Level of detail for granular audio-graphic rendering : representation, implementation, and user-based evaluation

Hui Ding

▶ **To cite this version:**

Hui Ding. Level of detail for granular audio-graphic rendering : representation, implementation, and user-based evaluation. Other [cs.OH]. Université Paris Sud - Paris XI, 2013. English. NNT : 2013PA112205 . tel-00913624

HAL Id: tel-00913624
https://theses.hal.science/tel-00913624

Submitted on 4 Dec 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE PARIS-SUD

ÉCOLE DOCTORALE : Ecole Doctorale Informatique Paris-Sud (ED 427)

Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur
(LIMSI CNRS) (UPR3251)

DISCIPLINE Informatique

THÈSE DE DOCTORAT

soutenue le 30/09/2013

par

Hui DING

Level of detail for granular audio-graphic rendering: representation, implementation, and user-based evaluation

Directeur de thèse :	Christian JACQUEMIN	Professeur (Université Paris-Sud)
Co-directeur de thèse :	Emmanuelle FRENOUX	Maître de Conférences (Université Paris-Sud)

Composition du jury :

Président du jury :	Anne VILNAT	Professeur (Université Paris-Sud)
Rapporteurs :	Venceslas BIRI	Maître de Conférences HDR (Université Marne la Vallée)
	Stéphane NATKIN	Professeur (CNAM/CEDRIC)
Examinateurs :	Samia BOUCHAFA	Professeur (Université d'Évry-Val-d'Essonne)

Table des matières

Résumé 3

Introduction 5

1. Motivation 5
2. Objectif et portée 6
3. Approche 7
4. Contribution 9
5. Thesis organization ..... **Erreur ! Signet non défini.**

Résumé

Simulation en temps réel de scènes audio-visuelles complexes reste difficile en raison du processus de rendu techniquement indépendant mais perceptivement lié à chaque modalité. En raison de la dépendance cross-modale potentiel de la perception auditive et visuelle, l'optimisation de graphiques et de rendu sonore, tels que le niveau de détail (LOD), doit être considéré de manière combinée, mais pas comme des questions distinctes. Par exemple, dans l'audition et de la vision, les gens ont des limites perceptives sur la qualité de l'observation. Techniques de LOD conduit par la perception pour les graphismes ont été grandement progressé depuis des décennies. Cependant, le concept de LOD est rarement pris en compte dans l'évaluation et le rendu crossmodal.

Cette thèse porte sur l'évaluation de la perception crossmodale sur le rendu LOD audiovisuel par des méthodes psychophysiques, sur lequel on peut appliquer une méthode fonctionnelle et générale, à terme, d'optimiser le rendu. La première partie de la thèse est une étude des problématiques. Dans cette partie, nous passons en revue les différentes approches LOD et discutons les issues, en particulier du point de vue au niveau de la perception crossmodale. Nous discutons également les résultats principaux sur le design, le rendu et les applications interactives des scènes audio et graphiques dans le cadre du projet ANR Topophonie dont la thèse a eu lieu. Une étude des méthodes psychophysiques pour l'évaluation de la perception audio-visuelle est également présentée afin de fournir une solide connaissance du design expérimentale. Dans la deuxième partie, nous nous concentrons sur l'étude de la perception des artefacts d'image dans le rendu LOD audio-visuel. Une série d'expériences a été conçue pour étudier comment la modalité audio supplémentaire peut influer sur la détection visuelle des artefacts produits par la méthode LOD d’imposteur. La troisième partie de la thèse est axée sur le X3D étendu que nous avons conçu pour la modélisation de LOD audio-visuel. Dans la dernière partie, nous présentons le design et l'évaluation du système original par le rendu LOD crossmodal. L'évaluation de

la perception audio-visuelle sur le système LOD crossmodal a été atteinte grâce à une série d'expériences psychophysiques.

Notre contribution principale est que nous offrons une compréhension originale de LOD crossmodal avec de nouvelles observations, et l'explorer par des expériences et des analyses perceptives. Les résultats de notre travail peuvent être, éventuellement, les preuves empiriques et des lignes directrices pour un système de rendu LOD crossmodale conduit par la perception.

Mots-clés: niveau de détail graphique, niveau de détail sonore, perception crossmodale, perception audiovisuelle, expérience perceptuelle, méthodes psychophysiques.

