

Lorraine

# Dynalips : les secteurs du jeu vidéo et de l'animation bouleversés

Il aura fallu plus de dix ans de recherche et développement au Pr Slim Ouni, informaticien, à Louis Abel et à Théo Biasutto-Lervat pour créer Dynalips dont la technologie DeepLipSync révolutionne le traitement de la parole dans les jeux vidéo et les films d'animation. Une pépite lorraine.

Il aura fallu une dizaine d'années de recherche et développement au Pr Slim Ouni, chercheur en informatique, à Louis Abel, doctorant en informatique, et à Théo Biasutto-Lervat, ingénieur en intelligence artificielle, pour créer la technologie DeepLipSync, portée par la start-up Dynalips. Lancé en cette première semaine de juin, ce logiciel permet de faire parler de manière réaliste et en quelques minutes, des personnages de jeux vidéo (première industrie culturelle mondiale avec un chiffre d'affaires en 2023 de 168 milliards d'euros) et de films d'animation. Une technologie de rupture qui vient bouleverser le monde de l'animation en lui simplifiant la tâche.

Les scientifiques, qui exercent leurs talents au sein du Laboratoire lorrain de recherche en informatique et ses applications (Loria) de Villers-lès-Nancy (54), une structure commune au CNRS, à l'Université de Lorraine, à l'Inria et à CentraleSupé-



Il aura fallu une dizaine d'années de recherche et développement au Pr Slim Ouni, à Louis Abel et à Théo Biasutto-Lervat pour créer Dynalips. Photo Thierry Fedrigo

lec, se sont positionnés sur une difficulté majeure de l'animation numérique : rendre des personnages numériques en 2D ou en 3D doués de parole, leur insuffler la vie.

## Shrek et Buzz L'Éclair

Shrek ou Buzz L'Éclair ont été rendus attachants parce qu'ils paraissent plus vrais que nature, incarnés. Or, pour atteindre ce niveau de vérité en images de synthèse, il faut des mois de pétrissage informatique. Donner vie à des dessins requiert des légions d'animateurs, qui passent des heures derrière leurs écrans à peaufiner détails, expressions, gestuelles. Seconde après seconde, un amas se transforme ainsi en ogre goguenard et un âne en un Jo hâbleur. La tchatte de Jo dans Shrek est une

performance. La synchronisation labiale (ou « lipsync » en anglais) est, en effet, le casse-tête des animateurs.

Dans un dessin animé ou un jeu vidéo de bonne tenue, il faut une journée pour faire articuler 30 secondes de texte à un personnage virtuel. Pour un film du niveau de Toy Story, une journée suffit à peine pour modéliser sept secondes de laïus. Le mouvement des lèvres est reproduit point par point. Ensuite viennent les expressions du visage : joie, tristesse, colère... Bref, aujourd'hui, Dynalips automatise cette part chronophage du labeur. La plateforme abat ce travail fastidieux en une volée de minutes.

## Soulager les animateurs

Il suffit d'enregistrer sa bande-

son, de la télécharger, de l'assortir de la transcription écrite, et de laisser les algorithmes mouliner. Une prouesse. Ensuite, on choisit son avatar virtuel dans une galerie de portraits. Dynalips est adossé aux modèles ultra-réalistes MetaHumans du studio de jeux vidéo Epic Games, qui lui a attribué une bourse Epic MegaGrants. Le résultat est bluffant. Pour une phrase de six secondes, Dynalips est capable de traiter 856 images et 900 paramètres en quelques secondes. La facturation se fait au temps d'utilisation. « Dynalips réduit ainsi le temps de production et donne la possibilité aux animateurs de se concentrer sur des tâches artistiques comme les expressions faciales et les caractéristiques uniques du personnage », indique Slim Ouni.

Le chercheur a eu l'idée de Dynalips en discutant avec un producteur de studio d'animation de la place luxembourgeoise. Du clic au premier clic, il a fallu dix ans de maturation et d'acharnement. Mais Slim Ouni planchait sur le sujet déjà bien avant : « Ça fait vraiment une vingtaine d'années que j'étudie le mécanisme de l'articulation, de la communication parlée, comment le son est généré à partir de l'articulatoire en passant par le mouvement de la langue, le mouvement des lèvres, jusqu'au signal acoustique ». Aujourd'hui, Dynalips est parfaitement opérationnel en trois

langues : anglais, français et allemand.

## Et après ?

Prochains axes de développement pour les créateurs de cette pépite de la French Tech, accompagnée par l'Incubateur Lorrain et déjà lauréate du Prix Pépite du Ministère de l'Enseignement supérieur : « La priorité, c'est la parole en temps réel automatisée. C'est-à-dire, gommer le temps de latence pour des applications dans les chatbots (robots conversationnels) ou les jeux vidéo où vous avez des joueurs qui parlent par l'intermédiaire d'un avatar. Ce n'est pas simple. On devrait commencer à travailler sur cette extension le mois prochain. On a eu une bourse pour développer cette partie-là en France. C'est un objectif à un an ».

Deuxième marche, le multilingue. « Notre objectif, poursuit Slim Ouni, c'est de passer de trois à plein de langues. On travaille sur une solution qui nous permettrait de réduire les coûts. Plutôt que d'intégrer les langues une par une, ce qui est très cher, on réfléchit à une autre façon de faire. » D'ores et déjà, la technologie DeepLipSync de Dynalips pourrait trouver des applications dans les « systèmes d'apprentissage des langues, les assistants virtuels ou les applications d'assistance aux malentendants ».

● Thierry Fedrigo