UNIVERSIDAD TECNOLÓGICA DE LA MIXTECA

División de Estudios de Postgrado

Maestría en Modelación Matemática

Dimensión fractal y transformada de Gabor en el problema de escalabilidad para la identificación biométrica del iris del ojo

TESIS

QUE PARA OBTENER EL GRADO DE:

Maestro en Modelación Matemática

PRESENTA:

JUAN FELIPE MONTESINOS IBÁÑEZ

DIRECTOR:

Dr. Tomás Pérez Becerra

CO-DIRECTOR:

Dr. Salvador Sánchez Perales

HUAJUAPAN DE LEÓN, OAXACA *OCTUBRE DE 2023*

Dedicatoria

A mis padres, por su amor, trabajo y sacrificio en estos años, gracias a ustedes he logrado llegar hasta aquí y convertirme en lo que soy. A mis hermanos por estar siempre presentes, acompañandome y por el apoyo moral, que me brindaron a lo largo de esta etapa de mi vida. A todas las personas que nos han apoyado y han hecho que el trabajo se realice con éxito en especial a aquellos que nos abrieron las puertas y compartieron sus conocimientos.

Agradecimientos

A mi familia, por haberme dado la oportunidad de formarme en esta prestigiosa universidad y haber sido mi apoyo durante todo este tiempo, a mis profesores quienes con la enseñanza de sus valiosos conocimientos hicieron que pueda crecer día a día como profesional, gracias a cada uno de ustedes por su paciencia, dedicación, apoyo incondicional y amistad.

De manera especial quiero expresar mi más grande y sincero agradecimiento al Dr. Tomas Pérez Becerra y Dr. Salvador Sánchez Perales, quienes con sus conocimientos, enseñanzas y colaboraciónes permitieron el desarrollo de esta tesis.

A la Universidad Tecnológica de la Mixteca, por haberme brindado tantas oportunidades y enriquecerme en conocimiento.

Índice general

Introducción **VII**

El iris como identificador biométrico VIII

La dimensión fractal en la biometría del iris del ojo X

Sobre las transformaciones que preservan la dimensión fractal XII

Redes neuronales en el reconocimiento biométrico XIV

1. Preliminares sobre transformaciones que conservan la dimensión fractal 1

1.1. Dimensión fractal 1

1.2. Dimensión de Hausdorff 3

1.3. Dimensión de Conteo de Cajas 5

1.4. Transformaciones bi-Lipschitz 9

1.5. Transformaciones DP 10

2. Preservación de la dimensión fractal bajo transformaciones y filtros de

Gabor **13**

2.1. Transformaciones reflexión, rotación y normalización 14

2.2. Filtros de Gabor y su preservación de la dimensión 21

2.2.1. Filtro de Gabor que no preserva la dimensión 24

2.3. Preservación de la dimensión bajo funciones campana de Gauss 26

2.3.1. Filtros de Gabor en espacios euclidianos 35

3. Modelo de reconocimiento biométrico con etapa de dimensión fractal 45

3.1. Aumento de datos (Data Augmentation) 46

3.2. Normalización de las imágenes 48

3.3. Filtrado de imágenes y su dimensión 49

3.4. Modelo de redes neuronales para el reconocimiento 53

3.4.1. Redes neuronales convolucionales 53

3.4.2. Entrenamiento por subclases 55

Conclusiones **59**

Referencias **63**

Introducción

La verificación de la identidad personal ha sido parte relevante en el desarrollo de toda sociedad y surge como una necesidad básica para las relaciones humanas. Cualquier individuo posee características fisiológicas y de comportamiento que lo identifican y lo hacen único. A lo largo de la historia de la civilización, el hombre ha empleado estas peculiaridades para realizar la identificación. La utilización del rostro y de las huellas dactilares para este propósito inició en el año 500 A.C., en ese entonces el reconocimiento constaba del simple hecho de recordar la cara de alguien. Las culturas China y Egipcia también emplearon las impresiones de los dedos ([40]).

El primer acercamiento a un procedimiento de identificación de personas se dio en Francia en 1880; Alphonse Bertillon, criminólogo francés, con el propósito de poder identificar a los delincuentes, creó un método científico que permite reconocer a criminales por medio de ciertas medidas fisiológicas, conocido como *sistema Bertillon* ([40]). Esta metodología se basa en el hecho de que, a pesar de que las dimensiones de un hueso de un individuo coincidan con la de otro, es muy poco probable que la estructura ósea de dos humanos sea exactamente igual. Tal idea da nacimiento a la antropología judicial (también llamada forense).

Luego, la primera aproximación científica al uso de las marcas dactilares como iden-

tificador biométrico fue realizada en 1880 por el médico británico Henry Faulds. Acompañando a un amigo a una excavación arqueológica, notó que se podían identificar las delicadas formas dejadas por artesanos en fragmentos antiguos de cerámica. Examinando sus propias líneas en los dedos y las de sus amigos, se convenció de que los patrones eran únicos en cada individuo. Tiempo después propuso utilizar tinta para registrarlas. Este método tenía dos grandes ventajas: los sujetos tienen huellas únicas e irrepetibles y estas permanecían inalterables toda la vida.

Desde principios del siglo XXI, el término biometría se ha utilizado de manera rigurosa para referirse a la identificación de personas en función de sus características biológicas. Se define como el estudio de métodos automáticos para verificar una identidad o reconocer a un individuo vía uno o más distintivos corporales, empleando técnicas matemáticas y estadísticas. Estos rasgos físicos se llaman identificadores biométricos, entre los cuales se destacan: las huellas digitales, el iris, el rostro, la retina, la voz y la geometría de la palma de la mano ([40]).

El iris como identificador biométrico

La utilización de patrones del iris como método de identificación fue propuesto en 1936 por el oftalmólogo Frank Burch. Para la década de 1980, la metodología ya había aparecido en películas de James Bond, pero permanecía siendo ciencia ficción. En 1985 los Doctores Leonard Flom y Aran Safir retomaron la idea y su investigación y documentación les concedió una patente en 1987, ellos fueron los pioneros en sugerir que dos iris no son iguales. En 1993, John Daugman ([12]) presentó el primer sistema de identificación del iris a través de la *transformada de ondas de Gabor*, su razonamiento se basó en transformar la imagen en un conjunto de 256 bytes que se denomina *Código de Iris*; luego, comparó la discrepancia de los códigos, mediante la *distancia de Hamming*, y observó la similitud entre dos imágenes ([14]). En 1996, Wildes sugirió el proceso de reconocimiento automático, el cual resultó ser un procedimiento de alta complejidad computacional ([46], [45]). Tiempo

después, Boles construye un algoritmo que utiliza una transformación de ondas ([3]). A su vez, Zhu, Tan y Sánchez también presentaron otras técnicas ([31], [28] y [35]). La mayoría de estos procesos se basan en el análisis multiescala, esto es, métodos que consideren diversas características para realizar la búsqueda de los datos correspondientes al usuario.

Algunas ventajas de utilizar el iris como identificador biométrico son:

- Se trata de un órgano interno que está protegido contra daños y el desgaste por una membrana transparente y altamente sensible (la córnea). Esto lo distingue de las huellas digitales, las cuales pueden ser difíciles de reconocer después de realizar trabajos manuales durante años.
- El iris es una estructura que presenta una elevada variabilidad en toda su superficie, ello hace que, analizando tan solo una parte del mismo, se pueda obtener una gran cantidad de información.
- El iris tiene una textura fina que, como las huellas dactilares, se determina al azar durante el proceso embrionario de gestación. Incluso individuos genéticamente idénticos tienen texturas completamente diferentes.
- Por lo general, su uso es bastante sencillo y cómodo para cualquier persona.
- En los artículos [9] y [18] se propone que utilizar el iris como identificador biométrico tiene un desempeño superior al de otros identificadores, por ejemplo, el rostro o la voz.

Por otro lado, el iris presenta algunas desventajas que dificultan su uso en sistemas de reconocimiento ([8],[27]):

- **Tamaño.** Es pequeño, aproximadamente 1 cm de diámetro.
- **Ubicación.** Es un objeto móvil localizado en una superficie curva, húmeda y reflexiva.

- **Oclusión.** Su imagen suele estar afectada por pestañas, párpados y reflejos de luz, limitando el área de interés la cual contiene la información del iris.

Se sabe que existe una gran aleatoriedad en el patrón del iris, por lo que algunos investigadores han realizado estudios para probar su unicidad (por ejemplo [16]). En [11] Daugman reportó que la probabilidad de que falle el reconocimiento es de 1 en 200 billones, este dato fue extraído de la investigación efectuada a la base de datos del sistema de seguridad para cruzar la frontera de los Emiratos Árabes Unidos, la cual cuenta con 632,500 registros de iris diferentes. Asimismo, indicó que no pueden existir dos iguales, inclusive los de un mismo individuo discrepan, aún entre hermanos gemelos. Esto genera una alta efectividad en la autenticación.

La dimensión fractal en la biometría del iris del ojo

En el artículo [22], Jampour y Naserasari mencionan que, a pesar de todas las ventajas que ofrece la identificación biométrica por iris de ojo, existe el problema de que los métodos propuestos hasta ese momento presentan una alta complejidad computacional para un banco de datos mayor de 100,000 registros, tal problema se denomina como problema de escalabilidad. Además, en ese mismo trabajo, sugieren que la dimensión fractal tiene el potencial para utilizarse en la mejora de la eficiencia en un reconocimiento biométrico, debido a que permite descomponer un catálogo relativamente grande de individuos en conjuntos cuya dimensión se encuentre dentro de un intervalo de longitud pequeña, por lo que la base se puede seccionar en tantas porciones como se requiera.

La dimensión fractal fue introducida por primera vez por Mandelbrot en [33], en ese mismo trabajo define a una cierta clase de figuras geométricas irregulares que no encajan en la teoría clásica de Euclides, bautizados como fractales. Estas formas son conjuntos con dimensión de Hausdorff estrictamente mayor que la dimensión topológica, esto es, es

un número racional no entero. Tal valor numérico es una de las características intrínsecas que posee una figura de este tipo.

La idea de utilizar la dimensión fractal en la identificación que utiliza al iris del ojo aparece en el trabajo [23], en el cual los autores de ese artículo extraen el disco del iris y eliminan cualquier característica adicional, por ejemplo la pupila, luego, lo transforman a coordenadas rectangulares (esta etapa se llama *normalización*) y es en este punto donde calculan la dimensión fractal para proponerla como una variable clasificatoria. En la Figura 1 se observa tal proceso.

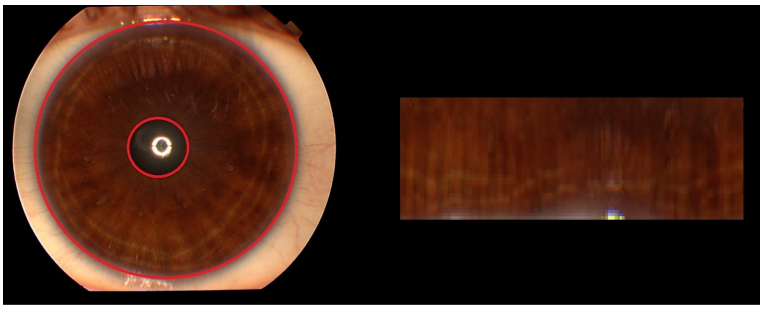

Figura 1: Extracción y normalización del iris.

Sin embargo, durante el desarrollo del proyecto *"Implementación de la dimensión fractal en métodos de identificación biométrica del iris del ojo"* con registro DEP/F-001 ante la División de Estudios de Postgrado de la Universidad Tecnológica de la Mixteca, se identificó que los métodos son más precisos si la clasificación por dimensión fractal se introduce en alguna etapa siguiente a la aplicación de una transformación para eliminar cualquier posible factor que afecte en la detección, comúnmente llamado *filtrado*, al contrario de lo propuesto por Jampour y Naserasari, quienes la proponen al inicio. En esta investigación se propone utilizar la función de Gabor bidimensional para realizar el filtrado de las imágenes posterior a su normalización (ver la Figura 2).

Figura 2: Imagen filtrada con la función de Gabor.

Sobre las transformaciones que preservan la dimensión fractal

Como se mencionó anteriormente, en el procesamiento de imágenes se utilizan algunas transformaciones para obtener de ellas la información necesaria. Estas pueden ser rotaciones, traslaciones, normalización, o la aplicación de ciertas funciones especiales, por ejemplo la función de Gabor paramétrica, definida como el producto de una Gaussiana con una exponencial compleja.

El filtro de Gabor se ha utilizado en varios aspectos, por ejemplo, para describir la respuesta de las células corticales, en la visión humana, análisis de documentos, mejora y compresión de imágenes, estudio y discriminación de texturas, entre otros (véase [6]).

Por otro lado, la dimensión fractal es un parámetro que se ha utilizado para explicar algunas características de varios fenómenos. Por mencionar algunos, en el estudio del universo (por ejemplo, en [5] se mencionan las erupciones solares, granulación solar, la morfología de cráteres, distribución de regiones específicas en galaxias y medio interestelar), reacciones electroquímicas ([19]), en la retina humana ([42]), biología ([4]), antropología ([30]) y medicina ([29]).

Los filtros de Gabor también se han implementado en el reconocimiento biométrico a través del iris del ojo. Por ejemplo, en [11] y [15] Daugman propone un sistema de identificación de usuarios basado en filtrar y transformar la imagen del iris, con lo que

registra la patente [13]. En [22] proponen una agrupación jerárquica mediante la dimensión fractal tras la normalización de la imagen obtenida del ojo. Para implementarlo, el proceso debe comenzar con la extracción del disco del iris y su transformación rectangular, luego se aplica algún filtro de Gabor para mantener sólo la información relevante para identificar al individuo, al final se introduce una clasificación fractal en la máquina o método de aprendizaje profundo para localizar al sujeto de quien se obtuvo la imagen del ojo. Sin embargo, una imagen al pasar por las distintas transformaciones puede perder el valor de la dimensión original, como se demuestra en la Sección 2.3.1 a través del Ejemplo 2.2.2 de esta tesis, el cual muestra un filtro que no preserva la dimensión fractal.

Modificar la dimensión fractal tal vez no es tan grave en muchos casos, pero en otros podría tener grandes repercusiones; por mencionar alguno, en análisis de suelos (recomendamos la referencia [44] para un contexto histórico sobre esta aplicación): en [39] se estudió la dimensión fractal de tres superficies porosas, se mostró que la dimensión de entropía de la porosidad de uno de ellos es de 0.85 y la dimensión fractal de 0.83, entonces plantean la pregunta "¿Es una coincidencia?". En ese artículo, finalmente, los autores determinan que la dimensión fractal puede considerarse como un indicador de las propiedades físicas fundamentales del suelo. Por otro lado, en [36] se presenta un esquema fractal para extraer y medir algunas características geométricas de la porosidad de yacimientos naturalmente fracturados del sureste de México, utilizando los valores medios de las dimensiones fractales extraídas de imágenes digitales obtenidas mediante tomografía computarizada de rayos X, y concluyen que todos mostraron una correlación estadísticamente significativa con la porosidad de los estratos geológicos, esto es, que la dimensión fractal coincide con la porosidad del sustrato. Para calcular el valor de la dimensión, la imagen se transforma en un gráfico de barras en tonos grises, si la dimensión se altera por este cambio, posiblemente cause que un medio se confunda con otro y acarrear serios problemas, por ejemplo, en exploración petrolera.

