

INTRODUCTION A L'INFORMATIQUE GRAPHIQUE 3D : MODÉLISATION DE FORMES VIRTUELLES ET SYNTHÈSE D'IMAGES NUMÉRIQUES

Tamy BOUBEKEUR

*Professeur en Science Informatique
Télécom ParisTech*

L'informatique graphique 3D est la science de la modélisation des formes, de leur apparence et de leur mouvement en 3D dimensions. Elle se situe à l'intersection des sciences informatiques, de la géométrie appliquée, du traitement du signal et de la physique (lumière, matière, mouvement). Elle trouve son application dans des domaines aussi divers que la conception assistée par ordinateur (CAO), les effets spéciaux, les jeux vidéo, l'architecture, l'archéologie et la médecine. Dans ce document, nous donnons un tour d'horizon des différents champs de cette science et mettons en perspective son rôle central dans le processus qui consiste à traduire le concept mental d'un objet (ou d'un monde) en un modèle numérique reproduit artificiellement dans une machine de calcul.

Modélisation géométrique

La modélisation géométrique consiste à représenter une forme à l'aide de nombres ; valeurs caractéristiques de la forme pour un certain modèle de représentation. On distingue en général les représentations de volumes et de surfaces (interface entre volumes). Pour l'ensemble des applications 3D, on s'intéresse plutôt aux surfaces, car elles concentrent la majorité des calculs et des simulations (réflexion de la lumière, détection de collision, habillage texturé).

Au cours des dernières années, les « maillages surfaciques » (un cas particulier de surface localement paramétrique) ont été largement adoptés dans l'industrie, essentiellement pour trois raisons : tout d'abord ils permettent de représenter des surfaces de topologie arbitraire ; ensuite les algorithmes de synthèse d'images sont en général restreints aux maillages ; enfin ils supportent tout une variété de traitements automatiques puissants (simplification, subdivision, filtrage, reconstruction, remaillage, compression) qui traduisent directement les concepts classiques du traitement du signal et sont critiques dans la chaîne de traitement des données 3D numérisées.

Les maillages surfaciques peuvent être générés de deux façons : automatiquement à l'aide d'une chaîne de capture 3D ou bien interactivement, à l'aide d'outils de modélisation interactive. Ces outils logiciels emploient souvent plusieurs modèles de représentation de surfaces, tels que les surfaces *splines* ou les « surfaces de subdivision », mais permettent toujours d'en extraire un « maillage », à savoir un ensemble de sommets (des points en trois dimensions) reliés entre eux par des polygones (triangles en général). La figure 1 donne une illustration de surface maillée par des triangles.

Traitement géométrique

Le traitement géométrique numérique consiste à transformer automatiquement une forme 3D afin d'en améliorer la qualité (selon des critères objectifs) ou à adapter à un scénario particulier. La plupart des méthodes de traitement géométrique sont conçues pour être appliquées sur les maillages polygonaux. Les plus utilisées sont la « simplification », la « subdivision », le « filtrage », la « reconstruction » et le « remaillage ».

