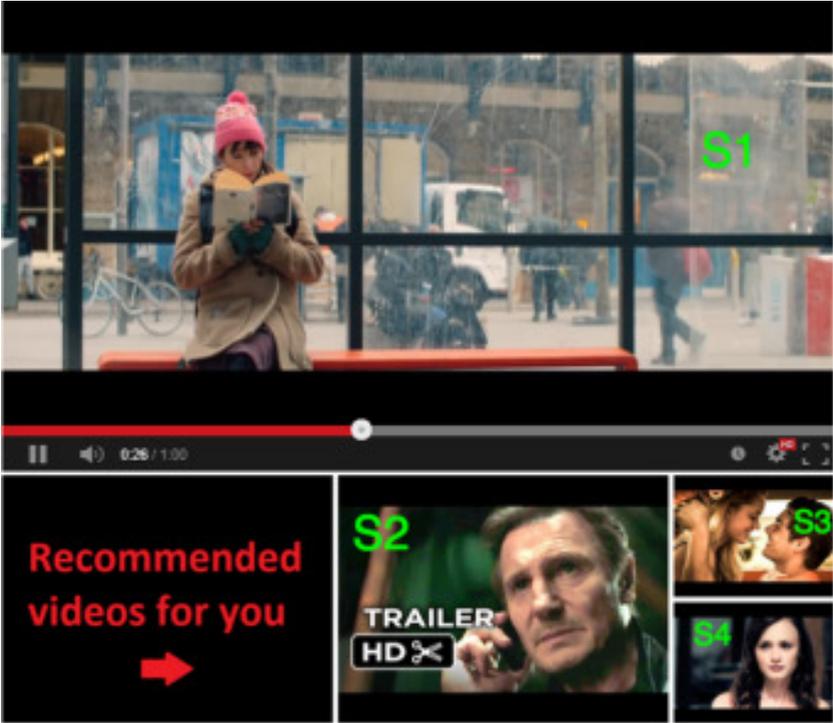
La transmission progressive (ou *streaming*) d'un contenu multimédia est délicate en cas d'accès aléatoires à différentes portions du contenu. Elle l'est également lorsque les ressources en bande passante sont variables, voire insuffisantes. On dit que le *streaming* est adaptatif lorsqu'un algorithme contrôle la transmission en fonction des accès et des ressources observés. Parmi les politiques de contrôle, nous nous intéressons à celles qui cherchent à précharger les portions d'un contenu susceptibles de correspondre aux requêtes futures d'un ou de plusieurs utilisateurs. Nous caractérisons ainsi des politiques de préchargement optimales, au sens de la résolution de *processus décisionnels de Markov* complètement ou partiellement observables qui minimisent des cumuls de latences. Nous caractérisons dans ce travail des politiques optimales séquentielles et proportionnelles qui généralisent celles, plus simples, de la littérature. Nous montrons finalement comment détecter différents profils d'utilisateurs et optimiser, simultanément, des politiques de préchargement pour chacun d'eux.

**Participants VORTEX :** Vincent CHARVILLAT.

**Participants extérieurs :** Cezar PLEZCA (Académie Technique Militaire de Bucarest), Wei Tsang OOI (National University of Singapore, Singapour).

**Publication associée :**

Cezar PLEZCA, Vincent CHARVILLAT, Wei Tsang OOI, *Multimedia prefetching with optimal Markovian policies*, Journal of Network and Computer Applications, volume 69, pages 40-53, juillet 2016.



**Exemple** - Contenu multimédia combinant plusieurs segments vidéo à précharger : une politique séquentielle peut consister à précharger  $S_3$  puis  $S_4$  pendant la lecture de  $S_1$ , si  $S_2$  a une faible probabilité d'être téléchargé.

# Projet REALISM

**Mots-clés :** RVB-D, réalité diminuée, *inpainting*, édition 3D.

La réalité diminuée est une branche de la réalité augmentée qui consiste généralement à supprimer des éléments réels d'un environnement numérisé. L'objectif est de pouvoir procéder à l'aménagement virtuel d'un environnement d'intérieur (suppression puis déplacement de meubles, modification des textures du sol ou des murs). Les premiers scénarios de réalité diminuée relèvent essentiellement du traitement d'image, avec notamment l'utilisation de *inpainting*, mais restent limités en termes de possibilités.