En el estudio de transformaciones que no alteran la dimensión fractal, se sabe que si es bi-Lipzchitz, entonces conserva la dimensión del conjunto de la preimagen (vea [1]). Esta

condición se debilita si el dominio se puede descomponer en una unión contable de conjuntos sobre los cuales es bi-Lipschitz. Por ejemplo, en el Teorema 2.3.8 proporcionamos una función no bi-Lipzchitz que preserva la dimensión fractal, por lo tanto, la condición bi-Lipzchitz es suficiente pero no necesaria.

A su vez, en el Ejemplo 2.2.2 se demuestra que la función filtro de Gabor no preserva la dimensión fractal. Por lo que uno de los objetivos de esta investigación es determinar algunas otras condiciones necesarias para que los filtros de Gabor no alteren este valor. Así, en la Sección 2.3.1 se realiza un análisis y se proporcionan tales las condiciones, las cuales son impuestas sobre los parámetros, con los que se asegura que se mantiene la dimensión.

Redes neuronales en el reconocimiento biométrico

Finalmente, al garantizar que la dimensión fractal no se alterará al realizarle las transformaciones mencionadas, se da paso al reconocimiento mediante algún método de aprendizaje profundo. Para esta investigación, se eligió las redes neuronales convolucionales, debido al buen desempeño que han mostrado en reconocimiento de imágenes. Por otro lado, se utilizará la base de fotografías *Phoenix* ([17]), la cual cuenta con 128 capturas de ojos tanto derechos como izquierdo de 64 personas, esta base se encuentra alojada en el sitio de la Univerzita Palackého v Olomouci en República Checa y es de acceso libre.

La propuesta de esta investigación es la de realizar el entrenamiento con un número pequeño de individuos y con una alta replicabilidad de cada iris, esto es, considerar al menos 50 imágenes de un solo ojo, para lograr esta cantidad, se toma la imagen de un ojo y se le aplica algunas técnicas de *Data Augmentation*, específicamente, la imagen original se rota, se acerca, se aleja, se refleja y se hacen combinaciones de estas transformaciones, este proceso se puede ver en la Sección 3.1. Posteriormente, el entrenamiento se realizará para un número fijo de iris con dimensión fractal relativamente "cercana" entre ellas, con

lo que el tiempo de ejecución se disminuirá al mínimo, esto será posible debido a que se dividirán cada grupo de imágenes respecto a su dimensión tal como lo sugiere Jampour y Naserasari, solo que con una variante mostrada a detalle en la Sección 3.4.2. Con esto se da una solución parcial al problema de escalabilidad, sin embargo, aún persiste esta problemática, debido a los tiempos que se requieren para realizar el entrenamiento de la red, este análisis detallado se encuentra a manera de discusión en la Sección 3.4.2.

En síntesis, la tesis se encuentra estructurada de la siguiente forma:

Capítulo 1. Se proporcionan los preliminares necesarios para el análisis de los capítulos posteriores. Entre ellos se encuentran la dimensión fractal, las transformaciones bi-Lipschitz y un estudio sobre las transformaciones que preservan la dimensión fractal.

Capítulo 2. Se realiza un análisis sobre la preservación de la dimensión fractal de las trasformaciones: reflexión; rotación; acercamiento; alejamiento y normalización. Se demuestra que los filtros de Gabor no preservan la dimensión en general y, de forma esquemática, se introducen las funciones campana reales y se extiende el estudio hasta los filtros de Gabor bi-dimensionales, como consecuencia de este proceso se determinan las condiciones necesarias respectivas para que se preserve la dimensión. En la investigación bibliográfica realizada no se hallaron los resultados 2.1.1, 2.1.2, 2.1.3,2.1.5,2.1.7,2.3.2, 2.3.3, 2.3.5, 2.3.6, 2.3.8, 2.3.9, 2.3.15 y 2.3.16, por lo que son originales y se encuentran en el artículo [38].

Capítulo 3. En este capítulo se muestra el proceso de *data augmentation* para crear nuevas imágenes, luego se analiza la función de la red neuronal convolucional en el reconocimiento biométrico posterior a su entrenamiento en una clase específica cuyos elementos comparten cierta "cercania" de la dimensión fractal. Se finaliza el capítulo con una discusión sobre el punto de escalabilidad alcanzado con la metodología propuesta en esta investigación.

Capítulo 1

Preliminares sobre transformaciones que conservan la dimensión fractal

El objetivo de este capítulo es introducir los fundamentos básicos que son utilizados en el desarrollo de la tesis, en particular, los relacionados con la dimensión de Hausdorff y la dimensión de conteo de cajas; en específico, la primera da las bases teóricas de la dimensión fractal y la segunda proporciona la técnica utilizada para su cálculo.

1.1. Dimensión fractal

Mandelbrot en [32] introduce la palabra fractal, a partir del adjetivo latino *fractus*, que significa 'interrumpido' o 'irregular', y la utiliza para referirse a las figuras que tienen una forma irregular cuya estructura se repite a diferentes escalas, pero principalmente a objetos con dimensión no entera, esto es, puede ser una fracción simple como $\frac{1}{2}$ o $\frac{5}{3}$, e incluso un número irracional, por ejemplo π, e o $\sqrt{2}$. Así, se puede decir que algunas curvas tienen dimensión entre 1 y 2, o que, para ciertas superficies su dimensión está entre 2 y 3, y finalmente definir "polvo" con dimensión entre 0 y 1. Cuando una dimensión no

es entera se le llama *dimensión fraccionaria.*

Se puede verificar que los puntos tienen dimensión topológica 0; las líneas, 1; las superficies, 2 y los volúmenes, 3. Sin embargo, en geometría fractal, las formas son tan complejas que casi «cubren» el espacio en el que se encuentran, por ejemplo, es posible que una línea sea prácticamente una superficie (ver Figura 1.1); por esta razón, se considera a la irregularidad como un incremento fraccionario de la dimensión topológica. En este contexto, es viable que una curva posea un área en un rango dimensional entre 1 y 2, así mismo, una superficie tenga volumen entre 2 y 3, a este valor se le conoce como la dimensión fractal y, formalmente, se define la siguiente manera:

Definición 1.1.1. Sea (X, τ) un espacio topológico y $E \subseteq X$. Consideremos $N(E, \epsilon)$ como el número mínimo de bolas abiertas de radio ϵ necesarias para cubrir el conjunto E. La *dimensión fractal* (o *de similitud*), $\dim_{sim}$, se define como

$$\dim_{sim}(E) = \lim_{\epsilon \to 0} \frac{\log N(E, \epsilon)}{\log \frac{1}{\epsilon}} = \lim_{\epsilon \to 0} -\frac{\log N(E, \epsilon)}{\log \epsilon},$$

siempre que el límite exista.

En [37] se aborda la idea intuitiva que sustenta la Definición 1.1.1 de dimensión fractal en forma de límite.

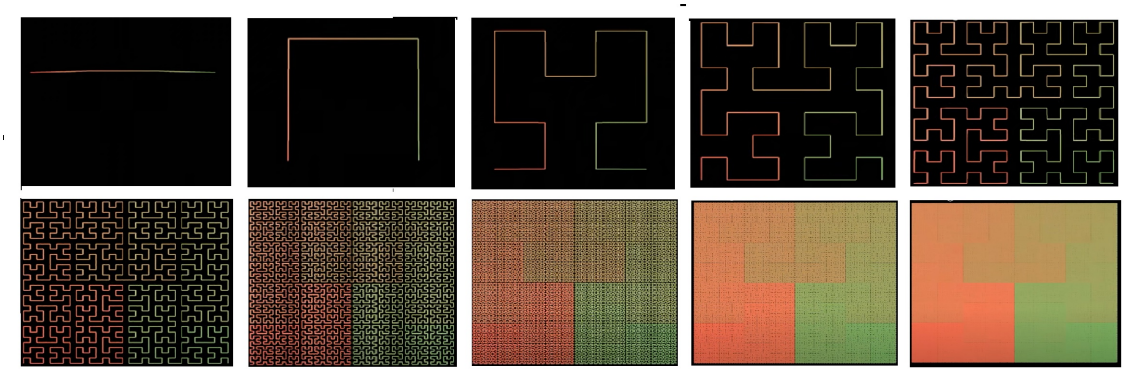

Figura 1.1: Distintas iteraciones de la curva de Hilbert, considerada una curva que tiene área. Imagen adaptada de *Encanto fractal: curvas que recubren el plano.* https://matemelga.wordpress.com/2018/07/25/curvas-fractales-que-rellenan-el-plano/

1.2. Dimensión de Hausdorff

Definición 1.2.1. Sea X un conjunto. Una función $d : X \times X \to \mathbb{R}$ es una *métrica* o *distancia* sobre X si cumple con las siguientes condiciones, para cualesquiera $x, y, z \in X$:

1. $d(x, y) \geq 0$.
2. $d(x, y) = 0$ si y solo si $x = y$.
3. $d(x, y) = d(y, x)$.
4. $d(x, z) \leq d(x, y) + d(y, z)$.

Al par (X, d) se le llama *espacio métrico* y al número $d(x, y)$ se le llama distancia de x a y.

Definición 1.2.2. Para un número real $p \geq 1$, la p-norma de $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$ está definida por

$$\|x\|_p = \sqrt[p]{|x_1|^p + |x_2|^p + \cdots |x_n|^p}.$$

Se sabe que estas normas son equivalentes entre sí. En esta investigación se considerará el espacio euclidiano dotado de la norma 2 ($p = 2$).

En lo subsecuente, se mostrará la construcción de la dimensión de Hausdorff. Se conoce que esta dimensión es equivalente a la fractal [37] y a su vez a la dimensión de conteo de cajas, por lo que en lo sucesivo al mencionar el concepto de dimensión fractal haremos referencia a la dimensión de Hausdorff. Referimos al lector a la tesis [37] para un estudio detallado de las equivalencias mencionadas.

Definición 1.2.3. Sean $X = \mathbb{R}^n$ con la métrica usual d y $U \subseteq X$ no vacío. El *diámetro* de U se denota por $|U|$ y se define como

$$|U| = \sup\{d(x, y) : x, y \in U\}.$$

Cuando $U = \emptyset$ se considera $|U| = 0$.

Definición 1.2.4. Sean (X, τ) un espacio topológico, K un subconjunto de X. Una familia $\mathcal{U} = \{U_\alpha : \alpha \in I\}$ es una *cubierta de* K, si $K \subseteq \bigcup_{\alpha \in I} U_\alpha$.

Definición 1.2.5. Sean $\delta \in \mathbb{R}$ y $E \subseteq \mathbb{R}^n$ tal que $\delta > 0$. Una colección $\{U_i\}_{i \in \mathbb{N}}$ de subconjuntos de $\mathbb{R}^n$ es un δ-*recubrimiento* de E, si $\{U_i\}_{i \in \mathbb{N}}$ es una cubierta de E y $|U_i| < \delta$ para todo $i \in \mathbb{N}$.

Definición 1.2.6. Sea X un conjunto. Se denota por 2^X a la colección de subconjuntos X.

Definición 1.2.7. Sean $X = \mathbb{R}^n$ con la métrica usual d, $s \in [0, \infty]$ y $\delta > 0$. Sea $E \subset X$, se define $H^s_\delta : 2^{\mathbb{R}^n} \to [0, \infty]$, como

$$H^s_\delta(E) = \text{ínf} \left\{ \sum_n |U_n|^s : \{U_n\}_{n \in \mathbb{N}} \text{ es un } \delta\text{-recubrimiento de } E \right\}.$$

Se sabe que H^s_δ es una medida exterior.

Definición 1.2.8. Sea $E \subset \mathbb{R}^n$ y s un número no negativo, la medida *s-dimensional de Hausdorff de* E es

$$H^s(E) = \lim_{\delta \to 0} H^s_\delta(E).$$

Proposición 1.2.9. Sean $q \in \mathbb{R}^+$ y $E \subset \mathbb{R}^n$. Luego existe un único $s \in [0, \infty]$ tal que

$$H^q(E) = \begin{cases} \infty, & \text{si } \ q \in [0, s); \\ 0, & \text{si } \ q \in (s, \infty). \end{cases}$$

Definición 1.2.10. Sea E un subespacio de $\mathbb{R}^n$. La *dimensión de Hausdorff de* E, denotada como $\dim_H(E)$, se define como

$$\dim_H(E) = \sup\{s : H^s(E) = \infty\} = \text{ínf}\{s : H^s(E) = 0\}.$$

Los siguientes teoremas se conocen, sin embargo, los recordamos debido a su relevancia en esta investigación.

Teorema 1.2.11. (Teorema del Valor Medio). Sean $a, b \in \mathbb{R}$ con $a < b$, $f : [a, b] \to \mathbb{R}$ una función continua en el intervalo cerrado $[a, b]$ y diferenciable en (a, b). Luego existe al menos algún punto $c \in (a, b)$ tal que

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Teorema 1.2.12. (Teorema de Rolle). Sean $a, b \in \mathbb{R}$ con $a < b$, $f : [a, b] \to \mathbb{R}$ una función continua en el intervalo cerrado $[a, b]$, diferenciable en (a, b) y $f(a) = f(b)$. Luego existe al menos algún punto $c \in (a, b)$ tal que

$$f'(c) = 0.$$

Teorema 1.2.13. Sea $f : \mathbb{R} \to \mathbb{R}$ una función continua en $\mathbb{R}$. Si para todo $x \in \mathbb{R}$ se cumple que $f(x) > 0$ y $\lim_{x\to\infty^+} f(x) = \lim_{x\to\infty^-} f(x) = 0$, entonces existe $x_0 \in \mathbb{R}$ tal que $f(x) \leq f(x_0)$, para todo $x \in \mathbb{R}$.

Teorema 1.2.14. (Teorema de Weierstrass). Sean $a, b \in \mathbb{R}$ con $a < b$ y $f : [a, b] \to \mathbb{R}$ una función continua en el intervalo cerrado y acotado $[a, b]$, Luego existen dos puntos $x_1, x_2 \in [a, b]$ tales que $f(x_1) \leq f(x) \leq f(x_2)$, para cualquier $x \in [a, b]$.

1.3. Dimensión de Conteo de Cajas

El método de conteo de cajas para obtener la dimensión fractal de una figura consiste en generalizar el proceso con el que se encuentra la longitud, el área y el volumen. Considérese como punto de partida a un segmento de longitud 1, un cuadrado de área 1 y un cubo con volumen 1, estos objetos tienen dimensión topológica igual a 1, 2 y 3, respectivamente. A continuación, divídase a cada uno de ellos en N partes iguales con tamaño δ (ver Figura 1.2).

Tamaño	$d=1$	$d=2$	$d=3$
$\delta = 1$	$N=1$	$N=1$	$N=1$
$\delta = \frac{1}{2}$	$N=2$	$N=4$	$N=8$
$\delta = \frac{1}{3}$	$N=3$	$N=9$	$N=27$

Figura 1.2: Distintas divisiones de correspondientes a cada figura.

