

FACULTAD DE INGENIERÍA DE LA
UNIVERSIDAD DE LA REPÚBLICA

PROCESAMIENTO DIGITAL DE SEÑALES DE
AUDIO

CURSO 2012

Beat Tracking

Autores:

Gonzalo GUTIÉRREZ

Matías TAILANIÁN



9 de diciembre de 2012

La métrica de la señal se puede pensar como una estructura de pulsos percibidos a diferentes escalas temporales en una pieza musical. Se consideran 3 niveles métricos básicos: el *tatum*, el *tactus* o *beat* y el *compás*. El *tatum* es el valor en tiempo más pequeño que puede encontrarse en una pieza musical, es la unidad atómica de la pieza. En general los otros valores de duración presentes en la pieza son múltiplos de *tatum*. El *tactus* o *beat* está más relacionado con el aspecto semántico, y está directamente vinculado con el **tempo** de la pieza. Por último el *compás* está vinculado con la tasa de cambios armónicos o la duración de un patrón rítmico.

El problema de la descripción automática de música es un problema todavía abierto donde se sigue intentando con gran actividad lograr mejorar los resultados del estado del arte. Tiene un atractivo muy importante en sí mismo pensando en aplicaciones como el acompañamiento automático, asistencia a la hora del editado, o estudios musicológicos, pero a su vez es imprescindible para lograr la transcripción automática de música.

El presente trabajo presenta un algoritmo de la detección y el seguimiento del *beat* de una pieza musical y se basa en el trabajo de “João Lobato Oliveira, Fabien Gouyon, Luis Gustavo Martins, Luis Paulo Reis”, titulado “IBT: A real time tempo and beat tracking system, presentado en la 11^a *International Society for Music Information Retrieval Conference, ISMIR*, en 2010” ([1]).

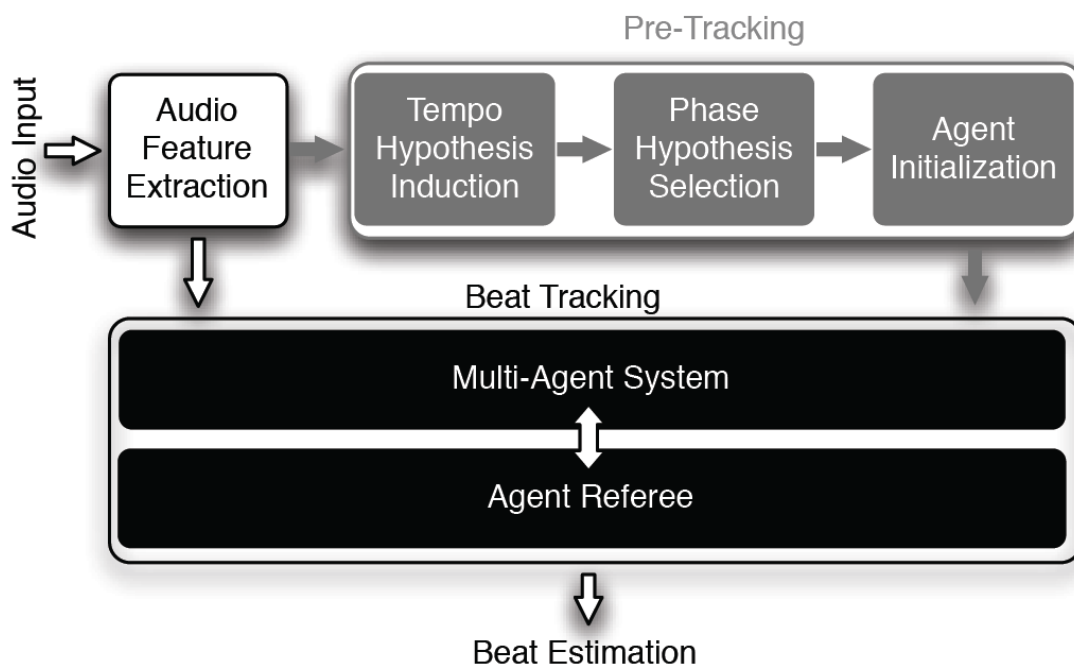


Figura 1: Diagrama de bloques

- Feature Selection
- Pre Tracking

- Período
- Fase
- Agente
- Beat tracking

0.1. Pre-Tracking

- Período

$$A(\tau) = \sum_{n=0}^m SF(n)SF(n+\tau)$$

SF es el flujo espectral suavizado para el *frame* n

$$\begin{cases} P_i = \arg \max_i \{A(\tau)\}, & i = 1, \dots, N \\ A(\tau) > \delta \frac{rms(A(\tau))}{M} \end{cases}$$

δ es un umbral determinado empíricamente y M es un rango de tiempos definido entre [50,250] BPM.

