

COMPARACIÓN DE UMBRALIZADORES GLOBALES PARA EL BINARIZADO EN IMÁGENES DE TEXTO PLANO

Ana Karen Angeles Gómez

Elimagdeth Anayatzin Rdríguez Carmona

Alvaro Anzueto Ríos

Instituto Politécnico Nacional

karenina.angeles@gmail.com

erodriguezc1003@gmail.com

alvaro.anzueto.rios@gmail.com

RESUMEN

En este trabajo se implementa y compara el desempeño de tres diferentes métodos de umbralización para imágenes. Siendo aplicados en imágenes de texto plano para obtener la segmentación y lograr separar el fondo del texto

Introducción:

El umbralizado en las imágenes es una técnica frecuentemente utilizada en el procesamiento de imágenes, las imágenes resultantes de este proceso pueden ser empleadas para el reconocimiento o extracción de características que ayuden a reconocer objetos y/o caracteres. En este trabajo se presentan tres métodos de umbralización, Método de Kittler, Método de Yen y Método de Otsu, y sus resultados son comparados para determinar el de mejor desempeño para el reconocimiento de texto.

Desarrollo:

Para la obtención de las imágenes se utiliza una cámara integrada en un celular, con una resolución de 1895 x 1080 pixeles; a una distancia de 19 cm se alcanza una imagen de 65mm de ancho x 130mm de alto. Ya que las imágenes obtenidas directamente de la cámara se encuentran en formato RGB, que significa que cada pixel en la imagen está compuesto por una combinación de tres valores, correspondientes al Rojo, Verde y Azul (RGB por sus siglas en inglés Red, Green y Blue), la unión de estos valores nos dará la tonalidad de color percibido, es necesario hacer una transformación de estos valores a escala de grises, por lo cual se le asignará, a cada pixel, un valor entre el 0 que corresponde al negro y el 255 que corresponde al blanco, dando así una gama de 256 tonalidades. El método utilizado para la

conversión de RGB a escala de grises es el incluido en la librería de Matlab® con la función `rgb2gray` que desarrolla la ecuación 1, para la asignación de valores en pixeles. Para poder observar como se comportan las imágenes con ésta única capa de pixeles en escala de grises se obtiene un histograma que es el número de ocurrencias de cada nivel de gris.

$$\text{Gris} = 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B \dots 1$$

El método de Kittler [1] se basa en calcular una función de densidad de probabilidad para dos zonas, el fondo (background) y el objeto (foreground), esto se calcula en la ecuación 2, y una ecuación de criterio dada en la ecuación 3.

$$p_f(t) = \frac{1}{\sigma_f \sqrt{2\pi}} e^{-\frac{(t-\mu_f)^2}{2\sigma_f^2}} \quad p_o(t) = \frac{1}{\sigma_b \sqrt{2\pi}} e^{-\frac{(t-\mu_b)^2}{2\sigma_b^2}} \dots 2$$

Función criterio:

$$J(t) = 1 + 2(P_f(t) \log \sigma_f(t) + P_b \log \sigma_b(t)) - 2(P_f(t) \log P_f(t) + P_b \log P_b(t))$$

$$T^* = \text{Min}\{J(t)\} \dots 3$$

El Método de Yen [2] usa la entropía como cálculo de variabilidad el cual indica: un valor alto de entropía dispersión alta en los pixeles, y una entropía baja agrupación de pixeles. Las ecuaciones que lo definen son 4 y 5, foreground y background respectivamente.

$$C_f(t) = -\ln\left(\sum_{g=0}^t \left(\frac{p(g)}{P_t}\right)^2\right) \dots 4 \quad C_b(t) = -\ln\left(\sum_{g=0}^t \left(\frac{p(g)}{P_t}\right)^2\right) \dots 5$$

El Método de Otsu [3] realiza el cálculo de varianza acumulada (ecuación 13), maximizando esta varianza (ecuación 14) se puede determinar el nivel de separación entre dos clases obteniendo así la función de criterio (ecuación 6). Las ecuaciones empleadas en el método son las ecuaciones 6 a 14.

