APÉNDICE A

FUNDAMENTOS MATEMÁTICOS PARA EL PROCESAMIENTO DE IMÁGENES

El objetivo del procesamiento de la imagen es conseguir que esta resulte más adecuada para una aplicación específica, tratando de destacar los aspectos más significativos de esta con el fin de facilitar el procesado. De esta manera se esperan obtener mejores resultados y que el proceso tenga un coste computacional bajo. Cabe destacar que cada técnica de mejoramiento está profundamente relacionada con una aplicación determinada, y puede no funcionar adecuadamente, o incluso estropear, las imágenes utilizadas en otras aplicaciones.

A.1. FUNDAMENTOS BÁSICOS

Las aproximaciones para el procesamiento de imágenes se pueden clasificar en dos categorías principales: **dominio del espacio** *y* **dominio de la frecuencia**. Las técnicas basadas en el dominio del espacio, que se refieren al plano espacial de la imagen como tal, se basan en la manipulación directa de los píxeles de la misma. Las técnicas del dominio de la frecuencia están basadas en la manipulación de la transformada de Fourier de la imagen. En algunas aplicaciones se pueden utilizar métodos híbridos basados en ambos dominios.

A.1.1. Métodos basados en el dominio del espacio. El término dominio espacial se refiere al conjunto de píxeles que componen una imagen, y los métodos basados en el dominio espacial operan directamente sobre estos píxeles. Las funciones de procesamiento de imágenes en el dominio del espacio pueden representarse como:

$$g(x, y) = T[f(x, y)] \tag{A.1}$$

donde *f(x, y)* es la imagen de entrada, *g(x, y)* es la imagen procesada y *T* es un operador de *f*, definido sobre los píxeles vecinos de *(x, y)*. Se puede definir la vecindad de *(x, y)* como una subimagen cuadrada o rectangular centrada en *(x, y)*, cuyo centro se desplaza de píxel a píxel aplicando el operador en cada desplazamiento para obtener el valor de *g(x, y)* en cada posición. Aunque se pueden utilizar subimágenes con otras formas, como aproximaciones a círculos, las subimágenes cuadradas o rectangulares son las más comunes, debido al carácter rectangular de la disposición de los píxeles se facilita su implementación.

La forma más simple de *T* corresponde a solo un elemento, con un entorno 1 x 1, con lo que el nuevo valor del píxel depende exclusivamente de sí mismo, y no de sus vecinos. De esta manera se convierte en una transformación de los niveles de gris de los píxeles, denominada **procesamiento de punto**, y expresada por la función:

$$s = T(r) \tag{A.2}$$

donde, por simplificación de notación, *r* y *s* denotan el nivel de gris de *f(x, y)* y *g (x, y)*, respectivamente, en todos los puntos *(x, y)*.

Con subimágenes mayores a un elemento, e tendrá una variedad de procesamientos que dependerán de la aplicación específica y van más allá de los que es netamente mejora. Es habitual, en estos casos, la utilización de *máscaras* (conocidas también como plantillas, ventanas, filtros o *kernels*), que son una matriz pequeña, 3 x 3, 5 x 5 o 7 x 7 elementos, cuyos valores determinan la naturaleza del proceso. A estas técnicas generalmente se las denomina *procesamiento por máscaras* o *filtrado espacial*.

A.1.2. Métodos basados en el dominio de la frecuencia. La base de los métodos basados en el dominio de la frecuencia es el teorema de la convolución. Sea *g(x, y)* una imagen formada por la convolución de una imagen *f(x, y)* y un operador lineal invariante con la posición, *h(x, y)*, o sea,

$$g(x,y) = h(x,y) * f(x,y) \tag{A.3}$$

Del teorema de la convolución, la relación en el dominio de la frecuencia quedará como:

$$G(u,v) = H(u,v)F(u,v) \tag{A.4}$$

donde *G*, *H* y *F* son las transformadas de Fourier de *g*, *h* y *f*, respectivamente. En la terminología de la teoría de los sistemas lineales a la transformada *H(u, v)* se llama la **función de transferencia del proceso**, en óptica se conoce como la **función de tranferencia óptica**.

Se pueden expresar numerosos problemas de mejora de imágenes mediante la ecuación A.4. En una aplicación típica de mejora de imágenes, se parte de *f(x, y)* y el objetivo es seleccionar una función de transferencia *H(u, v)* tal que en la imagen deseada, se resalten ciertos aspectos de *f(x, y)*, dependiendo de *H (u, v).* Por ejemplo, los bordes de *f(x, y)* pueden resaltarse utilizando una función de transferencia que destaque las componentes de alta frecuencia de *F(u, v)*.

Suponiendo que se desconoce *h(x, y)*, y que se aplica una función impulso unitario (un punto de luz) al sistema. La transformada de Fourier de un impulso unitario es simplemente 1, por lo que de la ecuación (A.4) se deduce que *G(u, v)* = *H(u, v)*. La transformada inversa de la salida *G(u, v)* será *h(x, y)*. Este resultado ya es conocido en la teoría de los sistemas lineales: un sistema lineal invariante con la posición queda completamente definido por su respuesta al impulso. Es decir, cuando se le aplica la transformada de Fourier de un impulso unitario a un sistema, su salida es precisamente la función de transferencia del sistema, *H(u, v*). O lo que es lo mismo, aplicando el impulso directamente al sistema se obtiene la función de transferencia, *h(x, y)* a la salida. Por esta razón, *h(x, y*), la transformada inversa de la función de transferencia, se denomina la *respuesta al impuls*o.

