La géométrie de l'information pour l'analyse et le traitement des flux audio

5e Journées Jeunes Chercheurs en Audition, Acoustique musicale et Signal audio (JJCAAS 2009)

Arnaud Dessein

IRCAM
1, place Igor Stravinsky
75004 Paris

Courriel : `arnaud.dessein@ircam.fr`
Page web : `http://imtr.ircam.fr/index.php/Arnaud_Dessein`

Introduction

- Objectifs :
  - Etudier la géométrie intrinsèque des flux audio.
  - Prendre en compte leurs natures temporelle et probabiliste.
  - Définir des structures d'analyse et de traitement alternatives.
- Cadre de travail :
  - La *géométrie de l'information* (Cramer, Rao, Chentsov, Amari, etc.).
  - Probabilités et information vues par la *géométrie différentielle*.
  - Espaces où chaque point est une *distribution de probabilités* représentant une trame du flux audio.
- Motivations :
  - Analyse des contenus audio.
  - Analyse, transformation, synthèse des sons.
  - Aide à la composition et à l'analyse musicale.
  - Improvisation assistée par ordinateur.

Géométrie différentielle élémentaire

- Variété topologique :
  - Une *variété topologique* est un espace localement semblable à $\mathbb{R}^n$.
  - Le couple (U, ϕ) est une *carte locale* de $\mathcal{M}$.
  - Une famille de cartes locales recouvrant $\mathcal{M}$ est un *atlas* de $\mathcal{M}$.

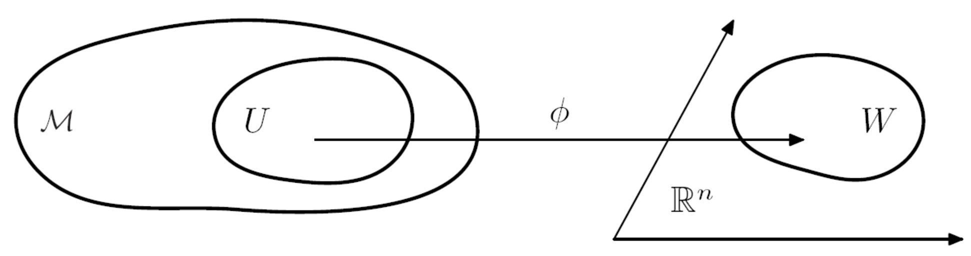

FIGURE 1 – Variété topologique et carte locale.

- Variété différentiable :
  - Une *variété différentiable* est une variété topologique possédant un atlas tel que les changements de coordonnées entre deux cartes locales s'intersectant sont différentiables.
  - L'*espace tangent* au point $p \in \mathcal{M}$ est une linéarisation de $\mathcal{M}$ en p.

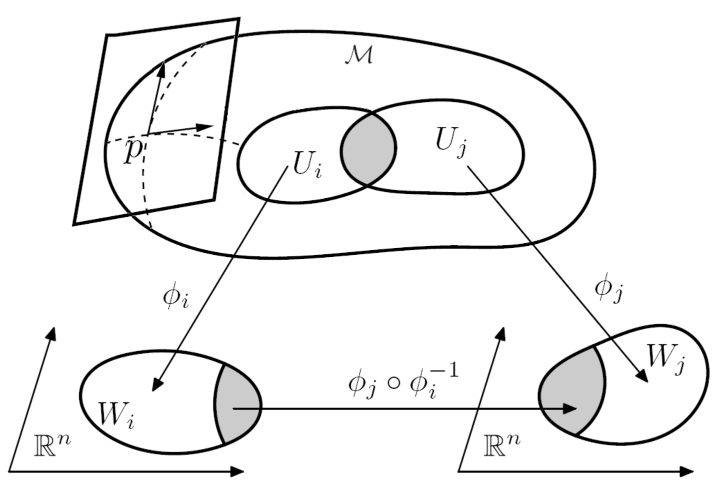

FIGURE 2 – Variété différentiable et espace tangent.

