



Condensé technique

Gelato

Rendu de production pour films



Présentation

Le logiciel de rendu NVIDIA® Gelato™ est un moteur de rendu de qualité production qui permet de dire adieux aux compromis. Ses propriétés et fonctions exceptionnelles en font le logiciel idéal pour le rendu des images finales dans la production de films. Les effets spéciaux et les longs-métrages d'animation figurant parmi les applications les plus exigeantes en matière de moteur de rendu, sa réussite dans ce domaine démontre une fois de plus que NVIDIA mène la danse de l'innovation en infographie.

Mais qu'est-ce que veut dire « dire adieux aux compromis » en parlant de moteurs de rendu ? Que faut-il pour effectuer le rendu d'un film ? Que signifie « qualité production » ? En quoi un moteur de rendu de films diffère-t-il d'un moteur de rendu de jeux, de graphiques en temps réel ou d'autres types de rendus ?

Les différences du rendu de films

Il est important de comprendre les différences capitales qui existent entre le rendu des jeux (et des autres applications graphiques en temps réel) et celui des images finales des films et des autres applications graphiques où la qualité prime.

Qualité d'image contre performance de rendu

Les marchés du film et des jeux ne dosent pas de la même façon les éléments du rapport qualité d'image/performance du moteur de rendu. Dans un film, les images sont synthétisées au préalable sur ordinateur et l'accent est mis sur la qualité et la complexité de l'image rendue finale. Toute hausse de qualité entraîne une satisfaction majeure des spectateurs tandis qu'une simple augmentation de la vitesse de rendu passe inaperçue aux yeux du public. Le contraire est vrai pour les jeux : le temps employé pour le rendu d'une image distrait le joueur. C'est pourquoi les moteurs de rendu de films comme Gelato sont principalement conçus pour gérer des scènes d'une complexité impressionnante plutôt que pour booster les performances en temps réel.

Dans les jeux comme dans les films, les scènes sont classées par ensembles d'images contiguës dans le temps et rendues consécutivement par des positions de caméra continues mathématiquement et un ensemble commun d'éléments artistiques. En ce qui concerne les jeux, ces ensembles de trames similaires sont appelés des

« niveaux » et durent plusieurs minutes de jeu. Pour les films, ils sont appelés des « plans » et ne durent généralement que quelques secondes à l'écran.

Compte tenu de cette distinction, les moteurs des jeux et la production de films amortissent de façon différente les frais de rendu. Dans un moteur de rendu de jeux, les textures sont chargées sur des dizaines de milliers d'images, et cette opération est habituellement effectuée avant le dessin de la première image et ne « rentre donc pas » en compte dans le calcul de la vitesse de rendu. Le rendu d'un film en revanche est effectué image par image. Mais même si des plans entiers étaient rendus simultanément pour maximiser la cohérence (sans compter que la complexité des scènes rendrait cela d'une difficulté prohibitive), les frais ne pourraient pas être amortis sur les quelques centaines d'images (au plus) qui composent un plan.

Utilisation des mappes de textures

Une autre différence clé est la façon dont chacun de ces supports utilise les mappes de textures. Les fichiers de texture que la production d'un film emploie dépassent plusieurs giga-octet tandis que ceux utilisés dans les jeux atteignent en général quelques dizaines de méga-octets. Le simple temps nécessaire pour lire les textures d'une image de film depuis un disque ou un réseau peut dépasser plusieurs minutes. Des coûts similaires sont engendrés par la lecture, le tri, l'organisation et la décomposition en polygones des giga-octets qui constituent la géométrie d'une scène.

Quelle que soit la rapidité du matériel graphique et quelle que soit l'ampleur du travail effectué par le matériel graphique à la place du CPU, la manipulation des données nécessaire pour ne serait-ce que préparer le rendu d'une image de film exclut le rendu en temps réel. Ce n'est pas une surprise : les jeux ne seraient pas non plus en temps réel s'il fallait charger de nouveaux niveaux toutes les 2 à 10 secondes de jeu (encore moins à chaque image). La limite inférieure de la vitesse du rendu des images finales d'un film est dictée par le transfert du disque et pas par d'autres facteurs tels que la vitesse de transformation, de remplissage ou d'ombrage.

La Loi de Blinn

Un moteur de rendu de films doit tenir compte de la loi de Blinn, qui dit qu'en présence d'une accélération de la vitesse de rendu, l'auteur augmentera la complexité, ce qui annulera en quelque sorte le temps gagné par le gain de vitesse. En d'autres mots, le temps nécessaire pour rendre une scène est une constante (en général 45 à 90 minutes) qui est fonction de la tolérance du studio en matière d'attente. Les studios répondent à l'introduction de ressources de rendu toujours supérieures par des scènes toujours plus complexes présentant davantage de géométrie, de lumières, des shaders plus sophistiqués et un éclairage réaliste plus coûteux, ils ne se contentent jamais de profiter simplement de l'accélération offerte. Une complexité accrue est un choix évident pour les studios cinématographiques car les images finales sont présynthétisées et insérées dans le film.