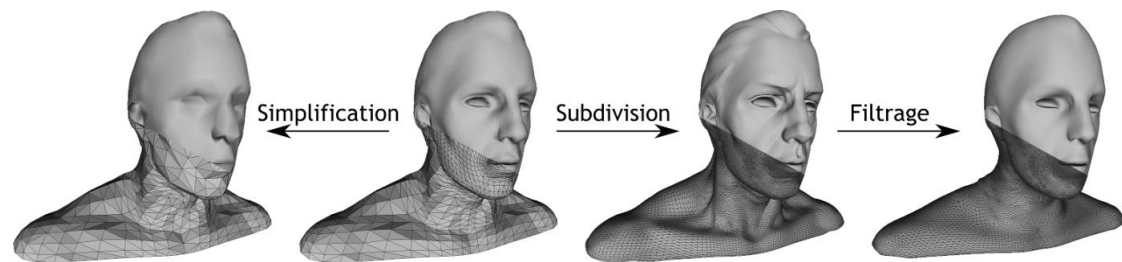


Figure 1 : Principaux traitements numériques de maillages surfaciques

La « simplification de surfaces » consiste à réduire le nombre de polygones employés pour modéliser une surface en essayant de préserver le plus possible la forme générale de l'objet. Elle est typiquement employée pour réduire la résolution des objets numérisés, souvent très denses, afin de pouvoir les manipuler, les transférer, les stocker et les afficher plus efficacement. Elle est également utile lorsqu'un contenu 3D doit être adapté à une plateforme de faible puissance (terminal mobile par exemple).

La « subdivision de surface » effectue l'opération inverse, en permettant d'augmenter la résolution d'un maillage de surface à l'aide de règles faisant tendre la surface obtenue vers une surface lisse. Différents « schémas de subdivision » existent et permettent de raffiner une surface en interpolant ses sommets d'origine, de subdiviser des maillages composés à la fois de quadrangles et des triangles ou encore de préserver les arêtes vives d'une forme. La subdivision de maillages est très employée en modélisation interactive de formes, et permet notamment aux concepteurs de contrôler la forme d'un objet à l'aide de peu de paramètres (maillage grossier de la forme) tout en ayant une surface finalement très dense. Le succès des méthodes de subdivision de surface dans le monde des effets spéciaux et de l'animation par ordinateur a poussé l'industrie du jeu vidéo à faire évoluer le *pipeline* graphique standard et son support matériel (GPU), en introduisant des fonctionnalités similaires pour les moteurs de jeux vidéo modernes. En particulier, les API (ou interface de programmation) de programmation de logiciels 3D temps-réel, telles que DirectX et OpenGL, proposent aujourd'hui une fonctionnalité spécifique permettant de raffiner à la volée le maillage affiché par l'application principale. Cette fonctionnalité, comme d'autres étages du *pipeline* graphique, est programmable : elle peut être redéfinie par le développeur d'application pour produire un maillage plus ou moins raffiné et épousant une forme plus ou moins complexe en fonction de l'application. La figure 1 donne un exemple de subdivision de maillage pour laquelle la résolution du maillage est augmentée et des détails géométriques sont insérés. Ces détails sont typiquement stockés dans une carte, et générés par un infographiste à l'aide d'outils de sculpture numérique. A l'aide de ce type de procédé, l'application

informatique principale peut manipuler un objet simple et économique en mémoire, en ayant néanmoins un objet complexe et détaillé à l'écran.

Enfin, le « filtrage de maillage » n'altère pas la structure sous-jacente des polygones, mais modifie la position des sommets de manière à « lisser » la géométrie, en retirant les variations de haute fréquence, tel que le bruit de capture ou les détails fins de la surface. C'est par exemple le cas du filtre laplacien qui déplace chaque sommet du maillage à la position moyenne de ses voisins. Certains filtres plus avancés, tel que le « filtre bilatéral » par exemple, permettent de préserver, dans une certaine mesure, les bords et arêtes vives de l'objet. Ceci permet de lisser certaines zones en préservant néanmoins les détails importants ou visuellement saillants. Dans une certaine mesure, lorsque le maillage est issu d'un processus de capture 3D, ces méthodes permettent de retirer le « bruit » lié à l'acquisition, à savoir la légère fluctuation erronée de la surface, due aux nombreuses approximations et limites des méthodes de numérisation.

