

– Sujet de Master en Informatique –

IRISA/D6 - Lannion

Septembre 2013

Titre : Étude de l’adaptation sur un ordinateur de type CUDA (GPGPU) des algorithmes de sélection d’unités pour la synthèse de la parole à partir du texte

Description du sujet : À partir d’une entrée sous forme textuelle, un système de synthèse vocale à partir du texte (dit *Text to speech* ou *TTS*) produit un signal de parole correspondant à une vocalisation de ce texte. Un tel système se décompose classiquement en deux grandes étapes : des traitements linguistiques suivis de traitements acoustiques dont l’objectif essentiel est de créer un signal parole respectant au mieux les consignes linguistiques. Concernant cette deuxième étape, la méthode dite de Synthèse Par Corpus (SPC) s’est imposée comme une référence depuis une vingtaine d’années. Cette approche consiste en la sélection d’unités sonores au sein d’un corpus de parole continue et en la juxtaposition de ces unités pour obtenir un signal de synthèse. Ces corpus de parole continue permettent d’avoir à disposition plusieurs réalisations sonores de la même unité phonologique.

Deux points importants influent sur la qualité de la parole de synthèse : d’une part le contenu du corpus de parole continue à partir duquel sont extraites les unités sonores et d’autre part les heuristiques utilisées pour déterminer un séquençement optimal de ces unités en vue de synthétiser une phrase. Ce problème de séquençement en aveugle (on ne connaît pas à l’avance la liste des unités les plus intéressantes) est fortement combinatoire. Avec un corpus de parole continue de taille importante cela peut devenir rédhibitoire. Par exemple, la phrase « Petit à petit, l’oiseau fait son nid », contenant 24 phones, donne naissance à un graphe d’environ 22 millions de nœuds à partir d’un des corpus de langue française exploités dans l’équipe. Il faut ensuite trouver un chemin optimal dans ce gigantesque graphe.

Pour ce travail de master, nous proposons, compte tenu de la complexité du problème, d’étudier la parallélisation ou la distribution sur une grille de calcul de type CUDA (GPGPU) d’un algorithme A* de sélection d’unités. Un protocole expérimental devra répondre à la question de la pertinence des choix de parallélisation mis en œuvre en terme de gain en efficacité comparés au chemin optimal trouvé par un algorithme séquentiel.

Mots-clés : Synthèse de la parole à partir du texte, Algorithmes de séquençement, Algorithme A*, Expérimentation, Calcul parallèle.

Contacts : Arnaud Delhay (arnaud.delhay@irisa.fr)

Bibliographie :

- [1] A. W. Black and K. A. Lenzo. Optimal data selection for unit selection synthesis. In *4th ISCA Tutorial and Research Workshop on Speech Synthesis*, 2001.
- [2] R. Inam. *A star Algorithm for Multicore Graphics Processors*. Master thesis, Chalmers University of Technology, 2010.
- [3] J. Nickolls, I. Buck, M. Garland, and K. Skadron. Scalable parallel programming with cuda. *Queue*, 6(2) :40–53, Mar. 2008.
- [4] H. Silen, E. Helander, J. Nurminen, K. Koppinen, and M. Gabbouj. Using robust viterbi algorithm and hmm modeling in unit selection tts to replace units of poor quality. In *Conference of the International Speech Communication Association (Interspeech)*, pages 166–169, 2010.