

Trabajo Práctico 0: El algoritmo de umbralización de Kittler

Ph. D. Saúl Calderón Ramírez
Instituto Tecnológico de Costa Rica,
Escuela de Computación
PAttern Recongition and MACHine Learning Group (PARMA-Group)

4 de abril de 2024

Fecha de entrega: Domingo 21 de Abril.

Entrega: Un archivo .zip con el código fuente LaTeX o Lyx, el pdf, y un script en jupyter, debidamente documentado. El script en jupyter debe estar escrito usando pytorch. A través del TEC-digital.

Modo de trabajo: Grupos de 3 personas.

En el presente trabajo práctico se introducirá el problema de la clasificación a través del estudio del concepto de máxima verosimilitud. Se realizará un repaso de la teoría básica relacionada con los fenómenos aleatorios con distribución Gaussiana, para facilitar el análisis de la función de verosimilitud. Posteriormente se visitará el problema de la segmentación de imágenes desde un enfoque de máxima verosimilitud, donde se desarrollará el algoritmo de Kittler [1]. El estudiante implementará tal algoritmo y analizará los resultados respecto a los planteos teóricos introducidos previamente.

1. Implementación del algoritmo de Kittler

1. Implemente el algoritmo de Kittler, y realice una prueba con la imagen de entrada provista, aplicando posteriormente el umbral óptimo obtenido.
 - a) **(20 puntos)** Implemente una función *calcular_momentos_estadisticos*(T , p) la cual reciba un umbral candidato T y una función de densidad p , y retorne todos los parametros de la función. Comente su implementación con detalle en este informe.
 - 1) Diseñe al menos 2 pruebas unitarias donde verifique el funcionamiento correcto. Detalle en este documento el diseño de tales pruebas y los resultados, indicando si son los esperados.
 - b) **(20 puntos)** Implemente la función *calcular_costo_J*(T) la cual calcule el costo del umbral candidato T . Comente su implementación con detalle en este informe.

- 1) Diseñe al menos 2 pruebas unitarias donde verifique el funcionamiento correcto. Detalle en este documento el diseño de tales pruebas y los resultados, indicando si son los esperados.
- c) **(20 puntos)** Basado en ambas funciones, implemente la función *calcular_T_optimo_Kittler(Imagen)* la cual retorne el T óptimo para umbralizar la imagen recibida, además de la imagen umbralizada.
- d) Aplique el algoritmo de Kittler en la imagen *cuadro1_005.bmp*, provista.
 - 1) Grafique el histograma normalizado de la imagen de entrada provista.
 - 2) Grafique la función $J(T)$, y documente el valor $T = \tau$ que logra el valor mínimo de $J(T)$, junto con las medias y varianzas de las dos funciones Gaussianas superpuestas. Son coherentes tales valores con el histograma graficado en el punto anterior?

a' El valor óptimo en el caso de esta imagen debe ser cercano a $\tau = 168$, con $\mu_1 = 149,45$, $\mu_2 = 219,49$ $\sigma_1^2 = 15,36$ y $\sigma_2^2 = 10,05$.
 - 3) Lograría el umbral óptimo τ obtenido umbralizar satisfactoriamente la imagen de prueba? Umbralice la imagen de entrada provista y muestre los resultados.

a' Asigne con valor de 255 los pixeles del cuadrado (clase *foreground*), y 0 los del fondo (clase *background*).
- e) **(25 puntos)** Pruebe la implementación del algoritmo de Kittler para detectar la actividad de voz humana en un audio, usando el audio de prueba provisto *contaminated_audio.wav*.
 - 1) Grafique el histograma del audio leído, y argumente si es apropiado usar el algoritmo de Kittler o no.
 - 2) Grafique la función $J(T)$, y documente el valor $T = \tau$ que logra el valor mínimo de $J(T)$, junto con las medias y varianzas de las dos funciones Gaussianas superpuestas. Son coherentes tales valores con el histograma graficado en el punto anterior?
 - 3) Lograría el umbral óptimo τ obtenido partir satisfactoriamente el sonido de prueba en 3 segmentos (2 de actividades de audio y 1 segmento sin actividad)? Umbralice el audio provisto y muestre los resultados.
2. **(15 puntos)** La distancia de Bhattacharyya compara dos funciones de densidad de probabilidad $p(x)$ y $q(x)$:

$$D_{JS}(p, q) = -\ln \left(\sum_{x \in X} \sqrt{p(x)q(x)} \right)$$

- a) Implemente la función *calcular_bhattacharyya_distance(p,q)*, para comparar las funciones de densidad estimada con el ajuste del modelo

mixto Gaussiano con Kittler $p(x)$ y la aproximación de la densidad con el histograma de los datos $q(x)$ para las dos pruebas realizadas con la imagen y el audio.

- b) Explique la relación entre la distancia de Bhattacharyya y el proceso de estimación de los parámetros óptimos implementado en el algoritmo de Kittler ¿Que sucede cuando la distancia de Bhattacharyya entre el histograma de los datos y el modelo estimado crece o decrece ?
3. **(10 puntos extra)** Calcule el umbral óptimo con el algoritmo de Kittler, y umbralice la imagen de *trackedCell15.tif* provista documentando los resultados. Muestre la imagen umbralizada y el histograma de la misma.
- a) Usando la imagen, su histograma, y la matriz de confusión para la clase *foreground*, explique el porqué del resultado obtenido.
 - b) Como modificaría el algoritmo de Kittler para mejorar el resultado de la umbralización? Puede usar recursos bibliográficos externos.

Referencias

- [1] Josef Kittler and John Illingworth. Minimum error thresholding. *Pattern recognition*, 19(1):41–47, 1986.