Introduction

1. Motivation

Au cours des dernières décennies, l'application de rendu audio-visuel de scènes 3D devient de plus en plus courante dans divers domaines tels que PC / jeux vidéo commerciaux, illustration médicale et l'éducation. En conséquence, les technologies de produits informatiques de percepts ont augmenté considérablement pour chaque modalité sensorielle de l'audition et de la vision. En effet, alors que la complexité et la réalité des graphismes 3D continuent de plus en plus pris en charge avec la croissance de la capacité de matériel informatique, l'étude du niveau de détail a également été considérablement élargie pour lutter avec le compromis entre la complexité et la performance. Au lieu de sélectionner manuellement LOD, la régulation simplification avec une métrique de fidélité devient la méthode de sélection LOD principalement utilisé dans la pratique des algorithmes de LOD, spécialement pour une grande base de données qui est souvent le cas dans les applications 3D. En conséquence, la façon de mesurer la fidélité est une question importante qui a été le plus demandé. Pour répondre à la question, les deux indicateurs d'erreur attributs géométriques et de surface sont généralement considérés dans LOD optimisation des processus. Toutefois, des preuves empiriques montrent que la mesure essentielle de la fidélité est la perception. En conséquence, les mesures de perception visuelle ont été étudiés pour guider les cadres de niveau de détail pour les deux dernières décennies. Le premier système LOD entraîné perceptuelle est proposé par Reddy à la fin des années 90 [Reddy, 1997]. Depuis lors, le défi de cette envergure est de comprendre les fondements de la psychologie de la perception et de les appliquer dans la LOD ou même d'autres techniques d'infographie.

De même, les technologies de rendu audio ont été considérablement développées pour le son 3D complexes, dans les aspects de la modélisation de la source, la simulation de propagation, traitement numérique du signal, spatialisation, etc. Entre-temps, il y a aussi

quelques essais intéressantes faites pour gérer les sources sonores massives en virtuel environnement en utilisant la perception auditive [Tsingos et al. 2004]. Ces croissances rapides des technologies de rendu audio et visuel ont plus ou moins impliqué l'utilisation de la perception humaine dans leur propre communauté de la recherche, respectivement. Cependant, très peu de travaux ont été réalisés dans l'exploration de l'efficacité de système de perception intermodale de l'audition et de la vision dans une tentative de donner une ligne directrice pour les applications d'affichage audio-visuels, comme la sélection LOD, qui peut être perçue par l'utilisateur, et la modélisation de scènes LOD audio - visuelles.

Au meilleur de notre connaissance, le sujet de LOD intermodale n'a été étudié par Bonneel et al. [2010] par rapport à la LOD graphique de la forme d'illumination et contacter simplification de son, mais une telle étude de perception intermodale n'a jamais été fait sur des simplifications basées sur l'image et à base de polygones (qui sont les méthodes les plus couramment utilisés dans les applications 3D) et l'environnement sonner simplification par Bonneel et al. [2010] par rapport à la LOD graphique de la forme illumination et contacter simplification de son, mais une telle étude de perception intermodale n'a jamais simplifications été fait sur base d'images et à base de polygones (qui sont les plus méthodes couramment utilisées dans les applications 3D) et la simplification de son environnement.

2. Objectif et portée

Dès la fin du siècle dernier, chercheurs de l'image / l'infographie ont commencé à étudier la connaissance interdisciplinaire afin d'étudier les phénomènes de perception intermodale pour justifier un certain degré d'audio et/ou visuel fidélité. De la même période, les chercheurs LOD ont également commencé l'enquête de l'orientation de principe pour la sélection LOD en utilisant des modèles de système de vision humain. L'exploitation du système perceptif n'est pas seulement avancée dans l'infographie de l'environnement virtuel, mais aussi dans la synthèse sonore. En 2004, Tsingos et al. [2004] ont d'abord