A.2. PROCESAMIENTO DE PUNTO

Las técnicas de procesamiento de punto solo tienen en cuenta el nivel de gris de cada píxel, ignorando el valor de sus vecinos. Algunas de las transformaciones más simples y comunes son:

¬ **Negativo de la imagen:** Se obtiene empleando la función de transformación que se encarga de invertir el orden de blanco a negro, de tal manera que la intensidad de la imagen obtenida disminuye a medida que la intensidad de la imagen de entrada aumenta. Si L es el número de niveles de gris, desde 0 hasta L–1, la transformación está de acuerdo a

$$s = (L-1) - r\,. \tag{A.5}$$

¬ **Aumento de contraste:** Las imágenes de bajo contraste se pueden originar por muchas causas: iluminación, rango dinámico limitado del sensor o apertura incorrecta de la lente. El objetivo es aumentar el rango dinámico de los niveles de gris de la imagen procesada con el fin de poder diferenciar con mayor facilidad los objetos dentro de la escena. En general, se toma *r1* • *r2 y s1* • *s2* para que se mantenga el orden de los niveles de gris, previniendo la creación de defectos en los niveles de intensidad de la imagen procesada. Es posible tomar la cantidad de puntos que sean necesarios para realizar el mejoramiento más adecuado a las necesidades de la aplicación.

¬ **Compresión del rango dinámico.** A veces el rango dinámico de las imágenes procesadas excede la capacidad del dispositivo de visualización, en cuyo caso solo las partes más brillantes de la imagen presentadas en el monitor o impresora. Una forma efectiva de reducir el rango dinámico de la imagen es realizar la siguiente transformación, donde c es una constante de escala y la función logarítmica realiza la compresión deseada,

$$s = c\log(1 + |r|)\,. \tag{A.6}$$

¬ **Fraccionamiento de los niveles de gris.** En ocasiones se desea resaltar un rango concreto de los niveles de gris de una imagen. Para ello hay una gran cantidad de métodos, pero en definitiva son variaciones de dos ideas básicas. Un primer método consiste en asignar un valor alto a todos los niveles de gris que se desea resaltar y un nivel bajo a los demás, con lo que se consigue una imagen en dos colores claramente diferenciados. Se pueden formular un gran número de variaciones de estas dos transformaciones, dando pie al gran número de métodos mencionados anteriormente.

¬ **Fraccionamiento de los planos de bits:** Si cada punto de la imagen está representado por 8 bits, es decir, un byte y la imagen se separa en sus respectivos planos, y en cada plano se coloca el bit n de cada punto, se tendrán 8 imágenes binarias, a partir de las cuales se podrán realizar análisis en función del bit que contengan. Hay que observar que únicamente los cinco planos de más peso contendrán información significativa visualmente, los otros planos contribuyen a los detalles más finos de la imagen.

A.2.1. Procesamiento del histograma. El histograma de una imagen digital con un rango de niveles de gris comprendido en el intervalo $[0\ldots(L\text{-}1)]$ es la función discreta $p(r_k) = n_k/n$, donde r_k es el nivel de gris de orden k, n_k es el número de píxeles en la imagen con ese nivel de gris y n es el número total de píxeles en la imagen. k puede tomar valores desde *0* hasta $L - 1$. De forma aproximada, se puede decir que $p(r_k)$ es la probabilidad de ocurrencia del nivel de gris r_k. De la observación del histograma se puede deducir si la imagen es brillante u oscura, si tiene mucho o poco contraste, pero no dice nada específico acerca del contenido de la misma. Únicamente informa de la posibilidad de mejorar el contraste de la imagen mediante la manipulación del histograma, ya que es una función unidimensional que contiene algo de la información extractada de una señal bidimensional.

A.2.1.1. Ecualización del histograma. Suponiendo que la variable r representa el nivel de gris de un punto de la imagen, y que es una cantidad continua normalizada, por lo que debe estar comprendida en el intervalo [0, 1], donde $r = 0$ representa al negro y $r = 1$ al blanco, entonces se puede aplicar una transformación del tipo descrito por la ecuación (A.2). Se supone que esta transformación cumple las condiciones:

a. $T(r)$ tiene un único valor y es monótonamente creciente en el intervalo 0 • r • 1 y
b. 0 • $T(r)$ • 1 para 0 • r • 1.

La condición *a* preserva el orden de blanco a negro en la escala de grises, mientras que la condición *b* garantiza un mapeado consistente en el rango de valores dado. La transformación inversa se denota como

$$r = T^{-1}(s) \qquad 0 \leq s \leq 1 \tag{A.7}$$

donde se supone que T -1 cumple también las condiciones *a* y *b* con respecto a la variable *s*. El nivel de gris en la imagen puede verse como una cantidad aleatoria en el intervalo [0, 1]. Si las variables son continuas, los niveles de gris original y transformado pueden ser caracterizados por su respectiva función de densidad de probabilidad $p_r(r)$ y $p_s(s)$. Si $p_r(r)$ y $T(r)$ son conocidas, y $T^{-1}(s)$ cumple la condición *a*, la función de probabilidad de densidad del nivel de gris transformado es

$$p_s(s) = \left[p_r(r) \frac{dr}{ds} \right]_{r=T^{-1}(s)} \tag{A.8}$$

Las técnicas siguientes están basadas en la modificación de la apariencia de la imagen mediante el control de la función de densidad de su nivel de gris a partir de la función de transformación *T(r)*. Sea la función de transformación

$$s = T(r) = \int_0^r p_r(w)dw \qquad 0 \leq r \leq 1 \tag{A.9}$$

donde *w* es una variable de integración ficticia. La parte de la derecha de la ecuación es la función de distribución acumulada de *r* (CDF). En este caso la función de transferencia cumple las condiciones *a* y *b*. De la ecuación (A.9) se deduce que la derivada de *s* con respecto a *r* es:

$$\frac{ds}{dr} = p_r(r) \tag{A.10}$$

Sustituyendo *dr/ds* en la ecuación (**A.8**):

$$p_s(s) = \left[p_r(r) \frac{1}{p_r(r)} \right]_{r=T^{-1}(s)} = [1]_{r=T^{-1}(s)} = 1 \,. \tag{A.11}$$