En revanche, les approches les plus récentes utilisent des informations telles que la position de la caméra, pour améliorer le rendu final et diversifier les scénarios possibles. Cependant, ces travaux permettent d'effectuer la suppression mais pas l'édition, car les informations 3D sont insuffisantes. Parallèlement, l'arrivée de capteurs « RVB-D » à faible coût, tels que la Kinect, permet d'obtenir une meilleure connaissance de l'environnement 3D. Plusieurs travaux utilisent les possibilités offertes par ces capteurs pour améliorer la compréhension de la scène.

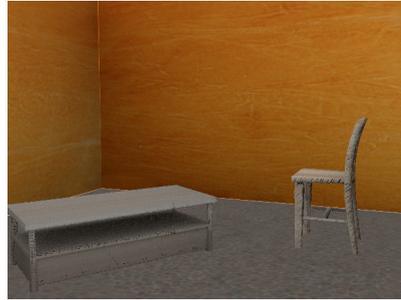
Dans le cadre du projet REALISM, nous avons commencé par montrer les limites des approches actuelles d'édition de scènes. Nous avons ensuite utilisé les différents travaux de recherche sur la compréhension de scènes 3D pour améliorer la chaîne de traitement, sous la forme de plusieurs contributions (rectification globale, critère de confiance et processus d'édition).

**Participants VORTEX :** Julien FAYER, Simone GASPARINI, Géraldine MORIN.

**Participants extérieurs :** Benjamin COUDRIN, Stéphane MERCIER (société InnerSense, Toulouse).

**Publication associée :**

Julien FAYER, Géraldine MORIN, Simone GASPARINI, Benjamin COUDRIN, *Réalité diminuée : “couper/coller” interactif pour l'aménagement d'intérieur*, Journées Françaises d'Informatique Graphique (JFIG 2016), Grenoble, décembre 2016.



**Exemple 1** - Déplacement dans une scène de synthèse. À gauche : position d'origine. À droite : translation de la table et rotation de la chaise.



**Exemple 2** - Déplacement d'un meuble dans une scène réelle. En haut à gauche : position d'origine. En haut à droite : translation. En bas : rotations de  $\pm 45^\circ$ .

# Projet *4D Cloud Embryos*

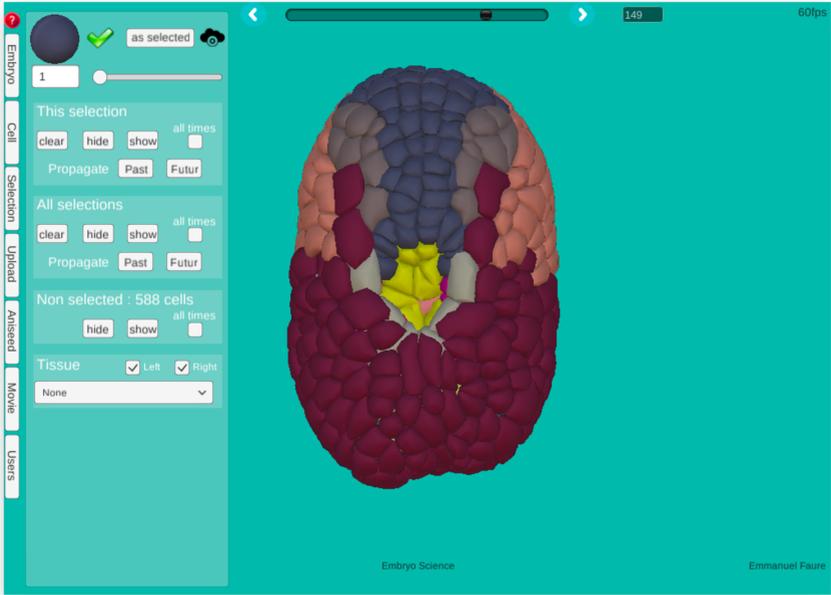
**Mots-clés :** interface de visualisation 3D, WebGL, biologie du développement.

