

Convertir canción en karaoke: quitar la voz

Objetivo: Aprender a obtener y manipular los valores (muestras) de los canales de sonido para eliminar la voz humana de éstos con el fin de obtener una canción de estilo *karaoke*.

Introducción:

La idea general supone que existe una diferencia bastante perceptible entre la música del canal derecho y la del izquierdo pero no así con la voz humana que es muy similar en los dos canales. Al obtener los vectores de ambos y restarlos, la voz humana se pierde en su mayor parte, no obstante buena parte de la música permanece.

Desarrollo:

La idea general de la introducción se puede visualizar con el siguiente gráfico, donde las muestras del canal izquierdo son usadas como “tamiz” o “coladera” que “filtra” los valores del derecho. Es decir, la figura 1 es una *forma de interpretar* la ecuación:

$$[\text{Vector } \textit{karaoke}] = [\text{Vector derecho}] - [\text{Vector izquierdo}]$$

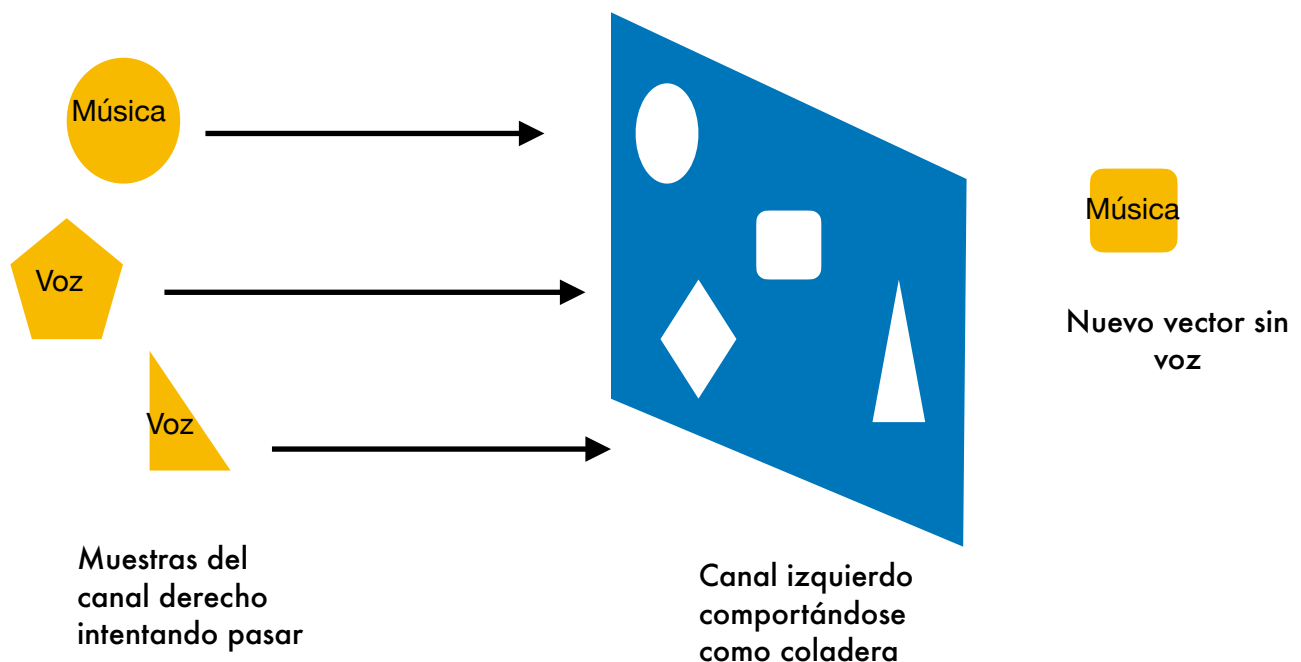
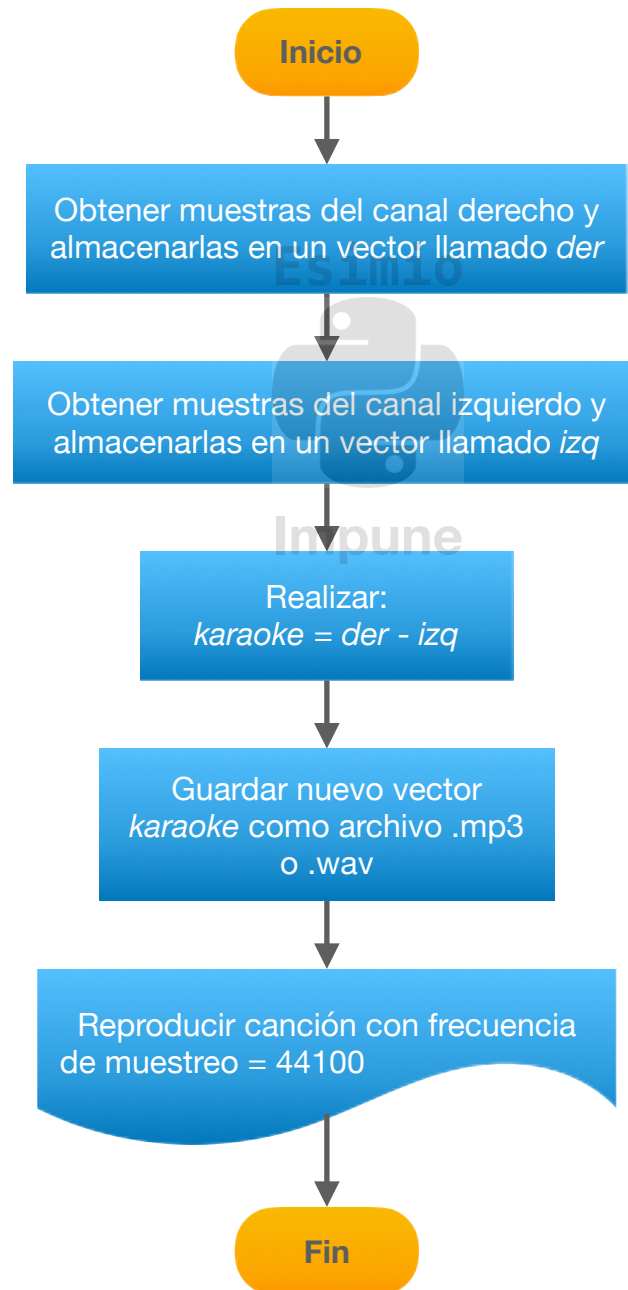