$$g(x, y) = \begin{cases} 1 & \leftrightarrow f(x, y) > T \\ 0 & \leftrightarrow f(x, y) \leq T \end{cases} \dots 6$$

$$C_1: \frac{p_1}{w_1(t)}, \dots, \frac{p_t}{w_1(t)} \dots 7 \quad C_2: \frac{p_{t+1}}{w_2(t)}, \frac{p_{t+2}}{w_2(t)}, \dots, \frac{p_L}{w_2(t)} \dots 8$$

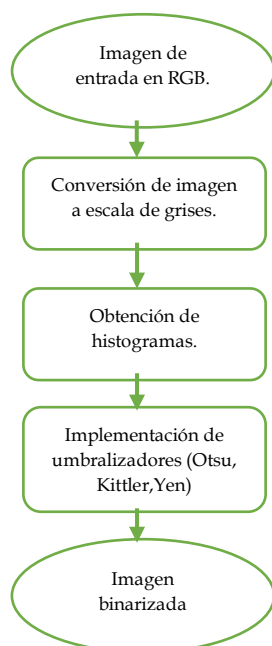
$$\omega_1(t) = \sum_{i=1}^t \frac{i \cdot p_i}{\omega_1(t)} \dots 9 \quad \omega_2(t) = \sum_{i=t+1}^L \frac{i \cdot p_i}{\omega_2(t)} \dots 10$$

$$\mu_1 = \sum_{i=1}^t \frac{i \cdot p_i}{\omega_1(t)} \dots 11 \quad \mu_2 = \sum_{i=t+1}^L \frac{i \cdot p_i}{\omega_2(t)} \dots 12$$

$$\sigma_B^2 = \omega_1(\mu_1 - \mu_T)^2 + \omega_2(\mu_2 - \mu_T)^2 \dots 13$$

$$t^* = \max_t \{\sigma_B^2(t)\} \quad 1 \leq t \leq L \dots 14$$

Diagrama de flujo del proceso de cálculo de umbral



Como primer bloque, en el diagrama de flujo, se tiene la adquisición de la imagen de entrada en formato RGB. Esta imagen es transformada a escala de grises empleando la ecuación 1, este proceso es realizado en el segundo bloque. El tercer bloque se encarga de obtener la frecuencia con que se presenta en la imagen cada uno de los niveles de grises esto equivale a obtener el histograma.

Con el histograma obtenido se pueden aplicar las ecuaciones de los tres métodos mencionados, cada uno de ellos arroja el parámetro de umbral que corresponde a un nivel de gris. Los niveles de grises que superan este valor (umbral) son reasignados con un valor binario correspondiente a uno o cero. El valor uno corresponde a una tonalidad blanca, el valor cero corresponde a una tonalidad negra.

Resultados:

Para realizar las pruebas a los tres métodos se tomaron dos fotografías de una página, que contiene únicamente líneas de texto, la primera fotografía cubre la mitad superior de la página y la segunda fotografía abarca la mitad inferior. Los tres métodos se aplicaron de manera global para obtener el valor de umbral. Las imágenes de partida se encuentran a color por lo que es necesario transformarlas.

A continuación, en las imágenes 1 y 3, se muestran las fotografías iniciales en formato RGB (superior e inferior de la hoja), las imágenes 2 y 4 muestran el resultado de aplicar la conversión a escala de grises.

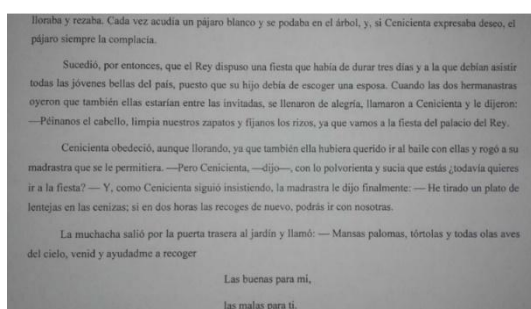


Imagen 1. Imagen RGB primer mitad de la página.

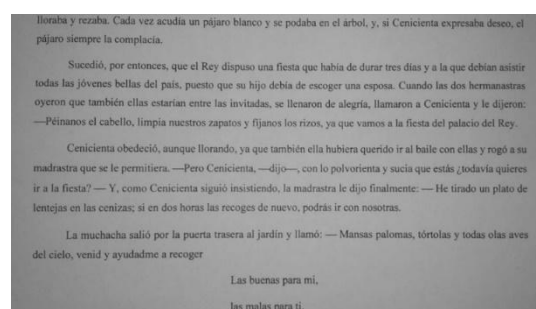


Imagen 2. Imagen primera mitad en escala de grises.

