



Escuela de  
Ingeniería y Arquitectura  
**Universidad** Zaragoza

30357

LABORATORIO DE SEÑAL Y COMUNICACIONES

---

# Estudio Previo Práctica 1

---

*Autores:*

Priscila Gómez Lizaga

Diego Cajal Orleans

*Nia:*

684061

620622

20 de febrero de 2018

## 1. Búsqueda bibliográfica DTW

A continuación se presentan algunos ejemplos de aplicaciones reales del DTW.

**Aplicación:** Clasificador de insectos voladores

**Descripción:** El objetivo es obtener el espectro del sonido del insecto que permite determinar el tipo de insecto se trata. Prevención sanitaria.

**Referencia:** Dynamic Time Warping Averaging of Time Series allows Faster and more Accurate Classification [2014]. Autores: F. Petitjean G. Forestier G.I. Webb A.E. Nicholson Y. Chen E. Keog.

**Aplicación:** Indexación de series temporales

**Descripción:** Mejora de la medida de la distancia para series temporales con respecto a la distancia euclídea.

**Referencia:** Exact indexing of dynamic time warping [2005]. Autores: Eamonn Keogh, Chotirat Ann Ratanamahatana.

**Aplicación:** Creación de índices de búsqueda

**Descripción:** Automatización de indexación de textos históricos con ruido para facilitar búsquedas.

**Referencia:** Word image matching using dynamic time warping [2003]. Autores: T.M. Rath, R. Manmatha

## 2. Distancia de Itakura-LPC

Se tiene que demostrar que la energía del error de predicción de un segmento  $x_0$  para un modelo de predicción lineal de  $P$  coeficientes es:

$$E_{x_0x_1} = a_1^T R_{x_0x_0} a_1$$

Para ello partimos de la expresión del error de predicción

$$e[n] = x[n] - \sum_{i=1}^P a_i x[n-i] \quad (1)$$

donde  $a_i$  son los coeficientes del filtro. Si modificamos el vector de coeficientes de forma que quede

$$a_1 = [1, -a_1, -a_2, \dots, -a_P]$$

podemos expresar el error de forma compacta

$$e[n] = \sum_{i=0}^P a_{1i} x[n-i]$$

que expresado en forma matricial queda

$$e = a_1^T X$$

La potencia del error de un LPC es la esperanza del error cuadrático

$$e^2 = (a_1^T X)(a_1^T X)^T$$

$$e^2 = a_1^T X X^T a_1$$

$$E[e^2] = E[a_1^T X X^T a_1]$$

$$E[e^2] = E_{x_0 x_1} = a_1^T E[X X^T] a_1$$

$$E_{x_0 x_1} = a_1^T R_{x_0 x_0} a_1$$

### 3. Respuesta frecuencial de un filtro IIR

La respuesta frecuencial de un filtro IIR es  $1/fft$  de los coeficientes del filtro. La transformada de Fourier hace un barrido sobre la circunferencia unidad en el espacio  $Z$ , obteniendo información de ceros. Como los coeficientes del filtro IIR están en el denominador (polos), se debe hacer el inverso multiplicativo.

## 4. Búsqueda bibliográfica Mel

De la misma manera que se ha hecho en el primer apartado, se presentan aplicaciones reales en las que se utiliza la escala perceptual auditiva Mel y/o los MFCC (Mel Frequency Cepstral Coefficients).

**Aplicación:** Reconocimiento de voz

**Descripción:** Se utilizan los MFCCs como método de extracción de características y DTW para comparar patrones.

**Referencia:** Voice Recognition Algorithms using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques [2010]. Autores: Lindasalwa Muda, Mumtaj Begam, I. Elamvazuthi

**Aplicación:** Detector automático de fases tempranas en el Parkinson.

**Descripción:** El objetivo es detectar la enfermedad de Parkinson a través del reconocimiento del habla, los análisis se llevan a cabo mediante parámetros a corto plazo y más precisamente mediante MFCC combinados con Modelos Mixtos Gaussianos. En el artículo se presenta el análisis durante cuatro tareas: vocales sostenidas, repeticiones de sílabas rápidas, libertad de expresión y lectura, donde se adapta la metodología clásica.

**Referencia:** Automatic detection of early stages of Parkinson's disease through acoustic voice analysis with mel-frequency cepstral coefficients [2017]. Laetitia Jeancolas, Habib Benali, Badr-Eddine Benkelfat, Graziella Mangone, Jean-Christophe Corvol, Marie Vidailhet, Stephane Lehericy, Dijana Petrovska-Delacrétaz.<sup>5</sup>

**Aplicación:** Identificador de sarcasmo

**Descripción:** Debido a que cuando se habla de una forma sarcástica se produce un cambio en el tono, pitch... se propone en primer lugar extraer el audio de interés que pasará, a continuación, por un módulo que se encargará de reconocer el sarcasmo.

**Referencia:** Understanding sarcasm in speech using mel-frequency cepstral coefficient [2017]. Abhinav Mathur, Vikas Saxena, Sandeep K Singh

**Aplicación:** Clasificador de llantos infantiles

**Descripción:** Reconocer el llanto de un niño a veces resulta complicado por lo que se ha desarrollado un sistema de clasificación utilizando MFCC y BNN (Backpropa-

gation Neural Network). Se clasifican en 3 clases: hambriento, cansado o molesto.

**Referencia:** Infant's Cry Sound Classification using Mel-Frequency Cepstrum Coefficients Feature Extraction and Backpropagation Neural Network [2016]. Yesy Diah Rosita y Hartarto Junaedi.

## 5. Cálculo eficiente de distancias

Código en *Matlab* para la creación de una matriz de distancias entre dos vectores o matrices de forma eficiente.

```
1 function [out2] = distance( x,y )
2
3 xlen = length(x);
4 ylen = length(y);
5 auxy = zeros(ylen);
6 mx_y = zeros(ylen,xlen);
7 mx_x = zeros(ylen,xlen);
8
9 %auxy = y(end:-1:1); Si la queremos con el minimo abajo a la izquierda
10 auxy = y;
11 mx_y = repmat(auxy', 1, xlen);
12 mx_x = repmat(x, ylen, 1);
13 out1 = mx_x - mx_y;
14 out2 = out1.^2;
15
16 end
```

```
1 function [D] = distance_matrix( X,Y )
2
3 [D1,N1]= size(X);
4 [D2,N2]= size(Y);
5 D = zeros(N1,N2);
6
7 for i = 1:N1
8     for j = 1:N2
9         x = X(:,i);
10        y = Y(:,j);
11        D(i,j) = (x - y)'*(x - y);
12    end
13 end
14
15 end
```