[[1]](#footnote-1)

Preparation of Papers for IEEE TRANSACTIONS and JOURNALS

First A. Author, Second B. Author, Jr., and Third C. Author, Member, IEEE

# Estimacion del tempo

## Defincion de tempo

El tempo de una canción suele dar una idea del ritmo de una canción. Por lo general, las canciones tienen un pulso subyacente de una frecuencia específica, este pulso se denomina ‘beat’. Para un oyente humano el hecho de identificar el beat de una canción es simple, normalmente se acompaña dicho beat con el pie. Sin embargo, es un problema difícil de definir para ser resuelto de manera automática por una computadora. El tempo de una canción puede medirse en BPM(beats per minute) que se define como la cantidad de veces que se repite este pulso subyacente en un intervalo de tiempo de un minuto.

Se estableció como objetivo la implementación de un algoritmo que pueda determinar el BPM de un archivo de música wav dentro de un rango tolerable de error(±4BPM) y ue dicho algoritmo funcione para el rango de 60 ≤ BPM ≤ 180.

## Algoritmo

Se implementó un algoritmo de detección de BPM simple basado en los métodos implementados en [1] y [2].

### Segmentacion

El algoritmo puede ser implementado en tiempo real por lo que se ejecuta cada vez que se obtienen un numero M de muestras de audio. Para nuestra implementación utilizamos por default M=375.

Estas M muestras de audio se denominará bloque y funciona como uno de los parámetros del algoritmo.

### Senal de eventos ritmicos

Para cada uno de estos bloques de M muestras, se calcula un único valor de potencia que se almacena en memoria. La potencia del bloque se calcula según la siguiente formula:

Donde α es un parámetro definido en el intervalo 0< α< 1. M es el valor de la cantidad de muestras tomadas por bloque, y los elementos dentro de la sumatoria son los cuadrados de las muestras de audio del bloque actual. Para valores más grandes de α se tiene que la señal se ve menos afectada ante nuevos valores de las muestras.

Los valores de potencia obtenidos dan una idea de cómo varia la energía de la señal con el tiempo. Se tienen picos de amplitud muy grandes en los instantes de tiempo en los que hay eventos rítmicos importantes (generalmente cerca de donde ocurre un beat).

El arreglo donde se almacenan los valores de potencia tiene un límite máximo de N valores. Cuándo se excede este máximo se descarta el valor más viejo y se desplazan todos los valores para hacer lugar al valor mas reciente.

### Calculo del espectro

Una vez que se calculó el valor de potencia se procede a calcular el espectro de la señal mediante la aplicación de la FFT (Fast Fourier Transform) a los N valores de potencia.

El espectro obtenido tiene una frecuencia de muestreo distinta al de la señal de audio original debido a que se tiene un solo valor de potencia por cada bloque analizado. Por lo que la nueva frecuencia de muestreo obtenida es:

Luego se tiene que la resolución del espectro calculado es:

### Calculo y estimación del tempo.

Existe la siguiente relación matemática entre la frecuencia y el BPM:

Debido a que solo nos interesa el rango 60 ≤ BPM ≤ 180. Solo son de interés los bins de frecuencia correspondientes a ese rango de BPM. Los bins correspondientes a los limites se obtienen mediante la relación:

Donde se redondea el bin al entero más próximo. Se busca el bin dentro del rango definido para el cual la señal toma su máximo valor y se estima que su BPM asociado es el BPM del bloque.

## Implementacion y resultados

El algoritmo fue implementado en python utilizando dos canciones distintas y variando los parámetros a fin de ver el impacto en el resultado final. Para el cálculo del promedio () se utilizó la función ‘fftconvolve’ del paquete scipy. Para el cálculo del espectro, se utilizó la función ‘rfft’ del paquete numpy ya que el arreglo con valores de potencia siempre tiene valores reales y ‘rfft’ es computacionalmente más rápido que ‘fft’.

