Universidad de Chile
Facultad de Ciencias Físicas y Matemáticas
Departamento de Ciencias de la Computación

Control 2

Integrantes: Adolfo Rojas V.
Profesor: Juan Manuel Barrios
Ayudantes: Andrés Calderón
Martina Navarro
Scarleth Betancurt
Sebastián Sáez

Fecha de realización: 23 de junio de 2025
Fecha de entrega: 23 de junio de 2025
Santiago de Chile

1. Semana 5

1) a) El audio dura 229 segundos con profundidad de 16 bits (2 bytes) y un sample rate de 22050 por lo que el tamaño de A.raw es de $229*22050*2 = 10098900$ bytes, ahora por el sampling theorem tenemos que la frecuencia máxima es $22050/2 = 11025$

   b) Peor, para consumo humano 32 bits of depth (precisión del volumen) no aporta, por otro lado el sample rate es tan bajo que se dejan de escuchar los agudos $5512,5Hz$

2) a) Calculamos los samples por ventana $48000 * 0.04 = 1920$ luego como solo estamos interesados en los coeficientes reales podemos ignorar los complementos por simetría con lo que nos quedamos con un total de $1920/2 = 960$ coeficientes

   b) La máxima que se puede llegar a escuchar es de $40/0.04 = 1$kHz

3) a) La escala agrupa las frecuencias de forma parecida a cómo las escuchamos nosotros, así se evita que cambios en el sonido que no notamos afecten la comparación

   b) Buena idea porque este coeficiente es solo la energía total del audio, que puede verse afectada por cambios de volumen/ruido, sin cambiar realmente el contenido

Tabla 1: Matriz de distancias DTW A-M

2	12	14	17
13	5	16	22
21	7	13	20
22	16	10	12
25	23	13	12

4) a) Costo total = 12

Tabla 2: Matriz de distancias DTW B-M

2	10	16	17
4	14	12	17
17	7	20	22
19	17	9	12
21	25	13	10
23	29	19	11

   b) Costo total = 11

   c) El ganador es B puesto que a DTW no le importan las dimensiones/ventanas sino que la similitud y B consigue la mejor alineación (menor coste) respecto a M

2. Semana 6

1) De los 3 canales de colores asumimos que cada uno tiene una profundidad de un byte (o sea que 4:2:0 ocupa 1.5 bytes por píxel y 4:4:4 ocupa 3, aunque si se quisieran simplificar los cálculos también se podría asumir una profundidad por canal de un tercio de byte)

   a) $pixels = 1920 * 1080 = 2073600$ luego se necesitan
   $2073600 \text{ pixels} * 1.5\frac{\text{bytes}}{\text{pixels}} * 60\frac{\text{frames}}{\text{s}} * 1\text{s} = 186{,}624{,}000 \text{ bytes} \approx 187mb$

   b) $1280 * 720 * 3 * 24 = 66{,}355{,}200$ bytes $\approx 66mb$

   c) $3180 * 2160 * 1.5 * 30 = 309{,}096{,}000$ bytes $\approx 309mb$

   d) $1920 * 1440 * 3 * 30 = 248{,}832{,}000$ bytes $\approx 249mb$

2) a) Puesto que los frames de tipo B dependen de P al decodificar habrán errores con los cuadros 5, 7 y 8

   b) Todos los frames que lo usan de referencia hasta el próximo intra-coded frame se verán afectados con lo que habrán errores desde el cuadro 10 hasta el 15

   c) No hay frame que dependa de este tipo así que ningún otro cuadro se ve afectado

   d) Igual al item a) se ven afectados los B-frames anteriores y posteriores a este cuadro, también el P-frame que lo usaba de referencia, es decir los cuadros desde el 17 hasta el 22

3. Semana 8

1) a) Calculamos el average precision por cada sistema

   - Sistema 1: los índices donde se encuentran las imágenes relevantes para Q_1 son $\{2,4,6,8\}$ por lo que $AP(Q_1) = (\frac{1}{2} + \frac{2}{4} + \frac{3}{6} + \frac{4}{8})/4 = 0.5$ y para Q_2 son $\{3,6,8\}$ por lo que $AP(Q_2) = (\frac{1}{3} + \frac{2}{6} + \frac{3}{8})/3 \approx 0.347$. Finalmente

   $$MAP_1 = \frac{0.5 + 0.347}{2} \approx 0.424$$

   - Sistema 2: los índices donde se encuentran las imágenes relevantes para Q_1 son $\{1,2,12\}$ por lo que $AP(Q_1) = (\frac{1}{1}+\frac{2}{2}+\frac{3}{12}+0)/4 = 0.5625$ y para Q_2 son $\{1,10\}$ por lo que $AP(Q_2) = (\frac{1}{1} + \frac{2}{10} + 0)/3 = 0.4$. Finalmente

   $$MAP_2 = \frac{0.5625 + 0.4}{2} \approx 0.481$$

   b) - Sistema 1:
      - Mean Reciprocal Rank: para Q_1 la primera imágen se encuentra en 2do lugar por lo que se tiene $RR_{Q1} = \frac{1}{2}$, luego para Q_2, $RR_{Q2} = \frac{1}{3}$. Finalmente

      $$MRR_1 = (\frac{1}{2} + \frac{1}{3})/2 \approx 0.42$$

- Precision@1: para Q_1 se tiene precision@$1_{Q1} = 0$, luego para Q_2, precision@$1_{Q2} = 0$. Finalmente
$$\text{precision@}1_1 = (0+0)/2 = 0$$
- Recall@12: para Q_1 se tiene recall@$12_{Q1} = \frac{4}{4}$, luego para Q_2, recall@$12_{Q2} = \frac{3}{3}$. Finalmente
$$\text{recall@}12_1 = (1+1)/2 = 1$$
- R-precision: para Q_1 se encuentran 2 imágenes en el top-$|Q_1|$ con $|Q_1| = 4$ por lo que se tiene R-precision$_{Q1} = \frac{2}{4}$, luego para Q_2, R-precision$_{Q2} = \frac{1}{3}$. Finalmente
$$\text{R-precision}_1 = (\frac{2}{4} + \frac{1}{3})/2 \approx 0.42$$