La función de densidad en el intervalo de definición de la variable transformada *s* es uniforme, independientemente de la función de transformación inversa, cuya determinación analítica no siempre es sencilla.

El desarrollo anterior indica que al usar una función de transformación igual a la distribución acumulativa de *r* se produce una imagen cuyo nivel de gris tiene una densidad uniforme. Desde el punto de vista del mejoramiento de imágenes, este resultado implica un aumento del rango dinámico de los píxeles, que puede tener un efecto considerable en la apariencia de la imagen.

Para poder utilizar los conceptos analizados sobre el histograma en esta sección, se deben formular de manera discreta para aplicarlos a las imágenes digitales. Para valores discretos de los niveles de gris se tiene

$$p_r(r_k) = \frac{n_k}{n} \qquad 0 \leq r_k \leq 1 \quad \text{y} \quad k = 0,1,\ldots,L-1 \tag{A.12}$$

donde L es el número de niveles, $p_r(r_k)$ es la probabilidad del nivel de gris k, n_k es el número de veces que este nivel aparece en la imagen y n es el número total de píxeles en la misma. La representación de $p_r(r_k)$ frente a r_k se denomina *histograma*, y la técnica usada para obtener un histograma uniforme es conocida como *ecualización del histograma* o *linealización del histograma*.

La forma discreta de la ecuación (A.9) se obtiene mediante la relación

$$s_k = T(r_k) = \sum_{j=0}^{k} \frac{n_j}{n} = \sum_{j=0}^{k} p_r(r_j) \qquad 0 \leq r_k \leq 1 \quad \text{y} \quad k = 0,1,\ldots,L-1 \tag{A.13}$$

La transformación inversa será:

$$r_k = T^{-1}(s_k) \qquad 0 \leq s_k \leq 1 \tag{A.14}$$

donde $T(r_k)$ y $T^{-1}(s_k)$ se supone que satisfacen las condiciones *a* y *b* anteriormente establecidas. La función de transformación $T(r_k)$ puede calcularse directamente de la imagen usando la ecuación (A.13). Aunque la función de transformación inversa (A.14) no se utiliza, representa un papel fundamental en la siguiente sección.

A.2.1.2. Especificación del histograma. Debido a que con el método de ecualización del histograma solo se genera una aproximación a un histograma uniforme, que es bastante útil, no es tan operativo en condiciones interactivas en la mayoría de las imágenes.

Algunas veces se desea especificar una forma de histograma concreta, capaz de realzar ciertos niveles de gris. Para ver cómo es posible esto, se retorna momentáneamente a los niveles de gris continuos. Sean $p_r(r)$ y $p_z(z)$ las funciones de densidad de probabilidad original y deseada, respectivamente. Si inicialmente se aplica la ecualización del histograma a la imagen original según la ecuación (A.9), y si la imagen deseada estuviera disponible, sus niveles también podrían ser ecualizados utilizando la función de transformación:

$$v = G(z) = \int_0^z p_z(w)dw \,. \tag{A.15}$$

Con el proceso inverso, $z = G^{-1}(v)$, se pueden calcular los niveles de gris, z, de la imagen deseada. Este planteamiento es hipotético, ya que los niveles z son precisamente lo que se está buscando. De todas formas, $p_s(s)$ y $p_v(v)$ tendrán densidades uniformes e idénticas porque el resultado final de la ecuación (A.9) es independiente de la densidad dentro de la integral. Por tanto, si en lugar de usar v en el proceso inverso, se utilizan los niveles uniformes s obtenidos de la imagen original, los niveles resultantes $z = G^{-1}(s)$ tendrán la función de densidad deseada. Suponiendo que $G^{-1}(s)$ tenga un único valor, el proceso puede ser resumido en los siguientes puntos:

a. Ecualizar los niveles de la imagen original utilizando la ecuación (A.9)
b. Especificar la función de densidad deseada, y obtener la función de transformación $G(z)$ utilizando la ecuación (A.15)

c. Aplicar la función de transformación inversa, $z = G^{-1}(s)$ a los niveles obtenidos en el punto *a*

Con este procedimiento se obtiene una versión procesada de la imagen original con los nuevos niveles de gris caracterizados por la densidad especificada por $p_z(z)$.

Aunque el método de la especificación del histograma implica dos transformaciones, $T(r)$ seguida de $G^{-1}(z)$, se pueden combinar los dos pasos en una función que permita obtener los niveles deseados partiendo de los píxeles originales. De la exposición anterior, se puede deducir:

$$z = G^{-1}(s) \tag{A.16}$$

Al sustituir la ecuación (A.9) en la ecuación (A.16) se obtiene la función de transformación combinada:

$$z = G^{-1}[T(r)] \tag{A.17}$$

que relaciona r con z. Cuando esta ecuación es igual a $T(r)$, la especificación del histograma es equivalente a su ecualización.