Obsérvese que se cumple la relación

$$N \cdot \delta^d = 1, \tag{1.3.1}$$

donde el exponente del valor δ coincide con la dimensión del espacio en el que se encuentra cada figura. Si se generaliza el proceso para cualquier dimensión d y, además, se considera a d como una variable dependiente de δ, por lo que es denotada por d_δ, entonces al despejarla de la igualdad (1.3.1) se obtiene la formula:

$$d_\delta = \frac{\log(N_\delta)}{\log(\frac{1}{\delta})}, \tag{1.3.2}$$

donde N_δ es el número de "cajas" correspondientes al valor δ. Algunos ejemplos del cálculo de d_δ a partir de la fórmula (1.3.2) se pueden observar en la Figura 1.3, en donde en la parte (1) se muestra una recta en la que

$$d_{\frac{1}{3}} = \frac{\log(3)}{\log(3)} = 1,$$

en (2) aparece un cuadrado con

$$d_{\frac{1}{3}} = \frac{\log(9)}{\log(3)} = 2,$$

y en (3) se tiene un segmento de la curva de Koch con

$$d_{\frac{1}{3}} = \frac{\log(4)}{\log(3)} = 1.2619.$$

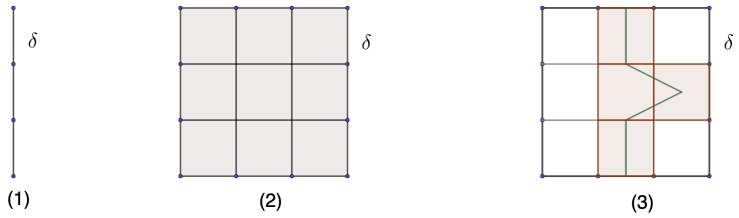

Figura 1.3: Cálculo del valor de d_δ para distintas figuras. (1) Recta con $d_{\frac{1}{3}} = 1$. (2) Cuadrado con $d_{\frac{1}{3}} = 2$. (3) Segmento de la curva de Koch con $d_{\frac{1}{3}} = 1.26$.

Para calcular la dimensión por el método de cajas de un conjunto acotado $E \subseteq \mathbb{R}^n$, se encierra este por una caja, donde cada uno de sus lados tenga una longitud fija L. Sin pérdida de generalidad se puede suponer que $L = 1$. Luego esta caja se divide en subcajas, donde los lados de cada subcaja tengan una longitud δ, y se cuentan las subcajas que cubren a E. El menor número de subcajas que cubren a E se denota por N_δ. A continuación, se realiza el cociente de logaritmos

$$d_\delta = \frac{\log(N_\delta)}{\log(\frac{1}{\delta})}.$$

Este procedimiento se repite considerando subdivisiones más pequeñas, es decir, para valores de δ progresivamente más pequeños. En consecuencia, la dimensión se obtiene en el límite $\lim_{\delta\to 0} d_\delta$ (siempre que exista), es decir,

$$\dim_B E = \lim_{\delta\to 0} \frac{\log(N_\delta)}{\log(\frac{1}{\delta})}. \tag{1.3.3}$$

Al valor $\dim_B(E)$ se le conoce como la *dimensión de conteo de Cajas.*

El método de conteo de cajas es un proceso infinito, realizarlo a mano carece de sentido, sin embargo, es posible llevarlo a cabo sólo para un número finito de fases y luego efectuar un ajuste lineal por el método de mínimos cuadrados. Por ejemplo, para la curva de Koch,

la Figura 1.4: (a) muestra que la curva está cubierta por $N_\delta = 8$ cajas, donde $\delta = 0.33$. (b) muestra que la curva está cubierta por $N_\delta = 20$ cajas para $\delta = 0.166$. (c) muestra que la curva está cubierta por $N_\delta = 40$ cajas para $\delta = 0.083$. Estos datos se muestran en el cuadro 1.3.1, el cual muestra los resultados obtenidos de realizar manualmente el conteo de cajas y los cálculos de los cocientes de los logaritmos.

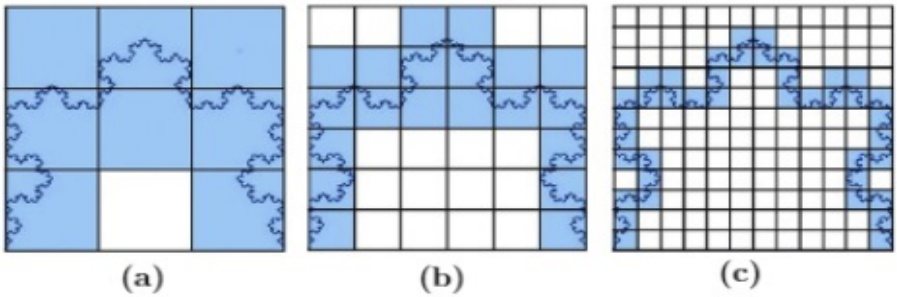

Figura 1.4: Distintos recubrimientos de la curva de Koch.

Fila	δ	N_δ	$\frac{1}{\delta}$	$\log(\frac{1}{\delta})$	$\log(N_\delta)$	$\log(N_\delta)/\log(\frac{1}{\delta})$
1	0.5	4	2	1	2	2
2	0.33	8	3	1.584962	3	1.892789
3	0.166	20	6	2.584962	4.321928	1.671950
4	0.083	40	12	3.584962	5.321928	1.484514

Cuadro 1.3.1: Datos obtenidos por medio de recubrimientos realizados a mano.

Los datos $\hat{x} = \log(N_\delta)$ y $\hat{y} = \log(\frac{1}{\delta})$, para cada valor que toman N_δ y δ en estas cuatro filas, son graficados en la Figura 1.5 como una serie de puntos que, al ajustarlos mediante el procedimiento de mínimos cuadrados, se obtiene una línea recta cuya ecuación está dada por $\hat{y} = 0.7752\hat{x} - 0.6492$. Observe que esta recta tiene la propiedad de que su pendiente es una aproximación de la dimensión de conteo de cajas, pues,

$$\dim_B = \lim_{\delta \to 0} \frac{\log(N_\delta)}{\log(\frac{1}{\delta})} \approx \lim_{\hat{x} \to \infty} \frac{0.7752\hat{x} - 0.6492}{\hat{x}} = 0.7752.$$

Por lo tanto, podemos concluir que la dimensión de conteo de cajas de la curva de Koch es el valor fraccionario $\dim_B \approx 0.7752$.

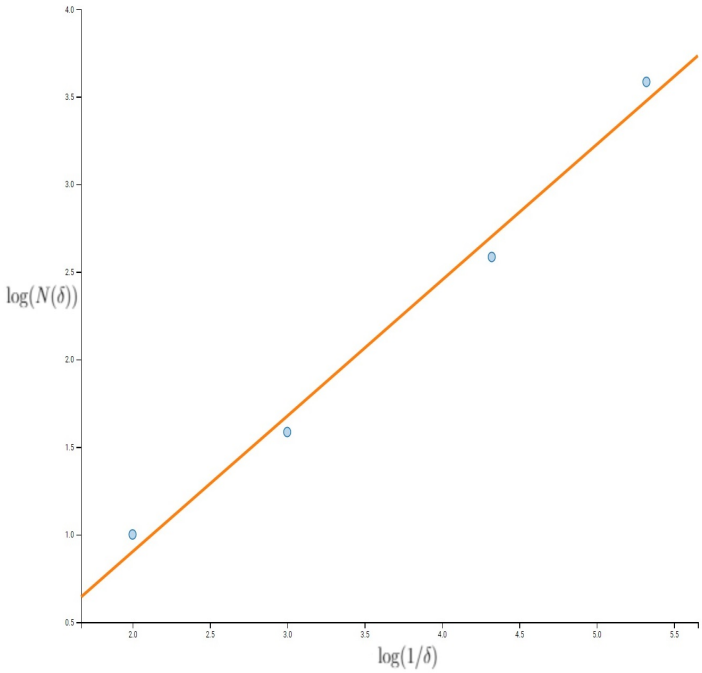

Figura 1.5: Gráfico de relación entre $\log(1/\delta)$ contra $\log(N(\delta))$.

En [37] se puede encontrar una demostración de que el método de conteo de cajas es equivalente a la dimensión de Hausdorff y se muestra que este proceso conduce a la obtención de la dimensión fractal.

1.4. Transformaciones bi-Lipschitz

Como se mencionó en la introducción, en el contexto de la búsqueda de condiciones para que una transformación no altere la dimensión fractal de una imagen se encuentra la clase de funciones bi-Lispchitz, por lo que en esta sección mostraremos este concepto y algunos resultados ligados a él.

Definición 1.4.1. Sean $E \subseteq \mathbb{R}^n$ y $f : E \to \mathbb{R}^m$ una función.

- Se dice que f es *Lipschitziana* sobre E, si existe $\lambda \in \mathbb{R}$ con $\lambda > 0$ tal que

$$\|f(x) - f(y)\| \leq \lambda \|x - y\|,$$

para cualesquiera $x, y \in E$ y la respectiva norma euclidiana o alguna equivalente a ella.

- Se dice que f es una transformación *bi-Lipschitz* sobre E, si existen $\lambda_1, \lambda_2 \in \mathbb{R}$ con $0 < \lambda_1 \leq \lambda_2 < \infty$ tales que

$$\lambda_1 \cdot \|x - y\| \leq \|f(x) - f(y)\| \leq \lambda_2 \cdot \|x - y\|, \tag{1.4.1}$$

para cualesquiera $x, y \in E$.

El Teorema 1.4.2 muestra algunas condiciones para que una función real sea Lipchitziana sobre intervalos arbitrarios [7].

Teorema 1.4.2. **S**ean I un intervalo de $\mathbb{R}$ y $f : I \to \mathbb{R}$ una función continua en I, derivable en el interior de I y tal que su derivada es acotada. Entonces f es Lipchitziana en I.

Proposición 1.4.3. La dimensión de Hausdorff tiene las siguientes propiedades ([20]):

I) Si $f : E \to \mathbb{R}^m$ es Lipschitziana, entonces $\dim_H(f(E)) \leq \dim_H(E)$.

II) Si $f : E \to \mathbb{R}^m$ es una transformación bi-Lipschitz, entonces $\dim_H(f(E)) = \dim_H(E)$.

III) $\dim_H(E) = \dim_H(\overline{E})$.

1.5. Transformaciones DP

Se conoce que las transformaciones que preservan la dimensión fractal, brevemente llamadas DP (por las siglas en inglés de *dimension preserving*), forman una clase más grande que las funciones bi-Lipschitz. En esta sección se muestran algunos resultados alrededor de ellas.

Definición 1.5.1. Una transformación f de $\mathbb{R}^n$ en $\mathbb{R}^m$ se dice que preserva la dimensión fractal en un conjunto $L \subset \mathbb{R}^n$, si para cualquier subconjunto $E \subset L$ se cumple la siguiente condición

$$\dim_H(E) = \dim_H f(E).$$

En este caso, f se llama una transformación DP sobre $\mathbb{R}^n$. Sea G_L la familia de todas las transformaciones de L que preservan la dimensión fractal en L. Es bien sabido que cualquier transformación bi-Lipschitz pertenece a este grupo ([26]). Si $L \subset M$, se sigue que $G_M \subset G_L$. También está claro que si f es una transformación DP, entonces para cualquier subconjunto E y para cualquier $s \neq \dim_H E$ tenemos que

$$H^s(E) = 0 \text{ si y solo si } H^s(f(E)) = 0;$$

$$H^s(E) = \infty \text{ si y solo si } H^s(f(E)) = \infty.$$

Los resultados 1.5.2 y 1.5.3 son los principales del artículo [1] y proporcionan el sustento teórico de esta investigación, por esta razón, son fundamentales en el Capítulo 2 de la presente tesis.

Lema 1.5.2. Si $f : \mathbb{R}^n \to \mathbb{R}^n$ es una transformación bi-Lipschitz, entonces es una transformación DP.

Teorema 1.5.3. Sea $f : \mathbb{R}^n \to \mathbb{R}^n$ una función. Si existe una descomposición numerable de $\mathbb{R}^n = \cup_i G_i$, y existen constantes positivas $\lambda_1^{(i)}$ y $\lambda_2^{(i)}$ tales que para todos los puntos $x \in G_i$, $y \in G_i$ y sus imágenes $f(x)$, $f(y)$, se tiene que

$$\lambda_1^{(i)}\|x - y\| \leq \|f(x) - f(y)\| \leq \lambda_2^{(i)}\|x - y\|,$$

entonces la transformación f es una transformación DP.

Capítulo 2

Preservación de la dimensión fractal bajo transformaciones y filtros de Gabor

El objetivo general de esta tesis es el proporcionar una metodología que utilice la dimensión fractal como una etapa clasificatoria en el reconocimiento biométrico, sin embargo, basados en una investigación bibliográfica realizada previamente a esta tesis, se observó que no existe un análisis sobre la preservación de la dimensión fractal al someter a una imagen a diversas transformaciones y filtrados, en particular, a los propuestos en esta investigación. Por lo que los resultados, ejemplos y contraejemplos que aparecen en este capítulo son aportaciones originales de esta tesis y se encuentran en proceso de publicación en el artículo [38].

2.1. Transformaciones reflexión, rotación y normalización

Para entrenar la red neuronal se crearán 49 nuevas imágenes a partir de una original. La técnica de data augmentation que se propone, utilizará reflexiones, rotaciones, acercamientos y alejamientos, además de ciertas combinaciones entre ellas. Los siguientes teoremas muestran que las reflexiones preservan la dimensión, en el primero se considera la reflexión sobre el eje central vertical (ver Figura 2.1) y en el segundo sobre el horizontal (ver Figura 2.2).

Teorema 2.1.1. Sean $a, b \in \mathbb{R}$ tales que $a, b \geq 0$ y $T : [0, a] \times [0, b] \to \mathbb{R}^2$ la transformación reflexión sobre el eje vertical definida como $T(x, y) = (a - x, y)$.T es una transformación DP.

Demostración. Nótese que

$$\begin{aligned} \|T(x_1, y_1) - T(x_2, y_2)\| &= \|(a - x_1, y_1) - (a - x_2, y_2)\| \\ &= \|(-x_1 + x_2, y_1 - y_2)\| \\ &= \sqrt{(-x_1 + x_2)^2 + (y_1 - y_2)^2} \\ &= \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \\ &= \|(x_1, y_1) - (x_2, y_2)\|. \end{aligned}$$

Para cualesquiera $\lambda_1 \leq 1$ y $\lambda_2 \geq 1$, se cumple que

$$\lambda_1 \|(x_1, y_1) - (x_2, y_2)\| \leq \|T(x_1, y_1) - T(x_2, y_2)\| \leq \lambda_2 \|(x_1, y_1) - (x_2, y_2)\|,$$

resulta que T es bi-Lipschitz, y por el Lema 1.5.2, T es DP. ■

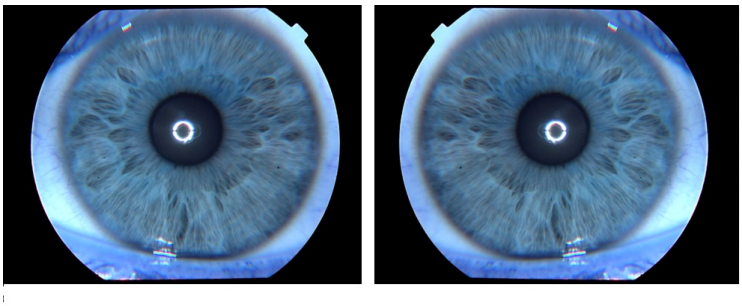

Figura 2.1: Reflexión de una imagen sobre el eje central vertical.

El teorema siguiente se demuestra de manera análoga al Teorema 2.1.1.