Modèles d'apparence

Alors que la modélisation géométrique donne une description numérique de la macrostructure d'un objet, son apparence est aussi régit par les méso et microstructures de sa surface. La méso-structure est un modèle artificiel permettant de représenter économiquement les détails fins d'une forme. Le plus souvent, elle s'appuie sur diverses cartes de valeurs plaquées sur la surface, stockant par exemple le relief local de, ou la variation, de couleurs le long de la surface. La microstructure de la surface donne quant à elle, une description statistique du relief microscopique de la surface. Ce relief détermine la manière dont un faisceau de lumière percutant la surface, verra son énergie redistribuée dans toutes les directions. On modélise numériquement cette distribution à l'aide d'une fonction, la BSDF (*Bidirectional Scattering Distribution Function*) qui permet, à l'aide de peu de paramètres, de modéliser des matériaux allant des miroirs (réflexion spéculaire parfaite, microstructure lisse) aux objets les plus mats (réflexion diffuse, microstructure hautement rugueuse). Il est intéressant de noter qu'à ce stade, la « couleur » d'un point sur une surface n'est pas, *a priori*, une donnée connue, mais plutôt une réponse qui varie en fonction de l'éclairage de la scène, du point de vue de l'observateur et du milieu.

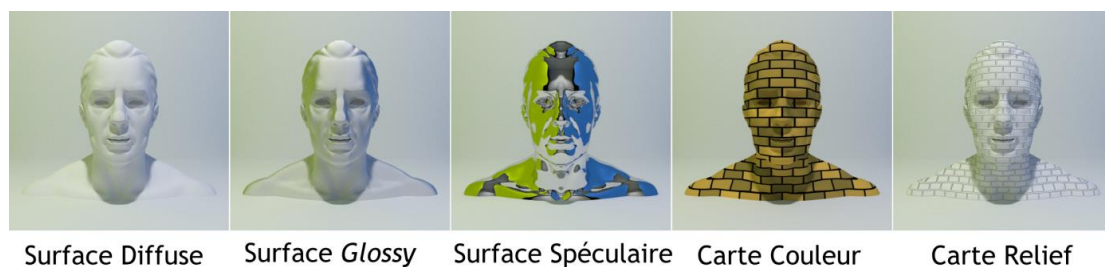


Figure 2 : Modèle d'apparence d'un objet

Synthèse d'images

La synthèse d'image (aussi appelée « le rendu ») vise à générer des images (2D, 3D, stéréo, etc) à partir d'une scène 3D. Cette scène 3D contient des modèles

géométriques équipés de leurs modèles d'apparence, un modèle d'illumination (ensemble de sources lumineuses) et un modèle de capteur (caméra virtuelle).

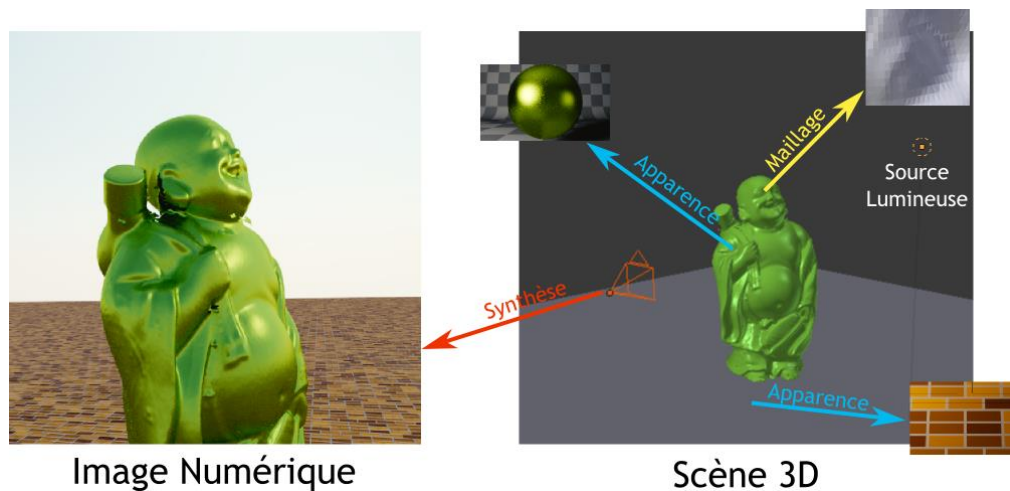


Figure 3 : composantes d'une scène 3D et image de synthèse produite

La synthèse d'image calcule une image en réalisant une simulation du transport de la lumière dans la scène, depuis les sources lumineuses vers le capteur. On peut classer les algorithmes de synthèse d'image en deux familles :

- les méthodes de rendu par « lancer de rayon », très employées pour le rendu physiquement plausible dans les effets spéciaux, les films d'animation ou bien encore la prévisualisation architecturale.
- les méthodes de rendu par « projection », très performantes car supportées par des processeurs spécialisés présents dans la plupart des ordinateurs (*graphics processor unit*, ou GPU), employées pour toutes les applications nécessitant une synthèse d'image à haute fréquence (rendu temps-réel) tels que les jeux vidéo ou la conception assistée par ordinateur.

