

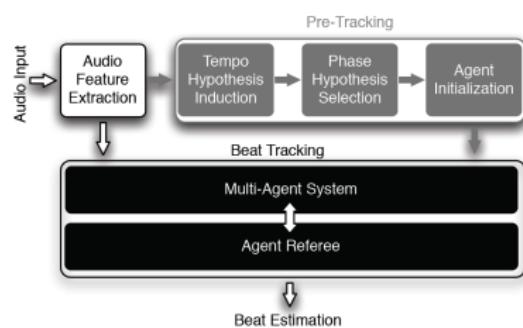
Seguimiento de Tempo

Procesamiento digital de señales de audio
Curso 2012

Gonzalo Gutiérrez, Matias Tailanián

26 de Diciembre del 2012

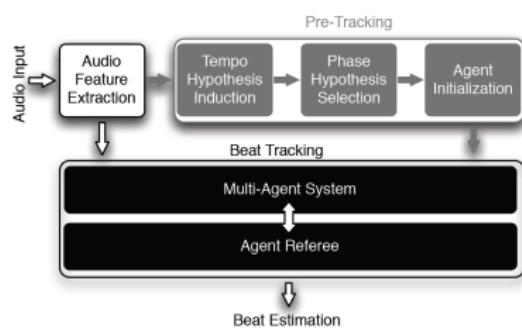
IBT: Real tempo and Beat-Tracking system



■ Feature Selection

Figura: Diagrama de bloques

IBT: Real tempo and Beat-Tracking system



- Feature Selection
- Pre Tracking

Figura: Diagrama de bloques

IBT: Real tempo and Beat-Tracking system

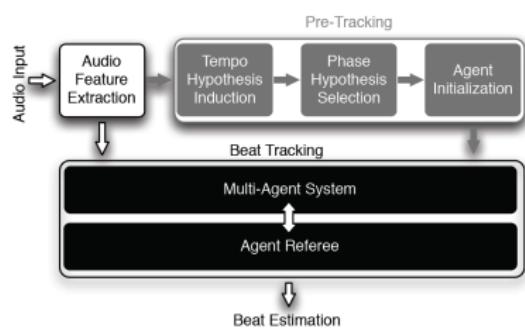


Figura: Diagrama de bloques

■ Feature Selection

■ Pre Tracking

- Período
- Fase
- Agente

IBT: Real tempo and Beat-Tracking system

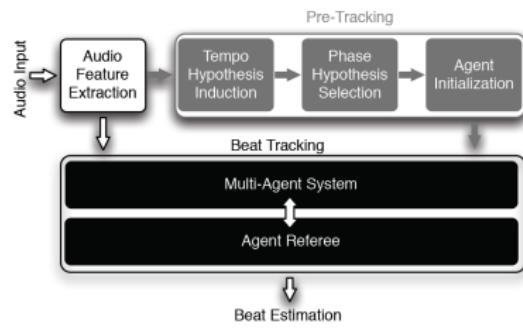


Figura: Diagrama de bloques

■ Feature Selection

■ Pre Tracking

- Período
- Fase
- Agente

■ Beat tracking

IBT: Real tempo and Beat-Tracking system

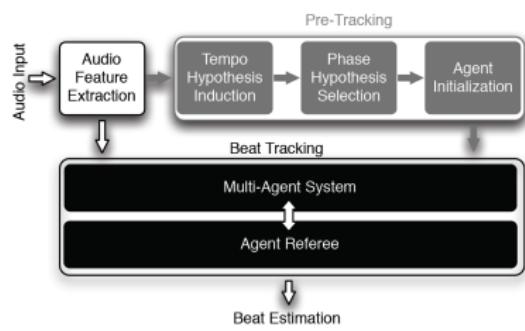
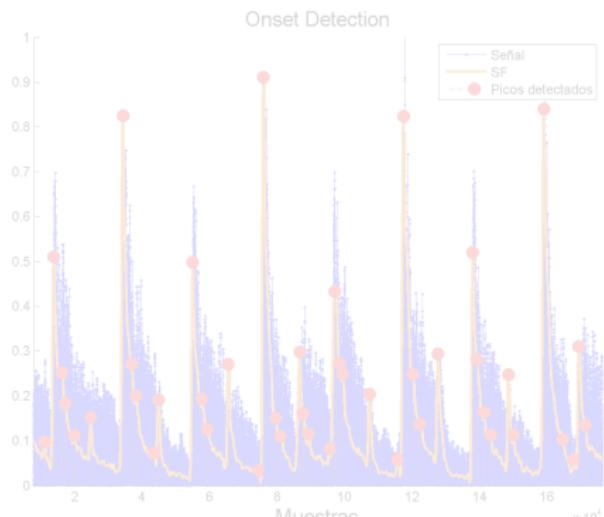


Figura: Diagrama de bloques

- **Feature Selection**
- **Pre Tracking**
 - Período
 - Fase
 - Agente
- **Beat tracking**
- **Ejemplos**
- **Conclusiones**

Feature Selection

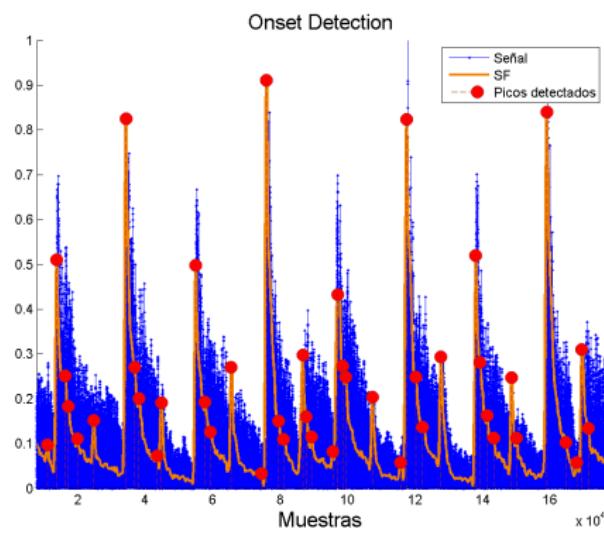
■ Onset detection



Feature Selection

■ Onset detection

$$SF(n) = \sum_{k=-\frac{\omega}{2}}^{\frac{\omega}{2}-1} HWR(|X(n, k)| - |X(n-1, k)|)$$