Teorema 2.1.2. Sean $a, b \in \mathbb{R}$ tales que $a, b \geq 0$ y $T : [0, a] \times [0, b] \to \mathbb{R}^2$ la transformación reflexión sobre el eje horizontal definida como $T(x, y) = (x, b - y)$ es una transformación DP.

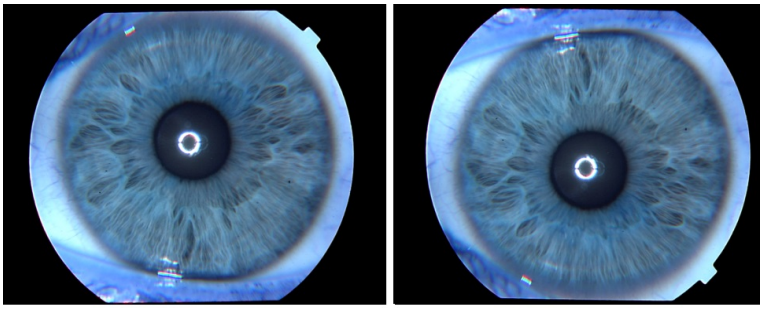

Figura 2.2: Reflexión de una imagen sobre el eje central horizontal.

Luego, a través del siguiente resultado, se demuestra que las rotaciones en cualquier ángulo preservan la dimensión.

Teorema 2.1.3. Sea $\theta \in [0, 2\pi]$. Se define la transformación rotación $T_\theta : \mathbb{R}^2 \to \mathbb{R}^2$, en el ángulo θ, como $T_\theta(x, y) = (x'_\theta, y'_\theta)$, con

$$x'_\theta = x\cos(\theta) - y\text{sen}(\theta),$$
$$y'_\theta = x\text{sen}(\theta) + y\cos(\theta).$$

El operador T_θ es una transformación DP.

Demostración. Se sigue de

$$\begin{aligned}
&\|T_\theta(x_1,y_1) - T_\theta(x_2,y_2)\|^2 \\
&= \|(x_1\cos(\theta) - y_1\text{sen}(\theta), x_1\text{sen}(\theta) + y_1\cos(\theta)) \\
&- (x_2\cos(\theta) - y_2\text{sen}(\theta), x_2\text{sen}(\theta) + y_2\cos(\theta))\|^2 \\
&= \| \left((x_1-x_2)\cos(\theta) + (y_2-y_1)\text{sen}(\theta), (x_1-x_2)\text{sen}(\theta) + (y_1-y_2)\cos(\theta)\right) \|^2 \\
&= (x_1-x_2)^2\cos^2(\theta) + 2(x_1-x_2)(y_2-y_1)\cos(\theta)\text{sen}(\theta) + (y_2-y_1)^2\text{sen}^2(\theta) \\
&+ (x_1-x_2)^2\text{sen}^2(\theta) + 2(x_1-x_2)(y_1-y_2)\text{sen}(\theta)\cos(\theta) + (y_1-y_2)^2\cos^2(\theta) \\
&= (x_1-x_2)^2\cos^2(\theta) + 2(x_1-x_2)(y_2-y_1)\cos(\theta)\text{sen}(\theta) + (y_2-y_1)^2\text{sen}^2(\theta) \\
&+ (x_1-x_2)^2\text{sen}^2(\theta) - 2(x_1-x_2)(y_2-y_1)\text{sen}(\theta)\cos(\theta) + (y_1-y_2)^2\cos^2(\theta) \\
&= (x_1-x_2)^2 + (y_1-y_2)^2 \\
&= \|(x_1,y_1) - (x_2,y_2)\|^2.
\end{aligned}$$

Con lo que se tiene que $\|T_\theta(x_1,y_1) - T_\theta(x_2,y_2)\| = \|(x_1,y_1) - (x_2,y_2)\|$. Nuevamente, para cualesquiera $\lambda_1 \leq 1$ y $\lambda_2 \geq 1$, se cumple que

$$\lambda_1\|(x_1,y_1) - (x_2,y_2)\| \leq \|T_\theta(x_1,y_1) - T_\theta(x_2,y_2)\| \leq \lambda_2\|(x_1,y_1) - (x_2,y_2)\|,$$

por tanto T_θ es bi-Lipschitz, y por el Lema 1.5.2, T es DP. ■

Los acercamientos (*zoom in*) y los alejamientos (*zoom out*) también preservan la dimensión, esto se observa en el Teorema 2.1.5.

Definición 2.1.4. Sea $\alpha \in (0,\infty)$ fijo. Se define la transformación $T_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$ como $T_\alpha(x,y) = \alpha(x,y)$:

- Si $\alpha < 1$, se le conoce como transformación acercamiento.
- Si $\alpha > 1$, es llamada transformación alejamiento.
- $\alpha = 1$, se le conoce como la identidad en $\mathbb{R}^2$.

Teorema 2.1.5. La transformación T_α es una transformación DP para cualquier $\alpha \in (0, \infty)$.

Demostración. Se tiene que

$$\begin{aligned}\|T_\alpha(x_1, y_2) - T_\alpha(x_2, y_2)\| &= \|\alpha(x_1, y_1) - \alpha(x_2, y_2)\| \\ &= \alpha\|(x_1, y_1) - (x_2, y_2)\|.\end{aligned}$$

Como $\|T_\alpha(x_1, y_1) - T_\alpha(x_2, y_2)\| = \alpha\|(x_1, y_1) - (x_2, y_2)\|$. Para $\lambda_1 = \frac{\alpha}{2}$ y $\lambda_2 = 2\alpha$, se cumple que

$$\lambda_1\|(x_1, y_1) - (x_2, y_2)\| \leq \|T(x_1, y_1) - T(x_2, y_2)\| \leq \lambda_2\|(x_1, y_1) - (x_2, y_2)\|,$$

por tanto T es bi-Lipschitz, y por el Lema 1.5.2, T es DP. ■

Al preservar la dimensión las transformaciones de los Teoremas 2.1.1, 2.1.2, 2.1.3 y 2.1.5, cualquier combinación de ellas también la preservará.

Ahora, el proceso de reconocimiento inicia con la normalización de la imagen del iris capturado, específicamente, el disco extraído se mapea en una forma rectangular (véase la Figura 2.3). El Teorema 2.1.7 afirma que esta transformación preserva la dimensión fractal de la imagen original.

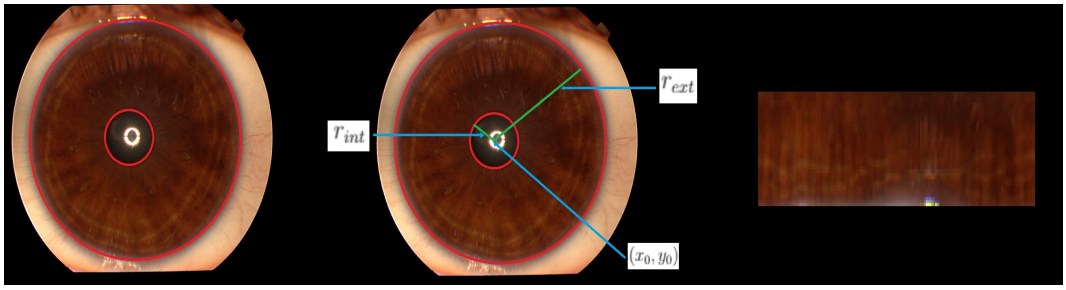

Figura 2.3: Normalización de la imagen del iris.

Definición 2.1.6. Sean (x_0, y_0) el centro de la pupila del ojo, r_{int} el valor del radio interior del iris y r_{ext} el radio exterior. Se define la normalización de la imagen como

la transformación $T : [r_{int}, r_{ext}] \times [0, 2\pi] \to \mathbb{R}^2$ definida como $T(r, \theta) = (x(r, \theta), y(r, \theta))$, donde

$$x(r, \theta) = x_0 + (r_{ext} - r(r_{ext} - r_{int})) \cos(\theta),$$
$$y(r, \theta) = y_0 + (r_{ext} - r(r_{ext} - r_{int})) \operatorname{sen}(\theta).$$

Teorema 2.1.7. La transformación definida en 2.1.6 es una transformación DP.

Demostración. Divídase el rectángulo $[r_{int}, r_{ext}] \times [0, 2\pi]$ como

$$\{[r_{int}, r_{ext}] \times [0, \pi]\} \cup \{[r_{int}, r_{ext}] \times [\pi, 2\pi]\}.$$

Se demostrará que T es bi-Lipschitz en $[r_{int}, r_{ext}] \times [0, \pi]$, es decir que existen constantes positivas λ_1 y λ_2 tales que:

$$\lambda_1 \|(r_1, \theta_1) - (r_2, \theta_2)\| \leq ||T(r_1, \theta_1) - T(r_2, \theta_2)|| \leq \lambda_2 \|(r_1, \theta_1) - (r_2, \theta_2)\|.$$

Esto es equivalente a probar que

$$\lambda_1^2 \|(r_1, \theta_1) - (r_2, \theta_2)\|^2 \leq \|T(r_1, \theta_1) - T(r_2, \theta_2)\|^2 \leq \lambda_2^2 \|(r_1, \theta_1) - (r_2, \theta_2)\|^2. \tag{2.1.1}$$

Desarrollando cada uno de los términos se tiene que

$$\begin{aligned}
\lambda_1^2 \|(r_1, \theta_1) - (r_2, \theta_2)\|^2 &= \lambda_1^2 (r_1 - r_2)^2 + \lambda_1^2 (\theta_1 - \theta_2)^2, \\
\|T(r_1, \theta_1) - T(r_2, \theta_2)\|^2 &= (r_{ext} - r_{int})^2 (r_1 - r_2)^2 \\
&\quad + 2[r_{ext} - r_1(r_{ext} - r_{int})][r_{ext} - r_2(r_{ext} - r_{int})][1 - cos(\theta_1 - \theta_2)], \\
\lambda_2^2 \|(r_1, \theta_1) - (r_2, \theta_2)\|^2 &= \lambda_2^2 (r_1 - r_2)^2 + \lambda_2^2 (\theta_1 - \theta_2)^2.
\end{aligned}$$

Primero, se demostrará que existe λ_2 que satisface la desigualdad del lado derecho de (2.1.1). Antes, se probará que para cada $x \in \mathbb{R}$, se tiene que

$$1 - \cos(x) \leq x^2. \tag{2.1.2}$$

En efecto, demostrar la desigualdad (2.1.2) es equivalente a demostrar que $x^2 - 1 + \cos(x) \geq 0$. Defínase $G(x) = x^2 - 1 + \cos(x)$. Se tiene que

$$G'(x) = 2x - \text{sen}(x),$$
$$G''(x) = 2 - \cos(x).$$

Obsérvese que la raíz de $G'(x)$ es $x = 0$. Evaluando $x = 0$ en $G''(x)$ resulta que $G''(0) = 1 > 0$ y por el criterio de la segunda derivada $G(x)$ tiene un punto mínimo en $x = 0$. Dado que $G(0) = 0$. Se concluye que $G(x) \geq 0$.

Considerando $x = \theta_1 - \theta_2$ en la desigualdad (2.1.2) se tiene que

$$\begin{aligned}(r_{ext} - r_{int})^2(r_1 - r_2)^2 + 2[r_{ext} - r_1(r_{ext} - r_{int})][r_{ext} - r_2(r_{ext} - r_{int})][1 - \cos(\theta_1 - \theta_2)] \\ \leq (r_{ext} - r_{int})^2(r_1 - r_2)^2 + 2[r_{int} + r_1(r_{ext} - r_{int})][r_{int} + r_2(r_{ext} - r_{int})](\theta_1 - \theta_2)^2.\end{aligned} \tag{2.1.3}$$

Como $r_1, r_2 \in [r_{int}, r_{ext}]$, se sigue que:

$$2[r_{ext} - r_1(r_{ext} - r_{int})][r_{ext} - r_2(r_{ext} - r_{int})] \leq 2[r_{ext} - r_{int}(r_{ext} - r_{int})]^2,$$

esto implica que

$$\begin{aligned}(r_{ext} - r_{int})^2(r_1 - r_2)^2 + 2[r_{ext} - r_1(r_{ext} - r_{int})][r_{ext} - r_2(r_{ext} - r_{int})](\theta_1 - \theta_2)^2 \\ \leq (r_{ext} - r_{int})^2(r_1 - r_2)^2 + 2[r_{ext} + r_{int}(r_{ext} - r_{int})]^2(\theta_1 - \theta_2)^2.\end{aligned} \tag{2.1.4}$$

Obsérvese que $(r_{int} + r_{ext})^2$ y $2[r_{ext} - r_{int}(r_{ext} - r_{int})]^2$ son constantes positivas. Ahora, fíjese $\lambda_2 = \text{máx}\{r_{int} + r_{ext}, \sqrt{2}[r_{ext} - r_{int}(r_{ext} - r_{int})]\}$. Así por (2.1.3) y (2.1.4), se concluye que:

$$\|T(r_1, \theta_1) - T(r_2, \theta_2)\|^2 \leq \lambda_2^2(r_1 - r_2)^2 + \lambda_2^2(\theta_1 - \theta_2)^2.$$

Ahora, se establecerá que existe λ_1 tal que satisface la desigualdad del lado izquierdo de (2.1.1). Dado que $r_1, r_2 \in [r_{int}, r_{ext}]$, se sigue que

$$2[r_{ext} - r_{ext}(r_{ext} - r_{int})]^2 \leq 2[r_{ext} - r_1(r_{ext} - r_{int})][r_{ext} - r_2(r_{ext} - r_{int})],$$

esto implica que

$$\begin{aligned}(r_{ext}-r_{int})^2(r_1-r_2)^2+2[r_{ext}-r_{ext}(r_{ext}-r_{int})]^2(\theta_1-\theta_2)^2\\ \leq (r_{ext}-r_{int})^2(r_1-r_2)^2+2[r_{ext}-r_1(r_{ext}-r_{int})][r_{ext}-r_2(r_{ext}-r_{int})](\theta_1-\theta_2)^2.\end{aligned} \tag{2.1.5}$$

Esto es, λ_1 debe cumplir que

$$\begin{aligned}\lambda_1^2(r_1-r_2)^2+\lambda_1^2(\theta_1-\theta_2)^2 &\leq (r_{ext}-r_{int})^2(r_1-r_2)^2\\ &+ 2[r_{ext}-r_{ext}(r_{ext}-r_{int})]^2[1-\cos(\theta_1-\theta_2)].\end{aligned}$$

Si $\lambda_1 \leq (r_{ext}-r_{int})^2$, entonces $\lambda_1(r_1-r_2)^2 \leq (r_{ext}-r_{int})^2(r_1-r_2)^2$. Resta determinar el valor de λ_1 que satisfaga

$$\lambda_1^2(\theta_1-\theta_2)^2 \leq 2[r_{ext}-r_{ext}(r_{ext}-r_{int})]^2\left[1-\cos(\theta_1-\theta_2)\right]^2,$$

esto es equivalente a

$$\lambda_1^2 \leq 2[r_{ext}-r_{ext}(r_{ext}-r_{int})]^2\frac{1-\cos(\theta_1-\theta_2)}{(\theta_1-\theta_2)^2}.$$

Dado que

$$\frac{1-\cos(\theta_1-\theta_2)}{(\theta_1-\theta_2)^2}$$

alcanza su valor mínimo en $-\pi$ y π. Se tiene que

$$\lambda_1^2 \leq 2[r_{ext}-r_{ext}(r_{ext}-r_{int}))]^2\frac{1-\cos(-\pi)}{\pi^2}=4\frac{[r_{ext}-r_{ext}(r_{ext}-r_{int})]^2}{\pi^2}.$$

Nuevamente, debido a que

$$(r_{ext}-r_{int})^2$$

y

$$4\frac{[r_{ext}-r_{ext}(r_{ext}-r_{int})]^2}{\pi^2}$$

son constantes positivas, se puede tomar

$$\lambda_1 = \text{máx}\left\{r_{ext}-r_{int}, \frac{2}{\pi}|r_{ext}-r_{ext}(r_{ext}-r_{int})|\right\}.$$

Luego por (2.1.3) y (2.1.4), se concluye que

$$\lambda_1^2(r_1 - r_2)^2 + \lambda_1^2(\theta_1 - \theta_2)^2 \leq ||T(r_1, \theta_1) - T(r_2, \theta_2)||^2.$$

Así, T es bi-Lipchitz en $[0, \pi]$. Análogamente se demuestra que es bi-Lipchitz en $[\pi, 2\pi]$. Por lo tanto, T es DP, como consecuencia del Teorema 1.5.3. ■

2.2. Filtros de Gabor y su preservación de la dimensión

Estas funciones fueron propuestas por Gabor en 1946 ([21]) para tratar las señales unidimensionales como un método para descomponer una señal en grupos de información, para su posterior codificación o tratamiento. Más recientemente, este tipo de funciones fueron introducidas en el campo de la visión por Marcelja ([34]) en el caso unidimensional y Daugman ([10]) en el bidimensional, en particular, se han tratado los campos receptivos de células en el córtex visual ([25], [41]), donde un campo receptivo es aquella zona en la que una neurona del sistema visual es sensible.