- Fase

Para cada P_i estimado se generan varias hipótesis para la fase: ϕ_i^j . Se supone fase y períodos constantes en cada ventana de análisis.

Se utiliza un *template* de tren de pulsos para ver cuál ajusta mejor.

- Agentes

Para cada pareja (P_i, ϕ_i) se computa la suma de errores entre el template de tren de pulsos y los máximos del flujo espectral

0.2. Beat-Tracking

Idea

Supervisar flujo de entrada y mantener balance entre inercia y rapidez de la respuesta

- $T_{in} \in [T_{in}^l, T_{in}^r]$

$$\begin{cases} P_i = P_i + 0,25 * error \\ \phi_i = \phi_i + 0,25 * error \end{cases}$$

- $T_{out} \in [T_{out}^l, T_{in}^l] \cup [T_{in}^r, T_{out}^r]$

El agente mantiene su período y fase y además crea 3 “hijos” variando dichos parámetros

Agent Release

Evalúa la distancia entre la predicción del *beat* (b_p) y el máximo local (m) de SF .

P_m es el máximo período permitido

$$\begin{cases} \Delta s = \left(1 - \frac{|error|}{T_{out}^r}\right) \frac{P_i}{P_m} SF(m), & \exists m \in T_{in} \\ \Delta s = -\left(\frac{|error|}{T_{out}^r}\right) \frac{P_i}{P_m} SF(m), & \exists m \in T_{out} \end{cases}$$

0.3. Referencias

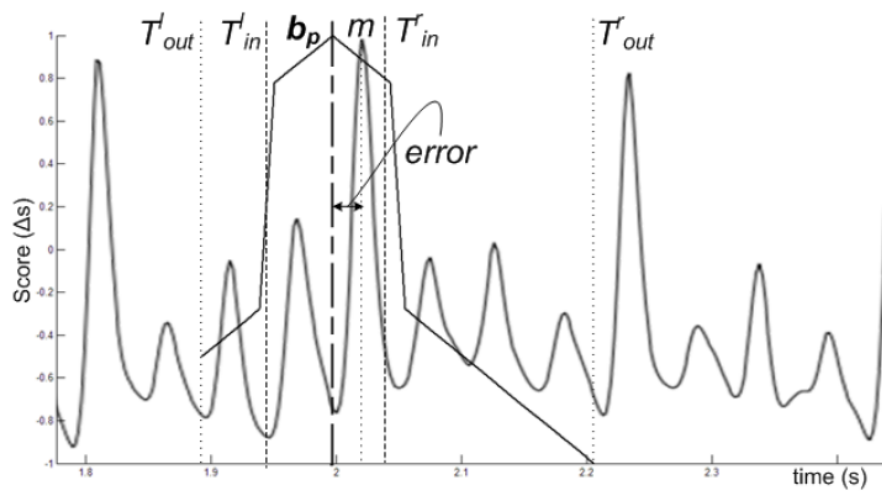


Figura 2: Niveles de tolerancia

Bibliografía

- [1] João Lobato Oliveira, Fabien Gouyon, Luis Gustavo Martins, Luis Paulo Reis, IBT: A real time tempo and beat tracking system, In *11th International Society for Music Information Retrieval Conference, ISMIR*, 2010.
- [2] S. Dixon. Automatic extraction of tempo and beat from expressive performances. In *Journal of New Music Research*, 30(1):39–58, 2001.
- [3] S. Dixon. Onset detection revisited. In *in Proceedings of the 9th International Conference on Digital Audio Effects*, pages 133–13, Montreal, Canada, 2006.
- [4] F. Gouyon, P. Herrera, and P. Cano. Pulse-dependent analyses of percussive music. In *AES 22nd International Conference on Virtual, Synthetic and Entertainment Audio*, 2002.