

OAM Extracteur Vocal

Explication Technique

Oliver Michalowski

2023 IV 16

Table des matières

1	Introduction	1
1.1	Definitions	1
1.2	Formules	2

1 Introduction

Le but de cet extracteur vocal est d'enlever la partie instrumentale d'un morceaux présentant une partie vocale à l'aide du fichier audio avec le morceaux complet (avec les voix) et la version instrumentale du morceau (sans les voix).

1.1 Definitions

Soient A et B des vecteurs à N dimensions ($N \mid N \in \mathbb{N}$).

$$A[n], B[n] \in [-1, 1] \forall n \in \mathbb{N} \setminus 0 \leq n < N$$

$$A[n] = 0, B[n] = 0 \forall n \in \mathbb{N} \setminus (n < 0) \cup (N \leq n)$$

Produit scalaire euclidien :

$$\langle A|B \rangle = \sum_{n=0}^{N-1} (A[n] \times B[n])$$

Norme (énergie) :

$$\|A\|^2 = \langle A|A \rangle$$

Puissance (énergie/échantillon) :

$$P(A) = \frac{\|A\|^2}{N}$$

Estimateur de la covariance :

$$\Gamma_{A|B}(p) = \frac{1}{N} \times \sum_{n=-\infty}^{+\infty} (A[n] \times B[n-p])$$

Quand $p = 0$:

$$\Gamma_{A|B}(0) = \frac{\langle A|B \rangle}{N}$$

Coefficient de corrélation :

$$r_{A|B} = \frac{\langle A|B \rangle}{\|A\| \times \|B\|}$$

$$r_{A|B}(p) = \frac{\Gamma_{A|B}(p)}{\|A\| \times \|B_{-p}\|} \quad \setminus \quad B_{-p}[n] = B[n-p]$$

1.2 Formules

$$VOIX = ORIGINAL - \beta \times INSTRU$$

$$\langle VOIX | INSTRU \rangle = 0$$

$$\Rightarrow \langle ORIGINAL | INSTRU \rangle = \langle VOIX + \beta \times INSTRU | INSTRU \rangle$$

$$\Rightarrow \langle ORIGINAL | INSTRU \rangle = \langle VOIX | INSTRU \rangle + \langle \beta \times INSTRU | INSTRU \rangle$$

$$\Rightarrow \langle ORIGINAL | INSTRU \rangle = 0 + \beta \times \|INSTRU\|^2$$

$$\Rightarrow \frac{\langle ORIGINAL | INSTRU \rangle}{\|INSTRU\|^2} = \beta$$