Le lancer de rayon (*ray tracing*) génère l'image en traçant un grand nombre de rayons « primaires » (un ou plusieurs par pixel de l'image) depuis le capteur vers la scène 3D. Pour chaque rayon, on cherche s'il y a intersection avec un objet de la scène. Si oui, on trace un rayon « d'ombre » vers chaque source lumineuse afin de déterminer si le point est éclairé par une ou plusieurs sources lumineuses. Si oui, la BSDF au point et les différents paramètres du modèle d'apparence permettent de calculer une couleur pour le pixel correspondant sur le capteur. On peut généraliser cet algorithme afin de prendre en compte les multiples rebonds de la lumière au sein d'une scène, où chaque surface devient une source lumineuse potentielle. On parle alors de tracé de chemins (*path tracing*) et cela consiste en l'exploration récursive de la scène en émettant de nouveaux rayons (secondaires) à chaque fois qu'il existe une intersection entre un rayon donné et un objet, jusqu'à un nombre limite de rayons fixé au départ. La manière dont les rayons secondaires sont distribués est régie par la BSDF au point d'intersection et à chaque rebond un rayon d'ombre est de nouveau tiré vers chaque source de lumière afin de mesurer la quantité d'énergie reçue directement au point en question. Le lancer de rayon permet également de simuler des phénomènes optiques complexes comme la « profondeur de champ » ou le « flou de mouvement ». Lorsque les surfaces sont hautement spéculaires, il peut être intéressant d'explorer la scène à la fois depuis le capteur et depuis les sources de lumière,

toujours en traçant de multiples chemins. On parle alors de tracer de chemins bidirectionnels (*bidirectional path tracing*). Ce type d'algorithme est coûteux en termes de calcul et s'appuie sur une structuration forte de la scène, notamment à l'aide de *structures hiérarchiques de partitionnement spatial*. Les plus populaires pour ce type de tâches sont les arbres kD (*kD-Tree*) qui permettent de trouver les intersections rayon-scène en temps logarithmique (au lieu de linéaire pour l'algorithme brute force)¹. On notera que le choix d'une représentation polygonale (triangles) de la scène est bien adapté au lancer de rayon, dans la mesure où il existe une formule analytique simple pour le calcul de l'intersection d'un rayon (demi-droite) et d'un triangle.

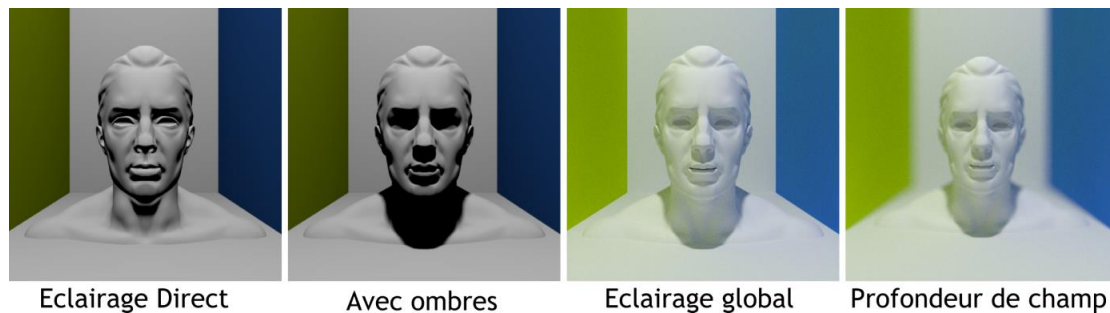


Figure 3 : Synthèse d'image avec simulation de différents phénomènes par tracer de chemins

