Ampliación de la Práctica 4

David Linde

Pau Guàrdia

Índice

1. INTRODUCCIÓN

El objetivo principal de esta memoria es presentar los resultados obtenidos en la mejora del sistema de 'Speak Recognition' ya creado anteriormente en el desarrollo de la Práctica 4.

En esta memoria se plasma el resultado del mejor sistema encontrado que nos permite mejorar con creces los resultados obtenidos en la primera parte de la práctica.

2. PRESENTACIÓN DE LOS RESULTADOS

En el antiguo sistema, nuestros resultados eran los siguientes:

LP = 8 coeficientes; LPCC = 8 coeficientes , 12 coef.cepstrales ; MFCC = 16 coeficientes

	LP	**LPCC**	**MFCC**
Tasa de error (%)	11.08	3.82	1.66

En los cuales el mejor 'Cost Detection' es para el MFCC con un valor 25,1.

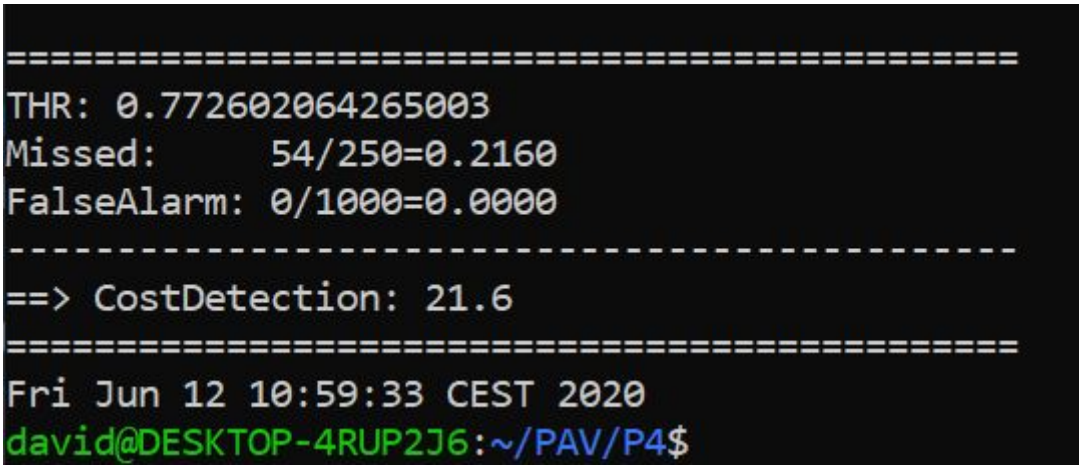
Hemos realizado una mejora a la hora de inicializar las GMM en vez de iniciar en formato `gmm.random_init(data,nmix);`, hemos realizado una inicialización `gmm.vq_lbg(data, nmix,em_iterations,em_threshold,verbose);` la cual es una cuantificación vectorial (VQ).

Tanto para el sistema antiguo como para el nuevo hemos usado lo siguiente para el GMM : Threshold = 0.00001 ; Iteraciones = 40 ; Gaussianas = 40.

Una vez realizado el cambio a inicialización VQ obtenemos estos resultados.

	LP	**LPCC**	**MFCC**
Tasa de error (%)	9.94	2.55	1.15

El mejor 'Cost Detection' es el logrado con el MFCC logrando un 21.6 seguido del LPCC con 32 y después con LP 85.2

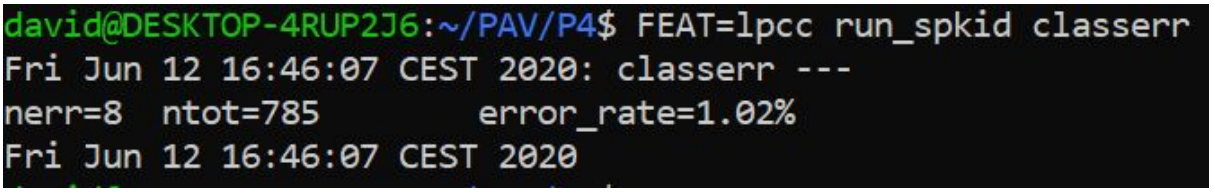


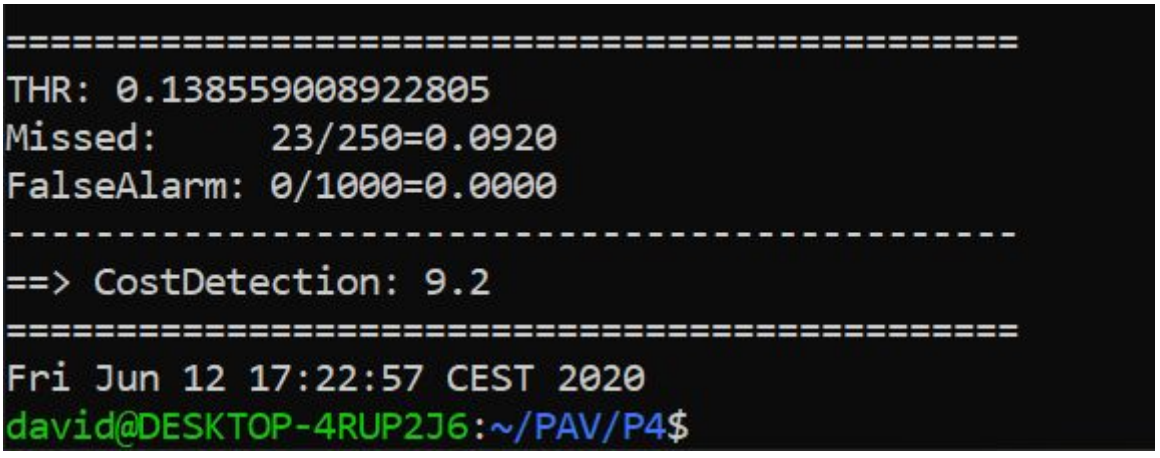
Como se puede observar en los resultados, hemos logrado mejorar los resultados obtenidos.

Una vez obtenidos estos resultados hemos buscado una forma de mejorar estos, hemos modificado el número de coeficientes del LPCC y también del LPCC cepstrum.

En esta nueva versión, usamos LPCC = 16 coeficientes y 20 coef.cepstrales.

Hemos obtenido estos resultados :





Todos los resultados correspondientes a la evaluación están en la carpeta work de la 'branch' Guardia-Linde. https://github.com/davilin98/P4/tree/Guardia-Linde

3. CONCLUSIÓN

Tal y como se muestra en el apartado anterior, los resultados del sistema se han mejorado notoriamente, donde pasamos de un coste de detección igual a 12 a ahora tener un coste de 9. Por lo que hace el porcentaje de error, este también mejora pasando de 1.66% a 1.02% es decir obteniendo un resultado 0.64% mejor.

Estos son los mejores resultados que hemos obtenido, con una tasa de error muy baja y con un coste de detección muy bajo en comparación con los otros métodos que hemos probado.