

12/12/2022

Auteurs :

CANIN Corentin - corentin.canin@etu.univ-grenoble-alpes.fr

WAGNER Samy - samy.wagner@etu.univ-grenoble-alpes.fr

NVIDIA Deep Learning Super Sampling

Résumé

Le marché vidéoludique s'est imposé ces dernières années comme le marché économique médiatique le plus important. Moteurs d'innovations dans le domaine de l'informatique, les jeux vidéo poussent les technologies matérielles à leurs limites pour offrir une qualité visuelle toujours plus impressionnante. Limités par ce matériel, les constructeurs ont dû développer des technologies alternatives au rendu graphique classique afin de repousser plus loin les limites du possible. On peut citer les technologies X^eSS d'Intel, FSR d'AMD ou encore la technologie DLSS de Nvidia.

Mots-clés

Jeux vidéo, rendu graphique, processeur graphique, mise à l'échelle, IA, apprentissage profond, fréquence de trame, performances.

Abstract

In the past years, the video game industry became the largest media industry. A driving force in computer innovations, video games push hardware to its limits to always offer the players better graphic quality. Restricted by this hardware, manufacturers had to develop alternative technologies to regular rendering to push back the limits of what is possible. We can mention the following technologies : X^eSS from Intel, FSR from AMD or DLSS from Nvidia.

Keywords

Video games, graphical rendering, Graphics Processing Unit, upscaling, AI, deep learning, frame rate, performance.

Contexte

Ces dernières années, la course à la qualité d'image fut plus intense que jamais. Avec l'apparition des écrans 1080p, 4K, 8K... la résolution maximale d'une vidéo YouTube a quadruplé en l'espace de seulement 7 ans entre 2009 et 2016, et le jeu vidéo est le premier secteur à souffrir de cette montée effrénée des exigences des utilisateurs.

Puisque le jeu vidéo est l'un des seuls domaines où la génération d'images par ordinateur doit être quasiment instantanée afin de limiter l'inconfort du joueur qui doit réagir en temps réel, et parce que les joueurs veulent profiter de la meilleure expérience visuelle durant leurs parties, il est vite devenu essentiel pour les constructeurs de trouver des solutions proposant les meilleurs performances en terme de graphismes et de fluidité. L'objectif de Nvidia, l'un des leaders du marché des processeurs graphiques était de développer une technologie permettant d'alléger la charge des cartes graphiques qui sont de plus en plus sollicitées par les nouveautés sur le marché vidéoludique comme le ray-tracing (que Nvidia propose également avec sa gamme RTX) ou la réalité virtuelle.

C'est en février 2019 que Nvidia sort officiellement la version 1.0 de sa technologie DLSS (*Deep Learning Super Sampling* ou *Suréchantillonnage par Apprentissage Profond*) utilisant l'intelligence artificielle pour soulager la charge du processeur de rendu graphique. Cette fonctionnalité est incluse dans leur nouvelle gamme de cartes graphiques RTX 20. Elle recevra deux mises à jour : la version 2.0 en août 2020 et plus récemment fin 2022, la version 3.0 exclusive à la gamme RTX 40. La raison pour laquelle cette technologie est exclusive aux gammes RTX de Nvidia est qu'elles embarquent des processeurs dédiés à l'exécution d'intelligence artificielle appelés Tensor Cores.

DLSS 1.0

Le jeu vidéo est un domaine du rendu graphique par ordinateur particulier, au sens où chaque trame composant l'image animée affichée à l'écran est générée en temps réel, réagissant aux actions du joueur sur la manette. On peut ainsi le placer en opposition au cinéma d'animation en images de synthèse par exemple où toutes les trames composant l'image sont pré-générées à l'avance et ne nécessitent pas un calcul en temps réel de la géométrie 3D. Il est donc très important, dans le cas du jeu vidéo, que le rendu graphique de chaque trame soit extrêmement rapide, idéalement durant moins d'1/60 secondes, afin d'offrir au joueur un jeu fluide et réactif. Les performances du processeur graphique sont alors pleinement sollicitées, et ne sont parfois pas suffisantes.

L'objectif de la technologie DLSS est d'alléger la charge du processeur graphique en effectuant le rendu graphique à une résolution inférieure à la résolution souhaitée, puis en agrandissant l'image rendue à la résolution souhaitée grâce à un modèle d'intelligence artificielle exécuté par des processeurs dédiés. Le processeur graphique peut ainsi effectuer le rendu de plus de trames pour une même durée, permettant d'augmenter la fréquence de trame du jeu (fréquence à laquelle une nouvelle trame est rendue et affichée) et ainsi la fluidité du jeu.