Dans ce projet, nous avons développé une interface en ligne de visualisation des données 3D+t de l'embryogenèse animale. Nous sommes partis du constat que le monde des jeux vidéo avait fait des avancées prometteuses grâce à la technologie WebGL, qui utilise l'accélération matérielle de rendu 3D par un simple navigateur web. Nous avons utilisé Unity, un des moteurs de jeux 3D les plus répandus dans l'industrie du jeu vidéo, qui permet notamment d'exporter très facilement son application en ligne. L'interface que nous avons développée permet notamment :

- De visualiser des reconstructions surfaciques 3D de membranes cellulaires d'embryons imagés en microscopie.
- D'interagir avec chacune des cellules, de manière à pouvoir les tourner, les isoler, les cacher, etc.
- De pouvoir effectuer des coupes dans l'embryon pour observer les comportements cellulaires à l'intérieur de l'embryon.
- De superposer aux membranes cellulaires différents codes de couleurs associant des propriétés physiques telles que le volume, le temps de vie de chaque cellule, son voisinage, etc. Ces propriétés, calculées en amont par des biophysiciens, peuvent être chargées à la demande par de simples fichiers de texte respectant une liste de formats préétablie.
- De pouvoir effectuer un ensemble de sélections sur des cellules, et de les propager dans le temps (futur ou passé) permettant ainsi de mieux caractériser l'évolution des clones.
- De créer des captures d'écran ou des films personnalisés exportables pour des présentations ultérieures. De plus, nous avons connecté cette interface de visualisation avec une base de données d'expression génétique, ce qui nous permet d'obtenir une visualisation 4D de la morphogenèse animale associée à l'expression des gènes.

**Participants VORTEX :** Axel CARLIER, Emmanuel FAURE, Sacha VAN-LEENE (ex-stagiaire VORTEX).

**Participants extérieurs :** Patrick LEMAIRE (CRBM, Montpellier), Christophe GODIN (projet Virtual Plants, INRIA, Montpellier).



**Illustration** - Interface de visualisation 4D de la morphogénèse animale accessible par un simple navigateur web.

# Projet POPART

**Mots-clés :** localisation par images, effets spéciaux.

Le projet européen POPART (*Previz for On-set Production - Adaptive Realtime Tracking*), qui a démarré en janvier 2015, s'est terminé en août 2016 par une évaluation et une démonstration à Grimstad (Norvège). Il a permis de développer un système pour la prévisualisation d'effets spéciaux en temps réel sur un tournage de film. L'ensemble des outils livrés permet d'incruster des éléments synthétiques dans des scènes naturelles pour donner au réalisateur du film 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 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*).

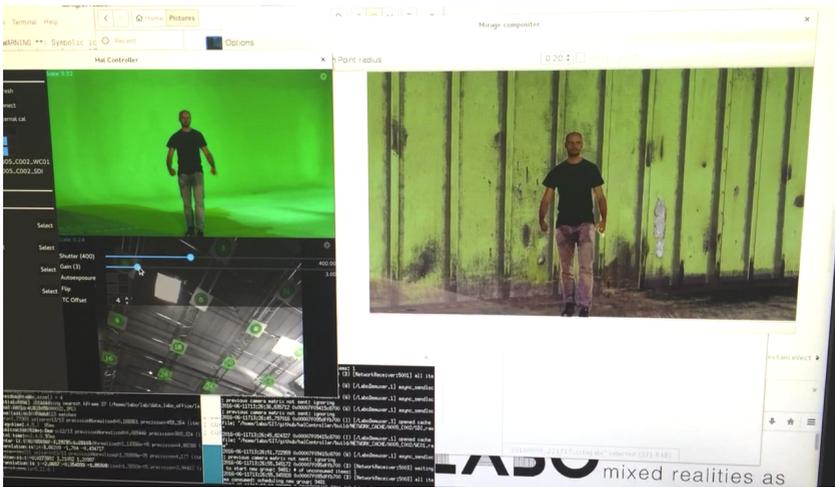
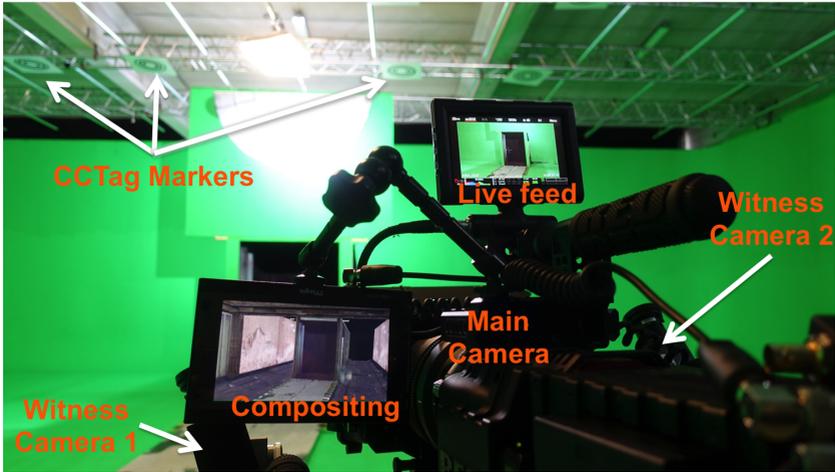
Les nombreux logiciels développés dans le cadre de ce projet sont librement accessibles sur le dépôt *open-source* AliceVision (<https://github.com/alicevision/>). En particulier, la bibliothèque de *structure-from-motion* OpenMVG (<https://github.com/alicevision/openMVG>) permet de créer une base de connaissances à partir d'un ensemble de photographies et d'interroger la base créée pour localiser de nouvelles images. La base de connaissances peut être créée en utilisant des détecteurs de points d'intérêt SIFT (PopSift, <https://github.com/alicevision/popsift>) et/ou des marqueurs concentriques (CCTag, <https://github.com/alicevision/cctag>).