Le rendu par projection (*rasterization*) traduit un raisonnement inverse. Plutôt que d'explorer la scène 3D depuis le capteur virtuel, on génère l'image en projetant la scène vers le capteur. La détermination efficace de la visibilité en un point de l'image est ici critique afin de s'assurer qu'un pixel est « rempli » par le polygone le plus proche le long de la ligne de vue passant par le pixel. Une solution simple à ce problème consiste à trier les polygones en fonction de leur distance au capteur, puis à projeter les polygones du plus éloigné au plus proche (algorithme *du peintre*). Néanmoins ce tri peut être très coûteux pour les scènes complexes et doit être effectué à chaque mouvement de la caméra ou changement dans la scène. Une solution plus efficace, implémentée matériellement dans les processeurs graphiques, consiste à maintenir une image fantôme stockant la profondeur de la scène en chaque pixel (*Z-Buffer*) : à chaque polygone projeté, on peut ainsi tester si les pixels qu'il recouvre ont déjà été recouverts par un polygone plus proche et dans ce cas, éviter de remplir les pixels en question.

D'une manière générale, la synthèse d'image est un processus de simulation visant, pour la plupart des applications, à générer des images physiquement « plausibles ». L'un des défis principaux tient souvent au besoin d'interactivité des systèmes l'exploitant, qui est en opposition avec le temps parfois très long nécessaire à la simulation lorsque des géométries, des apparences ou des phénomènes complexes sont présents dans la scène. Une solution partielle à ce problème consiste à approximer la solution physique, parfois grossièrement, à l'aide de modèles ne produisant une trop grande erreur perceptuelle, en profitant ainsi des limites naturelles de la vision humaine.

Animation

L'animation par ordinateur ajoute plusieurs modèles numériques de mouvements aux modèles géométriques et aux modèles d'apparences de la scène 3D.

Ils peuvent être très variés, allant de modèles de locomotion pour les bipèdes aux modèles de simulation de fluides, en passant par les modèles de mouvements de foules et les systèmes de particules. On distingue en général trois champs au sein de l'animation par ordinateur :

- la « modélisation interactive » du mouvement où un opérateur humain contrôle quelques paramètres (par exemple un squelette d'animation) inférant le déplacement et la déformation d'un modèle 3D.
- la « simulation physique » qui permet de générer des transformations complexes de la scène (fracture, collision, fluide, etc.) en calculant une approximation de la solution physique d'un ensemble de contraintes géométriques et physiques explicitement données dans la scène numérique.
- la « capture de mouvement » (ou *motion capture*) qui permet, à l'aide de capteurs spécialisés ou de caméras numériques, de suivre le mouvement d'un grand nombre de points dans une scène réelle. Ces trajectoires peuvent ensuite être employées dans le monde virtuel pour déplacer et déformer les modèles numériques.

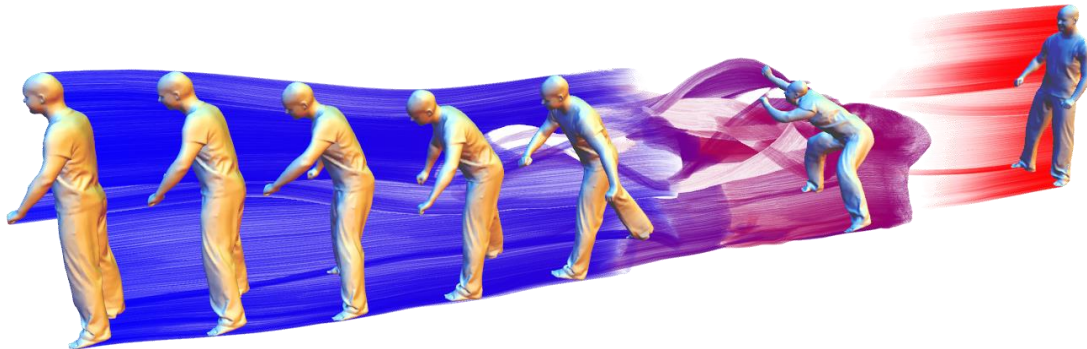


Figure 4 : Modèle de mouvement 3D reconstruit à partir d'une séquence de « performance capture ».