La première itération de la technologie DLSS se décompose en deux étapes reposant sur un réseau neuronal auto-encodeur convolutif : la première étape consiste en un renforcement des arêtes sur l'image et l'application d'un filtre d'anticrénelage (plus connu sous le terme

anglais *antialiasing*) sur la dernière trame rendue. Cette étape utilise comme données d'entrée la dernière trame rendue par le processeur graphique ainsi que des vecteurs de mouvement fournis directement par le moteur de jeu.

La seconde étape consiste en une mise à l'échelle de l'image obtenue après la première étape à la résolution souhaitée. Cette étape n'utilise en entrée du réseau neuronal que la trame à traiter. Cette quantité si réduite d'informations implique que le réseau doit générer une grande quantité d'informations nouvelles pour créer l'image à la résolution souhaitée, menant ainsi souvent à la création d'artéfacts ou d'erreurs incohérentes avec l'image qu'on aurait pu obtenir si on avait choisi un rendu natif à la résolution finale.

DLSS 1.0 entraînait son modèle neuronal sur le superordinateur Nvidia Saturn V, par suréchantillonnage. Malgré le nombre élevé d'échantillons utilisés pour l'entraînement (64 par pixel) et le nombre de variations appliquées (bruit, rotations, différents niveaux d'éclairage, résolutions, paramètres graphiques etc.), la technologie reçoit un accueil mitigé de la part des consommateurs déplorant la présence d'artéfacts sur l'image finale obtenue, un effet secondaire résultant du nombre réduit d'entrées au moment de la mise à l'échelle, l'empêchant de s'entraîner correctement sur tous les scénarios et cas limites. Un autre défaut critiqué du DLSS 1.0 est le peu de jeux supportés par celui-ci, dû au fait qu'il devait être entraîné individuellement sur chaque jeu.

DLSS 2.0

La majeure différence entre DLSS 2.0 et DLSS 1.0 est la nature des données d'entrée du réseau de convolution. DLSS 2.0 récupère beaucoup plus de données directement depuis le moteur de jeu que DLSS 1.0 afin d'affiner ses résultats. Il utilise notamment des vecteurs de mouvements, des informations de profondeur ainsi que de luminosité.

Contrairement à DLSS 1.0 qui n'utilise que la trame actuelle en entrée, DLSS 2.0 effectue du temporal antialiasing (anticrénelage temporel), il prend donc en entrée non seulement la trame actuelle mais également les trames précédentes afin de reconstruire des détails avec une meilleure précision. Cette technique était déjà couramment utilisée dans plusieurs moteurs de jeu modernes, cependant elle était implémentée de façon manuelle, avec des techniques de correction de certains problèmes tels que le scintillement ou des images rémanentes entraînant une perte en détails sur l'image.

DLSS 2.0, lui, implémente ces corrections via l'intelligence artificielle, ce qui donne un résultat beaucoup plus précis que les méthodes manuelles habituelles, bien que toujours imparfait. Les performances de DLSS 2.0 du point de vue de la correction restent cependant meilleures que celles d'un antialiasing classique. On peut noter que DLSS 2.0 peut être utilisé uniquement pour son aspect antialiasing, sans mise à l'échelle à posteriori si on le souhaite afin d'obtenir une meilleure qualité de rendu.

Concernant la mise à l'échelle, DLSS 2.0 se sert des données supplémentaires à sa disposition pour effectuer bien plus qu'une mise à l'échelle classique comme un filtre bilinéaire, les réseaux ESRGAN ou DLSS 1.0 qui fournissent souvent des résultats flous car ils ne travaillent qu'avec une seule image d'entrée. Puisque DLSS 2.0 a à sa disposition les trames précédentes, il peut réutiliser ces informations plutôt que de les "réinventer" à chaque trame.

En raison de tous ces changements entre DLSS 1.0 et DLSS 2.0, DLSS 2.0 obtient de bien meilleurs résultats concernant la rétention de détails ainsi que sa rapidité d'exécution. De plus, l'entraînement du modèle est maintenant généralisé et ne nécessite plus un entraînement spécifique par jeu, faisant grimper le nombre de jeux compatibles de 6 à 216, sans compter la compatibilité de DLSS 2.0 avec les deux moteurs de jeu 3D les plus populaires du marché : Unreal Engine 4 et Unity. DLSS 2.0 fut ainsi bien mieux accueilli que sa version 1.0 et représente une réelle révolution dans l'industrie en termes d'amélioration des performances et de la qualité visuelle.