Figura 1

Diagrama de flujo para entender el algoritmo que produce un sistema karaoke a partir de una canción en formato .mp3 o .wav:



Para llevar a cabo las operaciones de las primeras dos cajas (i.e. obtener muestras de los canales derecho e izquierdo) se debe de contar con algún software o lenguaje de programación capaz de procesarlas (e.g. Matlab, Python etc).

Pruebas: Valores ficticios de muestras y un ejemplo práctico del algoritmo de procesamiento de las muestras del sistema karaoke:

Suponiendo un vector (obtenido del canal derecho):

$$V_{\text{derecho}} = \langle 0.98, 0.37, 0.23, 0.46 \rangle$$

Y otro obtenido del canal izquierdo:

$$V_{\text{izquierdo}} = \langle 0.01, 0.43, 0.12, 0.42 \rangle$$

El vector resultante de la resta es:

$$V_{\text{karaoke}} = V_{\text{derecho}} - V_{\text{izquierdo}}$$

n / Vector	0	1	2	3
Muestras canal derecho	0.98	0.37	0.23	0.46
Muestras canal izquierdo	0.01	0.43	0.12	0.42
Vector <i>karaoke</i> restando el izquierdo	0.97	-0.06	0.11	0.04
Vector <i>karaoke</i> restando el derecho	-0.97	0.06	-0.11	-0.04

Tabla 1

En componentes:

$$V_{\text{karaoke}} = \langle 0.97, -0.06, 0.11, 0.04 \rangle$$

Conclusiones: El algoritmo, si bien mejorable, es eficaz en la labor encomendada. A pesar de que inevitablemente, parte de la música se pierde en el proceso, la mayor parte de ésta logra pasar “la barrera”. Resulta evidente el hecho de que restar el vector izquierdo al derecho es preferible a la acción inversa debido a que las muestras en éste último son por lo general mayores que las de aquél.