Les thématiques de recherche du projet POPART auront une suite dans le nouveau projet européen LADIO, qui a démarré en décembre 2016 (<http://www.ladioproject.eu>).

**Participants VORTEX :** Clément AYMARD (ex-ingénieur VORTEX), Simone GASPARINI, Pierre GURDJOS.

**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>



**Illustrations** - En haut : système POPART pour la prévisualisation, avec les deux caméras *witness*, le flux vidéo de la caméra principale et l'incrustation en temps réel des objets virtuels. En bas : interface logicielle pour la localisation de la caméra avec les marqueurs détectés (à gauche) et l'incrustation de la vidéo (à droite).

# Projet SCANBAT

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

L'obtention d'un modèle numérique de bâtiment est un processus complexe, qui nécessite l'intervention d'un professionnel muni d'outils adéquats, et donc extrêmement coûteux. Pourtant, de nombreuses applications utiles à des particuliers pourraient découler de l'existence d'un modèle numérique de bâtiment, même simplifié.

Le projet SCANBAT cherche à traiter ce problème, à travers trois questionnements majeurs :

- **Reconstruction** - Comment permettre à un utilisateur non expert d'obtenir un modèle numérique de bâtiment simple en utilisant des capteurs grand public ?
- **Application** - Comment utiliser ce modèle simplifié de bâtiment pour proposer à l'utilisateur une application d'aide à la maîtrise de sa consommation énergétique ?
- **Motivation** - Quels ressorts socio-économiques peuvent inciter l'utilisateur à adopter cette solution ?

Pour ce projet, l'équipe VORTEX était chargée de la partie **Reconstruction**. Nous avons fait le choix d'utiliser des vues prises par des capteurs de type tablette ou *smartphone* (cf. figure ci-contre), puis de formuler le problème en termes de co-segmentation interactive. La principale difficulté du projet réside dans la non expertise des utilisateurs, et donc dans la mise au point d'interactions à la fois simples et intuitives, mais suffisamment informatives pour obtenir de bons résultats.

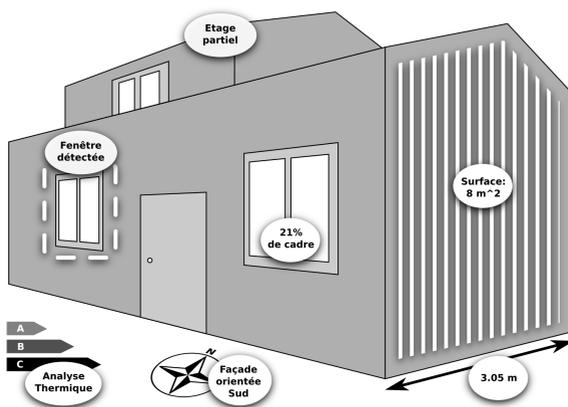
La société Efficiencia, qui conçoit et développe des solutions logicielles pour l'efficacité énergétique des bâtiments, a travaillé principalement sur la partie **Application**, alors que la société Orchis a apporté ses compétences pour la réalisation de la partie **Motivation**.

