

Informe de resultados: señales reales

Leandro Perren - Matias Meurzet

20 de octubre de 2015

1. Introducción

En este documento se presenta los resultados de obtener señales grabadas en ambientes reales cotidianos, tanto de ronquidos como de otros sonidos (aquí los definiremos como ruidos) que no son ronquidos y que pueden estar presentes en las señales obtenidas en entornos reales cotidianos, como por ejemplo:

1. Ambientales
2. Tos, estornudos
3. Vehículos
4. Voz
5. Música

El objetivo es poder evaluar características temporales (energía, zcr, rms, factor cresta) en las señales grabadas en entornos reales cotidianos. Una vez extraídas las características necesarias se podrá, entrenar y evaluar distintos modelos/clasificadores con la intención de construir una aplicación para dispositivos móviles capaz de detectar episodios de ronquidos y diferenciarlo de otras interferencias (tos, vehículos, voces, música, etc.) que pudieran acoplarse en la grabación del sueño de la persona.

2. Procedimiento

Para realizar las tareas antes mencionadas se procedió primeramente a grabar tanto ronquidos como ruidos por medio de los micrófonos de dispositivos móviles con sistema operativo Android (Tablet y celular) y utilizando la aplicación llamada «Easy Voice Recorder». Estas señales se grabaron con una frecuencia de 16000Hz, luego se utilizó el software Audacity para editar dichas señales extrayendo manualmente episodios de ronquidos como así también ruidos. En el caso de estos últimos se grabaron y editaron manualmente distintos posibles ruidos que podrían presentarse durante una grabación en ambientes normales. Una vez obtenidas estas señales se extrajeron las características temporales (energía,

zcr, rms, factor cresta, energía/zcr), tanto de los ronquidos como de los ruidos para poder obtener resultados y estadísticas que nos permitan evaluar las mismas como características representativas. Luego se generó una base de datos de características con las que se entrenarán los clasificadores.

2.1. Análisis de ronquidos

Durante el proceso de obtención de las señales (ronquidos) se llevaron a cabo distintas extracciones, las cuales fueron enumeradas como extraccion 1,2,3,etc, cada una de estas extracciones consiste de una grabación de una persona mientras duerme con el fin de captar en ella los ronquidos que pudieran producirse. Posteriormente cada una de estas extracciones fue analizada manualmente mediante el software Audacity y se extrajeron algunos segmentos de ronquidos presentes en la grabación creando distintos archivos llamados «ronquido_16000Hz_nro» todos de formato .wav los cuales contienen un episodio de ronquido cada uno.

1. Datos de la extracción 1 Esta primera extracción consiste de una grabación con una frecuencia de muestreo de 16000Hz de una persona de sexo masculino. De la misma se extrajeron 5 episodios los cuales fueron guardados en formato .wav y fueron nombrados como «ronquido_16000Hz_1», «ronquido_16000Hz_2», «ronquido_16000Hz_3», «ronquido_16000Hz_4» y «ronquido_16000Hz_5».
2. Datos de la extracción 2 Esta segunda extracción consiste de una grabación con una frecuencia de muestreo de 16000Hz de una persona de sexo masculino. De la misma se extrajeron 3 episodios los cuales fueron guardados en formato .wav y fueron nombrados como «ronquido_16000Hz_6», «ronquido_16000Hz_7» y «ronquido_16000Hz_8».

2.2. Análisis de ruidos

De igual manera a lo realizado en el caso de los ronquidos para la obtención de ruidos se llevaron a cabo extracciones que consistieron en grabar una serie de sonidos ambientales que no son ronquidos y que comúnmente se pueden encontrar en señales grabadas en entornos cotidianos. Se realizaron grabaciones a una persona hablando en una radio, una persona tosiendo y el sonido de un teléfono. De estas grabaciones se extrajeron fragmentos de dichos sonidos y se guardaron los mismos con los siguientes nombres, «ruido_16000Hz_habla», «ruido_16000Hz_tos», «ruido_16000Hz_telefono».