Systèmes interactifs et infographie

Dans le cadre de la conception de formes et de monde virtuels, l'utilisateur a accès à un écosystème d'outils et de logiciels lui permettant de créer rapidement et intuitivement des formes en 3D, en les habillant de paramètres d'apparence et de mouvement. Le cas particulier de la modélisation de formes a grandement évolué depuis le début des années 80 : la conception industrielle passait par un ajustement laborieux de nombreux « points de contrôle » de la forme, alors que les systèmes modernes se rapprochent plus d'outils de « sculpture » numérique, masquant les nombreux paramètres numériques sous-jacents, et donnant simplement accès à des « brosses », « burins » et autres « ciseaux » virtuels pour progressivement générer la forme imaginée. Ces « modeleurs interactifs » sont des systèmes informatiques capables de gérer des masses de données géométriques importantes (jusqu'à plusieurs milliards de points) tout en assurant une visualisation en temps-réel.

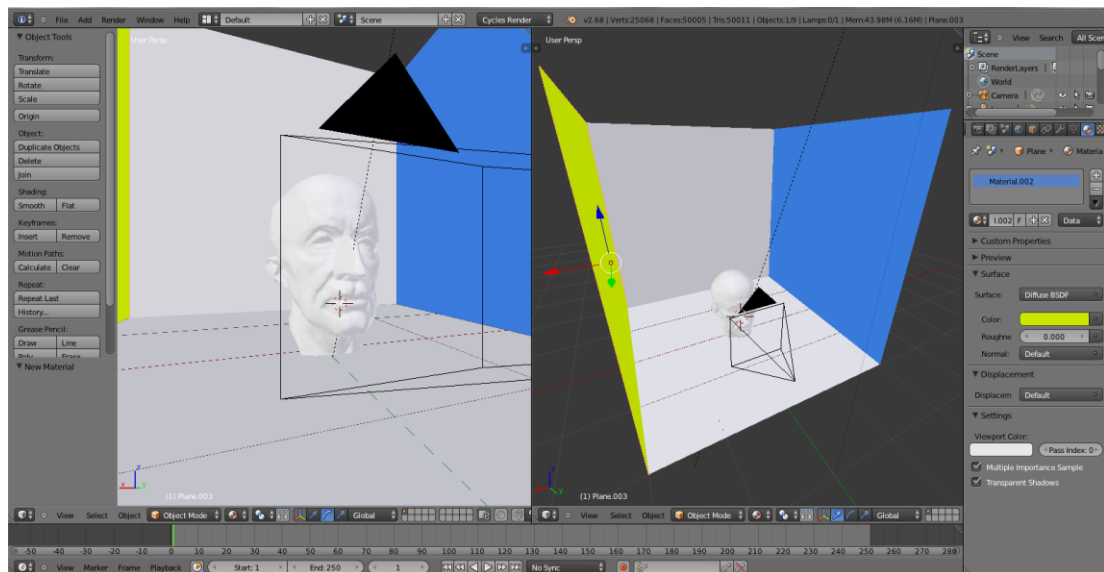


Figure 5 : Le logiciel libre Blender (www.blender.org), un programme interactif de conception 3D.

La multiplication des sources de données 3D (numérisation 3D, créations 3D par les communautés amateurs, modélisation 3D par stéréovision automatique à partir de grandes bases de photos, etc.) ouvre aujourd'hui un nouveau champ d'exploration pour la création de formes et de mondes virtuels. En fournissant un accès instantané à de larges collections d'objets existants, certains outils et plateformes permettant d'économiser un temps conséquent dans la phase de conception, en évitant de réinventer la roue et en permettant de se concentrer principalement sur la création de nouveaux objets. Parmi les problèmes techniques à résoudre pour ce type de scénarios d'applications, la recherche de contenus 3D dans de grandes masses de données 3D nécessite souvent une analyse poussée de modèles stockés. Elle fait appel à plusieurs méthodes issues de la vision par ordinateur, notamment la reconnaissance de formes et la mise en correspondance géométrique.