Formalmente, una función de Gabor compleja $G : \mathbb{R}^2 \to \mathbb{C}$ es el producto de una señal sinusoidal y una gaussiana, y se define como

$$G(x, y) = s(x, y)w_r(x, y), \tag{2.2.1}$$

donde $s(x, y)$ es una función trigonométrica compleja, denominada "sinusoide" o "portadora", y $w_r(x, y)$ es una gaussiana, también llamada "envolvente". A continuación se detalla cada una de estas componentes, así como se exhiben algunas de sus respectivas propiedades para el caso bidimensional, debido a su posterior aplicación en las imágenes del iris del ojo.

El primer factor que compone la función de Gabor, según [25], se denomina como la función ***sinusoide compleja*** y es la función $s : \mathbb{R}^2 \to \mathbb{C}$ definida de la forma

$$s(x, y) = e^{i(2\pi(u_0 x + v_0 y) + P)},$$

donde (u_0, v_0) y P son la frecuencia y la fase de la sinusoide, respectivamente.

Por otro lado, la ***envolvente gaussiana*** es la función $w_r : \mathbb{R}^2 \to \mathbb{R}$ definida como:

$$w_r(x, y) = Ke^{-\pi(a^2(x-x_0)_r^2+b^2(y-y_0)_r^2)},$$

donde (x_0, y_0) son las coordenadas del máximo de la función, a y b son parámetros de escala de la gaussiana, K es un factor de escala y el subíndice r denota una operación de rotación definida como:

$$\begin{aligned}(x - x_0)_r &= (x - x_0)\cos(\theta) + (y - y_0)\,\mathrm{sen}(\theta),\\ (y - y_0)_r &= -(x - x_0)\,\mathrm{sen}(\theta) + (y - y_0)\cos(\theta),\end{aligned}$$

donde θ es el ángulo de la rotación, que también es un parámetro que caracteriza a la función.

En síntesis, la función de Gabor se define por los siguientes parámetros:

- Para la sinusoide:
  1. (u_0, v_0): Frecuencia.
  2. P : Fase.
- Para la envolvente gaussiana:
  1. K: Escala de su magnitud.
  2. (a, b): Escala de los dos ejes.
  3. θ: Ángulo de rotación.
  4. (x_0, y_0): Coordenadas del máximo.

Por esta razón, la expresión de la función de Gabor $G : \mathbb{R}^2 \to \mathbb{C}$ es

$$G(x, y) = Ke^{-\pi(a^2(x-x_0)_r^2+b^2(y-y_0)_r^2))}e^{i(2\pi(u_0x+v_0y)+P)}. \tag{2.2.2}$$

Ya que se puede considerar a la sinusoide como dos funciones reales situadas en la parte real e imaginaria de una función compleja, de la siguiente forma:

$$Re(s(x,y)) = \cos\left(2\pi(u_0x + v_0y) + P\right),$$
$$Im(s(x,y)) = \operatorname{sen}\left(2\pi(u_0x + v_0y) + P\right),$$

esto es, se puede considerar a la función de Gabor como una función con dominio y codominio en $\mathbb{R}^2$, en esta investigación se considera a la función de Gabor como en la Definición 2.2.1.

Definición 2.2.1. La función de Gabor bidimensional es la función $G : \mathbb{R}^2 \to \mathbb{R}^2$ definida como

$$G(x,y) = (G_1(x,y), G_2(x,y)),$$

donde las funciones $G_1,\ G_2 : \mathbb{R}^2 \to \mathbb{R}$, están dadas por

$$G_1(x,y) = Ke^{-\pi(a^2(x-x_0)_r^2 + b^2(y-y_0)_r^2)}\cos(2\pi(u_0x + v_0y) + P)$$

y

$$G_2(x,y) = Ke^{-\pi(a^2(x-x_0)_r^2 + b^2(y-y_0)_r^2)}\operatorname{sen}(2\pi(u_0x + v_0y) + P).$$

Las funciones G_1 y G_2 se encuentran desfasadas 90° como se muestra en la Figura 2.4.

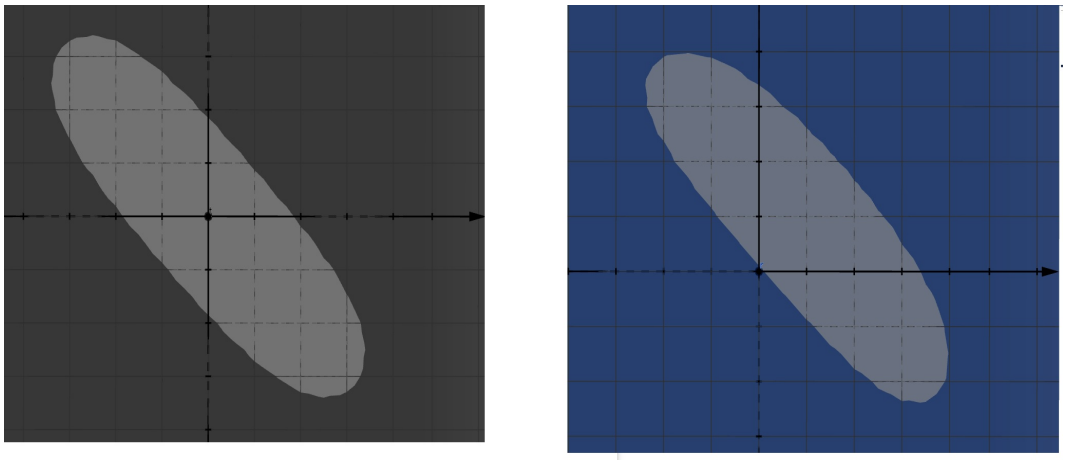

Figura 2.4: Partes real e imaginaria de una función de Gabor compleja en el dominio espacial; 128×128 píxeles; parámetros: $x_0 = y_0 = 0$, $a = 1/50$ píxeles, $b = 1/40$ píxeles, $\theta = -45°$, $P = 0°$.

Debido a que la aplicaciones principales de la función de Gabor es extraer la información relevante de imágenes, también se le denomina *filtro*. Además, al ser dependiente de los valores de sus parámetros, es común hablar de familias de filtros o simplemente denominarlos como *filtros de Gabor*.

2.2.1. Filtro de Gabor que no preserva la dimensión

Ejemplo 2.2.2. De manera experimental, se descubrió que las funciones de Gabor, en general, no conservan la dimensión fractal, para demostrarlo, se aplicaron dos filtros a una misma imagen, en la Figura 2.5 muestran los resultados a través de tres imágenes: la primera es la original y la segunda y tercera son imágenes resultantes de aplicar los filtrados. Para obtener la segunda imagen se aplicó un filtro Gabor con los siguientes parámetros:

- $(x_0, y_0) = (0, 0)$,
- $a = \frac{1}{8\sqrt{2\pi}}$,
- $b = 0$,
- $u_0 = \frac{\cos(\pi/4)}{10}$,
- $v_0 = \frac{\text{sen}(\pi/4)}{10}$,
- $P = 0.5$.

La tercera fue generada por un filtro con los siguientes parámetros:

- $(x_0, y_0) = (0, 0)$,
- $a = \frac{1}{\pi}\sqrt{\frac{2}{\pi}}$,
- $b = 0$,

- $u_0 = \cos(\pi/4)$,
- $v_0 = \text{sen}(\pi/4)$,
- $P = 0$.

A la imagen original se le calculó la dimensión mediante el método de conteo de cajas y se obtuvo un valor de 1.943. Con el mismo proceso se determinó que la dimensión de la segunda es igual, en este caso el filtro Gabor mantuvo la dimensión, sin embargo, la tercera imagen tiene una dimensión de 1.915, es decir, este último filtrado no preservó la dimensión.

Figura 2.5: Comparación de dos filtros Gabor distintos. Superior: imagen original. Centro: figura obtenida al aplicar el primer filtro a la imagen original. Inferior: imagen obtenida aplicando el segundo filtro a la imagen original. La dimensión fractal de estas imágenes es 1.943, 1.943 y 1.915, respectivamente.

Como consecuencia, se plantea como objetivo de esta tesis la búsqueda de las condiciones bajo las cuales los filtros de Gabor preserven la dimensión. La cual inicia en la siguiente sección.

2.3. Preservación de la dimensión bajo funciones campana de Gauss

Como se mencionó en la Sección 2.2, los filtros de Gabor son producto de una gaussiana por una sinusoidal. En este estudio se inicia el análisis por separado de cada componente, esto debido a que se observó que, para determinar las condiciones bajo las cuales los filtrados preserven la dimensión, se requerían de diversos resultados previos relacionados con cada una de ellas. Por esta razón, se comienza el estudio de casos particulares de cada una, se partirá de las funciones campana debido a su fuerte relación con la gaussiana. Los resultados 2.3.2, 2.3.3, 2.3.5, 2.3.6, 2.3.8, 2.3.9, 2.3.15 y 2.3.16 de esta sección son aportaciones originales de esta tesis y se pueden consultar en el artículo [38].

Definición 2.3.1. La función campana de Gauss se define como

$$f(x) = ke^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

donde k, μ y σ son constantes reales ($\sigma > -1$). El parámetro k es el valor del punto más alto de la campana, μ es la posición del centro de la campana y σ controla el ancho de la campana (Figura 2.6).

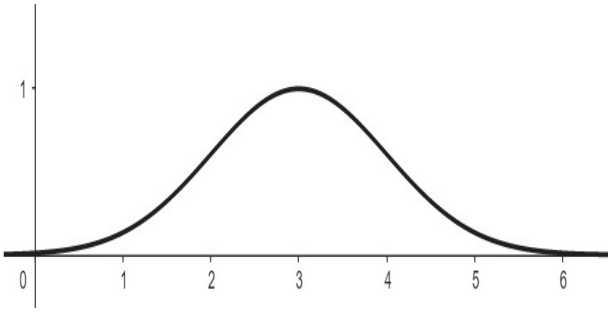

Figura 2.6: Función campana de Gauss o gaussiana con $k = 1$, $\sigma = 1$ y $\mu = 3$.

La siguiente proposición será utilizada en resultados posteriores.

Proposición 2.3.2. Si $f : \mathbb{R} \to \mathbb{R}$ está dada por $f(x) = \left| xe^{-\frac{(x-\mu)^2}{2\sigma^2}} \right|$, entonces existe $x_0 \in \mathbb{R}$ tal que $f(x) \leq f(x_0)$, para todo $x \in \mathbb{R}$.

Demostración. Dado que

$$\frac{\lim_{x\to\infty^+} x}{\lim_{x\to\infty^+} e^{\frac{(x-\mu)^2}{2\sigma^2}}} = \frac{\infty}{\infty},$$

se sigue de la regla de l'Hopital que

$$\lim_{x\to\infty^+} f(x) = \lim_{x\to\infty^+} \frac{x}{e^{\frac{(x-\mu)^2}{2\sigma^2}}} = \lim_{x\to\infty^+} \frac{1}{\frac{x-\mu}{\sigma^2} e^{\frac{(x-\mu)^2}{2\sigma^2}}} = 0.$$

De igual manera se muestra que $\lim_{x\to\infty^-} f(x) = 0$. Por la continuidad de f en $\mathbb{R}$ y el Teorema 1.2.13, existe $x_0 \in \mathbb{R}$ tal que $f(x) \leq f(x_0)$, para todo $x \in \mathbb{R}$. ■

En el siguiente resultado se establece que la función campana de Gauss es bi-Lipschitz sobre los compactos que no contengan al parámetro μ.

Proposición 2.3.3. Sea $f : \mathbb{R} \to \mathbb{R}$ una función campana de Gauss definida por $f(x) = ke^{-\frac{(x-\mu)^2}{2\sigma^2}}$. Entonces f es Lipschitziana en $\mathbb{R}$. Además, f es bi-Lipschitz en cada intervalo compacto $[a, b]$ que no contenga a μ.

Demostración. Nótese que

$$f'(x) = -\frac{k(x-\mu)}{\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \tag{2.3.1}$$

De este modo

$$\begin{aligned} |f'(x)| &\leq \left| \frac{kx}{\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right| + \left| \frac{k\mu}{\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right|, \\ &= \left| \frac{k}{\sigma^2} \right| \left| xe^{-\frac{(x-\mu)^2}{2\sigma^2}} \right| + \left| \frac{k\mu}{\sigma^2} \right| \left| e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right|. \end{aligned}$$

Por la Proposición 2.3.2, existe $x_0 \in \mathbb{R}$ tal que $\left| xe^{-\frac{(x-\mu)^2}{2\sigma^2}} \right| \leq \left| x_0 e^{-\frac{(x_0-\mu)^2}{2\sigma^2}} \right|$, para todo $x \in \mathbb{R}$. Además, $e^{-\frac{(x-\mu)^2}{2\sigma^2}} < 1$, con $x \in \mathbb{R}$. Tómese $L = \text{máx}\left\{ \left| x_0 e^{-\frac{(x_0-\mu)^2}{2\sigma^2}} \right|, 1 \right\}$, entonces

$$|f'(x)| \leq \left| \frac{k}{\sigma^2} \right| L + \left| \frac{k\mu}{\sigma^2} \right| L,$$

para todo $x \in \mathbb{R}$. El Teorema 1.4.2 implica que f es Lipschitziana en $\mathbb{R}$.

Sea $[a,b] \subseteq \mathbb{R}$ tal que $\mu \notin [a,b]$, sin pérdida de generalidad suponga que $\mu < a$. Debido a la continuidad de f' en $[a,b]$, por el Teorema de Weierstrass 1.2.14, existe $x_0 \in [a,b]$ tal que $f'(x_0) \geq f'(x)$, para todo $x \in [a,b]$. De (2.3.1), se sigue que $f'(x) < 0$, para todo $x > \mu$, esto implica que $f'(x_0) < 0$.

