

Projet de Programmation

Alexandre Kervadec, Guillaume Verdugo, Jeremy Arrestier, Thibaut Fabre

17 février 2015

Table des matières

1	Étude préalable des références et de l'existant	3
1.1	Compréhension du sujet	3
1.2	Références sur la transformation de la voix	3
1.3	Références sur la faciale	3
1.4	État de l'art	4
1.5	Programmation avec du contenu vidéo	4
1.6	Ergonomie	4
1.7	Base de données	4
	Bibliographie	4

Chapitre 1

Étude préalable des références et de l'existant

1.1 Compréhension du sujet

Afin de comprendre au mieux la prosodie, l'article [4] nous permet de mettre une définition sur ce terme technique.

L'article [3] nous donne des notions sur les émotions que l'on peut passer avec la voix, alors que le livre [6] nous aide sur les expressions faciales.

Enfin, le document [7] nous offre une approche sociétale du problème avec une étude sur la prosodie attitudinale pour la langue japonaise.

1.2 Références sur la transformation de la voix

Le cours de Ricardo Gutierrez-Osuna [9] et l'article de l'IEEE Signal Processing Letters [10] sont des descriptions physiques de la modulation et modification de la voix.

L'extrait de la IEEE International Conference on Acoustique de 1998 [1] décrit deux possibles modifications prosodiques.

Le document de Véronique Aubergé [2] nous offre une approche par la méthode Gestalt de la prosodie.

Dans cet article du journal Voice transformation using PSOLA technique [19], un système de conversion de voix utilisant PSOLA et un module pour les transformations spectrales sont étudiés.

1.3 Références sur la faciale

Le document [5] expose une approche sur la définition de règles de synchronisation des expressions faciales et particulièrement buccales par rapport à un certain discours.

Le document [15] est utile dans le sens qu'il porte sur l'étude de la prosodie faciale, et ce afin de pouvoir détecter l'ethnie du protagoniste, ce qui nous est utile car le projet porte sur trois langues : le français, l'anglais (américain) et le japonais.

1.4 État de l’art

L’article [16] nous donne une ligne directrice pour tout ce qui touche la reconnaissance des différents dialectes arabes. En d’autres termes, malgré une proximité géographique et linguistique, la prosodie permet la différenciation. L’article de la revue en ligne Alsic [14] nous apporte un point sur l’état actuel de l’art en matière de logiciel pour l’apprentissage de la prosodie, ce qui est proche de ce que l’on cherche à développer.

1.5 Programmation avec du contenu vidéo

Les articles [8] et [11] nous ont permis de mieux comprendre le concept d’encodage des vidéos et l’utilisation de *Codecs* pour pouvoir décoder ces vidéos afin de les lire.

1.6 Ergonomie

L’article [13] est une étude sur la “collaboration entre Ergonomie, Design et Ingénierie”, qui nous offre une façon de procéder afin de réaliser une ergonomie des plus efficace.

1.7 Base de données

Durant ce projet, il a fallu trouver le système de base de données le plus adapté à nos besoins. Pour ce faire, nous nous sommes appuyés sur l’article [18], disséquant l’architecture *NoSQL* (Not Only SQL). Nous tournant finalement vers le *SGBDR* (Système de Gestion de Base de Données Relationnelles) *SQLite*, nous nous sommes informés sur l’utilisation et les finalités de ce logiciel avec le livre [12].

Bibliographie

- [1] A. Acero. Source-Filter Models for Time-Scale Pitch-Scale Modification of Speech. In *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing*, Seattle, WA (USA), May 1998.
- [2] Véronique Aubergé. A gestalt morphology of prosody directed by functions : the example of a step by step model developed at icp. In *Speech Prosody 2002, International Conference*, 2002.
- [3] Véronique Aubergé. Prosodie et émotion. *Actes des deuxiemes assises nationales du GdR I*, 3 :263–273, 2002.
- [4] Jo-Anne Bachorowski. Vocal expression and perception of emotion. *Current directions in psychological science*, 8(2) :53–57, 1999.
- [5] Jonas Beskow. Rule-based visual speech synthesis. In *EUROSPEECH*, 1995.
- [6] Paul Ekman and Wallace V Friesen. *Unmasking the face : A guide to recognizing emotions from facial clues*. Ishk, 2003.
- [7] Dominique Fourer, Marine Guerry, Takaaki Shochi, Jean-Luc Rouas, Jean-Julien Aucouturier, and Albert Rilliard. Analyse prosodique des affects sociaux dans l'interaction face à face en japonais. In *XXXèmes Journées d'études sur la parole*, Le Mans, France, June 2014.
- [8] Mohammed Ghanbari. Video coding : an introduction to standard codecs. Institution of Electrical Engineers, 1999.
- [9] Ricardo Gutierrez-Osuna. Prosodic modification of speech. Introduction to Speech Processing.
- [10] Jesper Haagen. Transformation and decomposition of the speech signal for coding. *IEEE signal processing letters*, 1(9) :136, 1994.
- [11] Yun He, Jörn Ostermann, Marek Domanski, Oscar C Au, and Nam Ling. Introduction to the issue on video coding : Hvc and beyond. *Selected Topics in Signal Processing, IEEE Journal of*, 7(6) :931–933, 2013.
- [12] Jay Kreibich. *Using SQLite*. " O'Reilly Media, Inc.", 2010.
- [13] Chloé Lenté, Soizick Berthelot, and Stéphanie Buisine. Scénariser l'usage pour améliorer la collaboration entre ergonomie, design et ingénierie. 2014.
- [14] Philippe Martin. Winpitch ltl, un logiciel multimédia d'enseignement de la prosodie. *Alsic [en ligne]*, 8(2), 2005.
- [15] David Matsumoto. American-japanese cultural differences in the recognition of universal facial expressions. *Journal of cross-cultural psychology*, 23(1) :72–84, 1992.

- [16] Jean-Luc Rouas, Melissa Barkat-Defradas, François Pellegrino, Rym Hamdi, et al. Identification automatique des parlers arabes par la prosodie. *Journées d'Etude sur la Parole*, 2006.
- [17] VideoLan Streaming Solutions. Vlc media player, 2006.
- [18] Christof Strauch, Ultra-Large Scale Sites, and Walter Kriha. Nosql databases.
- [19] Hélène Valbret, Eric Moulines, and Jean-Pierre Tubach. Voice transformation using psola technique. *Speech Communication*, 11(2) :175–187, 1992.