L'exception qui confirme la règle est celle de l'éclairage interactif où la production a besoin de vitesses interactives pour obtenir de nombreux rendus similaires

successifs. Dans le cadre de l'éclairage interactif, la géométrie, la caméra et les shaders sont fixes, seuls les paramètres d'éclairage changent d'image en image. Les réalisateurs préfèrent voir les résultats obtenus par le moteur de rendu d'images finales au lieu de ceux de différents outils de prévisualisation de sorte à avoir une idée nette de ce à quoi ressemblera l'image finale. Dans ce cas, un degré de qualité moindre est acceptable. NVIDIA prévoit de lancer un outil d'éclairage interactif pour Gelato dans un futur proche.

Cycles de développement et de production

Une autre différence entre les jeux et les films est liée à leurs cycles de développement et de production respectifs. Les jeux sont joués par des millions de personnes et, pour chaque copie lancée, des centaines de milliers d'images sont rendues. La qualité de l'image, la clarté du code et la possibilité de le réutiliser ainsi que le temps du directeur technique/artistique s'échangent sans arrière-pensée contre une fréquence de défilement garantie.

Les images destinées aux films, en revanche, ne sont rendues qu'une seule fois à pleine résolution. Le temps passé par l'infographiste devant son ordinateur pour rendre les objets est souvent minime face à la quantité et au coût du temps nécessaire à un directeur technique pour écrire le shader. Cela a de nombreuses implications au niveau du développement d'API et explique pourquoi les API et les langages appropriés au temps réel (DX, OGL, Cg) ne le sont pas toujours au rendu hors ligne et vice versa.

Propriétés essentielles du rendu de production de films

Primitives géométriques

Les jeux, les graphiques en temps réel et même les progiciels « professionnels bas de gamme » emploient des polygones (principalement des triangles) pour modéliser les objets. Pour les films en revanche, tout se base en général sur des patches bicubiques, des NURBS (souvent découpées) et, toujours plus, sur des surfaces de subdivision. Les grands ensembles de polygones étant rares dans les images de films, ce sont les NURBS et les surfaces de subdivision qui doivent être optimisées.

Les points (pour les particules) et les courbes (pour le poil) sont également importants, et doivent être rendus efficacement par millions. De plus, les primitives courbes doivent effectuer un découpage adaptatif basé sur la surface de l'écran et sa sphéricité, de sorte à ne *jamais* donner d'artefacts de tessellation.

Complexité géométrique

La composante géométrique d'une unique image de film peut facilement consister en plusieurs giga-octets (il n'est pas rare qu'elle atteigne 25 Go). C'est pourquoi le premier point du cahier des charges d'un moteur de rendu de films doit être sa capacité à gérer davantage de géométrie que la RAM ne peut en contenir à un instant donné.

Gelato utilise pour cela tout un éventail de stratégies qui vont de la segmentation de la mémoire à la mise en cache sur disque en passant par le tri, le tri sélectif ou *culling*, le tests d'occlusion et la géométrie procédurale.

Complexité des textures et qualité

Une image de film requiert souvent des centaines ou des milliers de textures, nécessitant des dizaines de giga-octets de stockage. Les moteurs de rendu de films les plus efficaces utilisent un système de cache où des éléments cohérents individuels sont paginés à partir du disque, en fonction des besoins. Il est impossible de lire toutes les textures dans la mémoire avant le rendu. Et puisque les noms des textures peuvent être générés de manière dynamique par le shader, les noms des fichiers de textures restent souvent inconnus jusqu'à l'exécution du shader.

Les recherches dans les tables de texturage doivent être d'une grande fidélité. Les cartes de textures, d'environnement et d'ombrage doivent se reproduire sans artefacts, saccades ou bruit. Et dans de nombreux cas, les tables doivent avoir un filtre dont la qualité dépasse le niveau du mip-mapping trilinéaire.

Motion blur (simulation d'objets en mouvement)

Les moteurs de rendu de films doivent être en mesure de produire un effet de mouvement ou *motion blur* de haut niveau, sans effet stroboscopique et avec un bruit réduit. En sus, cet effet doit associer un flou de transformation (changement de position, d'orientation et d'échelle) à un flou de déformation géométrique. Les meilleurs éléments de cette catégorie de moteurs de rendu, à l'instar de Gelato, supportent le *motion blur* multi-segment pour un nombre arbitraire de nœuds.