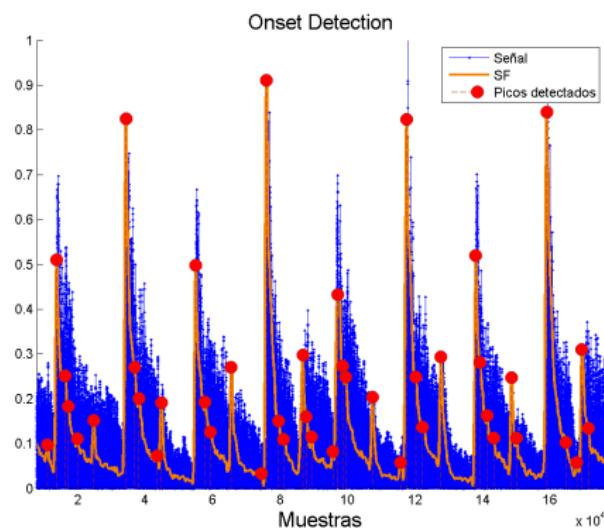


Feature Selection

■ Onset detection

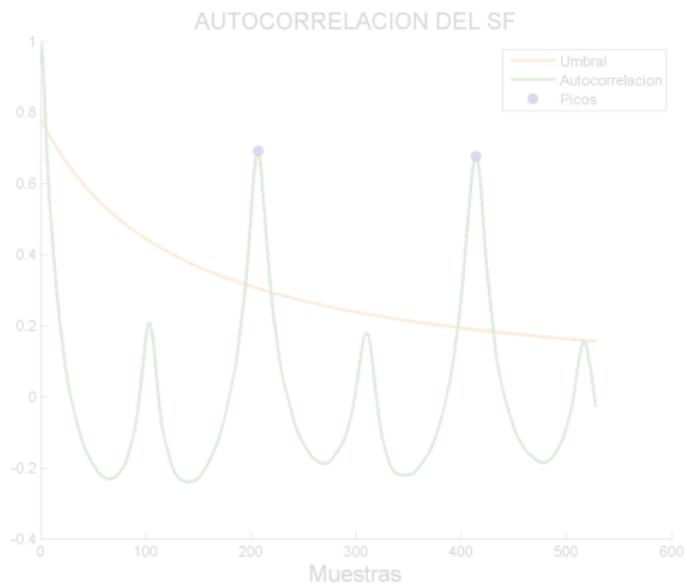
$$SF(n) = \sum_{k=-\frac{\omega}{2}}^{\frac{\omega}{2}-1} HWR(|X(n, k)| - |X(n-1, k)|)$$

$HWR(x) = \frac{x+|x|}{2}$ es la rectificación de media onda de x , solo nos interesa las variaciones crecientes en el espectro.