présenté une approche pour optimiser les sources sonores dans un grand environnement virtuel en utilisant des critères de perception audio. Récemment, des chercheurs ont commencé à étudier intermodale LOD par la perception. Une étude pilote [Bonneel et al. 2010] a étudié l'efficacité de intermodale LOD par la perception de la similitude de matériel en particulier, et en fonction de leur constatation psychophysique, ils tirent une fonction de la qualité perçue des intermodale LOD de sorte que l'on peut choisir la combinaison LOD appropriée du son et des graphiques pour le meilleur qualité perçue des matériaux . Cependant, il y a encore un manque d'enquête de perception intermodale globale dans le domaine de LOD. Il est dommage qu'il y a très peu de références qui peuvent nous guider dans la création, la gestion et représentant une scène LOD audio-visuel. La thèse présentée est dédié à cet important secteur en pleine croissance. Dans une tentative pour enquêter sur la limite fondamentale de détail intermodale qu'un système audio -visuel doit rendre, cette thèse présente un travail qui se concentre sur la perception des effets audio - graphiques introduites par différents simplifications, et la perception globale d'une scène LOD intermodale. Pour être clair, nous avons étudié la perception intermodale des artefacts les plus communes introduites par la base d'images système de rendu LOD pour modalité visuelle. Par ailleurs, nous avons également étudié la perception intermodale globale sur le système LOD intermodale. Enfin, nous avons introduit des concepts et des normes pour représenter une scène audio-visuel (LOD).

3. Approche

Dans notre travail, nous avons utilisé des méthodes psychophysiques qui analysent le système perceptif en étudiant les effets perçus sur l'expérience utilisateur en modifiant les propriétés / facteurs de stimuli selon une ou plusieurs dimensions physiques [Cunningham et Wallraven 2011a]. Les choses qui nous intéressent sont les relations entre le LOD stimulus unimodal / multimodal et des sensations humaines qui peuvent être touchés, et, finalement, une ligne directrice perceptuelle pour les applications audio - visuelles. Pour tenter d'enquêter sur cette relation, nous avons créé différentes scènes audio - visuels avec

des méthodes de génération LOD couramment utilisés comme base d'images LOD et à base de polygones LOD, et a conçu une série d'expériences de perception appropriées dans le respect des règles fondamentales psychophysiques. En analysant les données brutes issues d'expériences, nous avons tenté de comprendre la capacité d'uni- et inter- modal LOD à travers la perception audio - visuelle globale. L'élaboration d'un schéma de principe de sélection LOD appropriée est l'une des raisons pour lesquelles les mesures de perception visuelle sont appliquées dans LOD pour les graphiques [Luebke et al. 2002]. En effet, les chercheurs et les développeurs espèrent réduire détails plus loin tout en acceptant les artefacts générés qui peuvent ne pas être perçu par un utilisateur en raison de la limite de la perception visuelle. Simultanément, les chercheurs constatent également que son pourrait influencer la qualité visuelle perçue sous certaines conditions [Storms et Usa 2000]. Cela nous fait penser que l'avantage de la perception visuelle des objets acceptant existe encore ou si il est encore plus évident lorsque sensations multimodales sont impliqués. La tenue de cette question, nous avons généré notre première scène audio - visuelle non - réaliste appliqué image - based/impostor LOD. Comme l'image / imposteur basée LOD inévitablement introduit divers objets, nous avons ensuite conçu une expérience pour étudier la perception des objets dans les conditions intermodale.

Pour enquêter intermodale LOD en modalité de l'audition et de la vision, nous avons introduit la notion de LOD dans le son et inventé un moteur de rendu pour générer un système LOD intermodale. Nous avons mis le LOD imposteur basée sur les deux modalités, puis effectué une expérience basée sur la méthode des limites afin d'explorer la relation entre la perception globale et intermodale LOD. Le LOD imposteur basée sur la méthode que nous appliquons pour le son, qui est l'état de l'art et la plus appropriée pour la synthèse sonore basé sur le corpus . Au contraire, il y a beaucoup plus d'algorithmes de LOD graphiques et leurs variantes ont été proposées au cours des trois dernières décennies. Nous n'avons pas essayé tous. Toutefois, dans le but de généralisation, nous avons mis un autre algorithme largement utilisé qui est basé sur décimation sommet, au-delà de la limite de détection imposteur basée . Une série d'expériences basées sur la tâche N -alternative-non- choix forcé pour le seuil de discrimination a été réalisé pour chaque épreuve / scène. Grâce à la série d'expériences de scène audio -visuel (s), nous tentons d' explorer la

possibilité de intermodale LOD (c.-à-d. LOD graphiques à base de polygones et son imposteur LOD-based) au moyen de la perception globale.

4. Contribution

Comme indiqué précédemment, beaucoup de recherches ont été menées sur la perception conduit LOD graphiques, ainsi que l'audio? Perception intermodale visuelle. Cependant, autant que nous le savons, l'étude de la perception LOD intermodale a juste commencé récemment et est seulement sur la perception du matériau. Au meilleur de nos connaissances, notre travail est le premier à étudier la intermodale LOD du point de vue perceptif global et fournir des lignes directrices pour les praticiens qui rendent intermodale scènes LOD.

