

Eliminación de sombras en fotos de documentos mediante operaciones espaciales y morfológicas

Aponte Núñez Ruben Dario¹, Mantilla Lobatón Lyn Harold Grimaldo², Orellana Martínez Joel³, Toscano Recalde David Sebastián⁴

*Master en Inteligencia Artificial
Universidad de la Rioja*

I. INTRODUCCIÓN

Existen dos ramas a la hora de hablar de la eliminación de sombras, una con respecto a la eliminación de sombras en ambientes externos o ambientes generales (Gong & Cosker, n.d.) (Guo, Dai, & Hoiem, 2013) y (Gryka, Terry, & Brostow, 2015) y la segunda rama de trabajos que centran su enfoque en remover sombras de imágenes de documentos como los descritos en (Yang, Tan, & Ahuja, 2012), (Brown & Tsoi, 2006), y (Bako et al., 2016). Actualmente es común realizar la captura de documentos con dispositivos móviles, pero este proceso de adquisición generalmente puede generarse problemas como sombras duras o suaves, es decir, sombras con perfil definido o perfil no definido respectivamente. El fin consecuente de este proceso de captura es para sencillamente almacenarlos o para aplicaciones posteriores como reconocimiento de texto para traducción, transcripción, entre otros.

La problemática más destacada en la bibliografía es que al eliminar sombras suaves en los documentos quedan residuos del perfil de la sombra, en este documento se estudia dos métodos propuestos, y de línea base se toma un método propuesto en la actividad 1 por uno de los integrantes, el propósito de estos nuevos métodos es utilizar operaciones espaciales y morfológicas para mejorar el algoritmo y comparar los resultados. Para validar los métodos se utiliza incluso una imagen ejemplo catalogada como retadora por (Bako et al., 2016).

II. DESCRIPCIÓN DEL PROBLEMA

Cuando se realiza captura de foto de un documento pueden existir sombras que impiden apreciar el texto apropiadamente en estas zonas oscuras, se proponen dos técnicas distintas para la resolución del problema aplicando distintos filtros en ordenes diferentes que permiten eliminar las sombras y realzar el texto para facilitar su lectura o posible procesamiento posterior, como extracción de características o reconocimiento de textos. Los métodos propuestos se basan en operaciones elementales punto a punto, modificación de histograma, uso de filtros espaciales y morfológicos y segmentación usando *threshold*.

III. SOLUCIÓN PROPUESTA

Se proponen 2 métodos robustos, sin embargo, antes de explicar los métodos se definen las operaciones utilizadas:

A. Funciones auxiliares

Contraste por percentil: es decir se recibe un valor de 1 a 100 (percentil) que en realidad tiene un mapeo a un valor de intensidad en función del histograma de la imagen de 0 a 255. Si se recibe una imagen con un percentil $P = 34$, significa que para pixeles por debajo de la intensidad asociada al percentil 34 se realiza una disminución de su valor, mientras que para valores que estén encima se realiza un aumento. Esta función esta

¹ rubendario.aponte552@comunidadunir.net

² lynharoldgrimal.mantilla858@comunidadunir.net

³ joel.orellana102@comunidadunir.net

⁴ david.toscano240@comunidadunir.net

descrita en la ecuación 1 donde I es el conjunto de pixeles que corresponden a una imagen en grises I_i es cada uno de los pixeles y P es un valor entre 1 y 100 que determina el límite de lo que es considerado negro o blanco.

$$C(I, P) = \begin{cases} 255, & \text{si } I_i > P \\ 0, & \text{si } I_i \leq P \end{cases} \quad (1)$$

Función de contraste óptimo por MSE (Minimum Square Error): Esta función utiliza una estimación del error definida en la ecuación 2. Cuando se aplica una función de contraste a la imagen y esta no es adecuada sucede que pixeles que pertenecen a las letras se aclaran, o, al contrario, pixeles que no pertenecen a las letras se oscurecen, así que para encontrar el punto ideal de contraste se selecciona utilizando el valor que produzca el menor error cuadrático de la diferencia entre la imagen ya realizada el contraste y la imagen antes de realizar el contraste.

$$error(I, P) = \sum (I_i - C(I_i, P))^2 \quad (2)$$

Función de división de imagen por mascara: Esta función solo se utiliza para el método base y A, la idea es conseguir una máscara que permita definir qué sectores de la imagen corresponden a una zona clara y cuales corresponden a una zona oscura, para ello se aplica un filtro de mediana que elimina las letras y deja una imagen con únicamente sombras, a estas sombras se le aplica un condicional con respecto a su media, si es mayor el pixel será blanco y si es menor será negro obteniendo esta mascara.