La importancia de la ecuación (A.17) radica en el hecho de que no se necesita explícitamente una ecualización del histograma de la imagen. Todo lo que se requiere es que $T(r)$ sea determinada y combinada con la función de transformación inversa $G^{-1}(z)$. El problema planteado al utilizar este método para variables continuas reside en la determinación analítica de la función inversa. En el caso discreto, este problema se reduce por el hecho de que el número de niveles de gris distintos es relativamente pequeño, y es factible de calcular y almacenar el valor para cada nivel de gris. Las técnicas para la formulación discreta de la especificación del histograma son paralelas a las ecuaciones (A.12) y (A.13).

En la práctica, la función de transformación inversa de s a z no siempre tiene valor único. Esta situación se presenta cuando quedan niveles de gris vacíos en el histograma especificado (que hacen que la función de distribución acumulada permanezca constante en estos intervalos) o en el proceso de redondeo de $G^{-1}(s)$ al nivel de gris más cercano. Generalmente, la solución más sencilla a este problema es asignar los niveles de tal forma que el histograma sea tan compacto como sea posible.

La principal dificultad en la aplicación de este método reside en la construcción de un histograma significativo. Se tienen dos soluciones a este problema: la primera es especificar una función de densidad particular (como una función de densidad gaussiana) y formar un histograma digitalizando la función dada. La segunda solución consiste en especificar una forma del histograma por medio de un dispositivo gráfico (como una pantalla interactiva) cuya salida alimente al procesador que realiza el algoritmo de especificación del histograma.

A.2.1.3. Mejora local. En ocasiones es necesario resaltar detalles sobre áreas pequeñas. En estos casos, el número de píxeles de dicha área puede tener una influencia insignificante en el cálculo de transformaciones a escala global en la imagen, por lo que los de métodos globales, como los vistos en las secciones anteriores, quedan descartados. La solución en estos casos es el uso de funciones de transformación basadas en el nivel de gris de la vecindad de cada píxel.

El procedimiento para adaptar las técnicas de procesamiento del histograma a la mejora local es sencillo. Basta con definir cuadrados o rectángulos centrados en cada uno de los píxeles y actuar en cada caso sobre el histograma de los mismos. Si se procesan los píxeles por orden de adyacencia, la obtención en cada caso del histograma es muy sencilla, ya que basta con añadir una fila o columna y eliminar otra. Una solución adicional para reducir el tiempo del proceso puede ser la utilización de regiones no superpuestas, pero este método generalmente produce efectos de borde indeseables.

A.2.2. Sustracción de imágenes. La sustracción o resta de imágenes se obtiene calculando la diferencia entre todos los pares correspondientes de los píxeles de dos imágenes. Esta operación tiene muchas aplicaciones importantes en la segmentación y en la mejora de imágenes. La diferencia entre dos imágenes, $f(x, y)$ y $h(x, y)$, se expresa como:

$$g(x, y) = f(x, y) - h(x, y). \tag{A.18}$$

Una aplicación clásica de la ecuación (A.18) se presenta cuando es necesario hacer estudios del movimiento ya que al restar dos imágenes consecutivas de una escena, los píxeles de valores altos indican las zonas donde los cambios fueron más drásticos. Se usa también para la eliminación de patrones de iluminación no uniforme o en el caso de que se quiera eliminar la información correspondiente al fondo de una imagen, si este es conocido.

A.2.3. Promediado de imágenes. Se considera una imagen $g(x, y)$ formada por la suma del ruido $n(x, y)$ a la imagen original $f(x, y)$, donde se supone que en cada par de coordenadas (x, y) el ruido es no correlacionado y que tiene un valor medio de cero. El objetivo de este método es reducir el efecto del ruido mediante la suma y promedio de una serie de M imágenes ruidosas, $\{gi(x, y)\}$. Se puede calcular, de esta forma el valor medio de $g(x, y)$ en cada punto (x, y), dado por

$$\overline{g}(x, y) = \frac{1}{M}\sum_{i=1}^{M} g_i(x, y) \tag{A.19}$$

de aquí que

$$E\{\bar{g}(x,y)\} = f(x,y) \tag{A.20}$$

y

$$\sigma^2_{\bar{g}(x,y)} = \frac{1}{M}\sigma^2_{n(x,y)} \tag{A.21}$$

donde E es valor de expectación del promedio de $g(x, y)$ y la ecuación (A.21) representa la varianza del valor medio de g y n. La desviación estándar en cada punto en la imagen media es por lo tanto

$$\sigma_{\bar{g}(x,y)} = \frac{1}{\sqrt{M}}\sigma_{n(x,y)} \tag{A.22}$$

Las ecuaciones (A.21) y (A.22) indican que tal y como M aumenta, la variación de valor del pixel en cada punto se decrementa, es decir, que el valor medio de g en el punto (x, y) se aproxima más a $f(x, y)$.

A.3. FILTRADO ESPACIAL

El uso de máscaras espaciales para el procesado de imágenes generalmente se denomina **filtrado espacial** en oposición al **filtrado en el dominio de la frecuencia** que usa la transformada de Fourier (que en el sentido estricto también es un filtrado espacial), y las máscaras se denominan **filtros espaciales**. En esta sección se consideran filtros lineales y no lineales para el mejoramiento de imágenes.

