

Transformaciones Puntuales para Imágenes de Texto en Escala de Grises

Cristian David Sánchez Hernández— Código N.º 20162006079*

Diego Alejandro Parra Guzmán — Código N.º 20062005047**

Julio David García López — Código N.º 20052005040***

13 de septiembre
de 2010

Resumen

En este documento se realiza un estudio sobre transformaciones puntuales aplicadas a imágenes de texto en escala de grises para obtenerlas de manera binarizada, para ello se aplican técnicas descritas en [1], [2], [3], [4]. Para lograr esto se analiza cada imagen de manera individual aplicando diferentes técnicas, mirando y comparando cuál de todas se ajusta de manera adecuada, además de esto, dentro de estas imágenes se encuentran problemas de iluminación lo cual no permite realizar una binarización directa y por lo tanto requieren de un análisis local que permita lograr una correcta detección de las letras de la imagen.

Palabras claves binarización, Transformaciones Puntuales, filtros lineales, transformaciones globales, transformaciones locales.

*Estudiante de Pregrado de la Facultad de Ingeniería Electrónica de la Universidad Distrital Francisco Jose de Caldas. Correo Electrónico: cdsanchezh@correo.udistrital.edu.co.

** Estudiante de Pregrado de la Facultad de Ingeniería Electrónica de la Universidad Distrital Francisco Jose de Caldas. Correo Electrónico: daparrag@correo.udistrital.edu.co.

*** Estudiante de Pregrado de la Facultad de Ingeniería Electrónica de la Universidad Distrital Francisco Jose de Caldas. Correo Electrónico: jdgarcial@udistrital.edu.co.

1. Introducción

Dentro del procesamiento digital de imágenes han surgido aplicaciones tal como el reconocimiento de características particulares dentro de una imagen contaminada con una fuerte de ruido, problemas de iluminación, bajo contraste, entre otros, para esto se encuentran técnicas que permiten la extracción de esas características de manera adecuada, A estas técnicas se le conocen como transformaciones puntuales [1], que no son las únicas.

Algunas de las características de las transformaciones puntuales se especifican en [1], [2]. Las cuales son independientes de la posición del pixel, dentro de estas transformaciones, se pueden destacar las transformaciones globales, y operaciones de recorte, contraste y umbralización, al igual que operaciones básicas como suma, resta, multiplicación, división, operaciones lógicas como or, and, xor, negación. Estas técnicas serán aplicadas en la búsqueda de una metodología que permita la binarización de las imágenes de texto estudiadas.

Se hace necesario aplicar transformaciones puntuales a las imágenes estudiadas debido a que se encuentran problemas de iluminación como se puede apreciar en la (Figura1); al aplicar transformaciones puntuales se busca mejorar las imágenes y obtener de manera clara que se resalte las letras y el fondo.

2. Formulación del problema

Para las imágenes de estudio (Figura1); se busca una binarización donde se destaquen claramente las letras del fondo, para esto se aplican técnicas de transformaciones puntuales como se describen a continuación.



Figura 1 Imágenes de texto en escala de grises a procesar, (a) imagen 1, (b) imagen 2, (c) imagen 3, (d) imagen 4.

Se realizó un estudio del histograma el cual es una representación gráfica de la distribución que existe de las diferentes tonalidades de grises con relación al número de píxeles o porcentaje de los mismos. Este análisis se realizó a cada una de las imágenes (Figura2), donde primeramente se realizó una expansión del contraste de la siguiente manera

$$y = \frac{255}{L_{max} - L_{min}} \cdot (im - L_{min}) \quad (1)$$

Dónde:

L_{max} : Magnitud máxima de intensidad presente en la imagen.

L_{min} : Magnitud mínima de intensidad de la imagen.

Y : Imagen nueva resultado de la expansión de contraste.

im : imagen a procesar

Además de la expansión del contraste, se aplicaron modelos estadísticos, Analizando factores como la media y la varianza, para encontrar un punto adecuado donde realizar la binarización de estas imágenes.