B. Métodos propuestos

Método Base:

Se separa la imagen en dos según la máscara, después se obtiene el histograma de cada una de las secciones, se obtiene la moda de cada zona, por supuesto la moda de la zona oscura tendrá un menor valor de intensidad con respecto a la moda de la zona clara, o al menos así se asume para este trabajo, el siguiente objetivo es aclarar la

zona oscura así que se realiza una operación de suma tal que las modas de ambas zonas tengan el mismo valor, luego se juntan la zona clara y la zona oscura corregida en una sola imagen a la cual se realiza el siguiente procesamiento: se le aplica un contraste a la diferencia entre la imagen y la imagen pasada por un filtro de mediana.

Método A:

El método A es idéntico al método Base salvo por una diferencia, la máscara fue dilatada por una acción de dilatación para reducir la sección de sombra que, al analizar los pasos intermedios, se evidencia que el método Base sobreestima el tamaño de la sombra, En la Figura 1 se muestra un diagrama que especifica el método Base y la diferencia crucial que es el uso de un filtro morfológico de erosión que se aplica solo para el método A.

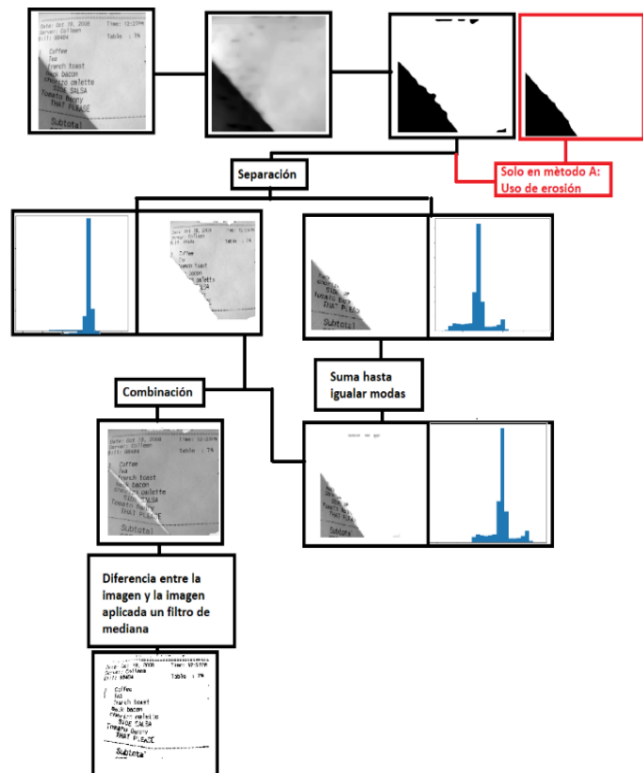


Figura 1. método Base.

Método B:

El método se compone de dos operaciones: La primera es aplicar con elemento estructural de $n \times m$, una operación morfológica de clausura (Closing), la cual se compone por definición de

aplicar una dilatación y posteriormente una erosión (todo en escala de grises). Esta operación constituye el “fondo” de la imagen y posteriormente se le resta a la imagen original (Chityala , R., & Pudipeddi, S., 2020, p. 209).



Figura 2. Operación de clausura con resta de fondo.
Fuente: (Chityala , R., & Pudipeddi, S., 2020).

Esta sucesión de operaciones permite rellenar los pequeños “huecos” en una imagen, conservando la forma y tamaño de objetos grandes de la imagen y también tiende a suavizar los contornos. Adicionalmente, en imágenes con sombra, permite cuasi eliminar las zonas oscuras, homogenizando la imagen (ver Figura 2).