Pre-Tracking

■ Período

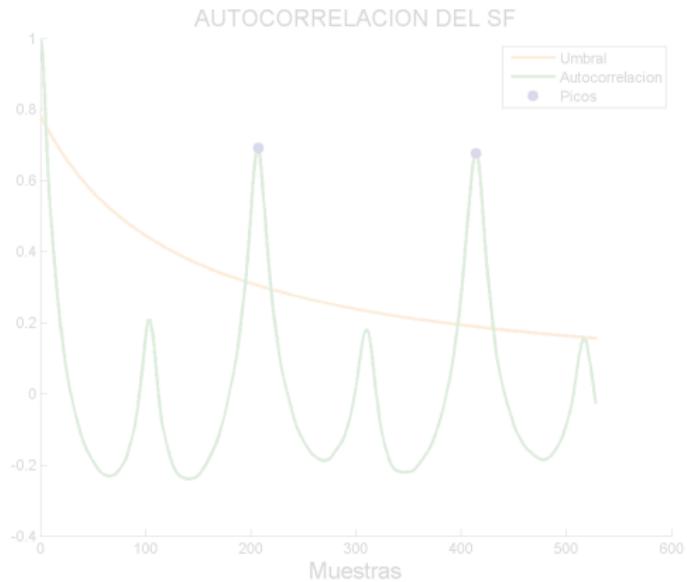


Pre-Tracking

■ Período

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

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

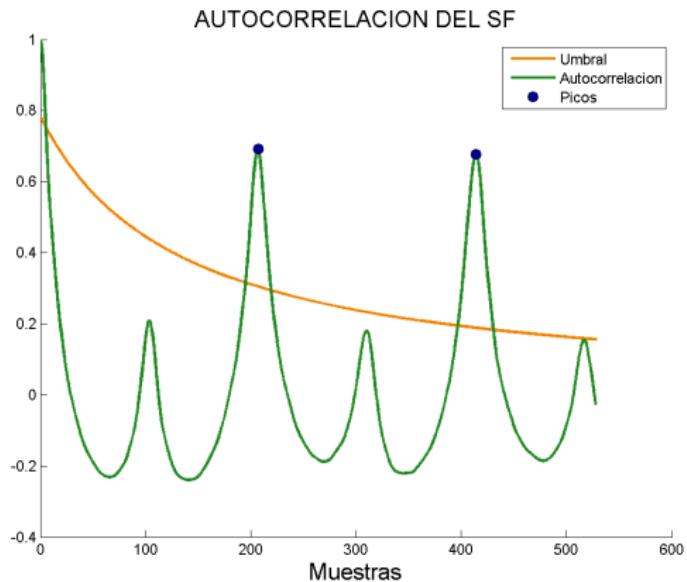


Pre-Tracking

■ Período

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

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



Pre-Tracking

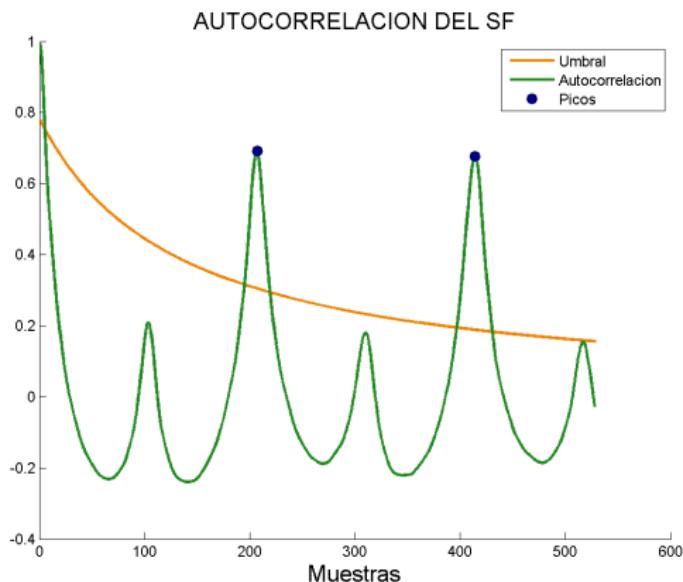
■ Período

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

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

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

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



Pre-Tracking

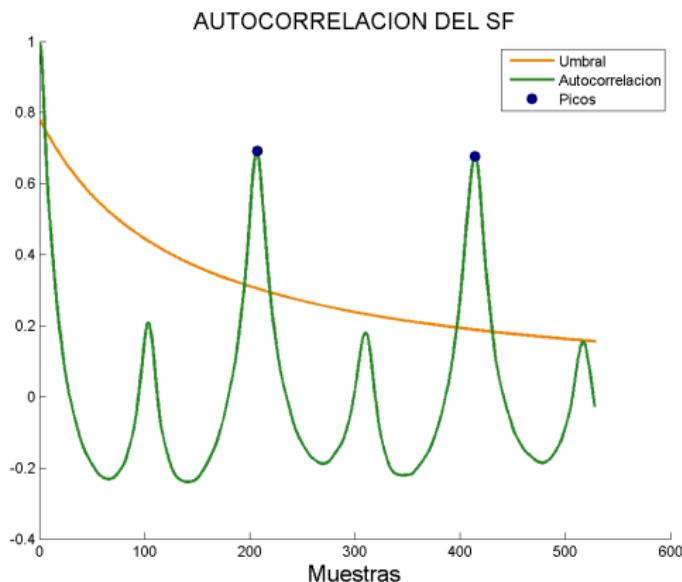
■ Período

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

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

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

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



Promedio de BPM

El pico mas alto corresponde a 120BPM

Pre-Tracking

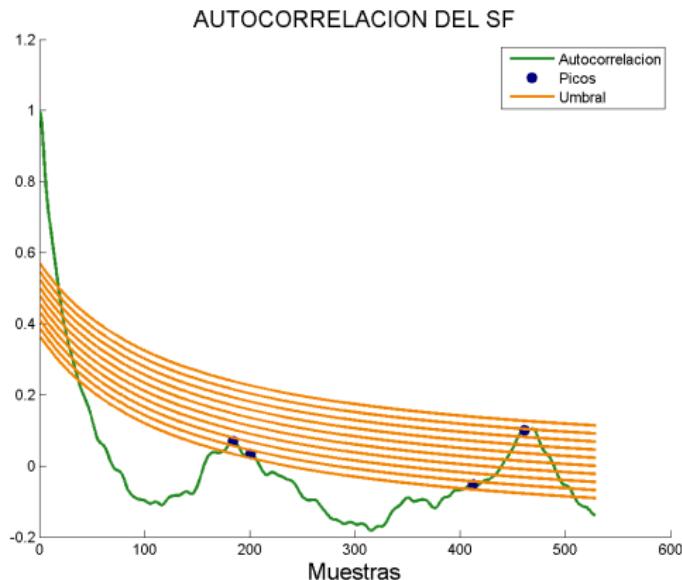
■ Período

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

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

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

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

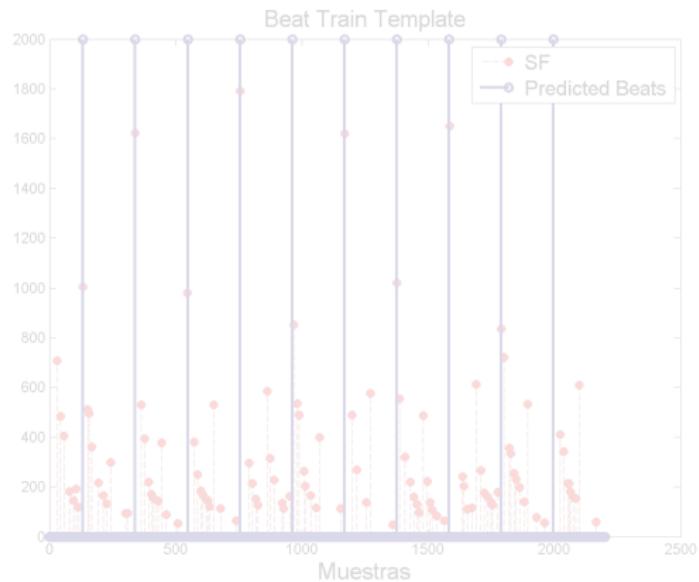


Umbral variable

Para garantizar la detección de picos

Pre-Tracking

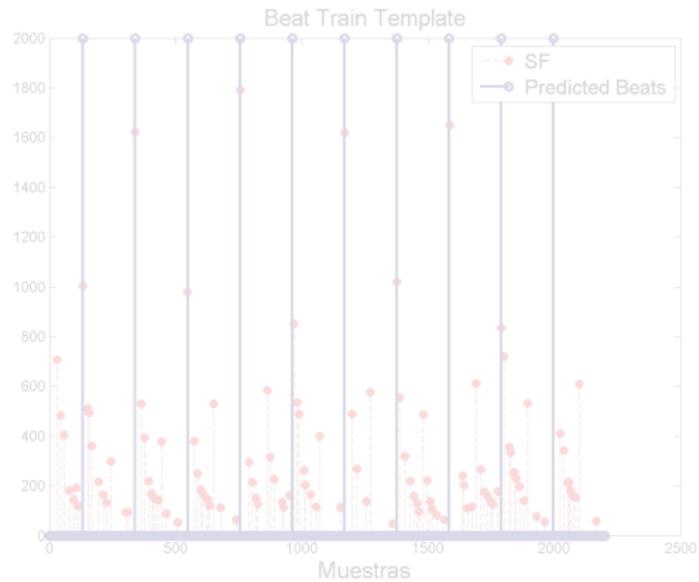
■ Fase



Pre-Tracking

■ Fase

Para cada P_i estimado se generan j hipótesis de 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.

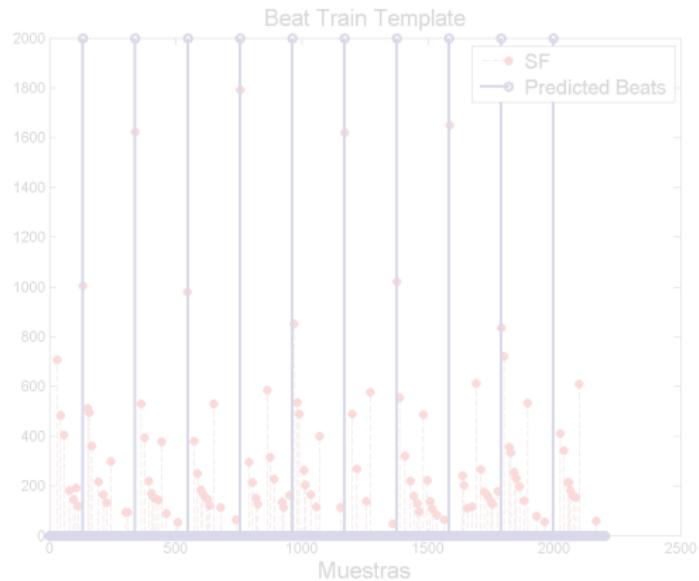


Pre-Tracking

■ Fase

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

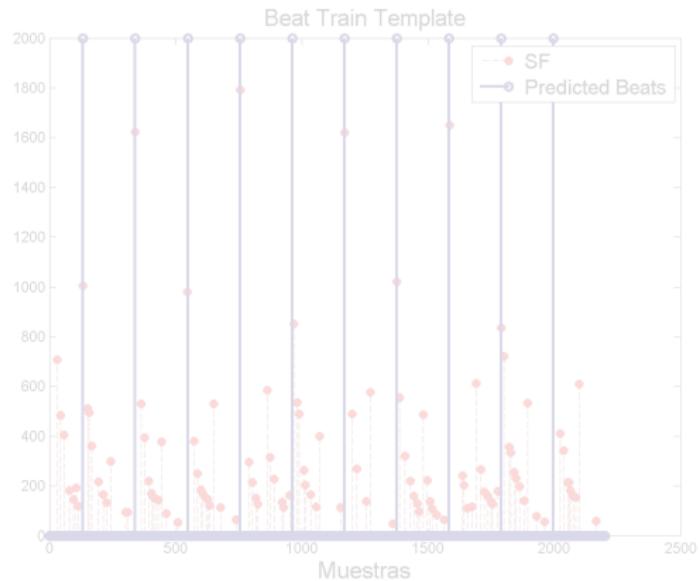
Se utiliza un *template* de tren de pulsos para ver cuál ajusta mejor.



