



UNIVERSIDAD
DE SANTIAGO
DE CHILE

Departamento de Matemática y Ciencia de la Computación

INFORME DEL PROYECTO ESTIMACIÓN DE ALTURA MUSICAL MÚLTIPLE 2

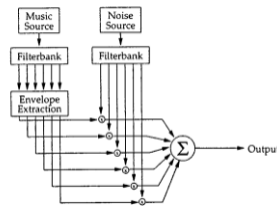
JENNIFER ROJAS GALVEZ

1. INTRODUCCIÓN

En este proyecto se presenta un método para usar un pequeño número de filtros de bandpass y bancos de filtros comb paralelos para analizar el tempo y extraer el ritmo de señales musicales. Este análisis se puede utilizar de manera predicativa para adivinar cuándo se producirán latidos en el futuro. El objetivo es construir un algoritmo computacional capaz de la extracción automática de pulso rítmico en un oyente humano. (También llamado seguimiento de ritmo y golpeteo de pies) El “ritmo” como concepto musical es intuitivo de entender, pero algo difícil de definir. Para nuestros propósitos, el ritmo de una pieza musical es la secuencia de impulsos de fenómenos equidistantes que definen un tempo para la música. Este proyecto solo se ocupa del ritmo y el tempo. Las agrupaciones y las relaciones fuerte/débil que definen el ritmo y la métrica no se consideran.

2. SIMPLIFICACIÓN PSICOACÚSTICA.

Observemos que el oído de los humanos posee características que nos permite hacer distintas simplificaciones en el procesamiento de señales musicales. Una de estas simplificaciones establece que el contenido rítmico de una señal puede extraerse a partir de las envolventes de las bandas de la señal musical, esta es la base para el algoritmo que se estudiará.

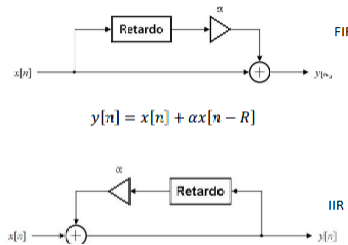


Date: 28-01-2022.

en la figura de arriba podemos observar que la amplitud de cada banda de la señal de ruido se ajusta con la envolvente de la banda correspondiente a la salida del banco de filtros musical y la señales de ruido resultantes se suman para así formar una señal de salida (por lo general tiene una percepción rítmica considerablemente igual a la señal musical original)

3. FILTRO COMB.

los filtros de comb se les llama así ya que la respuesta de amplitud se parece a la forma de un peine. son un tipo de filtros recursivos muy sencillo pueden ser del tipo FIR o IIR , su tarea es combinar la señal de entrada o salida con una versión retrasada.



En particular, si estimulamos un filtro comb con retardo T con un tren de pulsos del lado derecho de altura A y período K , obtenemos resonancia si $T=K$. Por otro lado, si T distinto K , la convergencia es a un valor más pequeño.

Así, un filtro con retraso que coincida (o divida uniformemente) con el período de un tren de pulsos tendrá una salida más grande (más enérgica) que un filtro con retardo no coincidente..

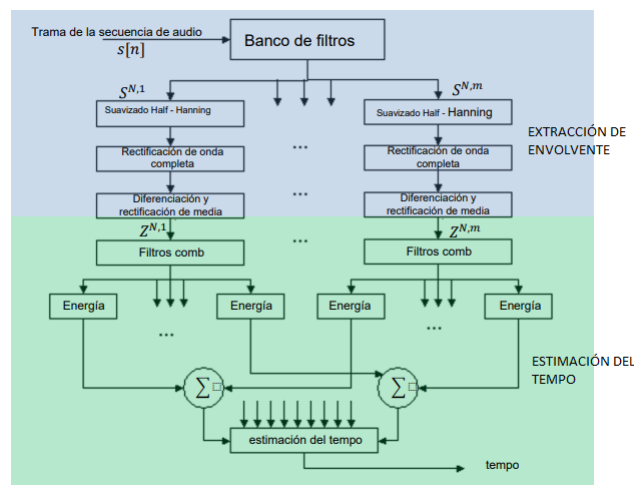
4. DESCRIPCIÓN DEL ALGORITMO.

El algoritmo propuesto por Eric scheirer se enfoca en encontrar el periodo fundamental en que aparecen impulsos en las pulsaciones periódicas de la canción. Para llevar esto a cabo se crean versiones filtradas a distintas bandas de señal musical para así obtener la envolvente de cada una de estas señales, diferenciándolas, rectificándolas y calculando la energía que produce la convolución de dichas señales con una serie de filtros de comb.