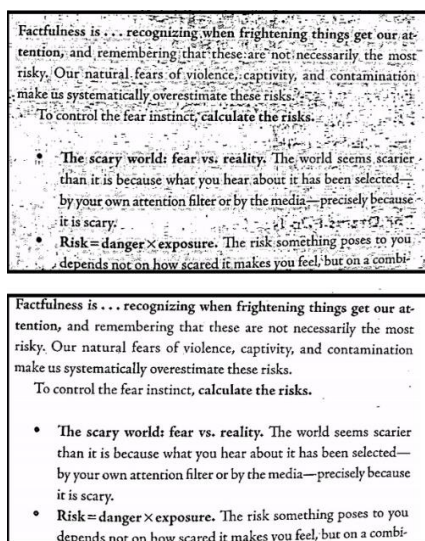


Figura 3. Operación de *Adaptive Threshold*. Fuente: (Schuck, 2020)

Finalmente, la segunda operación es aplicar una segmentación por límite variable (*Adaptive Threshold*) el cual implica definir una ventana para aplicar una segmentación binaria (con umbral C) basada en una suma ponderada gaussiana de los vecinos de cada pixel. Esto permite lograr mayor contraste en la imagen, pero perdiendo algo de definición en los bordes (Schuck, 2020).

IV. RESULTADOS

El método Base funciona solo para sobras duras, el método A funciona parcialmente para sombras suaves, mientras el método B funciona apropiadamente para todo tipo de sombras.



Figura 4. Resultado para los 3 métodos. Columnas: a) Imágenes originales. b) Método Base. C) Método A y d) Método B.

V. DISCUSIÓN

En la bibliografía se comenta bastante sobre la capacidad computacional de los algoritmos, en

este estudio debido al fin educativo lo hemos dejado sin tratar.

VI. CONCLUSIONES

El método base tiene un problema, que fue especificado anteriormente y que se confirma en los diferentes ejemplos mostrados en resultados, que al encontrarse frente a sombras suaves genera un perfil visible, mientras que el método A en general reduce el problema, pero no lo elimina por completo, mientras el método B lo reduce completamente, incluso aplicándolo para la imagen catalogada como retadora mencionada en (Bako et al., 2016)

REFERENCIAS

Bako, S., Darabi, S., Shechtman, E., Wang, J., Sun Kavalli, K., & Sen, P. (2016). Removing Shadows from Images of Documents. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 10113 LNCS, 173–183. doi: 10.1007/978-3-319-54187-712.

Brown, M. S., & Tsoi, Y. C. (2006, jun). Geometric and shading correction for images of printed materials using boundary. IEEE Transactions on Image Processing, 15(6), 1544–1554. doi: 10.1109/TIP.2006.871082.

Chityala, R., & Pudipeddi, S. (2020). Image processing and acquisition using Python. Chapman and Hall/CRC.

Gong, H., & Cosker, D. (n.d.). Interactive Shadow Removal and Ground Truth for Variable Scene Categories. Gryka, M., Terry, M., & Brostow, G. J. (2015, nov). Learning to Remove Soft Shadows. ACM Transactions on Graphics (TOG), 34(5). Retrieved from <https://dl.acm.org/doi/10.1145/2732407> doi: 10.1145/2732407

Guo, R., Dai, Q., & Hoiem, D. (2013). Paired regions for shadow detection and removal. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(12), 2956–2967.

doi: 10.1109/TPAMI.2012.214.

Schuck, J. Enhance a Document Scan using Python and OpenCV. (2020). Analytics Vidhya - Medium. (n.d.). Retrieved from <https://medium.com/analytics-vidhya/enhance-a-document-scan-using-python-and-opencv-9934a0c2da3d>

Yang, Q., Tan, K. H., & Ahuja, N. (2012). Shadow removal using bilateral filtering. IEEE Transactions on Image Processing, 21(10), 4361–4368. doi: 10.1109/TIP.2012.2208976.