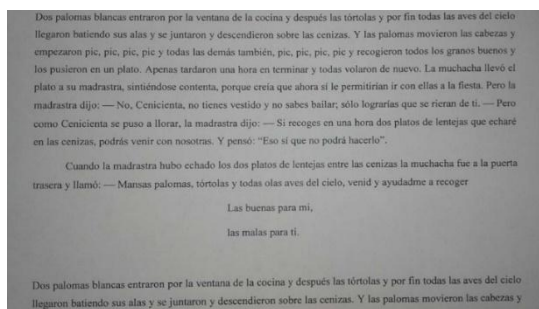


Imagen 3. Imagen RGB segunda mitad de la página.

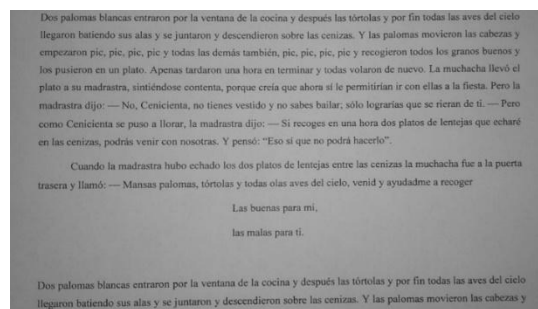


Imagen 4. Imagen segunda mitad en escala de grises.

En las imágenes 5 y 6 se presentan las imágenes bitonales empleando umbral igual a 125 y 123 respectivamente (ver Tabla 1), obtenido por el método de Yen. En las imágenes 5 y 6 se pueden observar las líneas correspondientes al texto, sin embargo, las esquinas presentan un exceso de píxeles de ruido, es decir, se observan las letras dilatadas, lo cual representa un resultado no satisfactorio.

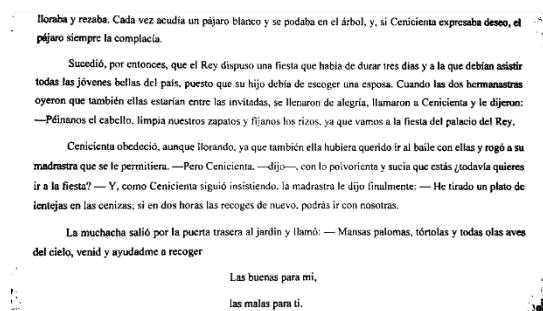


Imagen 5. Umbralizado obtenido por medio del método de Yen primer mitad de página.

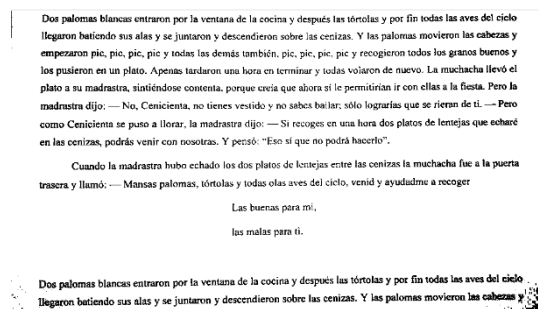


Imagen 6. Umbralizado obtenido por medio del método de Yen segunda mitad de página.

En las imágenes 7 y 8 se presentan las imágenes bitonales empleando un umbral igual a 152 y 148 respectivamente (ver Tabla 1), obtenido por el método de Kittler. De estas imágenes se puede comentar que los valores obtenidos no son los óptimos para la visualización total del texto ya que se observa una concentración de luz en el centro de las imágenes contrario a los extremos.

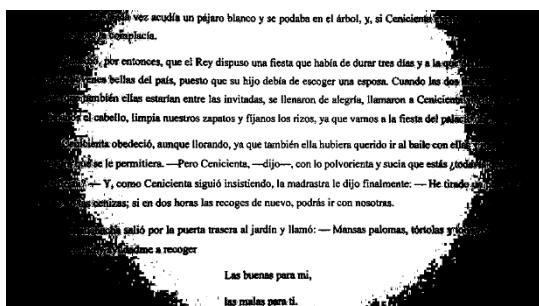


Imagen 7. Umbralizado obtenido por medio del método de Kittler primer mitad de página.

