

## Resum

---

El processament d'àudio multicanal ha experimentat un gran desenvolupament en els darrers anys i com a conseqüència s'ha produït un augment notable de la complexitat computacional en les noves aplicacions. Actualment, es pretén que la telecomunicació ofereixi una sensació de proximitat, compartint fins i tot el mateix entorn entre usuaris distants. És el que s'anomena: Esquemes d'Àudio Immersiu. En aquest fenomen intervenen diversos efectes acústics: so espacial 3D, compensació de sales, cancel·lació crosstalk, localització de fonts sonores, entre altres. Però, per dur a terme qualsevol d'aquests efectes en un sistema real, es necessita una alta capacitat computacional, el que representa una severa limitació quan es tracta d'executar aquestes aplicacions en temps real.

L'augment de la capacitat computacional ha anat històricament unida al nombre de transistors en un xip. Actualment, les millores en la capacitat computacional estan íntimament lligades al nombre d'unitats de procés que té un ordinador, el que permet un alt grau de paral·lelisme en computació. Aquest és el cas de les Unitats de Processament Gràfic (GPUs, *Graphics Processing Units*), que posseeixen actualment milers de nuclis computacionals. Les GPUs s'han relacionat tradicionalment amb la computació gràfica o el tractament d'imatges, però amb l'aparició de nous entorns de programació per a GPUs ( CUDA o OpenCL ) moltes aplicacions d'altres camps científics han pogut ser accelerades mitjançant la seua implementació a les GPUs. Aquesta tesi té com a objectiu desenvolupar aplicacions d'àudio que necessiten gran quantitat de recursos computacionals, demostrant amb això que les GPUs són eines totalment vàlides per dur-les a terme. Amb aquest objectiu, s'han implementat i avaluat diferents aplicacions del camp de processament de senyals d'àudio sobre l'entorn de programació CUDA. També s'han analitzat i resolt les possibles limitacions sorgides durant el procés d'implementació, tant des del punt de vista acústic com des del punt de vista computacional.

En la tesi s'han abordat els següents problemes:

La primera operació a implementar a la GPU era l'operació fonamental en el processament d'àudio: la convolució, ja que la majoria d'aplicacions d'àudio multicanal estan basades el filtrat massiu. Primerament, s'ha desenvolupat la convolució com un nucli computacional que posteriorment

s'ha utilitzat per desenvolupar una aplicació que combina concurrentment múltiples convolucions: cancel·lació generalitzada i equalització. La implementació proposta pot gestionar dues situacions comuns en el filtrat multicanal: buffers de mostres d'àudio de major tamany que el dels filtres, o per contra buffers de mostres d'àudio de menor tamany que el dels filtres.

S'han desenvolupat dues aplicacions d'àudio espacial a partir del filtrat massiu multicanal que fan servir les GPUs com a co-processadors. La primera aplicació gira entorn al so binaural. Aquesta aplicació presenta dues característiques importants: 1) Pot sintetitzar fonts sonores en posicions espacials que no estan incloses a les bases de dades dels filtres HRTFs, i 2) genera moviments continus entre diferents posicions. Aquestes característiques s'han obtingut en la implementació després de diverses proves tant objectives com subjectives. Posteriorment, s'ha estudiat el màxim nombre de fonts sonores que poden ser gestionades per diferents arquitectures GPU. El mateix estudi s'ha dut a terme a un sistema de síntesi d'ona Wave Field Synthesis (segona aplicació de so espacial) compost per 96 altaveus. La implementació d'aquest sistema en GPU pot reduir els efectes de sala durant la reproducció.

Un altre problema que s'ha abordat en aquesta tesi és la localització de fonts sonores en entorns sorollosos i amb molta reverberació. Per a aquest problema s'ha proposat una implementació basada en l'algorisme de localització *Steered Response Power with Phase Transform* (SRP-PHAT) en un sistema multi-GPU. L'exactitud en la localització de les fonts sonores està íntimament lligat a una malla espacial de punts on es busca la font i el nombre de micròfons utilitzats en l'algorisme. En aquesta tesi, s'han avaluat les capacitats de les GPUs quan aquestes implementen l'algorisme SRP-PHAT sota condicions de temps real, tenint en compte diferents paràmetres: tamany de la malla, nombre de micròfons, reverberació a la sala, i relació senyal a soroll.

Finalment, aquesta tesi tracta el problema del filtrat massiu multicanal quan els filtres presenten una resposta a l'impuls infinita (*Infinite Impulse Response*, IIR). S'han analitzat dos casos particulars: 1) Filtres IIR compostos de múltiples seccions paral·leles d'ordre dos, i 2) Filtres IIR que presenten una resposta plana en freqüència (allpass filters). Ambdues estructures es fan servir per a desenvolupar i accelerar dues aplicacions d'àudio diferents: 1) implementar múltiples equalitzacions a un sistema WFS, i 2) reduir el marge dinàmic a les senyals d'àudio.

**Paraules Clau:** Filtrat Massiu Multicanal, Síntesi d'Ona, Localització de

Fonts Sonores, So Espacial, Unitats de Processament Gràfic.