- Variété riemannienne :
  - Sous certaines conditions, on peut munir les espaces tangents de produits scalaires qui forment une *métrique riemannienne*.
  - Une *variété riemannienne* est une variété différentiable munie d'une métrique riemannienne.
  - Les structures engendrées par les espaces tangents et la métrique riemannienne sont locales et donnent accès à la notion de longueur d'une courbe.
  - On peut munir une variété riemannienne d'une *connexion affine* qui relie ces structures locales et donne accès aux notions de torsion, de courbure et de géodésique.

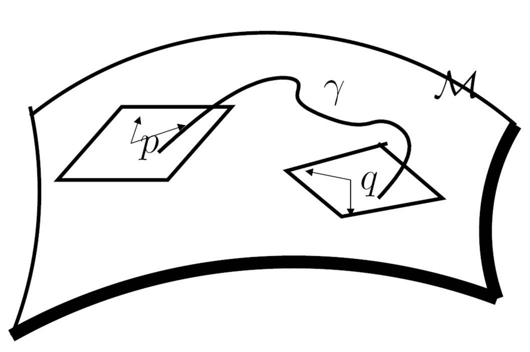

FIGURE 3 – Variété riemannienne et connexion affine.

Structure géométrique des modèles statistiques

- Modèles statistiques :
  - Un *modèle statistique* est une famille de *distributions de probabilités*
$$\mathcal{S} = \{p_\xi = p(x;\xi) : \xi = [\xi^1, \ldots, \xi^n] \in \Xi\}$$
  - Sous certaines conditions, $\mathcal{S}$ est une variété différentiable alors appelée *variété statistique*.
  - Exemple : $\mathcal{S}$ est la variété statistique des distributions normales
$$p(x;\xi) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \quad \text{avec } \xi = [\mu, \sigma]$$
- Métrique d'information de Fisher :
  - La *matrice d'information de Fisher* de $\mathcal{S}$ en ξ est la matrice $G(\xi)$ telle que
$$g_{ij}(\xi) = \int \frac{\partial}{\partial \xi_i} \log p(x;\xi) \cdot \frac{\partial}{\partial \xi_j} \log p(x;\xi) \cdot p(x;\xi) \cdot \mathrm{d}x$$
  - Sous certaines conditions, $G(\xi)$ engendre l'unique métrique riemannienne g sur $\mathcal{S}$, g est alors appelée la *métrique d'information de Fisher*.
- Connexions affines alpha :
  - Il existe une famille de connexions affines $\{\nabla^{(\alpha)}\}_\alpha$ sur $(\mathcal{S}, g)$. Cette famille est paramétrable par $\alpha \in \mathbb{R}$ et est unique. On appelle les $\{\nabla^{(\alpha)}\}_\alpha$ les *connexions affines* α.
  - Les connexions $\nabla^{(\alpha)}$ et $\nabla^{(-\alpha)}$ sont des *connexions affines duales* par rapport à g. La connexion duale d'une connexion ∇ est notée $\nabla^\star$.
- Divergences :
  - Une *divergence* sur $\mathcal{S}$ est une distance généralisée $\mathcal{D} : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ telle que pour tous $p, q \in \mathcal{S}$, on a : $\mathcal{D}(p \parallel q) \geqslant 0$, et $\mathcal{D}(p \parallel q) = 0$ ssi $p = q$.
  - La *divergence duale* d'une divergence $\mathcal{D}$ est la divergence $\mathcal{D}^\star$ définie par $\mathcal{D}^\star(p \parallel q) = \mathcal{D}(q \parallel p)$ pour tous p, q.
  - Les notions de connexions duales $\nabla, \nabla^\star$ et de divergences duales $\mathcal{D}, \mathcal{D}^\star$ sont étroitement liées et permettent de définir la notion de similarité.

Exemple applicatif

- Oracle audio :
  - Thèse d'Arshia Cont, article IEEE TASLP en cours de révision.
  - Automate pour la segmentation et l'apprentissage de structures.
  - Similarité entre états par la géométrie de l'information de $(\mathcal{S}, g, \nabla, \nabla^\star)$.
  - Détection incrémentale des répétitions et dépendances long-terme.
- Application :
  - Modélisation du contenu spectral par une distribution multinomiale.
  - Découverte de la structure du flux audio en temps-réel.
  - Accès direct aux répétitions grâce à l'apprentissage par automate.
  - Représentation parcimonieuse par une matrice de similarité.