Los filtros lineales están basados en los conceptos definidos en la sección A.2. Los denominados **filtros paso bajo** atenúan o eliminan las componentes de alta frecuencia del dominio de Fourier, sin modificar las de baja frecuencia. Las componentes de alta frecuencia caracterizan los bordes y otros contornos de la imagen, por lo que el efecto de este filtro es un difuminado de la imagen. Análogamente, los **filtros paso alto** atenúan o eliminan las componentes de baja frecuencia, sin modificar las de alta frecuencia, con lo que se resaltan los bordes y contornos de la imagen. Un tercer tipo de filtrado, denominado **filtro pasa banda**, que atenúan o eliminan todas las frecuencias que no están incluidas en su región o banda.

Independientemente del tipo usado, los filtros lineales se reducen a realizar la suma de los productos entre los coeficientes de la máscara y la intensidad de los píxeles por debajo la máscara en una zona específica de la imagen. Denotando el nivel de gris de los pixeles bajo la máscara en cualquier lugar como *z1*, *z*2, ..., *z*9, la respuesta de la máscara lineal será

$$R = w_1 z_1 + w_2 z_2 + \ldots + w_9 z_9 \qquad \text{(A.23)}$$

Si el centro de la máscara está situado en el punto *(x, y)* de la imagen, el nivel de gris de dicho píxel se sustituye por R. Seguidamente la máscara se mueve hasta el siguiente píxel de la imagen y se repite el proceso.

Los filtros no lineales actúan sobre la vecindad de cada píxel. En general, sus operaciones se basan directamente en el valor de los píxeles en la vecindad bajo consideración, y no usan explícitamente los coeficientes en la forma descrita en la ecuación (A.23). Se puede lograr una reducción efectiva del ruido con un filtro no lineal cuya función básica sea calcular la mediana de los niveles de gris en la vecindad de cada píxel. Otro ejemplo es el filtro de *valor máximo* (con una respuesta $R = max\ \{z_k \mid k = 1, 2, ..., 9\}$), que se usa para la detección del punto más brillante de la imagen, y el filtro de *valor mínimo*, que se utiliza para encontrar el punto más oscuro de la imagen.

A.3.1. Filtros de suavizado. Se utilizan para difuminar (emborronar) la imagen o reducir el ruido. La difuminación se utiliza durante el preprocesado. Con este se eliminan detalles pequeños de una imagen antes del procesado de los elementos grandes, también se corrigen discontinuidades en líneas y curvas. Se puede conseguir la eliminación del ruido tanto con filtros lineales como con filtros no lineales.

A.3.1.1. Filtros pasa bajas. La forma de la respuesta al impulso necesaria para implementar un filtro pasa bajas espacial indica que el filtro debe tener todos los coeficientes positivos. Para un filtro de 3 x 3 elementos, la máscara más sencilla es aquella que tiene todos sus elementos con un valor de +1. De todas formas, en la ecuación (A.23) se observa que la respuesta será la suma de los nueve niveles de gris, por lo que R puede estar fuera del rango permitido de niveles de gris. La solución más sencilla es escalar la suma dividiendo R por 9. Se debe notar que en este caso la respuesta de R será el valor medio de todos los píxeles que se encuentran en el área de la máscara. Por esta razón, la aplicación de máscaras con esta configuración se denomina comúnmente **promediado en el entorno**.

A.3.1.2. Filtrado de mediana. Uno de los principales inconvenientes del método de suavizado descrito anteriormente es la difuminación de bordes y otros detalles marcados. Si el objetivo es conseguir una reducción del ruido, un sistema alternativo es la utilización de los filtros que utilizan la mediana, que consiste en la sustitución del nivel de gris de cada píxel por la mediana de los niveles de gris del entorno, en lugar de utilizar el valor medio. Este método es particularmente efectivo cuando el patrón del ruido está formado por picos muy pronunciados y se desea preservar la definición de los bordes. Tal y como se ha indicado antes, estos filtros son no lineales.

A.3.2. Filtros de realce (*sharpen*). El objetivo principal de este tipo de filtros es realzar los detalles finos de una imagen, o mejorar los detalles que han sido difuminados, por error o por efecto de la adquisición de la imagen.

A.3.2.1. Filtros pasa altas básicos. La forma de la respuesta al impulso necesaria para implementar un filtro pasa altas muestra que debe tener positivos todos los coeficientes adyacentes al centro, y negativos todos los de la periferia. Para una máscara de 3 x 3 elementos se cumple esta condición si se elige un valor positivo en su punto central y valores negativos para el resto de los puntos.

En este caso la suma de los coeficientes es cero. De este modo, cuando la máscara está sobre un área de niveles de gris constantes o con pequeñas variaciones, la salida de la máscara es cero o muy pequeña. Se observa también que este filtro elimina el término de frecuencia cero, por lo que el valor medio de los niveles de gris de la imagen se hace cero, reduciendo considerablemente el contraste global de la imagen.

La reducción del valor medio de la imagen a cero implica que la imagen obtenida debe contener algunos niveles de gris negativos. Como solo se opera con niveles de gris positivos, en el resultado del filtro pasa bajas se debe realizar una conversión o recorte de escala para que el resultado final esté dentro del rango de valores permitidos. Una opción no válida en este caso es tomar el valor absoluto de los niveles de gris del filtrado, ya que los valores muy negativos se convertirían en puntos brillantes.