También se modificó su histograma aplicando funciones lineales donde se buscaba a través de técnicas de realce mejorar la apariencia de la imagen y poder realizar una binarización más adecuada a los requerimientos del problema, el objetivo es mejorar el contraste, ruido, escala de grises, distorsión, luminosidad, falta de nitidez, etc. [4]

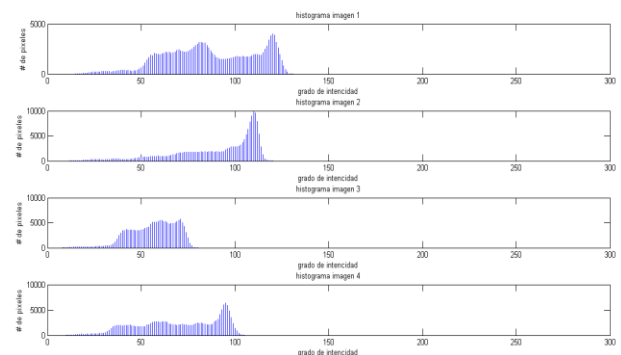


Figura 2 Histogramas originales para cada una de las imágenes a procesar (a) imagen 1, (b) imagen 2, (c) imagen 3, (d) imagen 4.

De los métodos de realce se destacan los métodos de realce en **el dominio espacial, y métodos de realce en el dominio de la frecuencia** [4], para este estudio se aplicaran las técnicas en el dominio espacial debido a que estas consisten en la manipulación directa de los píxeles de la imagen, específicamente se analizaron las imágenes aplicando **realce radiométrico**. [4], de donde se realizó una ecualización del histograma este consiste en encontrar una transformación con la cual el histograma tenga una representación uniforme, es decir, con la ecualización se trata de igualar en lo más posible el histograma de una imagen a un histograma ideal. Para lograr este objetivo se necesita tener en cuenta dos restricciones ya que sin estas se pueden obtener resultados erróneos.

Sea r una variable que representa los valores de intensidades de gris que para nuestro caso están en el intervalo cerrado $[0,255]$, donde $r=0$ representa el nivel de negro y $r=255$ el nivel blanco.

Sea s una transformación $T(r)$ la cual produce un nivel para cada valor del pixel r de la imagen original como

$$s=T(r), \quad 0 \leq r \leq 255 \quad (2)$$

Donde $T(r)$ debe ser **monótona creciente**, esto indica que la pendiente de la función no puede ser negativa, debido a que si así lo fuera la imagen se vería afectada en su totalidad, porque lo que se estaría haciendo sería invertir sus tonalidades, es decir que:

$$\frac{ds}{dr} \geq 0 \quad (3)$$

Además $T(r)$ debe tener inversa y esta debe ser **monótona creciente**, para esto se debe suponer que r y s son variables aleatorias y continuas, y además $P(r)$, $P(s)$ son funciones de densidad de probabilidad de r , s .

Luego de esto se aplicó una transformación local, lo que busca esta transformación es la de resaltar características de las imágenes, esto se hace por que la imagen no tiene intensidades de grises constantes en toda la imagen, y por lo tanto no puede ser analizada de manera global haciendo una transformación uniforme sobre cada pixel. Al aplicar la transformación local se busca hacer mucho más oscuro el contorno de las letras de la imagen, es decir, que cada pixel se modifica de acuerdo a su

vecindad (Figura 3), este proceso consistirá entonces en remplazar el pixel actual por el valor mínimo de su vecindad, si se representa la imagen en escala de grises como una función $f(x,y)$, donde los píxeles están definidos por las coordenadas (x,y) podemos entonces:

$$f(x,y) = \min(f(x,y), f(x+1,y), f(x,y+1), \dots, f(x+i,y+j)) \quad (2)$$

Dónde:

$f(x,y)$: valor del pixel actual.

i : valor máximo horizontal de la imagen.

j : valor máximo vertical de la imagen.