- Depuis, personne n'a jamais présenté un concept ou système de son LOD pour la scène audio-visuel, non seulement nous présentons et nous réalisons cette notion nouvelle dans une application de synthèse sonore up-to-date, mais aussi le combiner avec des graphismes LOD rendu comme un pipeline pour généraliser le système LOD audio-visuelle à travers un son et un moteur graphique LOD.
- Nous vous proposons une étude sur la perception multi-sensorielle psychophysique et des uni-LOD système intermodal.
- En conséquence, le guide basé sur le résultat d'expériences à travers la méthodologie de la perception psychophysique est prévue pour les praticiens qui ont besoin de générer un système audio-visuel en particulier avec les universités? ou technologies LOD cross-modal.
- Nous offrons également un nouveau concept pour les formaliser la modélisation du système audio-visuel (LOD), en particulier dans le format X3D.

Certains de nos résultats ont été publiés dans:

Hui Ding, Diemo Schwarz, Christian Jacquemin, and Roland Cahen. Spatial audiographic modeling for x3d. In Proceedings of the Sixteenth Annual International Conference on 3D Web Technology, 2011.

Hui Ding and Christian Jacquemin. Palliating visual artifacts through audio rendering. In Proceedings of Smart Graphics, pages 179-183, 2011.

Diemo Schwarz, Roland Cahen, Hui Ding, and Christian Jacquemin. Sound level of detail in interactive audiographic 3D scenes. In Proceedings of the International Computer Music Conference, 2011.

5. Organisation de la thèse

Le reste de la thèse est organisée en trois parties comprenant six chapitres et contient la conclusion de la thèse, une liste de glossaire, une bibliographie et une annexe, comme suit.

- Partie I, nous discutons le fond et les travaux connexes.
    - Chapitre 1, nous donnons un aperçu des formes de simplification de graphiques techniques LOD et de discuter du fond lié au sujet de son notions LOD.
    - Chapitre 2, nous présentons le rôle de ce travail de thèse dans le projet Topophonie qui a porté sur les techniques mises à jour dans le dépôt de rendu audio-visuel.
    - Chapitre 3, nous présentons un aperçu des uni-et perception multi-sensorielle et-uni et la gestion LOD cross-modal.
- Partie II: Chapitre 4 Nous présentons un développement d'un imposteur scène GLOD base, et une expérience sensorielle utilisée pour évaluer la capacité de perception sur les objets de scène statique / image à travers sensations intermodale.

- Partie III: Chapitre 5 Le concept de la modélisation d'un système audio-visuel est présenté dans le chapitre 5, et la spécification du format X3D prolongée ainsi qu'un guide d'utilisation de la méthode de modélisation sont fournis.
- Partie IV: Chapitre 6, nous introduisons le moteur audio-visuel LOD qui comprend des graphiques LOD le plus commun et up-to-date de son techniques LOD. Le cadre d'un rendu de la scène intermodale basée sur imposteur basée LOD est présentée, et l'expérience et les résultats psychophysiques sont décrits.

Bibliography

[Reddy 1997] Martin Reddy. Perceptually Modulated Level of Detail for Virtual Environments. PhD thesis, University of Edinburgh, 1997.

[Tsingos et al. 2004] Nicolas Tsingos, Emmanuel Gallo, and George Drettakis. Perceptual audio rendering of complex virtual environments. In ACM SIGGRAPH 2004 Papers, SIGGRAPH '04, pages 249258, New York, NY, USA, 2004. ACM.

[Bonneel et al. 2010] Nicolas Bonneel, Clara Suied, Isabelle Viaud-Delmon, and George Drettakis. Bimodal perception of audio-visual material properties for virtual environments. ACM Trans. Appl. Percept., 7:1:11:16, January 2010.

[Cunningham et Wallraven 2011a] Douglas Cunningham and Christian Wallraven. Experimental Design: From User Studies to Psychophysics. Taylor & Francis, November 2011a.

[Luebke et al. 2002] David Luebke, Benjamin Watson, Jonathan D. Cohen, Martin Reddy, and Amitabh Varshney. Level of Detail for 3D Graphics. Elsevier Science Inc., New York, NY, USA, 2002.

[Storms et Usa 2000] Russell L. Storms and Maj Usa. Auditory-visual cross-modal perception. In ICAD, 2000.