3. Extracción de características

Estos archivos creados cuyo contenido son ronquidos y otros sonidos fueron utilizados para extraerles características temporales con el fin determinar si las mismas son representativas de los ronquidos y que nos permitan diferenciarlos

de otros sonidos (voz, musica, tos, vehiculos, etc). Las características temporales extraídas en este trabajo son energía, zcr, rms, factor cresta, energía/zcr. Estas características fueron calculadas tomando para cada archivo ventanas deslizantes de 100ms cada una con un solapamiento del 50 %. A continuación se detalla la manera en que se calcula cada una de estas características y los resultados obtenidos para cada una de ellas.

3.1. Energía

El cálculo de la energía se realizó de acuerdo a la siguiente fórmula:

$$e_n = \sum_{i=1}^N x_i^2 / N \quad (1)$$

Donde e_n representa la energía del frame n correspondiente a una ventana de 100ms, x_i representa la amplitud de la muestra i-esima del frame n y N es la cantidad de muestras que tiene el frame en cuestión.

Como se puede observar en las figuras 1 y 2 se calculó la energía para los cinco ronquidos de la extracción 1 (ver Figura 1.) y para los 3 de la extracción 2 (ver Figura 2.), correspondientes ambos a una persona de sexo masculino. En la figura 3 se puede observar los gráficos de la energía para las señales de sonidos no deseados, las cuales corresponden a una persona tosiendo, un sonido de teléfono y por último el sonido de personas hablando. En todos estos gráficos se representó la energía para cada frame usando ventanas deslizante de 100ms con solapamiento de 50 %.

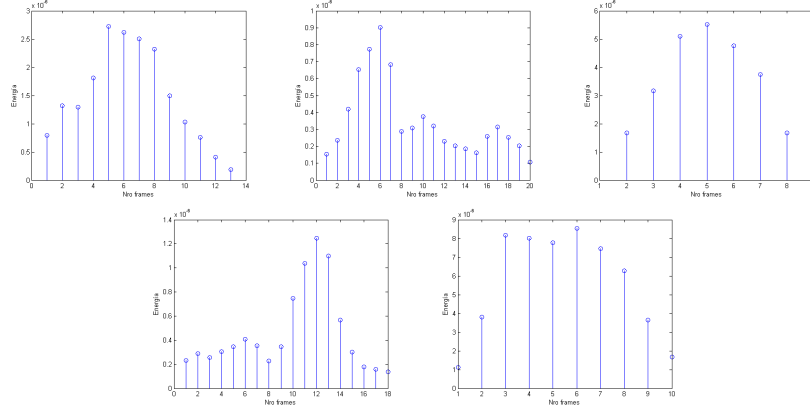


Figura 1: Energía de los ronquidos de la extracción 1

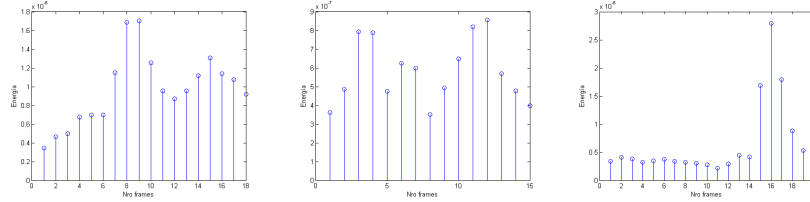


Figura 2: Energía de los ronquidos de la extracción 2

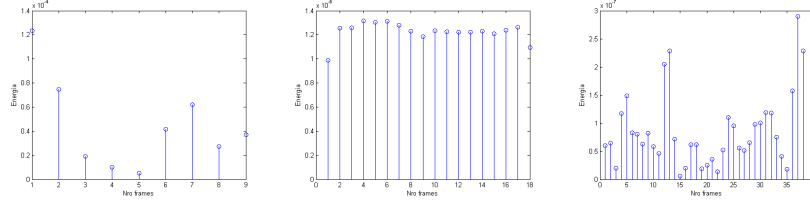


Figura 3: Energía de los ruidos: tos, telefono, habla

3.2. ZCR

La tasa de cruces por ceros (zero-crossing rate) es una medida de la cantidad de cambios de signos que sufre una señal, o lo que es igual el numero de veces que la señal toma el valor cero, y por lo tanto nos sirve de medida de la frecuencia que tiene la misma. El calculo del ZCR se realizó usando la siguiente formula:

$$zcr = \sum_{i=2}^N x_i * x_{i-1} \quad (2)$$

En las figuras 4 y 5 se encuentran graficados los resultados del calculo del ZCR para los cinco ronquidos de la extracción 1 (ver Figura 4.) y para los 3 de la extracción 2 (ver Figura 5.). En la figura 6 se pueden observar los gráficos

del ZCR para las señales de sonidos no deseados, las cuales corresponden a una persona tosiendo, un sonido de teléfono y por último el sonido de personas hablando. En todos estos gráficos se representó el ZCR para cada frame usando ventanas deslizante de 100ms con solapamiento de 50 %.