A.3.2.2. Filtro *high-boost.* Se puede aplicar un filtro pasa altas a una imagen realizando la resta entre la imagen original y la imagen filtrada con un filtro paso bajo, esto es,

$$PasaAltas(HighBoost) = Original - PasaBajas \; . \tag{A.24}$$

Al filtro high-boost (estímulo de las altas) se le conoce también como filtro de énfasis de las frecuencias altas y hay varias maneras de aplicarlo al ponderar las diferentes componentes de los operadores, obteniendo diferentes énfasis en la imagen obtenida.

A.3.2.3. Filtros diferenciales. Al determinar la media de un píxel en una región se tienden a difuminar los detalles de la imagen. Establecer la media es equivalente a una integración, puede suponerse que la operación contraria, una diferenciación, que tendrá el efecto contrario y se conseguirá un filtrado que aumenta la nitidez y define la imagen.

El método más común de diferenciación en las aplicaciones de procesamiento de imágenes es la utilización del gradiente. Para una función $f(x, y)$, el gradiente de f en las coordenadas (x, y) se define como el vector

$$\nabla \mathbf{f} = \begin{bmatrix} \dfrac{\partial f}{\partial x} \\ \dfrac{\partial f}{\partial y} \end{bmatrix}. \tag{A.25}$$

La magnitud de este vector es la aproximación para varias propuestas en la diferenciación de imágenes y esta dada por

$$\nabla f = mag(\nabla \mathbf{f}) = \left[\left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2 \right]^{1/2}. \tag{A.26}$$

La ecuación **(A.26)** puede aproximarse en el punto *z5* de varias formas, pero la más sencilla es utilizando la diferencia *(z5 – z8)* en la dirección *x,* y la diferencia *(z5 – z6)* en la dirección *y*, de la siguiente forma:

$$\nabla f \approx \left[(z_5 - z_8)^2 + (z_5 - z_6)^2 \right]^{1/2} \tag{A.27}$$

En lugar del uso de cuadrados y raíces cuadradas, se puede obtener el mismo resultado con la utilización de valores absolutos,

$$\nabla f \approx |z_5 - z_8| + |z_5 - z_6|. \tag{A.28}$$

Otra proposición para la aproximación de la ecuación (A.26) es la utilización de diferencias cruzadas:

$$\nabla f \approx \left[(z_5 - z_9)^2 + (z_6 - z_8)^2\right]^{1/2}, \tag{A.29}$$

o con sus respectivos valores absolutos como en (A.28).

Como las máscaras de tamaños pares son difíciles de implementar, se buscan otras aproximaciones a la ecuación (A.26) utilizando máscaras de 3x3 elementos, así

$$\nabla f \approx \left|(z_7 + z_8 + z_9) - (z_1 + z_2 + z_3)\right| + \left|(z_3 + z_6 + z_9) - (z_1 + z_4 + z_7)\right| \tag{A.30}$$

La diferencia entre la tercera y la primera fila de la región de 3x3 elementos se aproxima a la derivada en la dirección x, y la diferencia entre la tercera y la primera columna se aproxima a la derivada en la dirección y. Las máscaras denominadas **operadores de Prewitt** pueden usarse para implementar la ecuación (A.30). esto también es válido para los **oereadores de Sobel** y **de Roberts**.

A.4. MEJORAMIENTO EN EL DOMINIO DE LA FRECUENCIA

El mejoramiento de imágenes en el dominio de la frecuencia, en principio, es un proceso sencillo. Simplemente se determina la transformada de Fourier de la imagen que debe ser mejorada, y se multiplica por el resultado por la función de transferencia del filtro. Posteriormente se determina la transformada inversa para generar la imagen mejorada. Las ideas de difuminar una imagen reduciendo su contenido de alta frecuencia, o aumentar la definición de los objetos mediante el incremento de las componentes de alta frecuencia en relación con las componentes de baja frecuencia, se obtienen de forma directa de la transformada de Fourier. De hecho, la idea de los filtros lineales es mucho más intuitiva en el dominio de la frecuencia.

En la práctica, las máscaras pequeñas se utilizan mucho más que la transformada de Fourier a causa de su simplicidad de implementación y rapidez de operación. De todas formas, la comprensión de los conceptos en el dominio de la frecuencia es fundamental para solucionar algunos de los problemas que no quedan bien definidos en el dominio del espacio.

A.4.1. Filtros pasa bajas. Como se indicó anteriormente, los bordes y otras transiciones marcadas en los niveles de gris de la imagen, como por ejemplo el ruido, contribuyen significativamente en las componentes de alta frecuencia de su transformada de Fourier. Por lo tanto, se puede conseguir una difuminación (suavizado) de la misma atenuando un rango específico de las componentes de alta frecuencia.

De la ecuación (A.4), $F(u, v)$ es la transformada de Fourier de la imagen que debe ser suavizada. El problema planteado es la selección de la función de transferencia del filtro, $H(u, v)$, tal que se atenúen las componentes de alta frecuencia de $F(u, v)$. En las secciones siguientes se considera que la función de transferencia de los filtros afecta por igual a la parte real y a la parte imaginaria de $F(u, v)$, por lo que estos filtros se pueden denominar *filtros de desplazamiento de fase cer*o, ya que no alteran la fase de la transformada.