Sean $\lambda_1 = |f'(x_0)|$ y $x, y \in [a,b]$ con $x \neq y$. Por el Teorema del Valor Medio 1.2.11, existe $c_{xy} \in (a,b)$ tal que $\frac{f(y)-f(x)}{y-x} = f'(c_{xy})$. Entonces

$$\frac{f(y)-f(x)}{y-x} \leq f'(x_0) < 0,$$

y así

$$\left|\frac{f(y)-f(x)}{y-x}\right| \geq \lambda_1 > 0.$$

Por lo tanto, $|f(y)-f(x)| \geq \lambda_1 |y-x|$ para todo $x, y \in [a,b]$. En la Proposición 2.3.3 se probó que f es Lipschitziana sobre $\mathbb{R}$. Entonces, f es bi-Lipschitz en $[a,b]$. ■

Observación 2.3.4. La Proposición 2.3.3 establece que las funciones campana son bi-Lipschitz sobre los intervalos compactos siempre y cuando no contengan a μ. Sin embargo, sobre el resto, es decir, sobre los intervalos I que contienen a μ, las funciones campana de Gauss resultan no ser bi-Lipschitz. En efecto, suponga que f es bi-Lipschitz en $\mathbb{R}$, entonces existe $\lambda_1 > 0$ tal que

$$\lambda_1 |x-y| \leq |k| \left| e^{-\frac{(x-\mu)^2}{2\sigma^2}} - e^{-\frac{(y-\mu)^2}{2\sigma^2}} \right|,$$

para todo $x, y \in \mathbb{R}$. Considere x arbitrario fijo y $y \to x$, con lo que se tiene que

$$\begin{aligned} \lambda_1 &\leq |k||f'(x)| \\ &= |k| \left| -\frac{k(x-\mu)}{\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right|. \end{aligned}$$

Ahora, con $x = \mu$ se tiene que $\lambda_1 = 0$, lo cual es una contradicción. Por lo tanto, las funciones campana no son bi-Lipschitz en intervalos que contienen a μ.

Hasta ahora, se tiene que las funciones campana de Gauss no son bi-Lipschitz, sin embargo, preservan la dimensión fractal como lo establece el Teorema 2.3.5. Esto comprueba la afirmación dada por Alberverio, Pratsiovytyi y Torbin en [1] de que la clase de funciones DP no se restringe sólo a las bi-Lipschitz.

Teorema 2.3.5. La función Campana de Gauss es una función DP en $\mathbb{R}$.

Demostración. Considérese un intervalo compacto $[a, b]$ tal que $\mu \in (a, b)$. Tómese $n_1, n_2 \in \mathbb{N}$ tal que $a < \mu - \frac{1}{n_1}$ y $\mu + \frac{1}{n_2} < b$. Entonces

$$\mathbb{R} = \left(\bigcup_{k=0}^{\infty}[a-k-1, a-k]\right) \cup \left[a, \mu - \frac{1}{n_1}\right] \cup \left(\bigcup_{k=0}^{\infty}\left[\mu - \frac{1}{n_1+k}, \mu - \frac{1}{n_1+k+1}\right]\right) \cup \{\mu\} \tag{2.3.2}$$

$$\cup \left(\bigcup_{k=0}^{\infty}\left[\mu + \frac{1}{n_2+k+1}, \mu + \frac{1}{n_2+k}\right]\right) \cup \left[\mu + \frac{1}{n_2}, b\right] \cup \left(\bigcup_{k=0}^{\infty}[b+k, b+k+1]\right). \tag{2.3.3}$$

Por la Proposición 2.3.3, f es bi-Lipschitz en cada intervalo compacto que no contenga a μ dado en (2.3.2) y (2.3.3). Sobre el conjunto $\{\mu\}$ la función f es bi-Lipschitz debido a que con cualquier valor $\lambda > 0$ la condición se cumple. Por lo tanto y por el Teorema 1.5.3, f es DP en $\mathbb{R}$. ■

La idea de la prueba del Teorema 2.3.5 fue descomponer el conjunto $\mathbb{R}$ como una unión de intervalos compactos donde la función campana de Gauss es bi-Lipschitz en cada uno de ellos. Esto requería que la derivada de f fuera distinta de cero, por lo que se evitó el punto μ de estos intervalos debido a la Observación 2.3.4. De la misma forma, para comprobar que el producto de una campana Gaussiana con otra función es bi-Lipschitz en un intervalo compacto es necesario evitar los ceros de la derivada del producto.

Teorema 2.3.6. Sean $f : \mathbb{R} \to \mathbb{R}$ una función campana de Gauss, $g : \mathbb{R} \to \mathbb{R}$ una función diferenciable, cuya derivada es continua, y $[a, b] \subseteq \mathbb{R}$ un intervalo no degenerado. Entonces, fg no es bi-Lipschitz en $[a, b]$ si y solo si existe $x \in [a, b]$ tal que

$$x = \frac{g'(x)\sigma^2}{g(x)} + \mu \quad \text{o} \quad g(x) = g'(x) = 0.$$

Demostración. Para cada $x \in [a,b]$,

$$(fg)'(x) = \left(-\frac{(x-\mu)}{\sigma^2}g(x) + g'(x)\right) ke^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Se sigue que,

$$(fg)'(x) = 0 \quad \text{si y solo si} \quad -\frac{(x-\mu)}{\sigma^2}g(x) + g'(x) = 0 \tag{2.3.4}$$

$$\text{si y solo si} \quad x = \frac{g'(x)\sigma^2}{g(x)} + \mu \text{ or } g(x) = g'(x) = 0, \tag{2.3.5}$$

para todo $x \in [a,b]$. Supongamos que existe $x \in [a,b]$ tal que $x = \frac{g'(x)\sigma^2}{g(x)} + \mu$ o $g(x) = g'(x) = 0$, y fg es bi-Lipschitz en $[a,b]$. Entonces $(fg)'(x) = 0$ y existe $\lambda > 0$ para el cual

$$\lambda|y-s| \leq |f(y)g(y) - f(s)g(s)|,$$

para todos los $s, y \in [a,b]$. Asúmase, sin pérdida de generalidad, que $x \in (a,b)$. De este modo,

$$\begin{aligned}\lambda &\leq \lim_{y\to x}\left|\frac{f(y)g(y) - f(x)g(x)}{y-x}\right| \\ &= |(fg)'(x)| \\ &= \left|-\frac{(x-\mu)}{\sigma^2}e^{-\frac{(x-\mu)^2}{2\sigma^2}}g(x) + g'(x)e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right|. \end{aligned} \tag{2.3.6}$$

Como $x = \frac{g'(x)\sigma^2}{g(x)} + \mu$ o $g(x) = g'(x) = 0$, se sigue por (2.3.6) que $\lambda \leq 0$. Esto es una contradicción, se concluye que, fg no es bi-Lipschitz en $[a,b]$.

Ahora, asúmase que por cada $x \in [a,b]$, $x \neq \frac{g'(x)\sigma^2}{g(x)} + \mu$ y $g(x) \neq g'(x)$ o $g'(x) \neq 0$. Entonces por (2.3.4), $|(fg)'(x)| > 0$, para todo $x \in [a,b]$.

Para probar que fg es bi-Lipschitz sobre $[a,b]$, se seguirá la idea de la prueba de la Proposición 2.3.3. Por la continuidad de $|(fg)'|$, existe $x_0 \in [a,b]$ tal que $|(fg)'(x)| \geq |(fg)'(x_0)|$, para todo $x \in [a,b]$. Establezca $\lambda_1 = |(fg)'(x_0)|$. Así $\lambda_1 > 0$. Tómese $x, y \in [a,b]$ con $x \neq y$, por el Teorema del Valor Medio, existe $c_{xy} \in (a,b)$ tal que

$$\frac{f(y)g(y) - f(x)g(x)}{y-x} = (fg)'(c_{xy}).$$

Entonces,

$$\left|\frac{f(y)g(y)-f(x)g(x)}{y-x}\right| = |(fg)'(c_{xy})| \geq |(fg)'(x_0)| = \lambda_1.$$

En consecuencia, $|f(y)g(y)-f(x)g(x)| \geq \lambda_1|y-x|$.

Finalmente, de

$$\begin{aligned}|(fg)'(x)| &= \left|-\frac{k(x-\mu)}{\sigma^2}e^{-\frac{(x-\mu)^2}{2\sigma^2}}g(x)+kg'(x)e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right| \\ &\leq \left(\left|\frac{k}{\sigma^2}\right|\left|xe^{-\frac{(x-\mu)^2}{2\sigma^2}}\right|+\left|\frac{k\mu}{\sigma^2}\right|\left|e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right|\right)|g(x)|+k\,|g'(x)|\left|e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right|,\end{aligned}$$

se sigue, por la continuidad de g y g' y por la Proposición 2.3.2, que $|(fg)'|$ está acotado en $[a,b]$. Entonces por el Teorema 1.4.2, existe $\lambda_2 > 0$ tal que

$$|f(y)g(y)-f(x)g(x)| \leq \lambda_2|y-x|,$$

para todo $x, y \in [a,b]$. En consecuencia, fg es bi-Lipschitz en $[a,b]$. ■

Ejemplo 2.3.7. Sea f una función campana de Gauss con $k=1$, $\mu=1$ y $\sigma=2$, y sea $g(x)=e^{-\frac{x}{4}(1-\frac{x}{2})}$. Entonces $h=fg$ no es bi-Lipschitz en cada intervalo no degenerado I. De hecho, cualquier punto x cumple la condición del Teorema 2.3.6.

Teorema 2.3.8. Sean $h_1, h_2 : \mathbb{R} \to \mathbb{R}$ funciones dadas por

$$h_1(x) = ke^{-\frac{(x-\mu)^2}{2\sigma^2}}\cos(\alpha x+\beta)$$

y

$$h_2(x) = ke^{-\frac{(x-\mu)^2}{2\sigma^2}}\text{sen}(\alpha x+\beta).$$

Entonces h_1 y h_2 son funciones DP.

Demostración. Sólo se demostrará que h_1 es una función DP, para h_2 se sigue de forma análoga.

Defínase $g(x)=\cos(\alpha x+\beta)$. Sea

$$E = \left\{x \in \mathbb{R} : x = \frac{g'(x)\sigma^2}{g(x)}+\mu \quad \text{o} \quad g(x)=g'(x)=0\right\}.$$

Obsérvese que $h_1(x) = 0$ si y sólo si $\cos(\alpha x + \beta) = 0$ si y sólo si $x = \frac{(2n+1)\pi}{2\alpha} - \frac{\beta}{\alpha}$, $n \in \mathbb{Z}$. Entonces, por el Teorema de Rolle y la equivalencia (2.3.4), E es un conjunto numerable y se puede escribir de la siguiente manera:

$$E = \left\{\cdots x_3^-, x_2^-, x_1^-, x_0, x_1^+, x_2^+, x_3^+ \cdots\right\}$$

donde

$$\cdots < x_3^- < x_2^- < x_1^- < x_0 < x_1^+ < x_2^+ < x_3^+ < \cdots,$$

$$\lim_{n\to\infty} x_n^+ = \infty$$

y

$$\lim_{n\to\infty} x_n^- = -\infty.$$

Tómese las sucesiones $(c_{n,i}^-)$, $(d_{n,i}^-)$, $(c_{0,i}^-)$, $(d_{0,i}^-)$, $(c_{0,i}^+)$, $(d_{0,i}^+)$, $(c_{n,i}^+)$ y $(d_{n,i}^+)$ tales que

$$\begin{aligned}
&x_{n+1}^- < \cdots < c_{n,3}^- < c_{n,2}^- < c_{n,1}^- < d_{n,1}^- < d_{n,2}^- < d_{n,3}^- < \cdots < x_n^-,\\
&x_1^- < \cdots < c_{0,3}^- < c_{0,2}^- < c_{0,1}^- < d_{0,1}^- < d_{0,2}^- < d_{0,3}^- < \cdots < x_0,\\
&x_0 < \cdots < c_{0,3}^+ < c_{0,2}^+ < c_{0,1}^+ < d_{0,1}^+ < d_{0,2}^+ < d_{0,3}^+ < \cdots < x_1^+,\\
&x_n^+ < \cdots < c_{n,3}^+ < c_{n,2}^+ < c_{n,1}^+ < d_{n,1}^+ < d_{n,2}^+ < d_{n,3}^+ < \cdots < x_{n+1}^+,\\
&c_{n,i}^- \to x_{n+1}^-, \quad d_{n,i}^- \to x_n^-,\\
&c_{0,i}^- \to x_1^-, \quad d_{0,i}^- \to x_0,\\
&c_{0,i}^+ \to x_0, \quad d_{0,i}^+ \to x_1^+,\\
&c_{n,i}^+ \to x_n^+, \quad d_{n,i}^+ \to x_{n+1}^+.
\end{aligned}$$

Entonces

$$\mathbb{R} = \left(\bigcup_{n\in\mathbb{N}} \{x_n^-\}\right) \cup \left(\bigcup_{n\in\mathbb{N}} \left[\left(\cup_{i=1}^{\infty}[c_{n,i+1}^-, c_{n,i}^-]\right) \cup [c_{n,1}^-, d_{n,1}^-] \cup \left(\cup_{i=1}^{\infty}[d_{n,i}^-, d_{n,i+1}^-]\right)\right]\right) \tag{2.3.7}$$

$$\cup \left(\cup_{i=1}^{\infty}[c_{0,i+1}^-, c_{0,i}^-]\right) \cup [c_{0,1}^-, d_{0,1}^-] \cup \left(\cup_{i=1}^{\infty}[d_{0,i}^-, d_{0,i+1}^-]\right) \cup \{x_0\} \tag{2.3.8}$$

$$\cup \left(\cup_{i=1}^{\infty}[c_{0,i+1}^+, c_{0,i}^+]\right) \cup [c_{0,1}^+, d_{0,1}^+] \cup \left(\cup_{i=1}^{\infty}[d_{0,i}^+, d_{0,i+1}^+]\right) \cup \left(\bigcup_{n\in\mathbb{N}} \{x_n^+\}\right) \tag{2.3.9}$$

$$\cup \left(\bigcup_{n\in\mathbb{N}} \left[\left(\cup_{i=1}^{\infty}[c_{n,i+1}^+, c_{n,i}^+]\right) \cup [c_{n,1}^+, d_{n,1}^+] \cup \left(\cup_{i=1}^{\infty}[d_{n,i}^+, d_{n,i+1}^+]\right)\right]\right). \tag{2.3.10}$$

Del Teorema 2.3.6 , h_1 es bi-Lipschitz en cada subintervalo de (2.3.7), (2.3.8), (2.3.9) y (2.3.10). Además, está claro que h_1 es bi-Lipschitz en conjuntos singulares. Luego, por el Teorema 1.5.3, h_1 es una transformación DP en $\mathbb{R}$. ■

Teorema 2.3.9. Sea $h : \mathbb{R} \to \mathbb{R}$ una función dada por

$$h(x) = ke^{-\frac{(x-\mu)^2}{2\sigma^2}} \cos(\alpha x^2 + \beta).$$

Entonces h es una función DP.