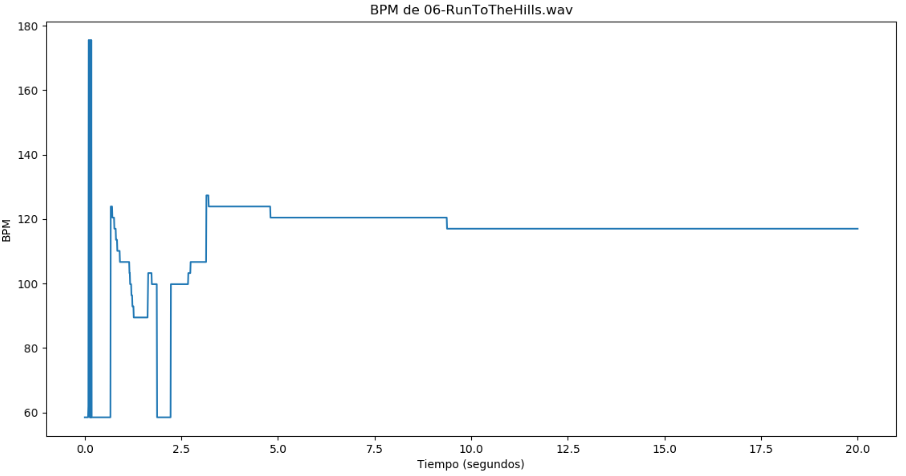


Figure 1.Grafica de BPM de los primeros 20segundos de 'Run To The Hills' de Iron Maiden

(α=0.85, M=375, N=2048, =44.1 KHz)

Como puede verse en ‘Figure 1’ el algoritmo obtiene un valor de BPM constante luego de un transitorio, en este caso los primeros 10 segundos de la canción. El valor de BPM en el que se establece es de 117 BPM lo que coincide con el BPM real de la canción.

El transitorio se debe al cálculo del espectro y también al de la potencia. El cálculo de potencia se basa en los valores calculados previamente, mientras mayor es el parámetro α mayor es la dependencia en los valores previos. Asimismo, el espectro se calcula siempre con N muestras de potencia, en el caso de la imagen para tener N muestras de potencia se tiene que:

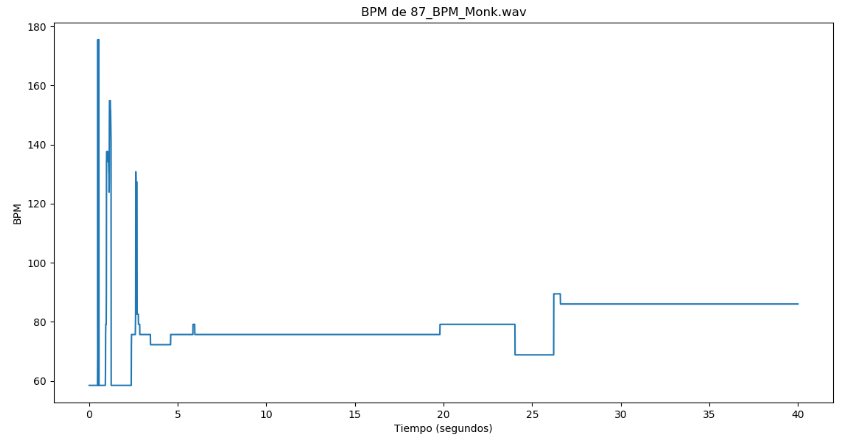


Figure 2.Grafica del BPM de 'Dinah' de Thelonius Monk

(α=0.85, M=375, N=2048, =44.1 KHz)

Luego se probó el algoritmo en una canción tocada únicamente por un piano. Esto se debe a que se buscó probar la validez del algoritmo en una canción un poco más compleja que no posea un instrumento de percusión que claramente índice los pulsos subyacentes a la canción. El resultado puede verse en ‘Figure 2’, luego del transitorio el resultado del BPM es de 86, lo cual es una buena estimación ya que se considera que el BPM real de la canción es 87. Se puede ver que para esta canción el transitorio duro el doble que para la primera canción que tenía una batería que indicaba el ritmo.

References

1. Beat Detection Algorithms in Signal Processors ETIN80. Luis Cavo, Siyu Tan, Adam Urga, 2016.
2. Design and implementation of a Beat Detector algorithm

(ETIN80 report). Jaime Gancedo, Sakif Hossain,

Wenpeng Song, 2018.

1. Streamlined Tempo Estimation Based on Autocorrelation and Cross-correlation With Pulses. Graham Percival, George Tzanetakis, 2014.

1. [↑](#footnote-ref-1)