Pre-Tracking

■ Fase

Para cada P_i estimado se generan j hipótesis de 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.

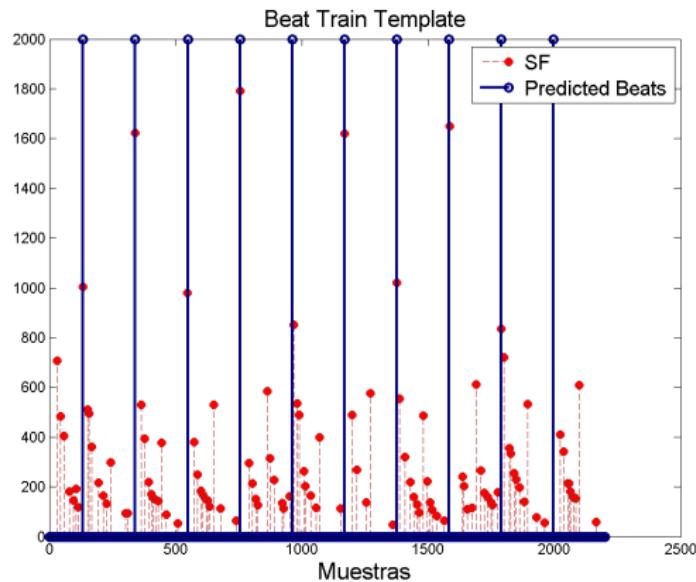


Pre-Tracking

■ Fase

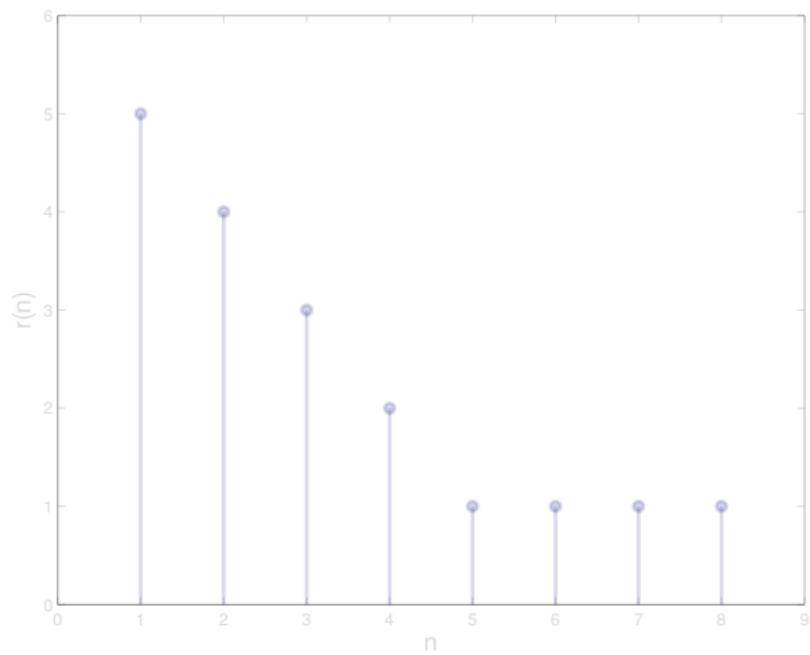
Para cada P_i estimado se generan j hipótesis de fase: ϕ_i^j . Se supone fase y períodos constantes en cada ventana de análisis.

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



Pre-Tracking

■ Fase

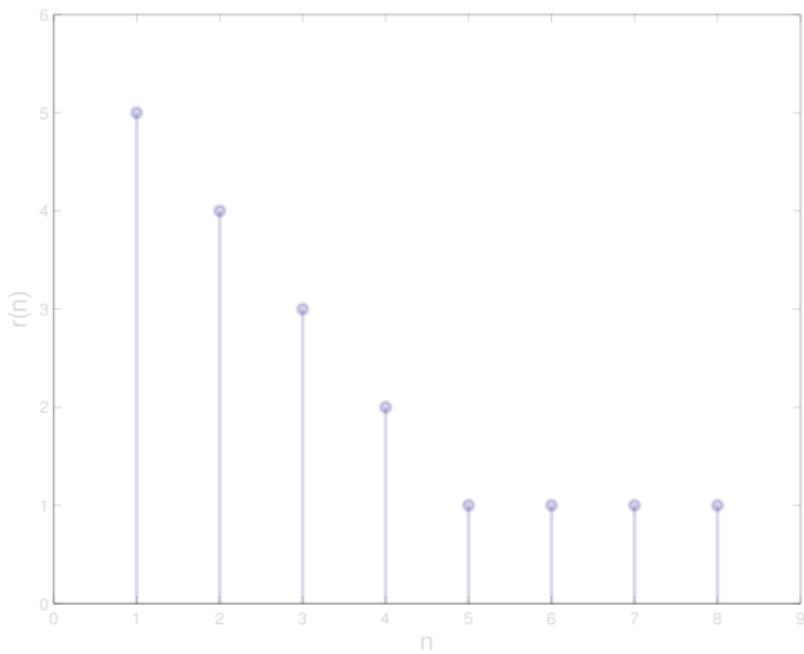


Pre-Tracking

■ Fase

Mejor fase

Se obtiene cuando al “mover” el tren de pulsos se alcanza un máximo en el puntaje que arroja la siguiente ecuación:



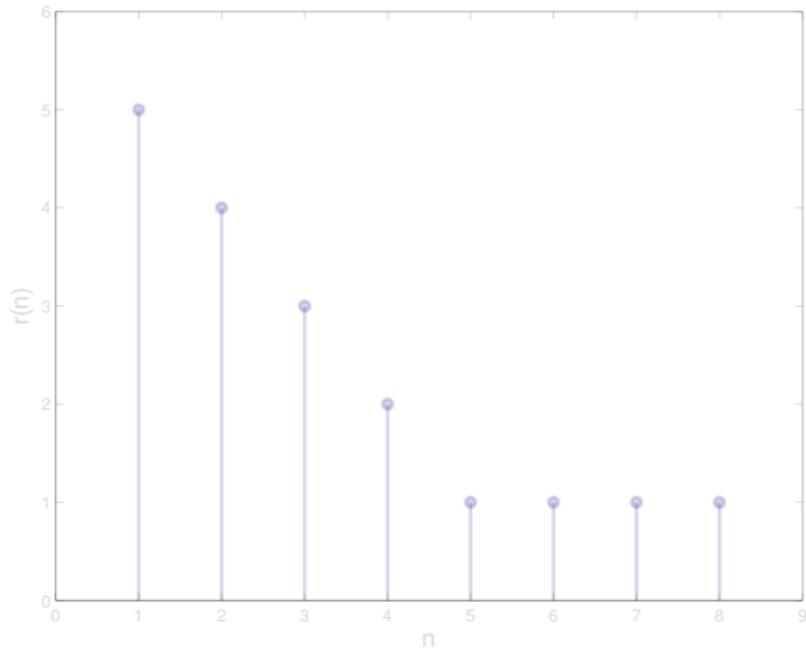
Pre-Tracking

■ Fase

Mejor fase

Se obtiene cuando al “mover” el tren de pulsos se alcanza un máximo en el puntaje que arroja la siguiente ecuación:

$$S_i = S_i^{rel} * \max(S_i^{raw})$$



Pre-Tracking

■ Fase

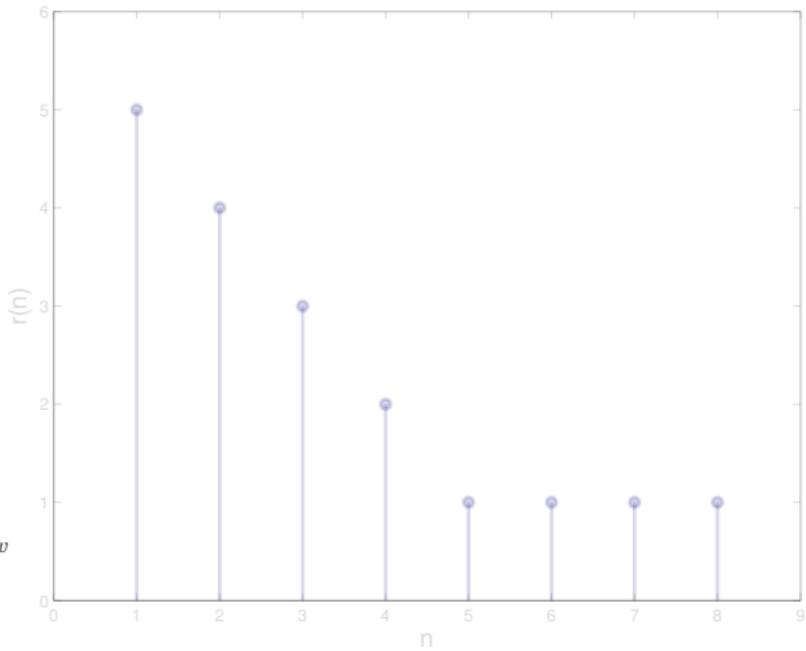
Mejor fase

Se obtiene cuando al “mover” el tren de pulsos se alcanza un máximo en el puntaje que arroja la siguiente ecuación:

$$S_i = S_i^{rel} * \max(S_i^{raw})$$

donde:

$$S_i^{rel} = 10 * S_i^{raw} + \sum_{j=0}^N r(n_{ij}) * S_j^{raw}$$



Pre-Tracking

■ Fase

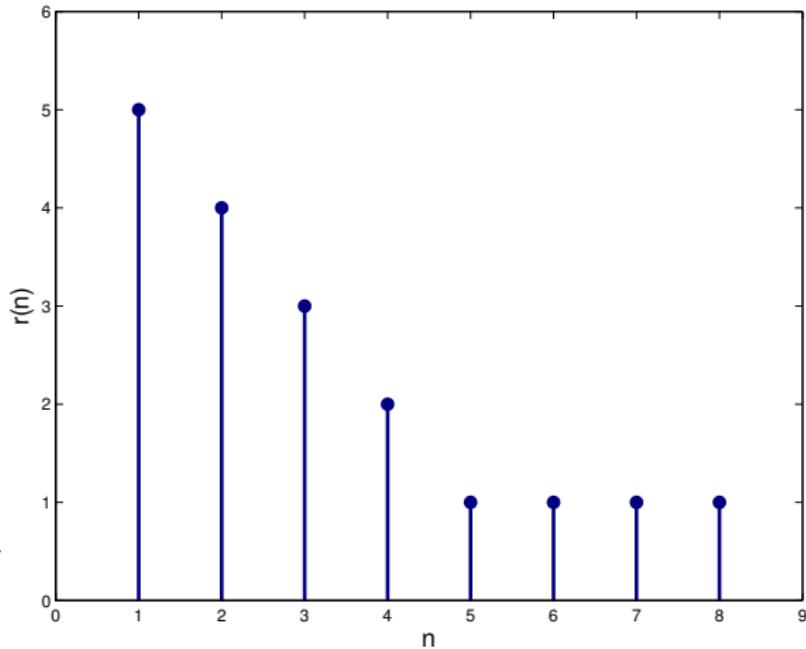
Mejor fase

Se obtiene cuando al “mover” el tren de pulsos se alcanza un máximo en el puntaje que arroja la siguiente ecuación:

$$S_i = S_i^{rel} * \max(S_i^{raw})$$

donde:

$$S_i^{rel} = 10 * S_i^{raw} + \sum_{j=0}^N r(n_{ij}) * S_j^{raw}$$



Pre-Tracking

■ Fase

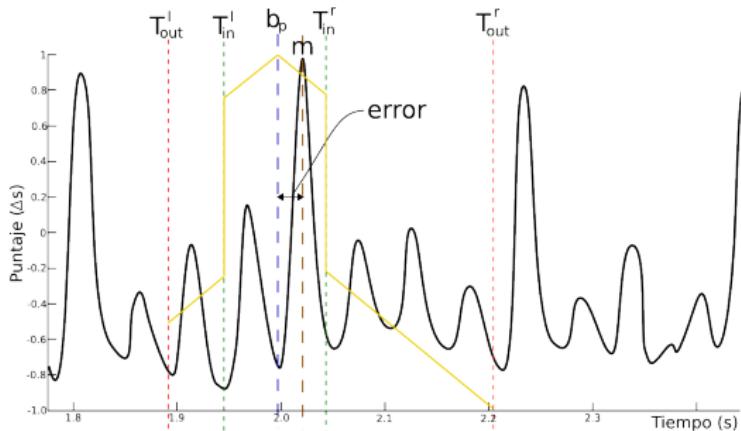
Mejor fase

Se obtiene cuando al “mover” el tren de pulsos se alcanza un máximo en el puntaje que arroja la siguiente ecuación:

$$S_i = S_i^{rel} * \max(S_i^{raw})$$

donde:

$$S_i^{rel} = 10 * S_i^{raw} + \sum_{j=0}^N r(n_{ij}) * S_j^{raw}$$



Pre-Tracking

■ Agentes

Pre-Tracking

■ Agentes

Las hipótesis estimadas

(P_i, ϕ_i, S_i) pueden ser ahora utilizadas para inicializar un conjunto de N agentes en el cual cada uno empezará a realizar su propio beat tracking

Beat Tracking

Agent Referee

Evalúa la distancia entre la predicción del beat (b_p) y el máximo local (m) de SF . P_m es el máximo período permitido

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

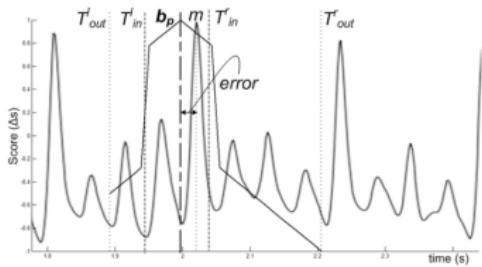
Beat Tracking

Agent Referee

Evalúa la distancia entre la predicción del beat (b_p) y el máximo local (m) de SF . P_m es el máximo período permitido