Luego de hacer varios experimentos se encontraron imágenes claramente resaltadas, a partir de esto se binarizaron las imágenes con un valor adecuado de donde se destacan algunos resultados.

Por otro lado se aplicó un filtro espacial para tratar de unificar los valores de las imágenes, y poder hacer la umbralización mucho más sencilla, estos filtros espaciales buscan modificar la contribución de determinados rangos de frecuencias de una imagen. El término espacial se refiere al hecho de que el filtro se aplica directamente sobre la imagen y no una transformación de la misma, es decir, en nivel de gris del pixel se obtiene directamente en función del valor de sus vecinos. Dentro de estos filtros se pueden considerar 2, de **tipo lineal** y de **tipo no lineal**, los filtros lineales se clasifican según las frecuencias que permiten pasar, por ejemplo los filtros pasa bajos atenúan o eliminan componentes de alta frecuencia, la convolución es una operación de filtraje. Este filtraje espacial tiene una alta relación con el filtraje en el dominio de la frecuencia.

En general se puede clasificar los filtros lineales según el resultado que se desee:

- Filtros paso-bajas (Smoothing Spatial Filters): Son utilizados en la reducción de ruido, suavizan y aplanan un poco las imágenes y en consecuencia se pierde la nitidez [2].
- Filtros pasa altas (Sharpening Spatial Filters): Estos filtros son utilizados para detectar cambios de luminosidad. Son utilizados en la detección de patrones como bordes o para resaltar detalles finos de la imagen [2].

- Filtros paso-banda: Son utilizados para detectar patrones de ruido. Ya que un filtro paso-banda son útiles para aislar los efectos de ciertas bandas de frecuencias seleccionadas sobre la imagen. De esta manera estos filtros ayudan a simplificar el análisis de ruido, razonablemente independiente del contenido de la imagen [2].

El filtro implementado para las imágenes se conoce con el nombre de **filtro sobel** el cual se basa en derivadas de gaussianas, estos filtros tienen un fundamento matemático y un comportamiento similar al de la visión humana (HVS); este filtro tiene la característica de la detección de cambios bruscos con los bordes.

La función gaussiana tiene la siguiente expresión

$$f(x)=N \cdot e^{-a \cdot x^2} \quad (4)$$

Donde N es una constante de normalización que depende de (a) pero es independiente del valor de (x). Hay varias posibilidades para el exponente (a), de manera estándar se puede definir a como

$$a=\frac{1}{2 \cdot \sigma^2} \quad (5)$$

Además de esto se tiene en cuenta las primeras derivadas de (4), para generar filtros pasa-altos, estos filtros son detectores de discontinuidades o cambios, basados precisamente en estas derivadas.

Este filtro está construido de manera continua el cual también tiene su representación para sistemas discretos, (figura 4) donde

$$h(z)=\delta(z+1)-\delta(z-1) \quad (6)$$

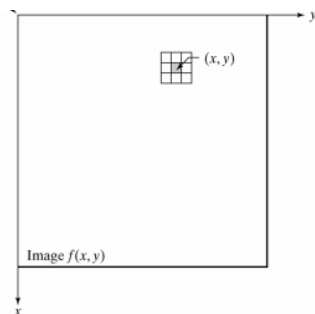


Figura 3 Representación de una vecindad espacial

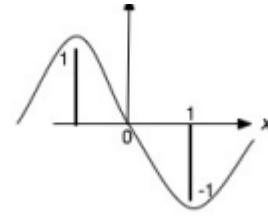


Figura 4 Filtro basado en derivadas de funciones gaussianas de tipo (4) y su representación discreta

Estos filtros detectores de discontinuidades, tiene regularmente aplicaciones en tareas previas a la segmentación o detección de bordes.

Para este caso la máscara aplicada para filtrar la imagen será de la siguiente forma:

$$\begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix} \quad (7)$$

La razón por la cual este filtro utiliza un peso de 2 al centro es para obtener algo de suavizado con lo cual se logra cierta reducción del ruido. Existen estudios que demuestran que este filtro corresponde a una aproximación binomial de primer orden para una primera derivada de una gaussiana, este filtro en particular se implementó en la dirección horizontal, existen versiones para versiones verticales.