Capture 3D

La capture 3D permet de créer rapidement des modèles numériques à partir d'objets réels. Elle se fait généralement grâce à un scanner laser, même si elle peut se limiter à quelques photographies numériques d'un objet. Une chaîne de traitement numérique va alors permettre l'assemblage d'un nuage dense de points 3D à partir de l'information capturée, puis le tissage d'une surface maillée entre les points. « Capture de forme » et « capture de mouvement » ont récemment été unifiés au sein de systèmes de « capture de performance », capables de suivre un très grand nombre de points en mouvement dans une scène. Ceci autorise la reconstruction d'un modèle géométrique complet, précis (un acteur par exemple) à chaque pas de temps. On notera également que le modèle d'apparence des formes numérisées peut être capturé : au-delà des matériaux purement diffus, il est souvent nécessaire de contrôler l'éclairage de la scène réelle pour numériser correctement les matériaux plus complexes (comme la peau humaine, les plastiques ou les métaux par exemple).

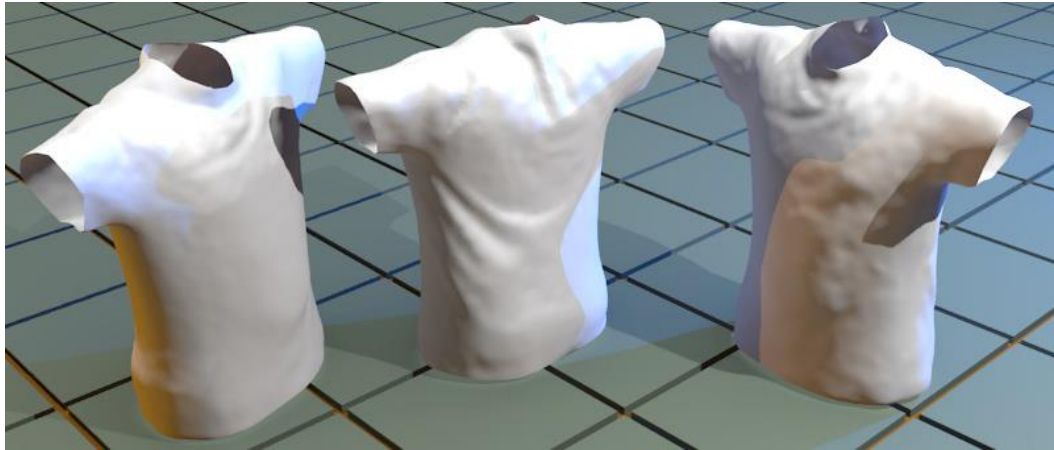


Figure 6 : Modèle de forme en mouvement (3D+temps)
d'un vêtement réel numérisé par capture de performance.

Impression 3D

La synthèse d'images et le contrôle de machines-outils ont longtemps été les principaux vecteurs d'exploitation des modèles 3D numériques. Or, le développement de la « fabrication numérique » (aussi appelée « impression 3D ») depuis le milieu des années 2000 permet aujourd'hui à tout un chacun de créer un objet physique à partir d'un modèle numérique 3D. En fonction de la technologie employée, l'apparence du modèle pourra être partiellement reproduite. Il est à noter que le retour au réel impose un certain nombre de contraintes sur la structure du modèle virtuel, afin de s'assurer de sa stabilité, de sa solidité ou simplement de sa faisabilité.