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

Niveles de tolerancia: $T_{in} \in [T_{in}^l, T_{in}^r]$ y
 $T_{out} \in [T_{out}^l, T_{in}^l] \cup [T_{in}^r, T_{out}^r]$



Beat Tracking

Agent Referee

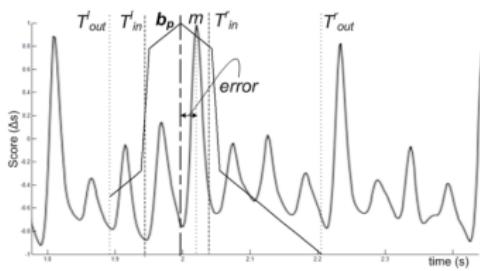
Evalúa la distancia entre la predicción del beat (b_p) y el máximo local (m) de SF . P_m es el máximo período permitido

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

Niveles de tolerancia: $T_{\text{in}} \in [T_{\text{in}}^l, T_{\text{in}}^r]$ y
 $T_{\text{out}} \in [T_{\text{out}}^l, T_{\text{in}}^l] \cup [T_{\text{in}}^r, T_{\text{out}}^r]$

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

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



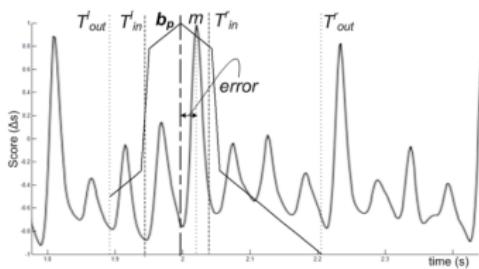
Beat Tracking

Agent Referee

Evalúa la distancia entre la predicción del beat (b_p) y el máximo local (m) de SF . P_m es el máximo período permitido

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

Niveles de tolerancia: $T_{in} \in [T_{in}^l, T_{in}^r]$ y $T_{out} \in [T_{out}^l, T_{in}^l] \cup [T_{in}^r, T_{out}^r]$



■ $T_{in} \in [T_{in}^l, T_{in}^r]$

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

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

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

Beat Tracking

■ Descarte de agentes

Beat Tracking

■ Descarte de agentes

- Replacement
- Redundancy
- Obsolescence
- Loss

Replacement

Si es el peor de un grupo de agentes que alcanzó su máximo, y su puntuación es inferior a una nueva creación de agente

Beat Tracking

■ Descarte de agentes

- Replacement
- Redundancy
- Obsolescence
- Loss

Redundancy

Para incrementar la eficiencia del algoritmo, se elimina un agente si se duplica el trabajo de un agente de score alto, si sus períodos difieren menos de 11,6ms y sus fases menos de 23,2ms

Beat Tracking

■ Descarte de agentes

- Replacement
- Redundancy
- Obsolescence
- Loss

Obsolescence

Se elimina un agente si la diferencia entre su puntuación y el mejor agente es mayor que el 80 %

Beat Tracking

■ Descarte de agentes

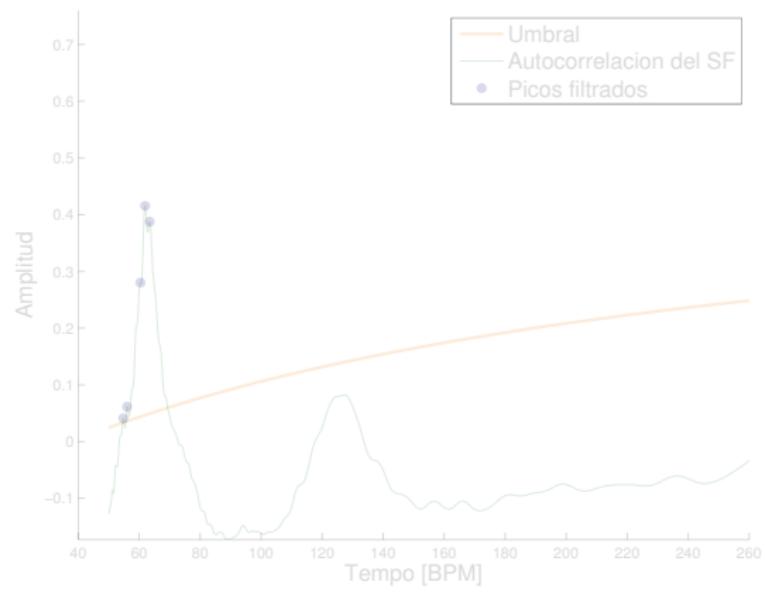
- Replacement
- Redundancy
- Obsolescence
- Loss

Loss

Se descarta una agente si parece estar “perdido”, es decir, si su predicción de beat cae 8 veces fuera de la ventana de tolerancia interior