Entonces el algoritmo propuesto posee 2 grandes partes : -La estimación del tempo: contiene además dos funciones: una función que extrae la envolvente, otra que estima el tempo y que está compuesta por un banco de filtros comb -Ubicación de las pulsaciones en una escala temporal

4.1. **Estimación del tiempo.** El bloque de estimación del tiempo contiene los siguientes sub-bloques tal cual se aprecia en la figura mas abajo:

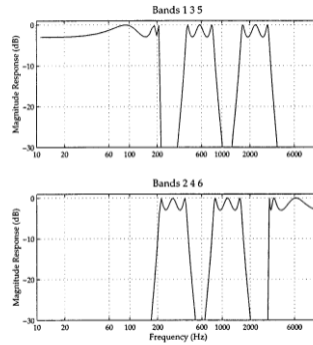
- Banco de filtros
- Filtrado con un pulso de half-Hanning
- Rectificación de onda completa
- Diferenciación
- Rectificado de medio onda
- Banco de filtros comb
- Estimación del tiempo



La señal pasará por las siguientes etapas antes de poder estimar su tiempo:

- Descomposición sub-banda
- Extracción de envolvente
- Diferenciador y Rectificación de media onda

4.1.1. *Descomposición sub-banda.* Se pasa la señal utilizando la transformada discreta de Fourier (DFT) y se obtiene la señal, luego esta señal pasa por un banco de filtros donde en el algoritmo tiene seis bandas y cada banda está claramente definida y cubre aproximadamente un rango de una octava., cada una de estas señales contienen un rango de frecuencias comprendidas: la banda mas baja 0-200 Hz; las sgtes cuatro bandas 200-400 Hz, 400-800 Hz, 800-1600 Hz y 1600-3200 Hz ;La banda más alta con frecuencia de corte a 3200 Hz. Cada filtro se implementa usando un filtro elíptico de sexto orden, con 3 dB de ondulación en el banda de paso y 40 dB de rechazo en la stopband



4.1.2. *Extracción de envolvente.* La envolvente de la señal es la tendencia general de la amplitud del sonido. Se convolucionará cada una de las señales con una ventana half-Hanning de longitud 0.2 segundos (200-ms half-Hanning-raised cosine window), llevando a cabo una integración de energía. Luego se toman cada una de las seis señales ya filtradas se les hace un rectificado de onda completa para reducir el contenido de alta frecuencia.

4.1.3. *Diferenciador y Rectificación de media onda.* Las señales en este bloque contienen seis rangos de frecuencias con una envolvente bien definida, pero es necesario derivarlas para recalcar los cambios de amplitud sonora. Estas seis señales luego de su derivación, se las vuelve a rectificar, pero sólo positivamente porque sólo interesa los incrementos de la amplitud.

4.1.4. *Banco de filtros comb.* Cada filtro de comb tiene la forma de

$$y_t = \alpha y_{t-T} + (1-\alpha) x_t$$

4.1.5. *Estimación de tempo.*

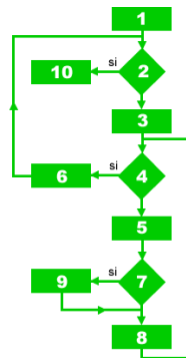


Figura 2.9. Diagrama de flujo del proceso de estimación del tempo

En el diagrama de flujo se puede observar mejor el funcionamiento del proceso de estimación del tempo.

Proceso 1:

-En primer lugar calculamos la DFT de la señal que ingresa.

-Establecemos los valores de bpm's que conforman el intervalo de analisis;
bpmmin=60, bpmmax=120

-Inicializamos la variable que almacenara el valor maximo de energia encontrada para cada banda entre 60 y 120 ; Maxe=0 , bpm=bpmmin

Condicional 2:

-Se inicia el primer bucle donde la condicion de salida es que bpm lleve a bpmmax; bpm<bpmmax

-Se repite el bucle bpmmax-bpmmin veces

Proceso 3:

-Se crea el filtro comb de acuerdo al actual bpm

-Inicializamos la variable e. e=0

-En la primera iteración su valor será e irá aumentado en una unidad hasta llegar a bpmmax.

-Inicializamos el valor de la variable m, donde su valor inicial es 1, y el final 6

Condicional 4:

-Iniciamos el segundo bucle donde la condición es que la banda sea igual a seis para finalizarlo.

Proceso 5:

-Se calcula la energía de la señal correspondiente a la banda m, al beat rate "bpm" y la almacenamos en la variable e.

Proceso 6:

-Se incrementa en una unidad el valor de bpm bpm=bpm+1 y volvemos a "3".

Condición 7:

-Comparamos si la variable "e" es mayor que el Maxe encontrado hasta el momento.

Proceso 8:

-Incrementamos el valor de m en uno m=m+1

Proceso 9:

-Guardamos el valor del bpm actual en la variable sbpm

-Retenemos el valor de e en la variable Maxe

Proceso 10:

-Una vez se haya recorrido todas las bandas y todo el intervalo de valores de los bpm's se entrega el resultado.

5. DETERMINACIÓN DE FASE

En esta fase se puede predecir el momento en que llegará el siguiente latido. El valor en la red de retardo se puede ver como "predicted output" en el caso de que la entrada sea cero. La implementación del modelo realiza la fase análisis cada 25 ms.

6. COMPARACIÓN CON MÉTODOS DE AUTO-CORRELACIÓN

- Ventajas: da información de fase, predice el siguiente pulso.
- Desventajas: menos eficiente en el uso de la memoria.

7. CONCLUSIONES

En este proyecto se pudo comprobar que la sensación de ritmo y tempo dependen principalmente de las variaciones de amplitud de las envolventes de las bandas de una señal con contenido musical.

Se ha descrito un algoritmo que puede realizar con éxito un seguimiento de audio digital que representa música de muchos tipos diferentes.

Una de las ventajas del sistema es que se puede implementar diseñando una serie de filtros sencillos y haciendo una búsqueda iterativa y sistemática que no tiene mayores complicaciones.