Con respecto a este filtro se hicieron algunas comparaciones con las propuestas anteriores, la variación del histograma y la variación de los píxeles a nivel local; se destacan algunos resultados.

3. Resultados

Como se adelantó en la sección anterior se implementaron tres tipos de procesamiento sobre las imágenes de manera independiente (Figura 1), en esta sección se publican los resultados más relevantes de cada una de estas.

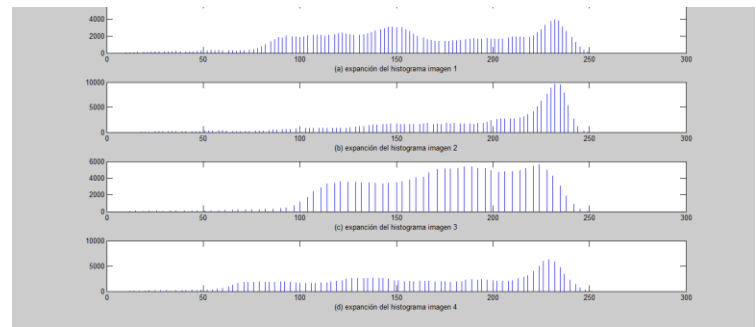


Figura 5 Expansión de histogramas para las imágenes procesadas (a) imagen 1, (b) imagen 2, (c) imagen 3, (d) imagen 4.

Inicialmente se hizo un análisis del histograma, realizando una expansión lineal sobre cada una de las imágenes (Figura 5), y se calculó un valor promedio para su binarización, este ejercicio, no arrojó resultados adecuados para la binarización de las imágenes (Figura 6), excepto por la imagen 1 donde se puede distinguir claramente las letras del fondo con pequeños detalles que la hacen imperfecta (Figura 7), esto es debido a la iluminación de esta imagen, la cual, es mayor que las demás (Figura 1).

Este resultado se dio ya que las imágenes no tienen una luminosidad uniforme, y por lo tanto, no se puede aplicar una técnica que modifique globalmente el valor de los píxeles, sin embargo se puede notar (Figura 6), que entre menor sea el área de la letra, menor será la posibilidad de que se distingan del fondo (figura 6-c), una alternativa que se expone más adelante es la de tratar de unificar el histograma, este proceso se conoce como **ecualización del histograma**.

Es importante destacar que este método cumple para valores de iluminación constante, lamentablemente para este problema no es así.

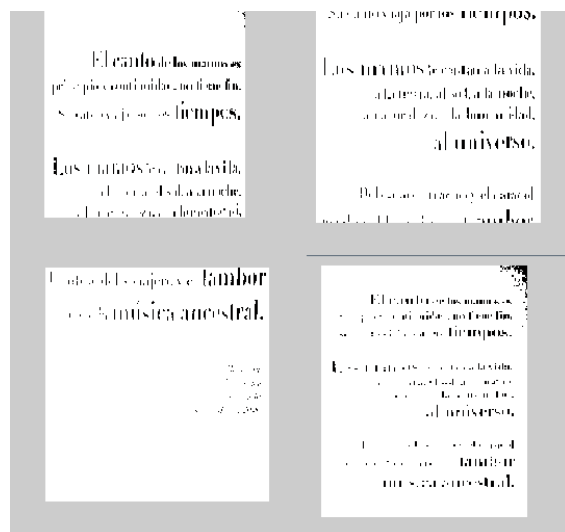


Figura 6 Binarización de imágenes con criterio de promedio y expansión lineal del histograma para las imágenes procesadas (a) imagen 1, (b) imagen 2, (c) imagen 3, (d) imagen 4.