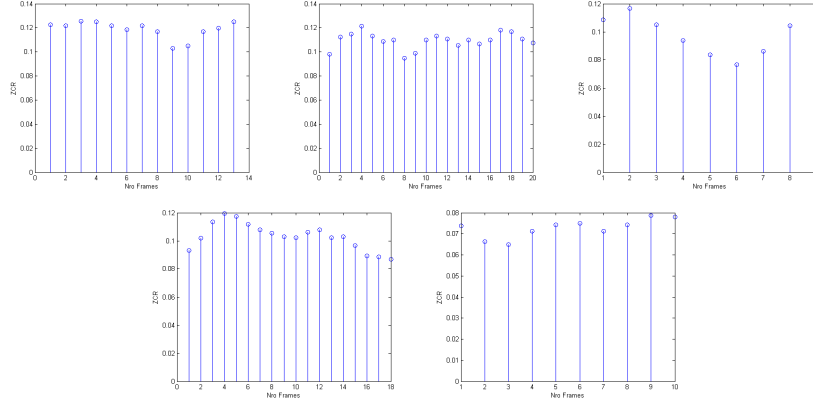


Figura 4: ZCR de los ronquidos de la extracción 1

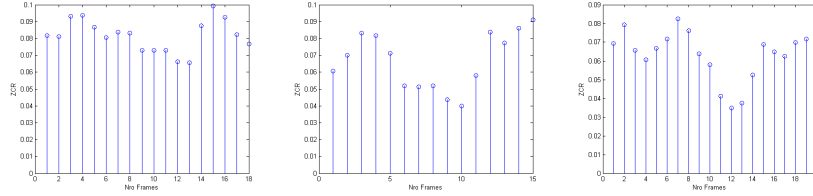


Figura 5: ZCR de los ronquidos de la extracción 2

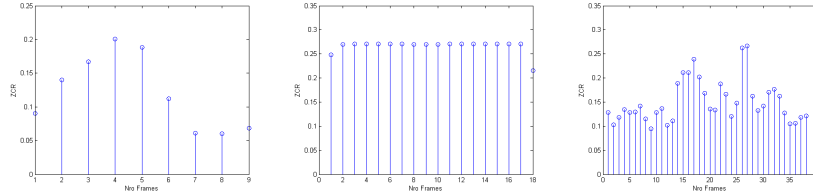


Figura 6: ZCR de los ruidos: tos, telefono, habla

3.3. Root Mean Square

El valor Root Mean Square se calcula tomando la raíz cuadrada de la media aritmética de los valores al cuadrado de cada muestra. Nos permite obtener un valor promedio de la señal sin tener en cuenta los efectos del signo de las muestras.

$$rms_n = \sqrt{\sum_{i=1}^N x_i^2 / N} \quad (3)$$

Donde rms_n representa el RMS del frame n correspondiente a una ventana de 100ms, x_i representa la amplitud de la muestra i-esima del frame n y N es la cantidad de muestras que tiene el frame en cuestión.

En la figura 7 y 8 se pueden observar los resultados del calculo del RMS para los ronquidos de las extracciones 1 y 2, y en la figura 9 podemos observar lo mismo para los tres ruidos.