A.4.1.1. Filtro ideal pasa bajas. Un filtro ideal pasa bajas en 2D es aquel cuya función de transferencia cumple la relación

$$H(u,v)=\begin{cases}1 & \text{si} \quad D(u,v)\leq D_0\\ 0 & \text{si} \quad D(u,v)> D_0\end{cases} \tag{A.31}$$

donde D_0 es una constante no negativa, y $D(u, v)$ es la distancia del punto (u, v) al origen del plano de la frecuencia, esto es

$$D(u,v) = \left(u^2 + v^2\right)^{1/2}. \tag{A.32}$$

Todas las frecuencias contenidas dentro del círculo de un radio determinado pasan sin atenuación, mientras que las frecuencias fuera del círculo se atenúan por completo. Este filtro es simétrico radialmente en torno al origen. Por tanto, se puede trazar una sección perpendicular al origen, en cualquiera de sus radios.

La **frecuencia de corte** para una sección de un filtro pasa bajas ideal, se define como el punto de transición entre $H(u, v) = 1$ y $H(u, v) = 0$. El concepto de la frecuencia de corte es muy importante para la definición de las características de un filtro, así como para la comparación del comportamiento entre diferentes filtros. La forma de la frecuencia de corte tan abrupto del filtro ideal no puede ser implementada con componentes electrónicos, aunque sí puede ser simulada en un computador.

A.4.1.2. Filtro de Butterworth (pasa bajas). La función de transferencia de un filtro pasa bajas de Butterworth de orden n y con una frecuencia de corte D_0 se define como la relación

$$H(u,v) = \frac{1}{1 + \left[D(u,v)/D_0\right]^{2n}} \tag{A.33}$$

donde $D(u, v)$ proviene de la ecuación (A.32). Diferente al filtro ideal, el filtro de Butterworth no tiene una discontinuidad abrupta que establezca claramente una frecuencia de corte. En este tipo de filtros se define un punto en el cual el valor de $H(u, v)$ es inferior a una fracción de su valor máximo. En el caso de la ecuación (A.33), se cumple que $H(u, v) = 0,5$ cuando $D(u, v) = D_0$. También es común el uso del valor $1/\sqrt{2}$, modificando la ecuación (A.33)de la forma

$$H(u,v)=\frac{1}{1+0.414\left[D(u,v)/D_0\right]^{2n}}\,. \tag{A.34}$$

A.4.2. Filtros pasa altas. Debido a que los bordes y otros tipos de cambios abruptos en el nivel de gris están asociados con las componentes de alta frecuencia de la imagen, se puede realizar un realce en los mismos utilizando un filtro pasa altas. Dicho filtro atenúa las componentes de baja frecuencia sin modificar la información de alta frecuencia contenida en la transformada de Fourier de la imagen. Igual que con los filtros pasa bajas, se consideran únicamente los filtros de desplazamiento de fase cero, que son radialmente simétricos y pueden definirse completamente mediante una sección, convirtiéndola en una función de la distancia al origen de la transformada de Fourier centrada.

A.4.2.1. Filtro ideal pasa altas. El filtro ideal pasa altas es aquel cuya función de transferencia cumple la siguiente relación:

$$H(u,v)=\begin{cases}0 & \text{si} \quad D(u,v)\le D_0\\ 1 & \text{si} \quad D(u,v)> D_0\end{cases} \tag{A.35}$$

donde D_0 es la distancia de corte medida desde el origen del plano de la frecuencia, y $D(u, v)$ se obtiene de la ecuación (A.32). Este filtro es el opuesto al filtro ideal paso bajo mostrado en la sección A.4.1.1, ya que atenúa completamente las frecuencias contenidas dentro del círculo de radio D_0, mientras que todas las frecuencias de fuera del círculo pasan sin atenuación. Igual que con el filtro paso bajo ideal, este filtro es físicamente irrealizable.

A.4.2.2.Filtro de Butterworth (pasa altas). La función de transferencia del filtro pasa altas de Butterworth de orden n y con una frecuencia de corte a una distancia D_0 del origen se define como la relación

$$H(u,v)=\frac{1}{1+\left[D_0/D(u,v)\right]^{2n}}\,, \tag{A.36}$$

donde *D(u, v)* se obtiene de la ecuación (A.32).

Se debe notar que cuando $D(u, v) = D_0$, el valor de *H(u, v)* está por debajo de la mitad de su valor máximo. Como en el caso del filtro paso bajo de Butterworth, la práctica más común es la selección de la frecuencia de corte de forma que el valor de *H(u, v)* quede por debajo de $1/\sqrt{2}$ de su valor máximo. La ecuación (A.36) se puede modificar de forma sencilla para que satisfaga esta condición, quedando de la siguiente manera:

$$H(u,v)=\frac{1}{1+0.414\left[D_0/D(u,v)\right]^{2n}} \tag{A.37}$$

A.5. GENERACIÓN DE MÁSCARAS ESPACIALES A PARTIR DE LAS ESPECIFICACIONES EN EL DOMINIO DE LA FRECUENCIA

Como ya se ha indicado, la velocidad y la simplicidad de implementación son características fundamentales de las máscaras espaciales utilizadas en el procesamiento de imágenes. En esta sección se expone un método para la generación de máscaras espaciales que se aproximan a un filtro definido en el dominio de la frecuencia.