Por otro lado se implementó un método radiométrico para tratar de unificar el histograma, este se conoce como **Ecualización del histograma**, el cual permitió mejorar la luminosidad de las imágenes pero al ser un método que afecta de manera global el valor de los píxeles no se obtuvo un buen reconocimiento de las letras respecto al fondo.

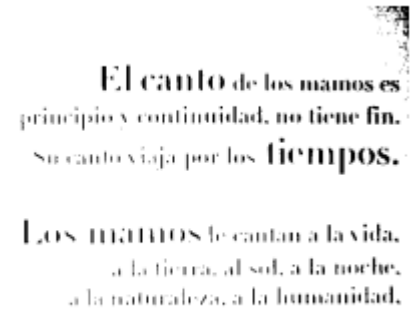


Figura 7 Mejor resultado al aplicar el método de análisis del histograma con promedio y expansión lineal del mismo, correspondiente a la imagen 1

Se puede destacar que este método mejora notablemente la binarización de las imágenes respecto al método de promedio y expansión del histograma, esto se puede notar al realizar una comparación entre la (Figura 8), y la (Figura 6), para esta técnica se pueden distinguir la binarización de las imágenes, imagen 1, imagen 2, (Figura 9), sin embargo para las letras que tienen un tamaño más pequeño, no se pueden distinguir y para este caso también se pierde su distinción.



Figura 8 Binarización de imágenes aplicando previamente la ecualización del histograma para las imágenes procesadas (a) imagen 1, (b) imagen 2, (c) imagen 3, (d) imagen 4.

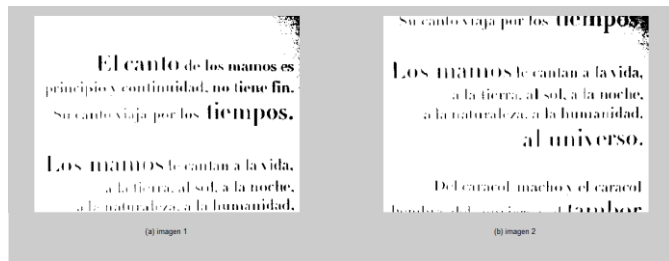


Figura 8 imágenes destacadas al aplicar ecualización del histograma (a) imagen 1, (b) imagen 2

Para terminar con este análisis, se puede hacer énfasis, en la aplicación de estos métodos no a nivel global, donde todos los píxeles se afectan de manera constante y por igual, si no, a nivel local donde se puedan, para este caso, lograr un comportamiento Mucho más adecuado, y hacer énfasis en las letras más pequeñas.

Como se explicó anteriormente no se puede realizar una transformación global de la imagen, es por esto que se intentó resaltar la vecindad de los píxeles de acuerdo con la ecuación (2), de donde se pudo notar un gran cambio en la imagen si se toma una vecindad mayor a 3x3 píxeles, casi, que las letras pierden sentido dentro de la imagen, este método, no mejora la luminosidad, pero destaca las letras dentro de la imagen, esto es deseable a la hora de realizar la binarización, Con este método se resaltan las letras pequeñas que en los casos expuestos anteriormente no se consiguió, sin embargo, que se destaque estas letras no significa que sean legibles, para este caso, las letras pequeñas se resaltan, pero, no son legibles.

Se puede demostrar al comparar las figuras 8, 9, 10, que este método fue superior a los 2 anteriores, resaltando en general los resultados Figuras 11, 12, 13, 14.