**Participants VORTEX :** Axel CARLIER, Vincent CHARVILLAT, Matthieu PIZENBERG.

**Partenaires :** sociétés Efficiencia (Toulouse) et Orchis.



**Principe** - Une vue du paradigme d'interaction envisagé dans SCANBAT, impliquant des prises de vue par un utilisateur au moyen d'une tablette et des interactions tactiles simples.



**Illustration** - Quelques exemples des caractéristiques du bâtiment que l'on cherche à obtenir.

# Projet MobVile

**Mots-clés :** recalage 2D/3D, courbure, multi-échelle, flou de défocalisation.

Le projet régional MobVile a pour but de créer des applications mobiles géolocalisées pour la collecte et la gestion visuelles d’alertes émises par les citoyens. Une des fonctionnalités proposées consiste à mettre à disposition des citoyens une application mobile pour déclencher une alerte via une simple photographie permettant d’identifier le problème. Les travaux développés portent sur la qualification visuelle du problème, grâce à un appariement entre des images 2D et des modèles 3D issus d’un inventaire.

Nous souhaitons qualifier l’état d’un élément à partir d’un modèle donné de cet élément, comme par exemple 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. Les images présentent des détails, par exemple des textures sur le fond ou sur l’objet, ou encore des contours occultants, qui ne sont pas présents dans le modèle 3D. Cela complique notablement la tâche visée.

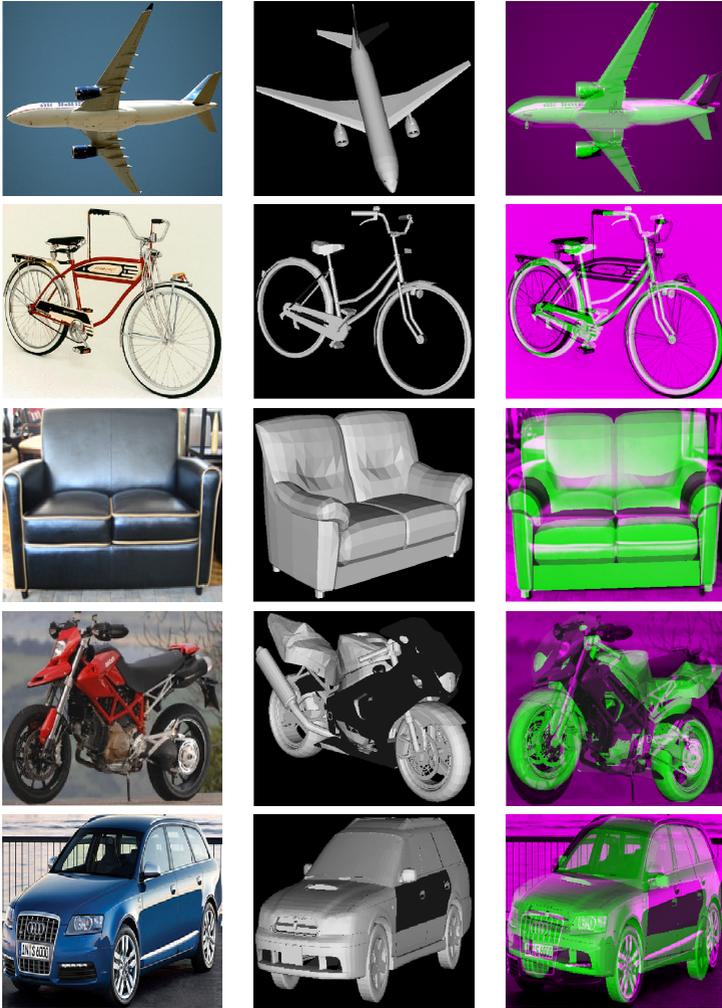
Notre approche consiste à projeter le modèle 3D suivant différents points de vue et à rechercher la projection la plus proche de l’image 2D acquise. Nous utilisons la notion de courbure pour détecter les points d’intérêt dans les deux modalités, en sélectionnant les maxima de courbure (en position et en échelle). Enfin, pour prendre en compte les difficultés liées aux textures et aux contours occultants, nous avons adapté au contexte de la détection de points d’intérêt 2D des méthodes permettant d’extraire la profondeur d’une scène en calculant des cartes issues de l’étude du flou de défocalisation.