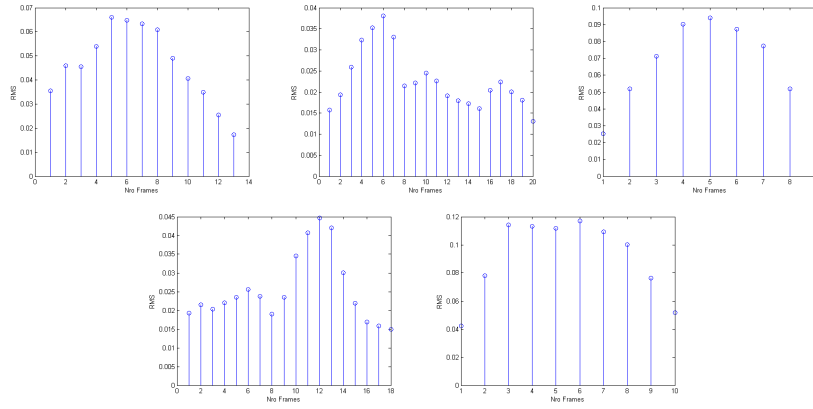


Figura 7: RMS de los ronquidos de la extracción 1

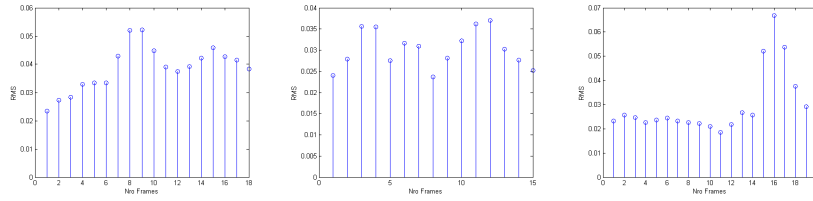


Figura 8: RMS de los ronquidos de la extracción 2

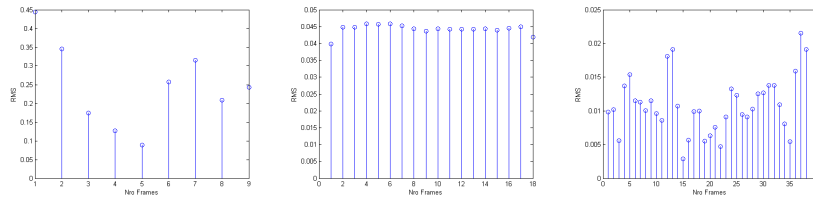


Figura 9: RMS de los ruidos: tos, telefono, habla

3.4. Factor cresta

El factor cresta es una medida obtenida haciendo el cociente entre el máximo valor absoluto de una señal dividido por su root mean square, dando un valor mayor en aquellas señales que tienen una serie de valores picos donde las magnitudes en esos valores tienden a ser mucho mayores que en el resto.

$$cresta_n = \max |x_n|/rms_n \quad (4)$$

Donde x_n representa cada muestra del frame n y rms_n es el root mean square del frame n.

En las figuras 10 y 11 se pueden observar los resultados del calculo del Factor Cresta para los ronquidos de las extracciones 1 y 2, y en la figura 12 podemos observar lo mismo para los tres ruidos.