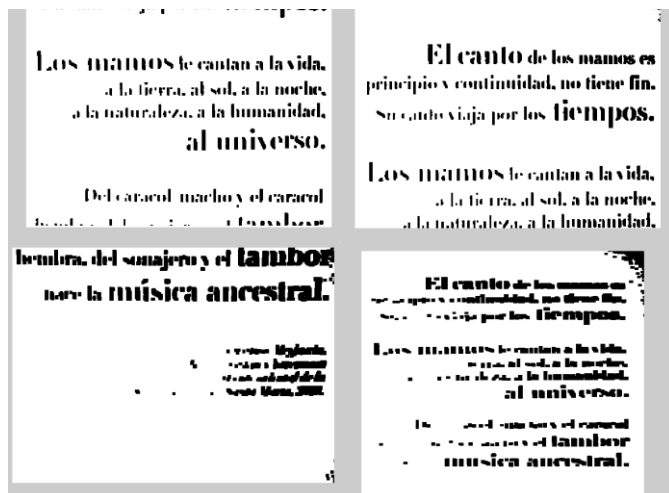


Figura 10 Imágenes binarizadas aplicando transformación de la vecindad

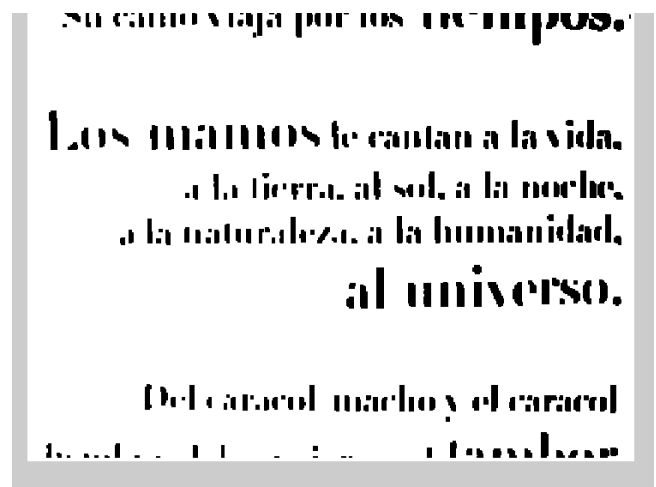


Figura 11 Imagen 1 Binarizada aplicando una transformación de la vecindad

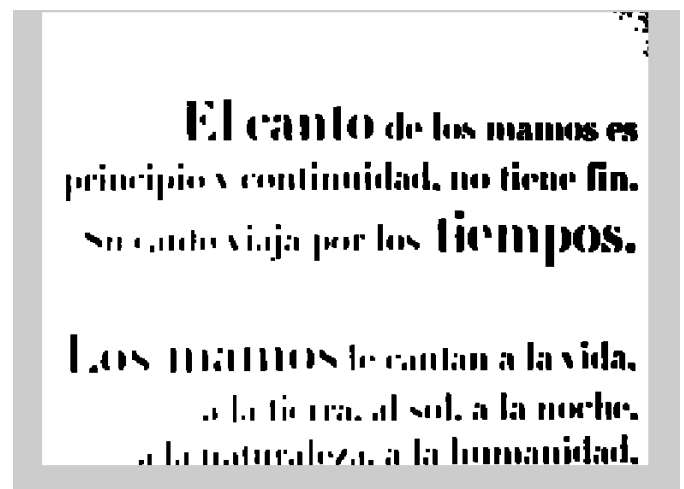


Figura 12 Imagen 2 Binarizada aplicando una transformación de la vecindad

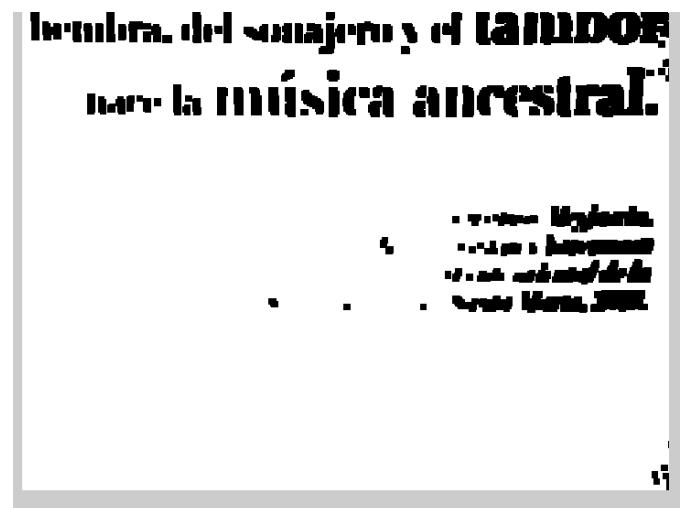


Figura 13 Imagen 3 Binarizada aplicando una transformación de la vecindad.