Ejemplos

■ Caso de buen funcionamiento

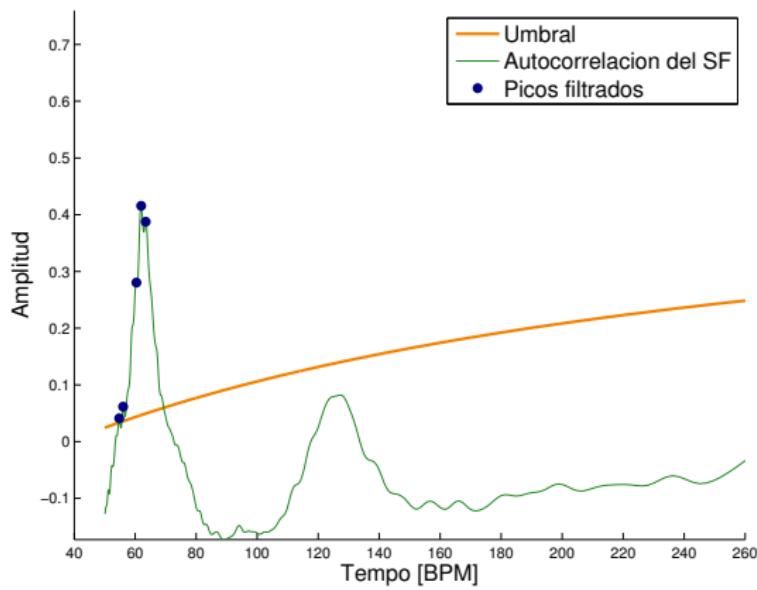


Ejemplos

■ Caso de buen funcionamiento

Pre-Tracking

El valor del mayor pico corresponde a 63BPM

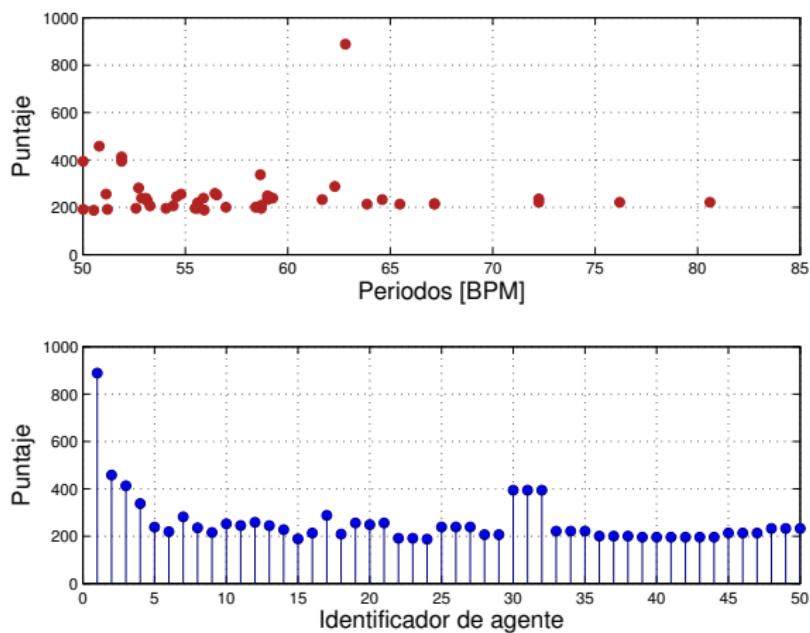


Ejemplos

■ Caso de buen funcionamiento

Beat Tracking

El 1er agente es el ganador para esta pieza

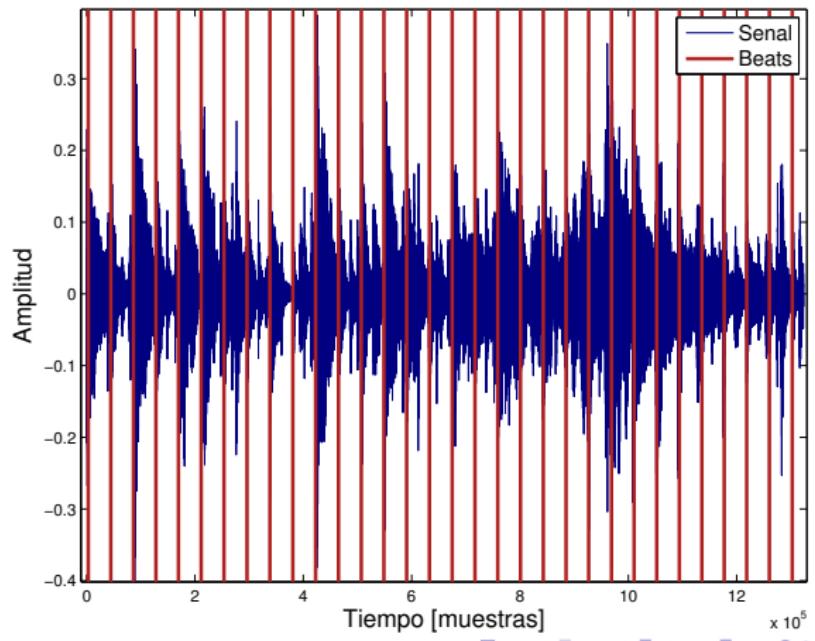


Ejemplos

■ Caso de buen funcionamiento

Agente 1

▶ PLAY



Ejemplos

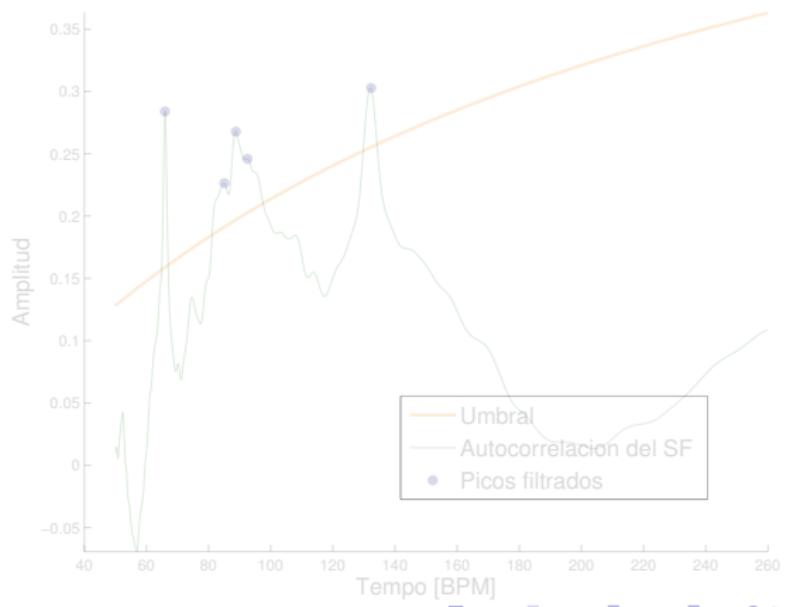
■ Caso de buen funcionamiento

Agente 1

Cont-Based:	cC: 100.00	cT: 100.00	aC:100.00	aT:100.00
F-Mesure:	f: 94.44	p: 94.44	r: 94.44	a: 89.47