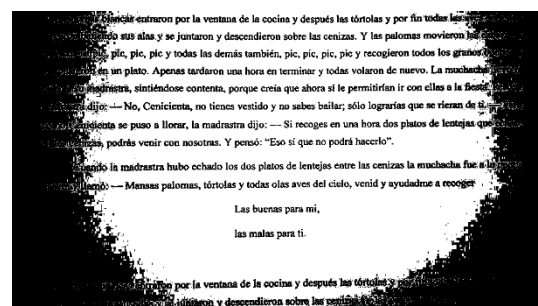


Imagen 8. Umbralizado obtenido por medio del método de Kittler segunda mitad de página.

En las imágenes 9 y 10 se presentan las imágenes bitonales empleando un umbral igual a 120 y 116 respectivamente (ver Tabla 1), obtenido por el método de Otsu. En las cuales se observa un mejor desempeño comparado con los métodos anteriores. En estas imágenes se observa la separación entre el texto y el fondo.

llenaba y rezaba. Cada vez acudía un pajarito blanco y se posaba en el árbol, y, si Cenicienta expresaba deseo, el pajarito siempre la complacía.

Sucedió, por entonces, que el Rey dispuso una fiesta que había de durar tres días y a la que debían asistir todas las jóvenes bellas del país, puesto que su hijo debía de escoger una esposa. Cuando las dos hermanastras oyeron que también ellas estarían entre las invitadas, se llenaron de alegría, llamaron a Cenicienta y le dijeron: —Péñanos el cabello, limpia nuestros zapatos y fíjanos los rizos, ya que vamos a la fiesta del palacio del Rey.

Cenicienta obedeció, aunque llorando, ya que también ella hubiera querido ir al baile con ellas y rogó a su madrastra que se le permitiera. —Pero Cenicienta, —dijo—, con lo polvorienta y sucia que estás ¿todavía quieres ir a la fiesta? —Y, como Cenicienta siguió insistiendo, la madrastra le dijo finalmente: —He tirado un plato de lentejas en las cenizas; si en dos horas las recoges de nuevo, podrás ir con nosotras.

La muchacha salió por la puerta trasera al jardín y llamó: — ¡Mansas palomas, tórtolas y todas las aves del cielo, venid y ayudadme a recoger

Las buenas para mí,
las malas para ti.

Imagen 9. Umbralizado obtenido por medio del método de Otsu

¡Dos palomas blancas entraron por la ventana de la cocina y después las tórtolas y por fin todas las aves del cielo llegaron batiendo sus alas y se juntaron y descendieron sobre las cenizas. Y las palomas movieron las cabezas y empezaron pic, pic, pic y todas las demás también, pic, pic, pic y recogieron todos los granos buenos y los pusieron en un plato. Apenas tardaron una hora en terminar y todas volaron de nuevo. La muchacha llevó el plato a su madrastra, sintiéndose contenta, porque creía que ahora sí le permitirían ir con ellas a la fiesta. Pero la madrastra dijo: — No, Cenicienta, no tienes vestido y no sabes bailar; sólo lograrías que se rieran de ti. — Pero como Cenicienta se puso a llorar, la madrastra dijo: — Si recoges en una hora dos platos de lentejas que echare en las cenizas, podrás venir con nosotras. Y pensó: “Eso sí que no podrá hacerlo”.

Cuando la madrastra hubo echado los dos platos de lentejas entre las cenizas la muchacha fue a la puerta trasera y llamó: — ¡Mansas palomas, tórtolas y todas las aves del cielo, venid y ayudadme a recoger

Las buenas para mí,
las malas para ti.

¡Dos palomas blancas entraron por la ventana de la cocina y después las tórtolas y por fin todas las aves del cielo llegaron batiendo sus alas y se juntaron y descendieron sobre las cenizas. Y las palomas movieron las cabezas y

Imagen 10. Umbralizado obtenido por medio del método de Otsu segunda mitad de página.