Les effets de profondeur de champ sont également importants.

Anticrénelage

Les films ne tolèrent aucun degré d'artefacts de crénelage visibles. Cet impératif est normalement respecté en prenant de nombreux points échantillons par pixel. Pour les scènes comportant des cheveux ou du poil, il est courant d'en prendre 64 à 100 par pixel. Le moteur de rendu doit reconstruire ces données sous-pixel en pixels avec un filtre spécifié par l'utilisateur de haute qualité dont la prise en charge soit supérieure à un pixel.

Taille, profondeur, format et données de l'image

Les films requièrent des résolutions d'image arbitraires (4 K pour les images finales, 8 K ou plus ne sont pas rares pour les cartes d'ombrage) ; des profondeurs de bits flexibles (8, 16, half, float) ; des formats d'images pratiques (Gelato prend en charge TIFF, ainsi qu'une API simple pour les programmes d'écriture d'images fournis par l'utilisateur).

Il est également capital que le moteur de rendu ne rende pas que les données de couleur, de profondeur et alpha, mais toutes les données calculées par les shaders. Cette dernière propriété est appelée « Arbitrary Output Variables » (variables de sortie arbitraires) ou AOV, dans l'industrie cinématographique tandis que dans le domaine des graphiques en temps réel elle est souvent appelée « Multiple Render Targets » (cibles de rendu multiples) ou MRT. Les images sont enregistrées sur disque ; effectuer le rendu directement dans la mémoire graphique n'est pas important.

Absence de limites

La production de films. ne tolère pas les limites arbitraires telles que celles relatives au nombre de textures, à la taille de la base des données géométriques, à la résolution, au nombre de canaux de sortie et au nombre de sources lumineuses.

Illumination globale

Les logiciels de rendu de films modernes comme Gelato prennent en charge les réflexions et les ombres basée sur le lancer de rayons, le transport de la lumière diffuse, les solvants, l'éclairage environnemental, l'occlusion ambiante et la diffusion des surfaces sous-jacentes.

Déplacement

Les déplacements de fréquence découpés de façon adaptative au niveau sous-pixel sont nécessaires dans un moteur de rendu et ne devraient pas utiliser davantage de ressources système que la géométrie sans déplacement .

Ombrage programmable flexible

Il suffisait jusqu'il y a peu de dire qu'un rendu de qualité cinématographique requérait un ombrage programmable (prouesse rare au niveau logiciel et impossible au niveau matériel). Mais, maintenant que l'ombrage programmable est monnaie courante — même dans le matériel — il convient d'énumérer les caractéristiques qui distinguent l'ombrage programmable pour films de ceux adaptés au rendu en temps réel :

- ❑ Des types de données riches tels que les vecteurs, couleurs, points, normales, matrices de longueur arbitraire.
- ❑ Les variables chaînes ; la manipulation de chaînes ; et la possibilité de faire référence à des textures, systèmes de coordonnées et données externes par un nom plutôt que par des pointeurs ou le transfert explicite de matrices sous forme de paramètres.
- ❑ La capacité de masquer les détails et les limites du matériel. Les directeurs techniques qui écrivent des shaders devraient pouvoir tout ignorer de la précision réduite des types de données, des limites de la mémoire et du nombre d'instructions ou encore des noms et des détails des registres matériels. Le moteur de rendu ne doit jamais créer d'artefacts ni perdre en précision à cause de ces limites matérielles.
- ❑ La possibilité d'appeler du code fourni par l'utilisateur sur l'hôte (fonctions intégrées ou *DSO shadeops*).
- ❑ Un bouclage indépendant des données, spécialement pour les applications telles que le *ay marching* pour les effets volumétriques, ou l'intégration numérique au sein des shaders.
- ❑ La compilation séparée des différents types de shaders et en particulier des ceux d'éclairage.
- ❑ La précision de calcul des dérivées et de l'anticrénelage. Les dérivées doivent utiliser la différenciation centrale et l'extrapolation sur les bords de la grille. Une simple différenciation vers l'avant donne des artefacts inacceptables.
- ❑ Une séquence de shaders sur une surface, avec un mécanisme flexible de superposition et de connexion des entrées et sorties. En rendant cela possible, Gelato repousse les limites de la technique.
- ❑ L'entrée et la sortie d'un nom arbitraire, d'un type de données et d'un nombre d'arguments de paramètres dans et hors des shaders.