Ejemplos

■ Caso de funcionamiento comprometido

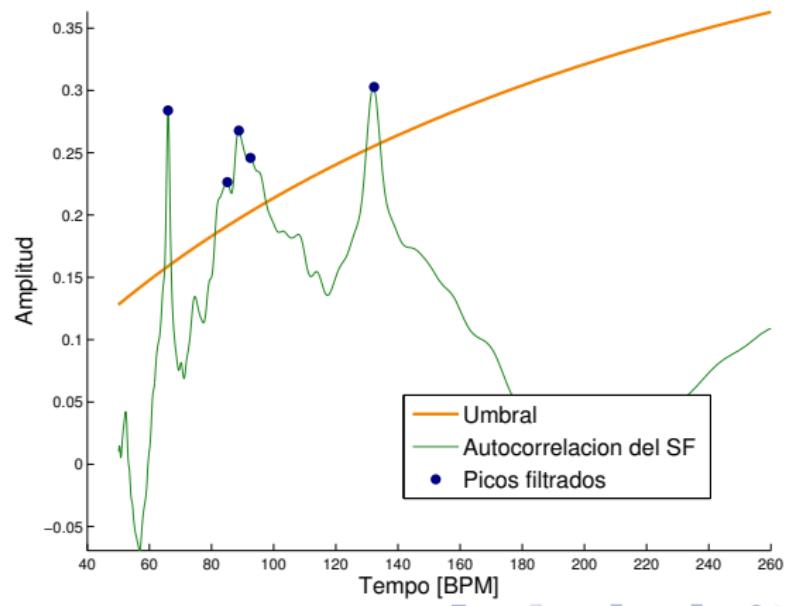


Ejemplos

■ Caso de funcionamiento comprometido

Pre-Tracking

El valor del mayor pico corresponde a 130BPM

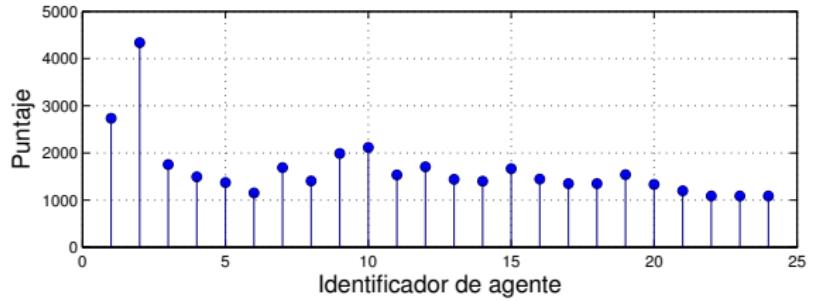
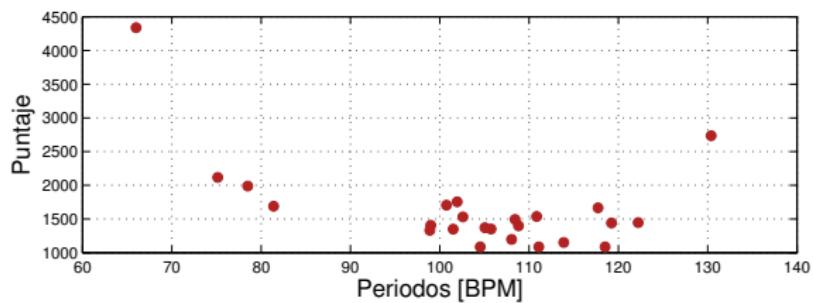


Ejemplos

■ Caso de funcionamiento comprometido

Beat Tracking

El 2do agente es el ganador para esta pieza

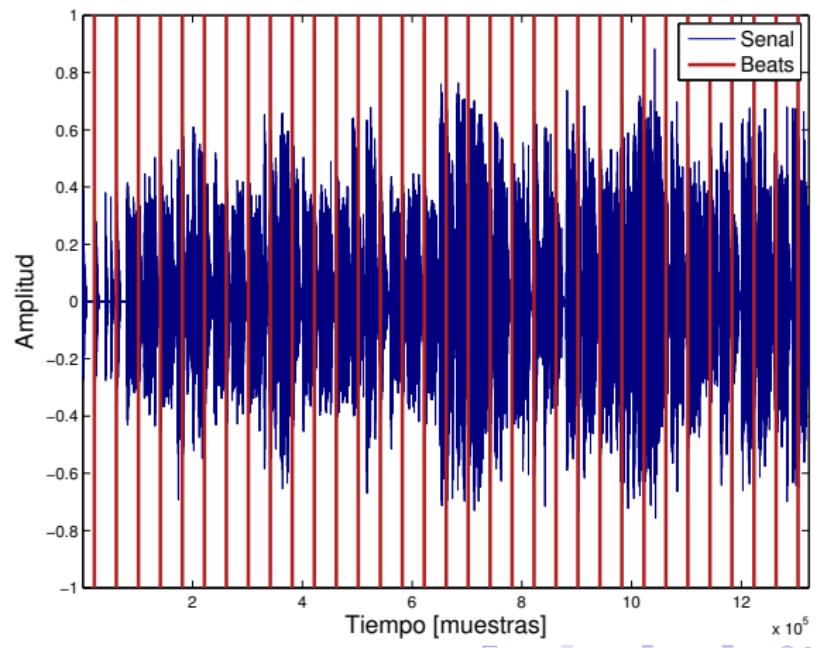


Ejemplos

■ Caso de funcionamiento comprometido

Agente 2

▶ PLAY

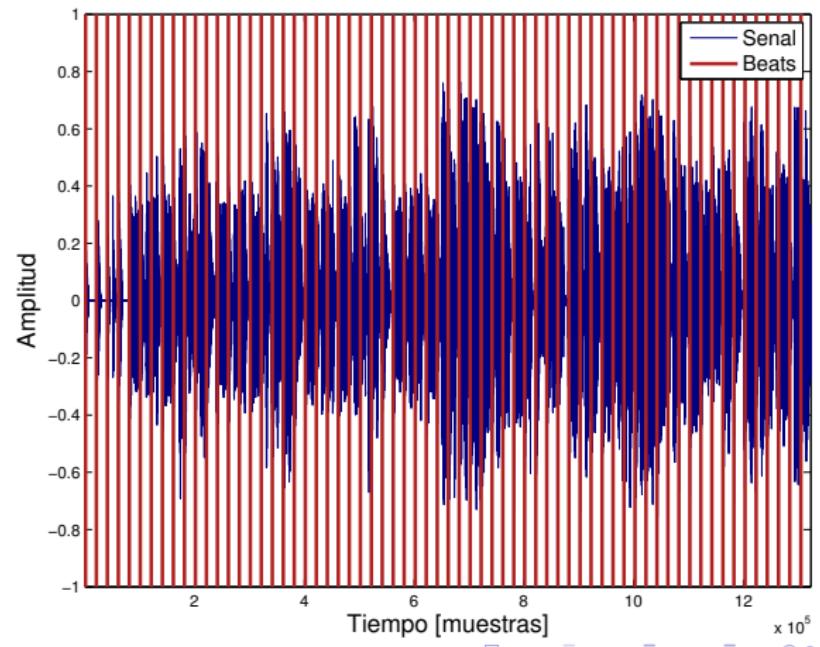


Ejemplos

■ Caso de funcionamiento comprometido

Agente 1

▶ PLAY



Ejemplos

■ Caso de funcionamiento comprometido

Agente 2

Cont-Based:	cC: 00.00	cT: 00.00	aC: 100.00	aT: 100.00
F-Mesure:	f: 67.47	p: 100.00	r: 50.91	a: 50.91

Ejemplos

■ Caso de funcionamiento comprometido

Agente 2

Cont-Based:	cC: 00.00	cT: 00.00	aC: 100.00	aT: 100.00
F-Mesure:	f: 67.47	p: 100.00	r: 50.91	a: 50.91

Agente 1

Cont-Based:	cC: 100.00	cT: 100.00	aC: 100.00	aT: 100.00
F-Mesure:	f: 100.00	p: 100.00	r: 100.00	a: 100.00

Conclusiones

- Desempeño aceptable

Conclusiones

- Desempeño aceptable
- Dudas de implementación

Conclusiones

- Desempeño aceptable
- Dudas de implementación
- Se puede mejorar la elección del agente ganador

Conclusiones

- Desempeño aceptable
- Dudas de implementación
- Se puede mejorar la elección del agente ganador
- Tiene robustez

Conclusiones

- Desempeño aceptable
- Dudas de implementación
- Se puede mejorar la elección del agente ganador
- Tiene robustez
- Buena relación entre inercia y rapidez de respuesta

Conclusiones

- Desempeño aceptable
- Dudas de implementación
- Se puede mejorar la elección del agente ganador
- Tiene robustez
- Buena relación entre inercia y rapidez de respuesta
- **Buenos resultados “auditivos”**

Referencias:

-  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.
-  S. Dixon. Automatic extraction of tempo and beat from expressive performances. In *Journal of New Music Research*, 30(1):39–58, 2001.
-  S. Dixon. Onset detection revisited. In *Proceedings of the 9th International Conference on Digital Audio Effects*, pages 133–13, Montreal, Canada, 2006.
-  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.