La imagen 12 y 13 muestran los histogramas para la fotografía de la imagen inicial que corresponde a la parte superior de la página caso similar en la imagen número 13 que corresponde a la parte inferior, de donde podemos observar que los histogramas son similares y se agrupan alrededor del nivel de gris 150.

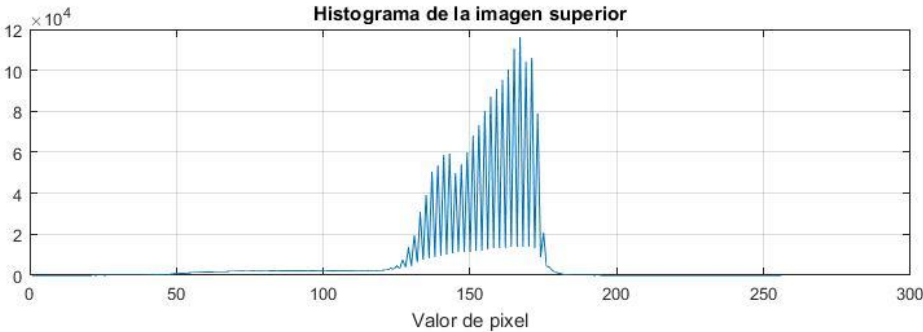


Imagen 12. Histograma de foto de primera mitad de página.

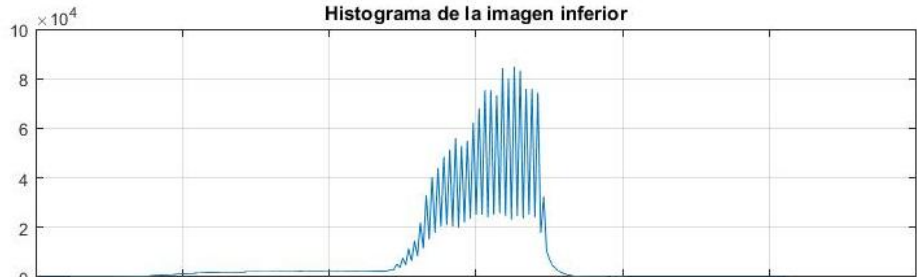


Imagen 13. Histograma de foto de segunda mitad de página.

Método	Umbral global	
	Mitad superior	Mitad inferior
Yen	125	123
Kittler	152	148
Otsu	120	116

Tabla 1. Valores de umbrales obtenidos.

De acuerdo a la literatura el método de Otsu se comporta mejor en imágenes con un histograma de distribución no bimodal, mientras que los métodos basados en logaritmo parecen tener un mejor rendimiento en histogramas bimodales. Por lo tanto, para las imágenes presentadas en este trabajo y por su tipo de histograma es de esperar que el método de Otsu ofrezca mejores resultados.

Conclusiones:

De acuerdo a los resultados obtenidos aplicando diferentes métodos de umbralización a las mismas fotografías y el comportamiento de estas observado en sus histogramas, se concluye que, el umbralizador global aplicado por el método de Otsu es el óptimo para este tipo de imágenes, como lo es el caso en el que se tengan fotografías de texto plano; pudiendo ser empleadas las imágenes resultantes para la extracción de texto.

Se puede confirmar que el método de umbralización de Otsu es adecuado cuando se tiene un histograma de distribución no bimodal; un caso donde se puede presentar esto es en una imagen de texto y fondo blanco. Sin embargo, en la literatura se dice que, para los casos de imágenes con histogramas de valles marcados, los otros métodos pueden tener un mejor desempeño.

Referencias:

1. J. Kittler and J. Illingworth (1986). Minimum error Threshholding. Great Britain. Pattern Recognition Vol. 19 No. 1.
2. Jui-Cheng Yen, Fu-Juay Changs and Shyang Chang (1995). A New Criterion for Automatic Multilevel Thresholding. IEEE Transactions on image processing, Vol 4.
3. Noboyuki Otsu. (1979). A Thresholding Selection Method from Gray-Level Histograms. IEEE Transactions on systems, man and cybernetics, Vol. smc-9, No.1.

Nota: Los autores del trabajo no tienen problemas en compartir los códigos implementados. En caso de requerir los códigos contactar a la dirección de correo: erodriguezc1003@gmail.com.