

الجزائرية الديمقراطية الشعبية الجمهورية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

المدرسة العليا للإعلام الآلي - 08 ماي 1945 - بسبدي بلعباس
Ecole Supérieure en Informatique
-08 Mai 1945- Sidi Bel Abbès



Mémoire de Fin d'étude

Pour l'obtention du diplôme d'ingénieur d'état

Filière : **Informatique**

Spécialité : **Intelligence Artificielle et Sciences de Données (IASD)**

Thème

Modeling Users' Behavior in 360-Degree Videos: A Framework for
Synthetic Data Generation

Présenté par :

ABDALLAOUI Walid

Soutenu le : **00/00/2024**

Devant le jury composé de :

- | | |
|------------------|-----------|
| - M/Mme/Mlle XXX | Président |
| - M/Mme/Mlle XXX | Encadreur |
| - M/Mme/Mlle XXX | Examineur |
| - M/Mme/Mlle XXX | Examineur |
| - M/Mme/Mlle XXX | Invité |

Année Universitaire : 2023 / 2024

Abstract

In recent years, Virtual Reality (VR) applications, including 360° videos, have garnered significant interest due to their immersive experiences, making them applicable across various domains such as education, medicine, and entertainment. A typical 360° video viewing setup involves users interacting with the scene through Head Mounted Devices (HMDs), with the field of view (FoV) being the region seen by the user at any given moment.

Despite their potential, 360° videos present unique challenges, primarily due to their high bandwidth requirements—up to six times that of traditional videos—and the necessity for low motion-to-photon latency to prevent cybersickness. Optimization models are crucial in addressing these challenges, Tile-based streaming, for instance, divides the video into smaller tiles, allowing for adaptive streaming of only the tiles within the viewer’s FoV. Accurately predicting a user’s FoV ensures the efficient streaming of content, reducing both bandwidth consumption and latency.

Most of the algorithms in the state of art are based on user’s viewing data history and are trained on limited data. Data scarcity prevents optimized streaming for videos with no watch history or limited viewing data. In this work we introduce a novel framework for streaming 360° videos that does not depend on user history. Our approach involves modeling user viewing behaviors and generating diverse synthetic viewing data. This data is used to train a predictive model that forecasts user motion trajectories in 360° streaming systems. Our method achieves prediction times under 6ms, significantly faster than existing approaches, and performs exceptionally well without relying on real user viewing histories.

We also implemented an interactive application for tailored dataset generation, providing researchers with the tools to create comprehensive datasets reflecting a wide range of user behaviors and video content. This framework facilitates optimized streaming, enhancing the quality and efficiency of 360° video experiences.

Keywords: VR, 360° Videos, HMD, Saliency, FoV, Streaming

Résumé

Ces dernières années, les applications de Réalité Virtuelle (VR), y compris les vidéos à 360°, ont suscité un intérêt considérable en raison de leurs expériences immersives, les rendant applicables dans divers domaines tels que l'éducation, la médecine et le divertissement. Une configuration typique de visionnage de vidéo à 360° implique que les utilisateurs interagissent avec la scène à travers des dispositifs montés sur la tête (HMD), le champ de vision (FoV) représentant la région vue par l'utilisateur à tout moment donné.

Malgré leur potentiel, les vidéos à 360° présentent des défis uniques, principalement en raison de leurs exigences élevées en bande passante, jusqu'à six fois supérieures à celles des vidéos traditionnelles, et de la nécessité d'une faible latence de mouvement à photon pour éviter le mal des transports. Les modèles d'optimisation sont cruciaux pour relever ces défis, notamment le streaming basé sur les tuiles, qui divise la vidéo en tuiles plus petites permettant un streaming adaptatif uniquement des tuiles dans le FoV du spectateur. Prédire avec précision le FoV d'un utilisateur garantit le streaming efficace du contenu, réduisant à la fois la consommation de bande passante et la latence.

La plupart des algorithmes de pointe sont basés sur l'historique des données de visionnage de l'utilisateur et sont formés sur des données limitées. La rareté des données empêche un streaming optimisé pour les vidéos sans historique de visionnage ou avec des données de visionnage limitées. Dans ce travail, nous présentons un nouveau cadre pour le streaming de vidéos à 360° qui ne dépend pas de l'historique de l'utilisateur. Notre approche consiste à modéliser les comportements de visionnage des utilisateurs et à générer diverses données de visionnage synthétiques. Ces données sont utilisées pour former un modèle prédictif qui anticipe les trajectoires de mouvement des utilisateurs dans les systèmes de streaming à 360°. Notre méthode atteint des temps de prédiction inférieurs à 6 ms, significativement plus rapides que les approches existantes, et offre des performances exceptionnelles sans dépendre des historiques de visionnage réels des utilisateurs.

Nous avons également mis en œuvre une application interactive pour la génération de jeux de données personnalisés, fournissant aux chercheurs les outils nécessaires pour créer des jeux de données complets reflétant une large gamme de comportements des utilisateurs et de contenus vidéo. Ce cadre facilite le streaming optimisé, améliorant la qualité et l'efficacité des expériences vidéo à 360°.

Mots-clés:RV, Vidéos 360°, HMD, Saillance, Champ de vision, Streaming