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

**Participants extérieurs :** Étienne LAMORT DE GAIL (société Imajing, Toulouse).

**Publication associée :**

Hatem RASHWAN, Sylvie CHAMBON, Pierre GURDJOS, Géraldine MORIN, Vincent CHARVILLAT, *2D-3D Registration based on Multiscale Curviness Saliency and Focus Curves* (soumis).



(a)

(b)

(c)

**Illustrations** - Quelques exemples de recalage entre (a) une image 2D et (b) un modèle 3D. (c) Superposition des données (image 2D en mauve, vue correspondante estimée dans le modèle 3D en vert).

# Projet VITI

**Mots-clés :** détection, contours, segmentation, papillons.

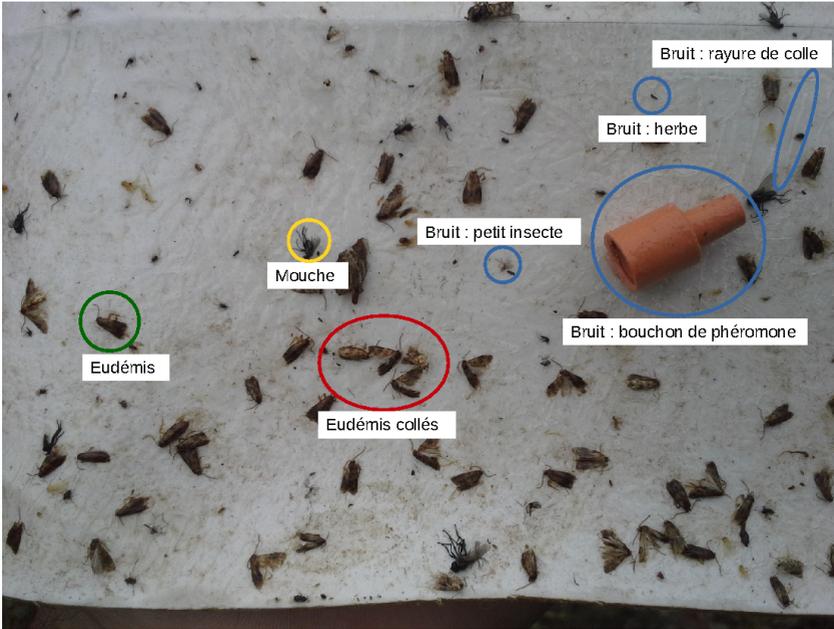
VITI est un projet FUI mené en collaboration avec la société SiConsult. Ce projet a pour objet la protection de la vigne, en étudiant le comportement de l'un de ses parasites, le papillon Eudémis. Plus précisément, le projet a pour objectif la comptabilisation de ces papillons par reconnaissance, à partir d'images acquises en conditions dégradées. La reconnaissance du papillon est difficile à cause du flou (dû à la condensation), du manque de contraste (dû aux conditions climatiques), de la présence de leurres, c'est-à-dire d'autres insectes, et du fait que le papillon Eudémis ne présente pas beaucoup de caractéristiques discriminantes, par rapport à d'autres insectes comme les mouches (cf. figure ci-contre). De plus, les images contiennent de nombreuses sources de bruit : insectes très petits, herbe, lignes de colle. Les insectes peuvent également être superposés dans le piège, ce qui complique les tâches d'identification et de comptabilisation.

L'approche envisagée pour réaliser le traitement consiste à :

1. Détecter les insectes avec précision, à l'aide d'un estimateur de contours robuste.
2. Éliminer les différents types de bruit, en effectuant une classification par l'algorithme des  $k$ -moyennes.
3. Séparer les insectes touchés, en utilisant un algorithme de segmentation fondé sur les régions.
4. Proposer un descripteur, robuste aux changements d'échelle et aux changements d'apparence, permettant de différencier les Eudémis des autres insectes pris au piège, grâce à une étape d'apprentissage.