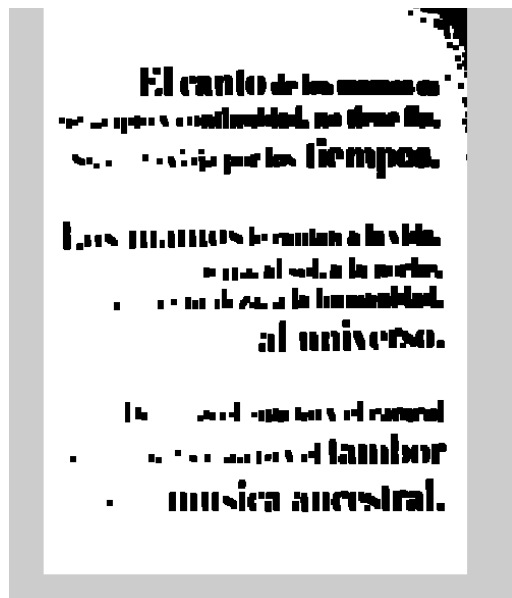


Figura 14 Imagen 4 Binarizada aplicando una transformación de la vecindad.

A pesar de que los resultados obtenidos con este análisis son buenos, no son los más adecuados y presenta dificultades como la distinción de letras pequeñas, es por esta razón que se intentó, buscar un método más elaborado para lograr mejores resultados.

El ultimo método implementado fue el de usar un filtro sobel como el expuesto en secciones previas.

Con la utilización de este filtro se logró mejorar de manera sorprendente la visualización de las imágenes, esto, es debido a que el filtro permite hacer la distinción de cambios fuertes de contraste dentro de la imagen, lo que es bastante aplicado en este tipo de problemáticas [4].

Se puede ver al comparar las imágenes 10 y 15 como se logra mejorar en gran medida los resultados previos, a pesar de ser un método espacial, los resultados son los expuestos al principio de este documento.

Se destacan de manera individual los resultados en las figuras 16, 17, 18, 19, esto es para notar la calidad del procesamiento.

A diferencia que en las secciones anteriores las letras pequeñas se distinguen claramente, además de ser legibles, se hace interesante combinar esta técnica con la técnica de realce de la vecindad, pero no es el objetivo de este documento.

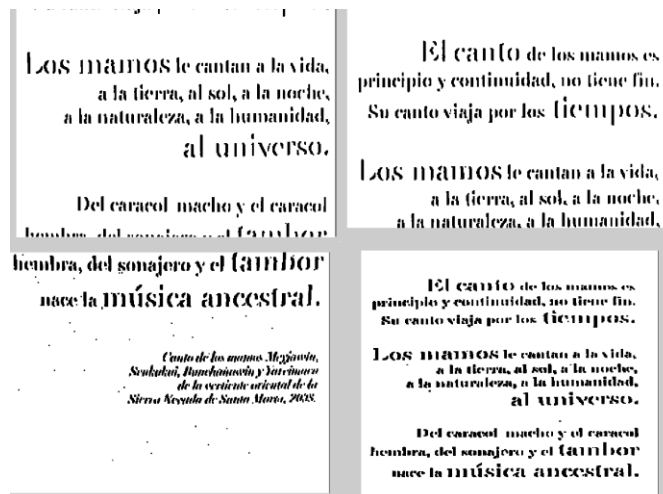


Figura 15 Imágenes binarizada aplicando un filtro sobel para las imágenes procesadas (a) imagen 1, (b) imagen 2, (c) imagen 3, (d) imagen 4.