Demostración. Nótese que

$$\begin{aligned} |h'(x)| &= \left| -\frac{k(x-\mu)}{\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \cos(\alpha x^2 + \beta) - 2\alpha x k e^{-\frac{(x-\mu)^2}{2\sigma^2}} \operatorname{sen}(\alpha x^2 + \beta) \right| \\ &\leq \left| \frac{k}{\sigma^2} \right| \left| x e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right| + \left| \frac{k\mu}{\sigma^2} \right| \left| e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right| + |2\alpha k| \left| x e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right|. \end{aligned}$$

Está claro que $e^{-\frac{(x-\mu)^2}{2\sigma^2}} < 1$, para todo $x \in \mathbb{R}$, y la Proposición 2.3.2 implica la existencia de $l_1 > 0$ tal que $\left| x e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right| < l_1$, para todo $x \in \mathbb{R}$. Entonces h' está acotado, por el Teorema 1.4.2, existe $\lambda_2 > 0$ tal que

$$|h(y) - h(x)| \leq \lambda_2 |y - x|, \tag{2.3.11}$$

para todo $x, y \in \mathbb{R}$. Resta por demostrar que h es Lipschitziana por abajo. Obsérvese que $h(x) = 0$ si y solo si $\cos(\alpha x^2 + \beta) = 0$. Entonces $x = \pm\sqrt{\frac{(2n+1)\pi}{2\alpha} - \frac{\beta}{\alpha}}$ son las raíces de h. Considérese sólo las raíces reales. Defina $x_n^{\pm} = \pm\frac{2n+1}{\pi}$, por lo que el conjunto de raíces de h es

$$\left\{ \cdots x_3^-, x_2^-, x_1^-, x_1^+, x_2^+, x_3^+ \cdots \right\}.$$

Sin pérdida de generalidad, tómese el intervalo $[x_n^+, x_{n+1}^+]$. Dado que $h(x_n^+) = h(x_{n+1}^+) = 0$, por el Teorema de Rolle existe $c_n^+ \in (x_n^+, x_{n+1}^+)$ tal que $h'(c_n^+) = 0$.

Tenemos que para cada intervalo $[x_n^{\pm}, x_{n+1}^{\pm}]$ existe una raíz de h'. Considérese

$$\left\{ \cdots c_3^-, c_2^-, c_1^-, c_1^+, c_2^+, c_3^+, \cdots \right\}$$

como el conjunto de todas las raíces de $h'(x)$.

Como en todo par de raíces consecutivas de h existe una raíz de h', entonces se puede considerar el siguiente conjunto

$$\left\{\cdots x_3^-, c_2^-, x_2^-, c_1^-, x_1^-, c_0^\pm, x_1^+, c_1^+, x_2^+, c_2^+, x_3^+ \cdots\right\}.$$

Obsérvese que

$$\mathbb{R} = \{c_0^\pm\} \cup (x_1^-, c_0^\pm) \cup (c_0^\pm, x_1^+) \cup_{n=1}^{\infty} [(x_{n+1}^-, c_n^-) \cup (c_n^-, x_n^-) \cup (x_n^+, c_n^+) \cup (c_n^+, x_{n+1}^+) \cup \{x_n^\pm\} \cup \{c_n^\pm\}].$$

Sobre cada conjunto unitario $\{x_n^\pm\}$, $n = 1, \ldots,$ se obtiene que

$$\lambda |x_n^\pm - x_n^\pm| \leq |h(x_n^\pm) - h(x_n^\pm)|,$$

para cualquier λ, lo mismo para cada $\{c_n^\pm\}$. En caso de existir puntos como los mencionados en el Teorema 2.3.6, serán considerados dentro del proceso anterior como conjuntos unitarios formados por cada punto, de esta manera, cualquier función es bi-Lipschitz sobre ellos. En las demostraciones subsecuentes se tomarán de esta manera.

Ahora, se probará que h es Lipschitziana por abajo en (c_n^+, x_{n+1}^+). En primer paso se descartará el caso $h(x) = h(y)$, para algunas $x, y \in (c_n^+, x_n^+)$, debido a que si existiera esta igualdad se divide el intervalo en dos partes y se repite el mismo proceso siguiente.

Considérese una sucesión $\{c_{n,i}^+\}$ de elementos en (c_n^+, x_{n+1}^+) tal que $c_{n,i}^+ \to c_n^+$. Divídase el intervalo en subintervalos de la forma

$$(c_n^+, x_{n+1}^+) = \cup_{i=1}^{\infty} (c_{n,i}^+, x_{n+1}^+).$$

La función $|h'(x)|$ es creciente en (c_n^+, x_{n+1}^+), por lo que en cada intervalo $(c_{n,i}^+, x_{n+1}^+)$ la derivada alcanza su valor mínimo en un extremo izquierdo $c_{n,i}^+$. Por consiguiente, tomamos $\lambda_{n,i}^\pm \leq |h'(c_{n,i}^+)|$. Entonces h es Lipschitziana por abajo en $(c_{n,i}^+, x_{n+1}^+)$. Análogamente se demuestra para el resto de los subintervalos y se obtiene una descomposición de $\mathbb{R}$, donde h es bi-Lipschitz y la conclusión se sigue del Teorema 1.5.3. ■

2.3.1. Filtros de Gabor en espacios euclidianos

Cuando se extienden los resultados de la subsección previa a espacios $\mathbb{R}^n$, se debe considerar la derivada de una función $G : \mathbb{R}^n \to \mathbb{R}^m$, $n \geq 1$, $m \geq 2$. Por ende, consideraremos el operador diferencial, el cual es la extensión natural del concepto de derivada.

Sea $L(\mathbb{R}^n, \mathbb{R}^m)$ el espacio de todas las transformaciones lineales de $\mathbb{R}^n$ a $\mathbb{R}^m$. Para $T \in L(\mathbb{R}^n, \mathbb{R}^m)$ la norma de T se define como

$$\|T\| = \sup_{\|x\|_2 \leq 1} \|Tx\|_2.$$

Es claro que $\|Tx\|_2 \leq \|T\|\|x\|_2$, para todo $x \in \mathbb{R}^n$, más aún, si $\|Tx\|_2 \leq \lambda\|x\|_2$, para todo $x \in \mathbb{R}^n$, entonces $\|T\| \leq \lambda$.

Definición 2.3.10. Sea E un conjunto abierto de $\mathbb{R}^n$. Una función $G : E \to \mathbb{R}^m$ se dice que es diferenciable en $x \in E$, si existe una transformación lineal $T_x : \mathbb{R}^n \to \mathbb{R}^m$ tal que

$$\lim_{h \to 0} \frac{\|G(x+h) - G(x) - T_x h\|_2}{\|h\|_2} = 0.$$

La transformación lineal T_x que satisface la Definición 2.3.10, si existe, es única. Esta es llamada la diferencial de G en el punto x y es denotada por $dG(x)$. Si G es diferenciable en todo punto de E, entonces G se dice ser diferenciable en E, en este caso dG define un operador de E a $L(\mathbb{R}^n, \mathbb{R}^m)$. El mapeo dG es acotado si existe $M > 0$ tal que $\|dG(x)\| \leq M$, para todo $x \in E$.

Teorema 2.3.11. Sean E un conjunto abierto de $\mathbb{R}^n$ y $G : E \longrightarrow \mathbb{R}^m$ diferenciable. Si existe un conjunto convexo A de E tal que dG es acotado en A por una constante positiva M, entonces

$$\|G(y) - G(x)\|_2 \leq M\|y - x\|_2,$$

para toda $x, y \in A$.

Demostración. Sea $x \in A$, dado que $dG(x)$ es un operador lineal y, por hipótesis, acotado, para $\epsilon > 0$, existe $h \in \mathbb{R}^n$, tal que

$$\lim_{\|h\|\to 0} \frac{\|G(x+h) - G(x) - dG(x)(h)\|}{\|h\|} = 0. \tag{2.3.12}$$

Consideremos y un punto de tal forma que $y - x$ esté en A, debido a que es convexo, y tal que $h = y - x$, entonces $y = x + h$. Luego,

$$\begin{aligned}
&\frac{\|G(y) - G(x)\|}{\|y - x\|} \\
&= \frac{\|G(x+h) - G(x+h) - dG(x)(h)\|}{\|h\|} \\
&\leq \frac{\|G(x+h) - G(x) - dG(x)(h)\|}{\|h\|} \\
&+ \frac{\|dG(x)(h)\|}{\|h\|}.
\end{aligned}$$

La igualdad (2.3.12) es equivalente a: Dado $\epsilon > 0$, existe $\delta > 0$ tal que si $\|h\| < \delta$ entonces

$$\frac{\|G(x+h) - G(x) - dG(x)(h)\|}{\|h\|} \leq \epsilon. \tag{2.3.13}$$

Se descompone A en conjuntos R que satisfacen que, para cada par de puntos $x, y \in R^n$,

$$\|x - y\| \leq \delta.$$

De aquí

$$\frac{\|G(x+h) - G(x) - dG(x)(h)\|}{\|h\|} + \frac{\|dG(x)(h)\|}{\|h\|} \leq \epsilon + M,$$

donde M es la cota de $dG(x)$, esto para cada par de elementos en R. Esto implica

$$\frac{\|G(x) - G(y)\|}{\|x - y\|} \leq M,$$

y se demuestra lo deseado. ■

El Teorema 2.3.12 es el de la función inversa para funciones sobre espacios euclideanos, en su formulación considérese $n \in \mathbb{N}$.

Teorema 2.3.12. (Teorema de la función inversa.) Sean $E \subset \mathbb{R}^n$ un conjunto abierto y $G : E \to \mathbb{R}^n$ una función de clase C^1. Sea $G(p) = q$ y supóngase que $dG(p)$ es invertible. Entonces existen conjuntos abiertos U y V de $\mathbb{R}^n$ tales que $p \in U$, $q \in V$ y $G : U \to V$ es biyectiva. Además, si $F : V \to U$ es la inversa de $G : U \to V$, entonces F es una función de clase C^1 en V y $dF(y) = (dG(x))^{-1}$ para todos $y \in V$ y $x \in U$ con $G(x) = y$.

Teorema 2.3.13. Sea $G : E \subset \mathbb{R}^n \to \mathbb{R}^n$ una función de la clase C^1 definida en un conjunto abierto $E \subset \mathbb{R}^n$. Si $p \in E$ es tal que $dG(p)$ es invertible, entonces hay una bola abierta B con centro en p tal que G es bi-Lipschitz en B.

Demostración. Sea $p \in E$ tal que $dG(p)$ es invertible. Sea $q = G(p)$, por el Teorema 2.3.12, existen conjuntos abiertos U y V de $\mathbb{R}^n$ tales que $p \in U$, $q \in V$, $G : U \to V$ es biyectiva y la inversa $F : V \to U$ de $G : U \to V$ es una función de clase C^1 en V y $dF(y) = (dG(x))^{-1}$, para todos $y \in V$ y $x \in U$ con $G(x) = y$.

Considérese $r > 0$ tal que $\overline{B(q,r)} \subseteq V$. Establézcase $W = F(B(q, \frac{r}{2}))$, luego $W \subseteq U$ es un conjunto abierto, pues $F : V \to U$ es un homeomorfismo. Primero se muestra que G es Lipschitz por abajo de W. La condición de que F sea de clase C^1 en V implica que $dF : V \to L(\mathbb{R}^n, \mathbb{R}^n)$ es una función continua. Por lo tanto, por la compacidad de $\overline{B(q,r)}$, hay $y_0 \in \overline{B(q,r)}$ tal que $\|dF(y)\| \leq |dF(y_0)\|$, para todo $y \in \overline{B(q,r)}$. Nótese que $0 < \|(dG(p))^{-1}\| = \|dF(q)\| \leq \|dF(y_0)\|$, es decir, $\|dF(y_0)\| > 0$. Por el Teorema 2.3.11,

$$\|F(y) - F(y')\|_2 \leq \|dF(y_0)\|\|y - y'\|_2, \tag{2.3.14}$$

para todo $y, y' \in \overline{B(q,r)}$. Sea $r, s \in W$. Luego $r, s \in U$ y

$$\|r - s\|_2 = \|F(G(r)) - F(G(s))\|_2.$$

Dado que G es biyectiva, existen $y_r, y_s \in B(q, \frac{r}{2})$ tal que $r = F(y_r)$ y $s = F(y_s)$. Se sigue que $G(r) = G(F(y_r)) = y_r$ y $G(s) = G(F(y_s)) = y_s$ esto implica que $G(r), G(s) \in B(q, \frac{r}{2})$, y así por (2.3.14),

$$\|F(G(r)) - F(G(s))\|_2 \leq \|dF(y_0)\|\|G(r) - G(s)\|_2.$$

Como consecuencia,

$$\|r-s\|_2 \frac{1}{\|dF(y_0)\|} \leq \|G(r)-G(s)\|_2. \tag{2.3.15}$$

para todo $r, s \in W$. Ahora, es claro que $p = F(G(p)) = F(q) \in F(B(q, \frac{r}{2})) = W$, de aquí que existe $r_2 > 0$ tal que $\overline{B(p, r_2)} \subseteq W$. Dado que G es de clase C^1 en E, existe $x_0 \in \overline{B(p, r_2)}$ tal que $\|dG(x)\| \leq \|dG(x_0)\|$, para todo $x \in \overline{B(p, r_2)}$. Asimismo, por el Teorema 2.3.11,

$$\|G(x)-G(x')\|_2 \leq \|dG(x_0)\|\|x-x'\|_2, \tag{2.3.16}$$

para todo $x, x' \in \overline{B(p, r_2)}$. Considérese $B_p = B(p, \frac{r_2}{2})$, luego $B_p \subseteq W$ y así por (2.3.15) y (2.3.16),

$$\|r-s\|_2 \frac{1}{\|dF(y_0)\|} \leq \|G(r)-G(s)\|_2 \leq \|dG(x_0)\|\|r-s\|_2,$$

para todos $r, s \in B_r$. Así pues, G es bi-Lipschitz en B_p. ■

Definición 2.3.14. Sea $G : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ una función cuyas derivadas parciales de primer orden existen en todo $\mathbb{R}^2$ y denote por G_1, G_2 a sus componentes escalares. La matriz jacobiana de G en un punto $(x, y) \in \mathbb{R}^2$ se define como:

$$JG(x,y) = \begin{pmatrix} \frac{\partial G_1}{\partial x}(x,y) & \frac{\partial G_1}{\partial y}(x,y) \\ \frac{\partial G_2}{\partial x}(x,y) & \frac{\partial G_2}{\partial y}(x,y) \end{pmatrix}. \tag{2.3.17}$$

Se sabe que si $G : \mathbb{R}^2 \to \mathbb{R}^2$, entonces la diferencial $dG(x, y)$ puede representar mediante la matriz jacobiana $JG(x, y)$; y $dG(x, y)$ es invertible si y solo si $\det JG(x, y) \neq 0$.