Del teorema de la convolución, se sabe que la ecuación (A.4) puede ser implementada en el dominio del espacio mediante la ecuación

$$g(x,y)=\sum_{i=0}^{N-1}\sum_{k=0}^{N-1}h(x-i,\,y-k)f(i,k) \tag{A.38}$$

donde $x, y = 0, 1, 2, \ldots, N - 1$. Para simplificar la notación, se utilizan imágenes cuadradas. En la ecuación (A.38), h es la representación espacial del filtro, f es la imagen de entrada y g la imagen de salida filtrada. Si la máscara es de tamaño $N \times N$, el resultado obtenido en la ecuación (A.38) es idéntico a la transformada inversa de Fourier de $G(u, v)$ en la ecuación (A.4). Como H es la transformada de Fourier de h, se puede afirmar que

$$H(u,v) = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} h(x,y)\exp\left[-j2\pi\frac{ux+vy}{N}\right] \tag{A.39}$$

para $u, v = 0, 1, 2, \ldots, N - 1$, Supóngase, por tanto, que $h(x, y)$ se restringe a cero para valores de $x > n$ e $y > n$, con $n < N$. Esta restricción crea una máscara de convolución $\breve{h}$ de $n \times n$ elementos, cuya transformada de Fourier es

$$\breve{H}(u,v) = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} \breve{h}(x,y)\exp\left[-j2\pi\frac{ux+vy}{N}\right] \tag{A.40}$$

para $u, v = 0, 1, 2, \ldots, N - 1$. El objetivo del proceso es encontrar los coeficientes de $\breve{h}$ tal que el error

$$e^2 = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1} \left|\breve{H}(u,v) - H(u,v)\right|^2 \tag{A.41}$$

se minimice, donde el operador de barras paralelas indica el módulo complejo.

La ecuación (A.40) se puede expresar de forma matricial

$$\breve{\mathbf{H}} = \mathbf{C}\breve{\mathbf{h}} \tag{A.42}$$

donde $\hat{\mathbf{H}}$ es el vector columna de orden N^2 que contiene los elementos de $\hat{H}(u,v)$, así mismo, $\hat{\mathbf{h}}$ el vector columna de orden n^2 que contiene los elementos de $\hat{h}(x,y)$, y $\mathbf{C}$ es una matriz de N^2 x n^2 de términos exponenciales cuya posición está determinada por el orden en $\hat{\mathbf{H}}$ y en $\hat{\mathbf{h}}$. Un procedimiento simple para generar los elementos de $\hat{H}(i)$, $i = 0, 1, 2, ..., N^2 - 1$, del vector $\hat{\mathbf{H}}$ a partir de $\hat{H}(u,v)$ es

$$\hat{H}(u,v) \Rightarrow \hat{H}(i) \tag{A.43}$$

para $i = uN + v$, con $u, v = 0, 1, 2, ..., N - 1$. Procesando cada fila de $\hat{H}(u,v)$, y asignando los valores a u y v, se forman los elementos a partir de las filas consecutivas. Los elementos de $\hat{\mathbf{h}}$, denotados como $\hat{h}(k)$, $k = 0, 1, 2, ..., n^2 - 1$, se obtienen de forma similar:

$$\hat{h}(x,y) \Rightarrow \hat{h}(k). \tag{A.44}$$

Finalmente, los elementos de la matriz $\mathbf{C}$, denotados como $C(i, k)$, se generan a partir de los términos exponenciales

$$\frac{1}{N}\exp\left[-j2\pi\frac{ux+vy}{N}\right] \Rightarrow C(i,k) \tag{A.45}$$

para $i = uN + v$ y $k = xn + y$, con $u, v = 0, 1, 2, ..., N - 1$, y $x, y = 0, 1, 2, ..., n - 1$. En notación matricial, la ecuación (A.41) se transforma en

$$e^2 = (\hat{\mathbf{H}} - \mathbf{H})*(\hat{\mathbf{H}} - \mathbf{H}) = \left\|\hat{\mathbf{H}} - \mathbf{H}\right\|^2 = \left\|\mathbf{C}\hat{\mathbf{h}} - \mathbf{H}\right\|^2 \tag{A.46}$$

donde * simboliza la traspuesta conjugada, el operador de barras dobles es el valor de la norma euclidiana compleja, y $\mathbf{H}$ es el vector formado por $H(u, v)$. Determinando la derivada parcial e igualándola a cero, se obtiene el valor mínimo de e^2 con respecto a $\hat{\mathbf{h}}$:

$$\frac{\partial e^2}{\partial \hat{\mathbf{h}}} = 2\mathbf{C}^*\left(\mathbf{C}\hat{\mathbf{h}} - \mathbf{H}\right) = 0 \tag{A.47}$$

y

$$\hat{\mathbf{h}} = \left(\mathbf{C}^*\mathbf{C}\right)^{-1}\mathbf{C}^*\mathbf{H} = \mathbf{C}^{\#}\mathbf{H} \tag{A.48}$$

donde la matriz $\mathbf{C}^{\#} = (\boldsymbol{C}^*\boldsymbol{C})^{-1}\ \boldsymbol{C}^*$, se denomina generalmente *matriz inversa generalizada de Moore-Penrose*. La ecuación (A.48) da los coeficientes necesarios para construir una máscara de convolución $\hat{h}(x, y)$ de *n x n* elementos con el mínimo error, a partir de una función de filtro, *H(u, v)* de *N x N* elementos, especificada en el dominio de la frecuencia. En general, los elementos de $\hat{H}(u, v)$ son cantidades complejas. De todos modos, si la función del filtro en el dominio de la frecuencia es real y simétrica, $\hat{h}(x, y)$ será también real y simétrico.

BIBLIOGRAFÍA

¬ González, R., Woods, R. (1996). Tratamiento digital de imágenes. Addison-Wesley/Días de Santos. Delaware.

¬ Castleman, K. (1996). Digital image processing. Prentice Hall, Englewood Cliffs. New Jersey.