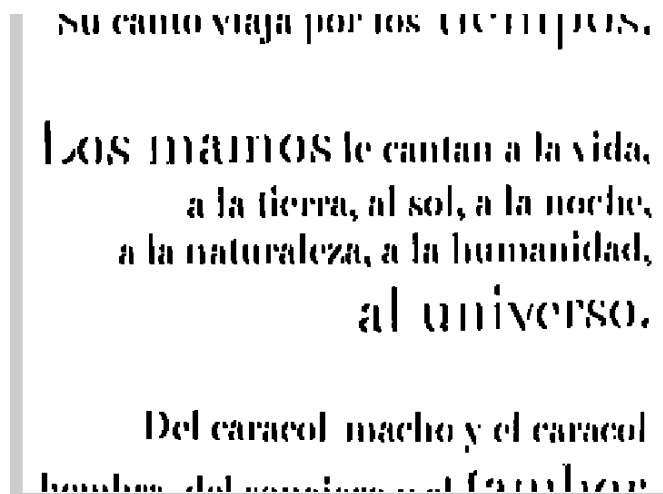


Figura 16 Imagen 1 binarizada aplicando un filtro sobel

**El canto de los manos es
principio y continuidad, no tiene fin.
Su canto viaja por los tiempos.**

**Los manos le cantan a la vida,
a la tierra, al sol, a la noche,
a la naturaleza, a la humanidad,**

Figura 17 Imagen 2 binarizada aplicando un filtro sobel

**hembra, del sonajero y el tambor
nace la música ancestral.**

*Canto de los manos Mezquía,
Sankulá, Hunchuawin y Yurimawa
de la vertiente oriental de la
Sierra Nevada de Santa Marta, 2023.*

Figura 18 Imagen 3 binarizada aplicando un filtro sobel

**El canto de los manos es
principio y continuidad, no tiene fin.
Su canto viaja por los tiempos.**

**Los manos le cantan a la vida,
a la tierra, al sol, a la noche,
a la naturaleza, a la humanidad,
al universo.**

**Del caracol macho y el caracol
hembra, del sonajero y el tambor
nace la música ancestral.**

Figura 19 Imagen 3 binarizada aplicando un filtro sobel

4. Discusión

De las diferentes propuestas realizadas para la solución del problema, se destacan dos la de transformación de vecindad y la de aplicación de un filtro sobel, dentro de estas, la del filtro sobel es mucho más elaborada, además de estar basada en la visión humana, esta técnica permitió obtener un resultado muy agradable, ya que se cumplió con el objetivo principal, el cual consistía en la binarización de imágenes de texto con problemas de iluminación.

Por otra al seguir una bibliografía previa se pudo ir avanzando gradualmente en la búsqueda de los objetivos, y por ende este resultado final.

Queda abierta la posibilidad de aplicar otro tipo de filtros buscando mejorar aún más estos resultados, al igual que analizar estas imágenes desde el análisis frecuencia, analizando por ejemplo su transformación de Fourier, e incluso aplicar la transformación wavelet para este análisis.

5. Conclusiones

Del análisis e implementación de los diferentes métodos para llegar a la binarización de las imágenes, se destaca una buena metodología, la cual permitió tener este tipo de resultados.

Por otro lado se puede agregar que las técnicas globales son bastante deficientes en comparación con técnicas locales para este tipo de aplicaciones, pero que son aceptables si no se desea un grado tan amplio de distinción de las letras.

Referencias bibliográficas

- [1] Gonzalez, R. C. , and Woods, P., *Digital Image Processing*, Addison Wesley, 2002.
- [2] Jhon C. Russ, *The image processing Handbook*, Fourth Edition, CRC Press LLC, Boca Raton, Florida, USA, 2002.
- [3] Erik Valdemar Cuevas Jimenes, Daniel Zaldivar N, *utilizando procesam* dor 02.
- [4] Marcos Fa e imagen y Alfaomega D.F , 2001. il de voz xico ,
- [5] Rafael C Gonzales and Richard E Woods, *Digital Image Processing*, Second Edition, Prentice-Hall, Inc New Jersey, USA, 2002.