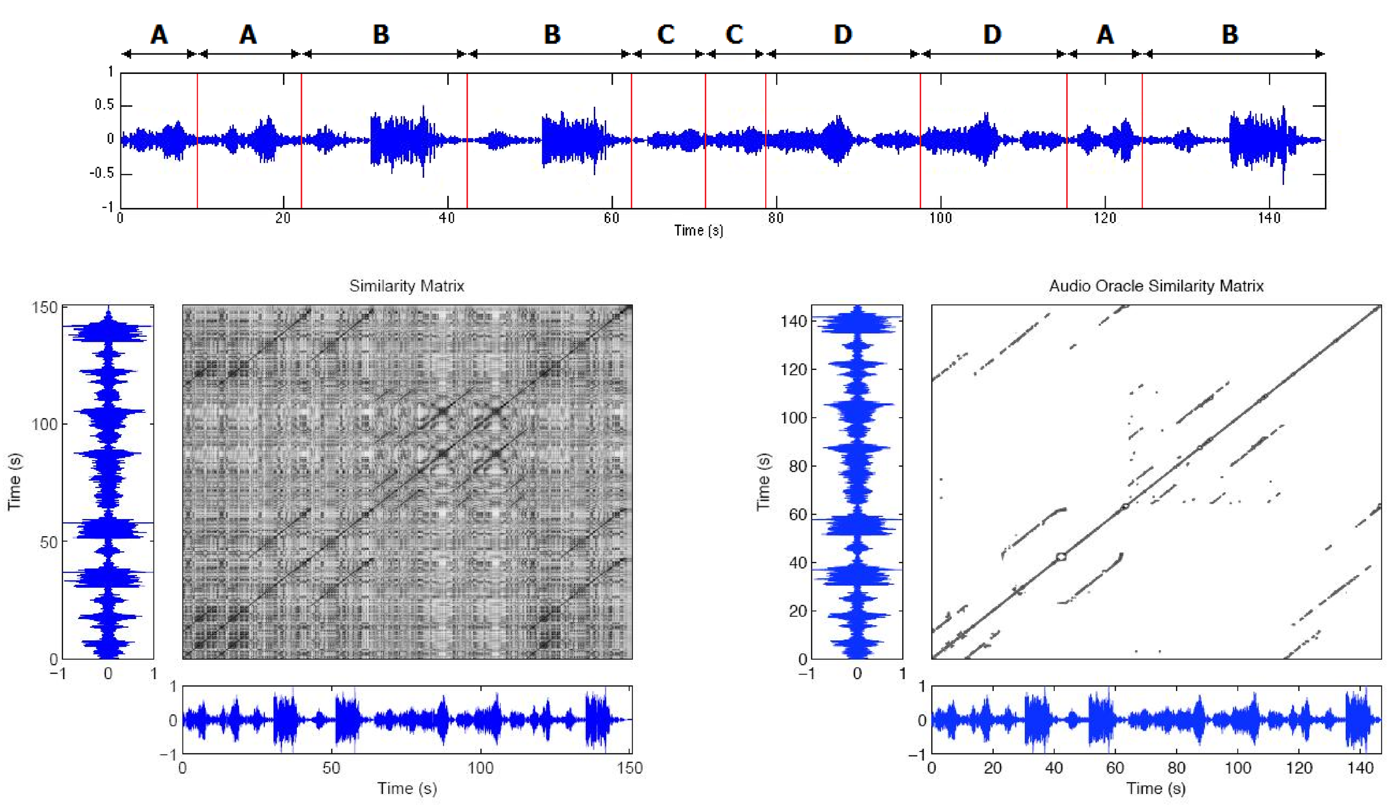

FIGURE 4 – Analyse d'un extrait de la *Sonate pour piano n°1* de Beethoven. Structure subjective sur la forme d'onde (en haut). Matrices de similarité classique (en bas à gauche) et avec l'oracle audio (en bas à droite).

- Ressources supplémentaires :
  - `http://imtr.ircam.fr/index.php/Music_Information_Geometry`
  - `http://imtr.ircam.fr/index.php/Audio_Oracle`