DLSS 3.0

Annoncé en septembre 2022, DLSS 3.0 est une évolution de la version 2.0. Il ajoute au procédé une analyse de flux optique, accéléré par des cœurs dédiés aux flux optiques dans l'architecture Ada utilisée dans les RTX 40. Cette analyse calcule un champ de vecteurs représentant le déplacement des pixels entre chaque image, permettant selon NVIDIA de prendre en compte les mouvements de plus fins détails comme les particules ou les ombres, là où le moteur de jeu fournit seulement des vecteurs pour le déplacement des objets et de la géométrie. Le réseau de convolution a ainsi 4 entrées, soit les trames actuelle et précédentes et les vecteurs du moteur du DLSS 2.0, et le flux optique calculé entre les deux dernières trames. L'ajout du flux optique permet la génération de nouvelles trames par IA entre celles mises à l'échelle par la technologie du DLSS 2.0 sans des anomalies visuelles qui seraient présentes si seuls les vecteurs du moteur étaient utilisés, telles que des ombres décalées. DLSS 3.0 augmente ainsi le nombre d'images par seconde en insérant des images qui ne sont pas calculées par les cœurs de rendu du GPU, dont le travail est déjà allégé par le rendu à plus basse résolution permis par DLSS 2.0 : selon NVIDIA, ce sont sept huitièmes des pixels affichés qui sont alors reconstruits par IA plutôt que d'être calculés.

Concurrents

Nvidia n'est pas le seul constructeur à proposer une technologie reposant sur la mise à l'échelle d'un rendu moins gourmand en ressources dans le but d'améliorer les performances de jeu pour une qualité graphique similaire.

AMD propose depuis juin 2021 la technologie FidelityFX Super Resolution (ou FSR), une alternative open source à DLSS. Celle-ci ne repose pas sur l'intelligence artificielle mais sur une version modifiée de l'algorithme de Lanczos. Sa version 2.0 utilise en entrée une image basse résolution sans filtre d'anticrénelage et des données moteur telles que des vecteurs de mouvement. Présentant des performances similaires, bien que très légèrement inférieures à DLSS 2.0, elle a l'avantage de ne pas être exclusive aux cartes AMD et est virtuellement compatible avec tous les jeux compatibles avec DLSS 2.0, tant que ceux-ci ne vérifient pas explicitement le processeur graphique ou le pilote utilisé.

L'autre leader du marché des processeurs, Intel, propose quant à lui depuis octobre 2022 le Xe Super Sampling (ou XeSS). Il s'agit d'une technologie basée sur l'intelligence artificielle, comme DLSS, qui est conçue pour tirer profit des processeurs IA XMX des cartes

graphiques Intel Arc. Bien que légèrement moins performante que DLSS 2.0, la technologie d'Intel a l'avantage d'être open-source et de ne pas être réservée aux cartes Intel, fonctionnant même sur la gamme GTX 10 de Nvidia.

Conclusion

La technologie DLSS et ses concurrents représentent une belle avancée pour l'industrie vidéoludique. Qu'elles soient basées sur l'intelligence artificielle ou non, elles profitent à tous, des joueurs modestes pour qui obtenir des performances convenables sur des jeux gourmands est difficile, aux joueurs plus aisés à la recherche d'une expérience en ultra haute définition. Même les développeurs y gagnent, ayant désormais plus de marge pour proposer des environnements toujours plus complexes, sans risquer de sacrifier la fluidité de jeu (le tristement célèbre Cyberpunk 2077 en profite largement).

On peut également saluer l'initiative d'AMD et d'Intel qui proposent des alternatives open-source et disponibles sur d'autres processeurs que leurs cartes propriétaires. Le DLSS est une technologie à surveiller, ses qualités d'amélioration de l'image pourraient bien être la pierre angulaire de l'industrie vidéoludique, voire même du rendu d'image par ordinateur, de demain.

Bibliographie

Contributeurs Wikipédia. (25 novembre 2022). *Deep learning super sampling*. Wikipédia.

https://en.wikipedia.org/wiki/Deep_learning_super_sampling

NVIDIA DLSS 2.0 : *A Big Leap In AI Rendering*. (s. d.). GeForce News | NVIDIA.

<https://www.nvidia.com/en-us/geforce/news/nvidia-dlss-2-0-a-big-leap-in-ai-rendering/>

GTC 2020 : *DLSS 2.0 - Image Reconstruction for Real-time Rendering with Deep Learning*. (10 juin 2020). NVIDIA Developer. <https://developer.nvidia.com/gtc/2020/video/s22698-vid>

Accès gratuit sur YouTube : <https://www.youtube.com/watch?v=d5knHzv0IQE>

Introducing NVIDIA DLSS 3. (s. d.). NVIDIA.

<https://www.nvidia.com/en-us/geforce/news/dlss3-ai-powered-neural-graphics-innovations/>

AMD. (17 mars 2022). *Introducing Radeon Super Resolution and FidelityFX Super Resolution 2.0*.

YouTube. <https://www.youtube.com/watch?v=JUQ8j-bpQ1Q>

Contributeurs Wikipédia. (8 novembre 2022). *GPUOpen*. Wikipédia.

<https://en.wikipedia.org/wiki/GPUOpen>

Roach, J. (5 octobre 2022). *What is Intel XeSS, and how does it compare to Nvidia DLSS ?* Digital

Trends. <https://www.digitaltrends.com/computing/what-is-intel-xess/>