– Sistema 2:

- Mean Reciprocal Rank: para Q_1 se tiene $RR_{Q1} = 1$, luego para Q_2, $RR_{Q2} = 1$. Finalmente
$$MRR_2 = (1+1)/2 = 1$$
- Precision@1: para Q_1 se tiene precision@$1_{Q1} = 1$, luego para Q_2, precision@$1_{Q2} = 1$. Finalmente
$$\text{precision@}1_2 = (1+1)/2 = 1$$
- Recall@12: para Q_1 se tiene recall@$12_{Q1} = \frac{3}{4}$, luego para Q_2, recall@$12_{Q2} = \frac{2}{3}$. Finalmente
$$\text{recall@}12_2 = (\frac{3}{4} + \frac{2}{3})/2 \approx 0.708$$
- R-precision: para Q_1 se tiene R-precision$_{Q1} = \frac{2}{4}$, luego para Q_2, R-precision$_{Q2} = \frac{1}{3}$. Finalmente
$$\text{R-precision}_2 = (\frac{2}{4} + \frac{1}{3})/2 \approx 0.42$$

c) Podríamos fundamentar erróneamente nuestra elección en base a la tabla 3 pero dado que solo nos interesa recuperar la mayor cantidad de imágenes similares en la primera pág que soporta 8 de estas podemos calcular rápidamente recall@8, de aquí fácilmente podemos concluir que el **sistema 1** es nuestra mejor opción

Tabla 3: Tabla resumen de métricas por sistema

	Sistema 1	Sistema 2
MAP	0.424	0.481
MRR	0.42	1
precision@1	0	1
recall@12	1	0.708
R-precision	0.42	0.42

2) a) El MAP de cada algoritmo ya está dado (corresponde a μ_i con $i \in \{1, 2, 3\}$)

- Algoritmo 1: con un (α) p-value de 5 % tenemos que $z \approx 1.96$ por lo que $IC_1 = 0.671 \pm 1.96 \cdot \frac{0.216}{\sqrt{20}} = 0.671 \pm 0.0947$. Finalmente

$$IC_1 = [0.576, \quad 0.766]$$

- Algoritmo 2: $IC_2 = 0.811 \pm 1.96 \cdot \frac{0.165}{\sqrt{20}} = 0.811 \pm 0.0723$. Finalmente

$$IC_2 = [0.739, \quad 0.883]$$

- Algoritmo 3: $IC_3 = 0.744 \pm 1.96 \cdot \frac{0.017}{\sqrt{20}} = 0.744 \pm 0.0075$. Finalmente

$$IC_3 = [0.737, \quad 0.752]$$

De aquí podemos concluir que el algoritmo 2 a pesar de no ser tan consistente como el algoritmo 3, sigue siendo el con mayor desempeño, por otro lado el peor es el algoritmo 1

b) - Algoritmo 1: $IC_1 = 0.671 \pm \frac{0.0947}{\sqrt{10}}$. Finalmente

$$IC_1 = [0.641, \quad 0.701]$$

- Algoritmo 2: $IC_2 = 0.811 \pm \frac{0.0723}{\sqrt{10}}$. Finalmente

$$IC_2 = [0.788, \quad 0.834]$$

- Algoritmo 3: $IC_3 = 0.744 \pm \frac{0.0075}{\sqrt{10}}$. Finalmente

$$IC_3 = [0.742, \quad 0.746]$$

Simplemente se estrecharon más los intervalos de confianza pero el ranking de desempeño entre algoritmos sigue siendo el mismo

3) a) En base a la tabla 4 podemos decir que el que más se parece a la realidad es el servicio 2 (o mejor dicho, el que menos se aleja de la realidad)

Tabla 4: Coeficiente de Spearman

	Real	Servicio 1	Servicio 2
España	1	3	4
Chile	2	4	1
Suiza	3	1	3
Honduras	4	2	2

$Diff_1^2$	$Diff_2^2$
4	9
4	1
4	0
4	4

Suma	16	14
Cf. Spearman	-0,6	-0,4

b) Para ambos servicios el total de pares posibles es 6

- Servicio 1:
  - (España, Chile): concuerda
  - (España, Suiza): no
  - (España, Honduras): no
  - (Chile, Suiza): no
  - (Chile, Honduras): no
  - (Suiza, Honduras): concuerda

  Finalmente $\tau = \frac{2-4}{6} = -0.33$
- Servicio 2:
  - (España, Chile): no
  - (España, Suiza): no
  - (España, Honduras): no
  - (Chile, Suiza): concuerda
  - (Chile, Honduras): concuerda
  - (Suiza, Honduras): no

  Finalmente $\tau = \frac{2-4}{6} = -0.33$

En base a esta métrica concluimos que somos indiferentes entre escoger qué servicio se acerca más a la respuesta real