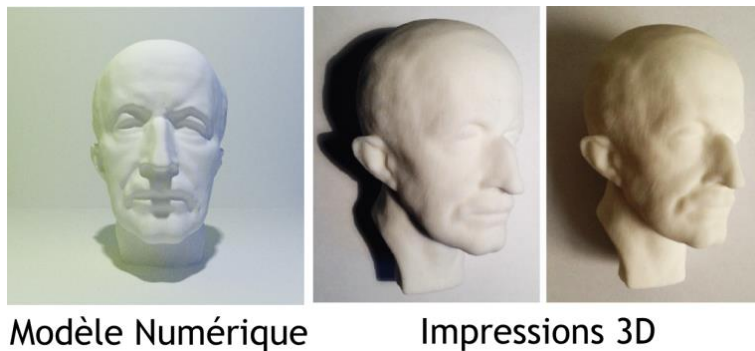


Figure 7 : Impression 3D à partir d'un modèle de forme maillé.

Réalité virtuelle

La modélisation, la synthèse et l'animation d'un monde artificiel fonctionnel et interactif ne sont en général pas suffisants : il est nécessaire d'équiper le système numérique de méthodes et de périphériques destinés à retranscrire fidèlement le monde virtuel à l'utilisateur. Par exemple, on peut utiliser un affichage multi-vues (pour la perception stéréoscopique 3D), reproduire un environnement sonore spatialisant ou encore retranscrire les contacts virtuels à l'aide de périphériques haptiques à retour de force. Certains travaux de recherche vont jusqu'à donner un retour olfactif et gustatif à l'utilisateur. La construction d'un monde virtuel autonome suppose également la simulation d'agent virtuels intelligents, dont le comportement pourra être « scripté » ou dynamique après une phase d'apprentissage à partir de

modèles de comportement humains. Les thèmes de l'intelligence artificielle et de la robotique sont ici explicitement liés à la réalité virtuelle.

Conclusion

L'informatique graphique offre de nombreux outils permettant de concevoir efficacement des mondes 3D numériques. Néanmoins, elle nécessite encore une certaine compréhension technique de la manière dont les outils logiciels fonctionnent en interne (notamment les paramètres des différents algorithmes énumérés tout au long de ce texte) pour être pleinement exploitée. À ce titre, de nombreux travaux de recherche récents visent à rendre la conception 3D aussi simple que peut l'être la génération d'autres médias (texte, son, image, vidéo), en combinant capture 3D simplifiée et interaction rapide sur les modèles obtenus. Malheureusement, un frein au développement plus rapide de ce type d'outils pour les « non-experts » tient au manque de formats standard pour l'échange simple de modèles, malgré les multiples tentatives des organismes de normalisation depuis une quinzaine d'années.

Notes Bibliographiques

Pour aller plus loin dans l'étude de l'informatique graphique, notamment sur ses aspects techniques et scientifiques, on pourra se référer aux séries « The GRAPHICS GEMS » et « GPU GEMS » (developer.nvidia.com) et aux ouvrages suivants :

AKENINE-MOLLER Tomas, HAINES Eric, HOFFMAN Naty, *Real Time Rendering*, A K Peters/CRC Press, 3^e édition, 2008.

BECHMANN Dominique, PEROCHE Bernard, (dir.), *Informatique Graphique, Modélisation Géométrique et Animation*, Paris, Hermès Sciences Publications, Lavoisier, IC2 : Traitement du signal et de l'image, 2007.

HUGUES John F., VAN DAM Andries, MCGUIRE Morgan (*et al.*), *Computer Graphics : Principles and Practice*, Upper Saddle River (NJ), Addison-Wesley, 3^e édition, 2013.

PEROCHE Bernard, BECHMANN Dominique, *Informatique Graphique et Rendu*, Paris, Lavoisier, 2007.

ⁱ En science informatique, on mesure le temps d'exécution des programmes et la taille des objets proportionnellement à une valeur décrivant, par exemple, la taille du problème à résoudre ou des données à traiter. On parle de « complexité ». Ici, un temps logarithmique signifie que pour N polygones dans la scène, le temps de recherche d'une intersection avec un rayon est proportionnel à log (N).