API et formats appropriés

Un logiciel de rendu de films doit s'insérer dans le pipeline de production d'un film, ce qui signifie au minimum :

- ❑ Une API procédurale nette (C ou C++). OpenGL et DX exposent de trop nombreux détails matériels ; ont une prise en charge limitée pour ce qui est des surfaces de niveau supérieur et l'attribution de shaders et l'association ; et sont

gênées par les larges parties de l'API consacrées à la fonctionnalité interactive qui n'est pas utilisée dans les films.

- ❑ Un format d'archivage de scènes pour gérer la géométrie. Gelato unifie le format d'archivage de scènes en écrivant des primitives procédurales avec le langage de scripts Python. Il est également possible d'écrire des plug-ins Gelato qui lisent directement tout format de scène, sans passer par aucun format intermédiaire arbitraire.
- ❑ La géométrie procédurale en tant que concept phare : par exemple, lecture de fichiers d'archives ; exécution de programmes produisant de la géométrie ; ou DSO appelant des API exécutées paresseusement comme demandé par le moteur de rendu.
- ❑ Un langage d'ombrage qui incorpore tous les éléments demandés par un film. Gelato a son propre langage d'ombrage (GSL) pour satisfaire ce besoin.
- ❑ La flexibilité de lire tout une variété de formats utilisés par les studios cinématographiques. Gelato a une API souple qui accepte pratiquement tous les formats et permet de tester facilement Gelato et de l'intégrer facilement dans les procédés existants d'un studio.
- ❑ Des plug-ins pour les principaux logiciels 3D. Gelato a un plug-in qui permet d'exporter des données de Maya à Gelato ; des plug-ins d'export pour d'autres systèmes de modélisation verront bientôt le jour.

Plate-forme appropriée

La plupart des grands studios utilisent Linux, Windows venant en général juste derrière. Gelato prend en charge ces deux systèmes d'exploitation et prendra également en charge Mac OS X si l'industrie du film le demande.

Conclusion

Le rendu de films diffère considérablement de celui en temps réel utilisé dans les jeux et ce, au niveau des exigences comme sur le plan du poids accordé aux différents éléments. Il est donc parfaitement compréhensible que les réalisateurs de films demandent un outil spécialement conçu à cette fin.

NVIDIA offre un outil unique qui répond aux exigences du marché. En lançant des produits comme NVIDIA Gelato, NVIDIA apporte la preuve concrète de ses efforts en faveur de l'industrie cinématographique, efforts qui trahissent une implication semblable à celle pour laquelle l'entreprise est connue dans le domaine des jeux vidéo.

NVIDIA.

Avis

L'ENSEMBLE DES SPÉCIFICATIONS DE CONCEPTION, CARTES DE RÉFÉRENCE, FICHIERS, DESSINS, DIAGNOSTICS, LISTES ET AUTRES DOCUMENTS NVIDIA (DÉSIGNÉS ENSEMBLE ET SÉPARÉMENT COMME LES « MATÉRIAUX ») SONT FOURNIS « EN L'ÉTAT ». NVIDIA NE FOURNIT AUCUNE GARANTIE, QU'ELLE SOIT EXPRESSE, TACITE, LÉGALE OU AUTRE, CONCERNANT LES MATÉRIAUX, ET EXCLUT EXPRESSÉMENT TOUTE GARANTIE IMPLICITE DE CONTREFAÇON, DE QUALITÉ MARCHANDE ET D'APTITUDE À UN USAGE PARTICULIER.

Les informations ci-incluses sont censées être précises et fiables. Toutefois, NVIDIA Corporation décline toute responsabilité quant aux conséquences de l'utilisation qui pourrait en être faite ou de la contrefaçon de brevets ou autres droits de tierces parties pouvant résulter de leur utilisation. Aucune licence n'est octroyée implicitement ou de quelque autre manière sous quelque brevet ou droit de brevet de NVIDIA Corporation. Les caractéristiques techniques mentionnées dans ce document peuvent être modifiées sans préavis. Cette publication annule et remplace toute information diffusée antérieurement. Les produits de NVIDIA Corporation ne peuvent en aucun cas être utilisés en tant que composants critiques pour des systèmes de survie sans l'accord préalable écrit de NVIDIA Corporation.

Marques

NVIDIA, le logo NVIDIA et NVIDIA nForce sont des marques ou des marques déposées de NVIDIA Corporation. Les autres noms de sociétés et de produits cités sont des marques commerciales de leurs sociétés respectives ou des sociétés auxquelles ils sont associés..

Droits d'auteur

© 2004 by NVIDIA Corporation. Tous droits réservés.



NVIDIA.

NVIDIA Corporation

2701 San Tomas Expressway

Santa Clara, CA 95050

www.nvidia.com