**Participants VORTEX :** Chafik BAKEY, Sylvie CHAMBON, Hatem RASHWAN.

**Participants extérieurs :** Christian LUBAT (société SiConsult, Toulouse).



**Exemple** - Image traitée, qui contient les papillons à détecter, mais également des leurres (mouches), et comporte les différentes sources de bruit (petits insectes, bouchon de phéromone, herbe) et les difficultés liées aux papillons collés.

## Publications 2016

### Revue internationale :

Viorica PĂTRĂUCEAN, Pierre GURDJOS, Rafael GROMPONE VON GIOI, *Joint A Contrario Ellipse and Line Detection*, IEEE Transactions on Pattern Analysis and Machine Intelligence, volume PP, numéro 99, 2016.

Rabih AMHAZ, Sylvie CHAMBON, Jérôme IDIER, Vincent BALTAZART, *Automatic Crack Detection on Two-Dimensional Pavement Images: An Algorithm Based on Minimal Path Selection*, IEEE Trans. on Intelligent Transportation Systems, volume 17, numéro 10, pages 2718-2729, septembre 2016.

Axel CARLIER, Kathryn LEONARD, Stefanie HAHMANN, Géraldine MORIN, Misha COLINS, *The 2D shape structure dataset: A user annotated open access database*, Computer & Graphics, volume 58, pages 23-30, août 2016.

Cezar PLEZCA, Vincent CHARVILLAT, Wei Tsang OOI, *Multimedia prefetching with optimal Markovian policies*, Journal of Network and Computer Applications, volume 69, pages 40-53, juillet 2016.

Minhui ZHU, Géraldine MORIN, Vincent CHARVILLAT, Wei Tsang OOI, *Sprite tree: An efficient image-based representation for networked virtual environments*, The Visual Computer, pages 1-18, juillet 2016.

Yvain QUÉAU, Richard MODRZEJEWSKI, Pierre GURDJOS, Jean-Denis DUROU, *A Full Photometric and Geometric Model for Attached Webcam / Matte Screen Devices*, Signal Processing: Image Communication, volume 40, pages 65-81, janvier 2016.

### Conférences internationales :

Kathryn LEONARD, Géraldine MORIN, Stefanie HAHMANN, Axel CARLIER, *A 2D Shape Structure for Decomposition and Part Similarity*, International Conference on Pattern Recognition (ICPR 2016), Cancún, Mexique, décembre 2016.

Bastien DURIX, Géraldine MORIN, Sylvie CHAMBON, Céline ROUDET, Lionel GARNIER, *Skeleton-Based Multiview Reconstruction*, IEEE International Conference on Image Processing (ICIP 2016), Phoenix, États-Unis, septembre 2016.

Hatem RASHWAN, Sylvie CHAMBON, Pierre GURDJOS, Géraldine MORIN, Vincent CHARVILLAT, *Towards multi-scale feature detection repeatable over intensity and depth images*, IEEE International Conference on Image Processing (ICIP 2016), Phoenix, États-Unis, septembre 2016.

Fotios LOGOTHETIS, Roberto MECCA, Yvain QUÉAU, Roberto CIPOLLA, *Near-Field Photometric Stereo in Ambient Light*, British Machine Vision Conference (BMVC 2016), York, Royaume-Uni, septembre 2016.

Laurent HOELTGEN, Yvain QUÉAU, Michael BREUSS, Georg RADOW, *Optimised photometric stereo via non-convex variational minimisation*, British Machine Vision Conference (BMVC 2016), York, Royaume-Uni, septembre 2016.

Tomislav PRIBANIĆ, Tomislav PETKOVIĆ, Matea DONLIĆ, Vincent ANGLADON, Simone GASPARINI, *3D Structured Light Scanner on the Smartphone*, International Conference on Image Analysis and Recognition (ICIAR 2016), Póvoa de Varzim, Portugal, pages 443-450, juillet 2016.

Lilian CALVET, Pierre GURDJOS, Carsten GRIWODZ, Simone GASPARINI, *Detection and Accurate Localization of Circular Fiducials Under Highly Challenging Conditions*, International Conference on Computer Vision and Pattern Recognition (CVPR 2016), Las Vegas, États-Unis, pages 562-570, juin-juillet 2016.