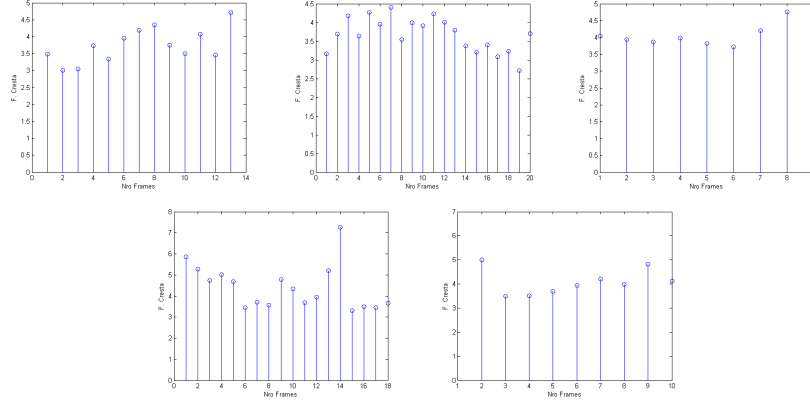


Figura 10: Factor cresta de los ronquidos de la extracción 1

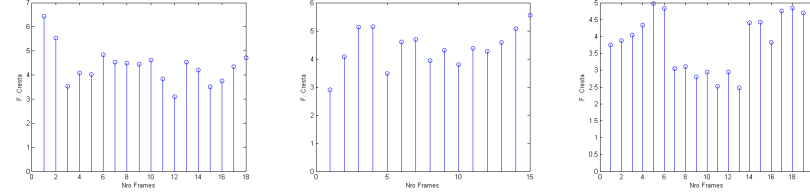


Figura 11: Factor cresta de los ronquidos de la extracción 2

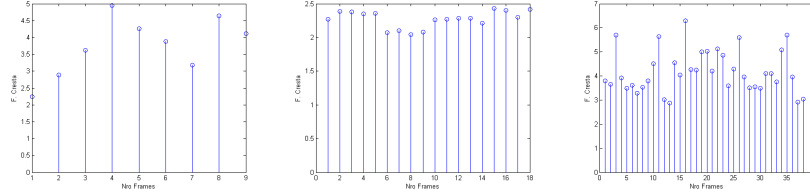


Figura 12: Factor cresta de los ruidos: tos, telefono, habla

3.5. Energía/ZCR

Esta medida es el resultado de calcular el cociente entre la energía de un frame y su ZCR. El uso de esta medida es muy comun en el análisis de voz para diferenciar sonidos sordos de sonoros, ya que los sonidos sordos se caracterizan por tener baja energía y un alto ZCR y lo inverso se da en los sonidos sonoros, por lo que este cociente refuerza este efecto.

$$EnergiaZCR_n = e_n / ZCR_n \quad (5)$$

Donde e_n y ZCR_n es la energía y el ZCR del frame n. En las figuras 13 y 14 se pueden observar los resultados del calculo de Energía/ZCR para los ronquidos de las extracciones 1 y 2, y en la figura 15 podemos observar lo mismo para los tres ruidos.

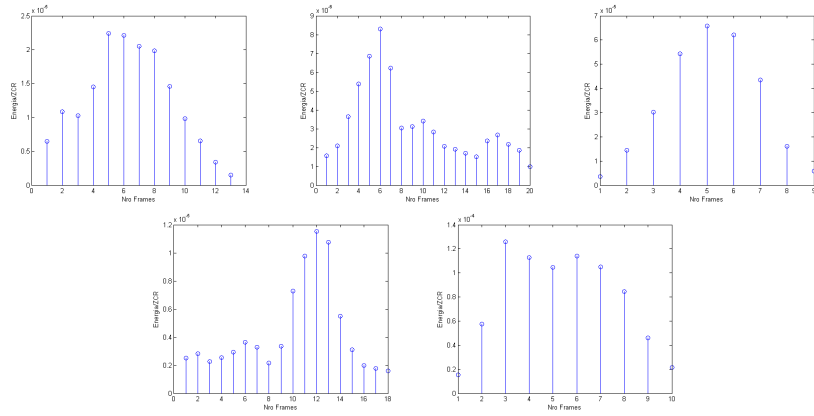


Figura 13: Energía/ZCR de los ronquidos de la extracción 1

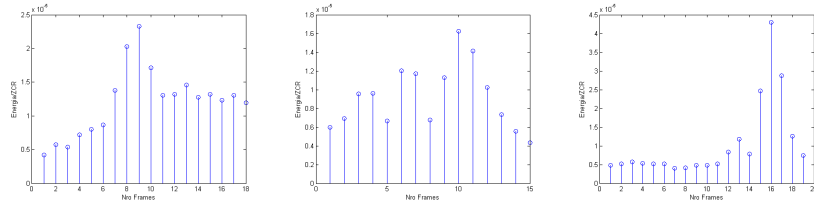


Figura 14: Energía/ZCR de los ronquidos de la extracción 2

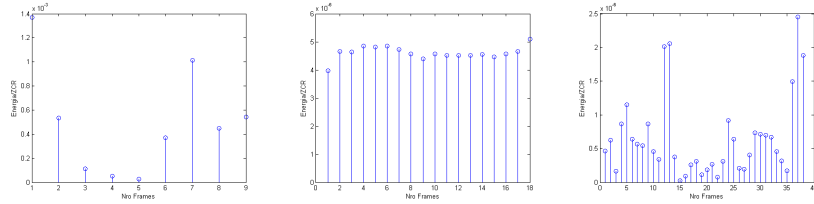


Figura 15: Energía/ZCR de los ruidos: tos, telefono, habla