Yvain QUÉAU, Roberto MECCA, Jean-Denis DUROU, *Unbiased Photometric Stereo for Colored Surfaces: A Variational Approach*, International Conference on Computer Vision and Pattern Recognition (CVPR 2016), Las Vegas, États-Unis, pages 4359-4368, juin-juillet 2016.

Roberto MECCA, Yvain QUÉAU, *Unifying Diffuse and Specular Reflections for the Photometric Stereo Problem*, Winter Conference on Applications of Computer Vision (WACV 2016), Lake Placid, États-Unis, mars 2016.

Michael BREUSS, Yvain QUÉAU, Martin BÄHR, Jean-Denis DUROU, *Highly Efficient Surface Normal Integration*, Algorithmy 2016, Podbanskè, Slovaquie, mars 2016.

### Conférences nationales :

Thibault BLANC-BEYNE, Géraldine MORIN, Kathryn LEONARD, Stefanie HAHMANN, Marie-Paule CANI, Axel CARLIER, *A 3D shape hierarchy based on the skeleton*, Journées Françaises d'Informatique Graphique (JFIG), Grenoble, novembre-décembre 2016.

Julien FAYER, Géraldine MORIN, Simone GASPARINI, Benjamin COUDRIN, *Réalité diminuée : "couper/coller" interactif pour l'aménagement d'intérieur*, Journées Françaises d'Informatique Graphique (JFIG 2016), Grenoble, décembre 2016.

Yvain QUÉAU, Bastien DURIX, Tom LUCAS, Jade BOUMAZA, Jean-Denis DUROU, François LAUZE, *Fusion de données RVB-D par stéréophotométrie colorée*, Congrès Francophone de Reconnaissance des Formes et Intelligence Artificielle (RFIA 2016), Clermont-Ferrand, juin-juillet 2016.

Bastien DURIX, Yvain QUÉAU, Tom LUCAS, Jade BOUMAZA, Jean-Denis DUROU, François LAUZE, *Étalonnage de sources lumineuses de type LED*, Congrès Francophone de Reconnaissance des Formes et Intelligence Artificielle (RFIA 2016), Clermont-Ferrand, juin-juillet 2016.

Bastien DURIX, Géraldine MORIN, Sylvie CHAMBON, Céline ROUDET, Lionel GARNIER, Olivier COTS, *Reconstruction basée squelette d'un objet 3D à partir de points de vue multiples*, Journées du Groupe de Travail en Modélisation Géométrique (GTMG 2016), Dijon, mars 2016.

### Thèses :

Marie-Anne BAUDA, *Compréhension de scènes urbaines par combinaison d'information 2D/3D*, thèse de l'Université de Toulouse, juin 2016.

# Sommaire

## Images et 3D

Reconstruction 3D multi-vue à partir du squelette . . . . .	3
Estimation de la réflectance à partir de données multi-vue . .	5
Analyse multi-résolution de formes 2D pour la segmentation et l'analyse de formes . . . . .	7
Segmentation hiérarchique de formes 3D basée sur le squelette	9
Compression des données 3D+t de la morphogenèse animale .	11
Stéréophotométrie différentielle . . . . .	13
Stéréophotométrie à deux images . . . . .	15
Habillage par <i>quads</i> d'un squelette . . . . .	17
Rectification métrique de l'image d'un marqueur à contour circulaire . . . . .	19
Détection et localisation de marqueurs circulaires en condi- tions critiques . . . . .	21

## Multimédia

Co-segmentation interactive par entourage . . . . .	23
Photomaton 3D . . . . .	25
Sélection de régions d'intérêt pour la transmission intelligente de contenus 3D . . . . .	27
Base de données de formes 2D segmentées par des utilisateurs	29
Métré d'une salle sur tablette . . . . .	31
Protocole et logiciel pour la capture et l'édition simplifiée de vidéos pédagogiques . . . . .	33
Détection de changements dans une vidéo . . . . .	35
Politiques optimales de préchargement de contenus multimédia	37

## Projets

Projet REALISM . . . . .	39
Projet <i>4D Cloud Embryos</i> . . . . .	41
Projet POPART . . . . .	43
Projet SCANBAT . . . . .	45
Projet MobVille . . . . .	47
Projet VITI . . . . .	49

## Publications 2016

51

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

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