Ahora, considérese el filtro de Gabor (2.2.2) como

$$G(x,y) = [G_1(x,y), G_2(x,y)],$$

con

$$G_1(x,y) = Ke^{-\pi(a^2(x-x_0)_r^2 + b^2(y-y_0)_r^2)} \cos(2\pi(u_0x + v_0y) + P)$$

y

$$G_2(x,y) = Ke^{-\pi(a^2(x-x_0)_r^2 + b^2(y-y_0)_r^2)} \text{sen}(2\pi(u_0x + v_0y) + P).$$

Las derivadas parciales de G_1 y G_2 son

$$\begin{aligned}
\frac{\partial G_1}{\partial x}(x,y) &= -2\pi k e^{-\pi(a^2(x-x_0)_r^2+b^2(y-y_0)_r^2)}\Big[\Big(a^2(x-x_0)_r\cos(\theta)\\
&\quad - b^2(y-y_0)_r\text{sen}(\theta)\Big)\cos(2\pi(u_0x+v_0y)+P)+u_0\text{sen}(2\pi(u_0x+v_0y)+P)\Big];\\
\frac{\partial G_1}{\partial y}(x,y) &= -2\pi k e^{-\pi(a^2(x-x_0)_r^2+b^2(y-y_0)_r^2)}\Big[\Big(a^2(x-x_0)_r\text{sen}(\theta)\\
&\quad + b^2(y-y_0)_r\cos(\theta)\Big)\cos(2\pi(u_0x+v_0y)+P)+v_0\text{sen}(2\pi(u_0x+v_0y)+P)\Big];\\
\frac{\partial G_2}{\partial x}(x,y) &= -2\pi k e^{-\pi(a^2(x-x_0)_r^2+b^2(y-y_0)_r^2)}\Big[\Big(a^2(x-x_0)_r\cos(\theta)\\
&\quad - b^2(y-y_0)_r\text{sen}(\theta)\Big)\text{sen}(2\pi(u_0x+v_0y)+P)-u_0\cos(2\pi(u_0x+v_0y)+P)\Big];\\
\frac{\partial G_2}{\partial y}(x,y) &= -2\pi k e^{-\pi(a^2(x-x_0)_r^2+b^2(y-y_0)_r^2)}\Big[\Big(a^2(x-x_0)_r\text{sen}(\theta)\\
&\quad + b^2(y-y_0)_r\cos(\theta)\Big)\text{sen}(2\pi(u_0x+v_0y)+P)-v_0\cos(2\pi(u_0x+v_0y)+P)\Big].
\end{aligned}$$

Por tanto, el determinante de la matriz jacobiana de G en un punto $(x,y)\in\mathbb{R}^2$ es

$$\begin{aligned}
\det JG(x,y) &= 4\pi^2k^2e^{-2\pi(a^2(x-x_0)_r^2+b^2(y-y_0)_r^2)}a^2\left[u_0\text{sen}(\theta)-v_0\cos(\theta)\right](x-x_0)_r\\
&\quad + 4\pi^2k^2e^{-2\pi(a^2(x-x_0)_r^2+b^2(y-y_0)_r^2)}b^2\left[u_0\cos(\theta)+v_0\text{sen}(\theta)\right](y-y_0)_r.
\end{aligned}$$

Si se redefine los parámetros

$$\begin{aligned}
c_1 &= a^2\left[u_0\text{sen}(\theta)-v_0\cos(\theta)\right],\\
c_2 &= b^2\left[u_0\cos(\theta)+v_0\text{sen}(\theta)\right],\\
c_3 &= c_1\cos(\theta)-c_2\text{sen}(\theta),\\
c_4 &= c_1\text{sen}(\theta)+c_2\cos(\theta),
\end{aligned}$$

entonces

$$\det JG(x,y) = 4\pi^2k^2e^{-2\pi(a^2(x-x_0)_r^2+b^2(y-y_0)_r^2)}[c_1(x-x_0)_r+c_2(y-y_0)_r]$$

y

$$\begin{aligned} c_1(x-x_0)_r + c_2(y-y_0)_r &= c_1((x-x_0)\cos(\theta) + (y-y_0)\text{sen}(\theta)) \\ &\quad + c_2((y-y_0)\cos(\theta) - (x-x_0)\text{sen}(\theta)) \\ &= (c_1\cos(\theta) - c_2\text{sen}(\theta))(x-x_0) \\ &+ (c_1\text{sen}(\theta) + c_2\cos(\theta))(y-y_0) \\ &= c_3(x-x_0) + c_4(y-y_0). \end{aligned}$$

De este modo,

$$\det JG(x,y) = 0 \Leftrightarrow c_4(y-y_0) = -c_3(x-x_0). \tag{2.3.18}$$

Defínase el conjunto

$$S = \{(x,y) \in \mathbb{R}^2 : c_4(y-y_0) = -c_3(x-x_0)\}. \tag{2.3.19}$$

El lugar geométrico de S es una línea cuando $c_4 \neq 0$ o $c_3 \neq 0$.

Proposición 2.3.15. Se cumplen las siguientes condiciones:

(1) Si $a = 0, b \neq 0$ y $u_0 \neq -v_0\tan(\theta)$, entonces $c_4 \neq 0$.

(2) Si $b = 0, a \neq 0$ y $v_0 \neq -u_0\tan(\theta)$, entonces $c_3 \neq 0$.

(3) Si $a \neq 0, b \neq 0$ y $v_0 \neq \pm u_0\sqrt{\frac{a^2\text{sen}^2(\theta)+b^2\cos^2(\theta)}{a^2\cos^2(\theta)+b^2\text{sen}^2(\theta)}}$, entonces $c_4 \neq 0$ o $c_3 \neq 0$.

Demostración. Sólo se demostrará (1) y (3). Supóngase que $a = 0$, $b \neq 0$, $u_0 \neq -v_0\tan(\theta)$ y $c_4 = 0$. Entonces $b^2(u_0\cos(\theta) + v_0\text{sen}(\theta))\cos(\theta) = 0$. Dado que $\tan(\theta)$ es un número real, se deduce que $\cos(\theta) \neq 0$, por ello $u_0\cos(\theta) + v_0\text{sen}(\theta) = 0$ y así $u_0 = -v_0\tan(\theta)$ lo cual es una contradicción y en consecuencia $c_4 \neq 0$.

Ahora, supóngase que $a \neq 0$, $b \neq 0$,

$$v_0 \neq u_0\sqrt{\frac{a^2\text{sen}^2(\theta) + b^2\cos^2(\theta)}{a^2\cos^2(\theta) + b^2\text{sen}^2(\theta)}}, \tag{2.3.20}$$

$$v_0 \neq -u_0\sqrt{\frac{a^2\text{sen}^2(\theta) + b^2\cos^2(\theta)}{a^2\cos^2(\theta) + b^2\text{sen}^2(\theta)}} \tag{2.3.21}$$

$$\tag{2.3.22}$$

y $c_4 = c_3 = 0$. Obsérvese que

$$\left(a^2\text{sen}^2(\theta) + b^2\cos^2(\theta)\right) u_0 = (a^2 - b^2)\text{sen}(\theta)\cos(\theta)v_0,$$
$$\left(a^2\cos^2(\theta) + b^2\text{sen}^2(\theta)\right) v_0 = (a^2 - b^2)\text{sen}(\theta)\cos(\theta)u_0.$$

Si $(a^2 - b^2)\text{sen}(\theta)\cos(\theta) = 0$, entonces $v_0 = 0 = u_0$ contradice (2.3.20). Es así como $(a^2 - b^2)\text{sen}(\theta)\cos(\theta) \neq 0$. Además, está claro que $u_0 \neq 0$ y $v_0 \neq 0$. Así

$$\left(a^2\text{sen}^2(\theta) + b^2\cos^2(\theta)\right)\frac{u_0}{v_0} = \left(a^2\cos^2(\theta) + b^2\text{sen}^2(\theta)\right)\frac{v_0}{u_0}.$$

De esto se desprende que

$$v_0^2 = u_0^2\frac{(a^2\text{sen}^2(\theta) + b^2\cos^2(\theta))}{(a^2\cos^2(\theta) + b^2\text{sen}^2(\theta))},$$

es decir,

$$v_0 = u_0\sqrt{\frac{(a^2\text{sen}^2(\theta) + b^2\cos^2(\theta))}{(a^2\cos^2(\theta) + b^2\text{sen}^2(\theta))}}$$

o

$$v_0 = -u_0\sqrt{\frac{(a^2\text{sen}^2(\theta) + b^2\cos^2(\theta))}{(a^2\cos^2(\theta) + b^2\text{sen}^2(\theta))}},$$

lo cual es una contradicción y se concluye que $c_4 \neq 0$ o $c_3 \neq 0$. ■

Teorema 2.3.16. Sea $G : \mathbb{R}^2 \to \mathbb{R}^2$ la transformación del filtro Gabor. Si se cumple una de las siguientes condiciones:

(1) $a = 0, b \neq 0$ y $u_0 \neq -v_0\tan(\theta)$;

(2) $b = 0, a \neq 0$ y $v_0 \neq -u_0\tan(\theta)$;

(3) $a \neq 0, b \neq 0$ y $v_0 \neq \pm u_0\sqrt{\frac{a^2\text{sen}^2(\theta)+b^2\cos^2(\theta)}{a^2\cos^2(\theta)+b^2\text{sen}^2(\theta)}}$;

entonces G es una transformación DP.

Demostración. Con cualquiera de las hipótesis, se deduce de la Proposición 2.3.15 que 2.3.19 es una recta. Supóngase sin pérdida de generalidad que $c_3 \neq 0$ y $c_4 \neq 0$ y defínase $c_5 = -\frac{c_3}{c_4}$.

Afirmación 1. Hay una familia numerable $\mathcal{B}$ de bolas abiertas en $\mathbb{R}^2$ tal que

$$\mathbb{R}^2 \setminus S = \cup\mathcal{B}$$

y para cada $B \in \mathcal{B}$, G es bi-Lipschitz en B.

De (2.3.18) y (2.3.19) se tiene que $\det JG(x,y) \neq 0$, para todo $(x,y) \in \mathbb{R}^2 \setminus S$. Tómese $\alpha > 0$ y considérense los siguientes conjuntos

$$\Psi_{1,n} = \{(x,y) \in \mathbb{R}^2 : c_5(x-x_0) + y_0 + \alpha + n - 1 \leq y \leq c_5(x-x_0) + y_0 + \alpha + n\},$$
$$\Psi_{2,n} = \left\{(x,y) \in \mathbb{R}^2 : c_5(x-x_0) + y_0 + \frac{\alpha}{n+1} \leq y \leq c_5(x-x_0) + y_0 + \frac{\alpha}{n}\right\},$$
$$\Psi_{3,n} = \left\{(x,y) \in \mathbb{R}^2 : c_5(x-x_0) + y_0 - \frac{\alpha}{n} \leq y \leq c_5(x-x_0) + y_0 - \frac{\alpha}{n+1}\right\},$$
$$\Psi_{4,n} = \{(x,y) \in \mathbb{R}^2 : c_5(x-x_0) + y_0 - \alpha - n \leq y \leq c_5(x-x_0) + y_0 - \alpha - n + 1\}.$$

Que son bandas en $\mathbb{R}^2$ paralelas a la línea S. Los conjuntos $\Psi_{2,n}$ y $\Psi_{3,n}$ están cerca de S cuando $n \to \infty$. Además,

$$\mathbb{R}^2 \setminus S = \left(\bigcup_{n\in\mathbb{N}} \Psi_{1,n}\right) \cup \left(\bigcup_{n\in\mathbb{N}} \Psi_{2,n}\right) \cup \left(\bigcup_{n\in\mathbb{N}} \Psi_{3,n}\right) \cup \left(\bigcup_{n\in\mathbb{N}} \Psi_{4,n}\right).$$

Para cada $k = 1,2,3,4$ y $n \in \mathbb{N}$ considérese $\{R_j^{k,n}\}_{j\in\mathbb{N}}$ una familia numerable de subconjuntos compactos de $\mathbb{R}^2$ tal que $\Psi_{k,n} = \bigcup_{j\in\mathbb{N}} R_j^{k,n}$. Sea $k,n,j \in \mathbb{N}$, para cada $(x,y) \in R_j^{k,n}$ existe según el Teorema 2.3.13 una bola abierta $B_{(x,y)}$ con centro en (x,y) tal que $B_{(x,y)} \subseteq \mathbb{R}^2 \setminus S$ y G es bi-Lipschitz en $B_{(x,y)}$. La familia $\{B_{(x,y)} : (x,y) \in R_j^{k,n}\}$ es una cubierta abierta de $R_j^{k,n}$. Por la compacidad de $R_j^{k,n}$, existe un subconjunto finito $M_j^{k,n}$ de $R_j^{k,n}$ tal que $R_j^{k,n} \subseteq \bigcup_{(x,y)\in M_j^{k,n}} B_{(x,y)}$. Defínase

$$\mathcal{B}_j^{k,n} = \{B_{(x,y)} : (x,y) \in M_j^{k,n}\}.$$

Entonces $\mathcal{B} := \bigcup_{k=1}^{4} \left[\bigcup_{n\in\mathbb{N}} \left(\bigcup_{j\in\mathbb{N}} \mathcal{B}_j^{k,n}\right)\right]$ es una familia numerable de bolas abiertas tal que $\mathbb{R}^2 \setminus S = \cup\mathcal{B}$ y por cada $B \in \mathcal{B}$, G es bi-Lipschitz en B.

Afirmación 2. Existe una familia numerable $\mathcal{V}$ de subconjuntos de S tal que

$$S = \cup\mathcal{V}$$

y para cada $V \in \mathcal{V}$, G es bi-Lipschitz en V.

Antes de probar esta afirmación, obsérvese que dado que las normas $\|\cdot\|_2$ y $\|\cdot\|_1$ son equivalentes, existen constantes positivas M_1 y M_2 tales que

$$M_1[|G_1(x,y) - G_1(x',y')| + |G_2(x,y) - G_2(x',y')|] \leq \|G(x,y) - G(x',y')\|_2 \leq M_2[|G_1(x,y) - G_1(x',y')| + |G_2(x,y) - G_2(x',y')|], \quad (2.3.23)$$

para todo $(x,y),(x',y') \in \mathbb{R}^2$. Defínase $H_1(x) = G_1(x, c_5(x-x_0)+y_0)$ y $H_2(x) = G_2(x, c_5(x-x_0)+y_0)$. Entonces

$$H_1(x) = ke^{-z_1(x-x_0)^2}\cos(z_2x + z_3),$$
$$H_2(x) = ke^{-z_1(x-x_0)^2}\text{sen}(z_2x + z_3),$$

donde $z_1 = \pi[a^2(\cos(\theta) + c_5\text{sen}(\theta))^2 + b^2(c_5\cos(\theta) - \text{sen}(\theta))^2]$, $z_2 = 2\pi(u_0 + c_5v_0)$ y $z_3 = 2\pi(v_0y_0 - c_5v_0x_0) + P$. De manera similar a la prueba del Teorema 2.3.8, se encuentra una sucesión numerable $\{E_i\}$ de subconjuntos de $\mathbb{R}$ tal que $\mathbb{R} = \underset{i\in\mathbb{N}}{\cup} E_i$ y H_1, H_2 son bi-Lipschitz en cada E_i. Entonces existen $\lambda_1, \lambda_2, \beta_1, \beta_2 > 0$ tales que

$$\lambda_1|x - x'| \leq |H_1(x) - H_1(x')| \leq \lambda_2|x - x'| \quad (2.3.24)$$

y

$$\beta_1|x - x'| \leq |H_2(x) - H_2(x')| \leq \beta_2|x - x'|, \quad (2.3.25)$$

para todo $x, x' \in E_i$. Defínase $V_i = \{(x,y) : x \in E_i \text{ y } y = c_5(x-x_0)+y_0\}$. Entonces

$$S = \underset{i\in\mathbb{N}}{\cup} V_i.$$

Ahora, se demostrará que G es bi-Lipschitz en V_i. Sea $(x,y),(x',y') \in V_i$. Luego $x, x' \in E_i$, $y = c_5(x-x_0)+y_0$ y $y' = c_5(x'-x_0)+y_0$. Las desigualdades (2.3.24) y (2.3.25) implican que

$$(\lambda_1 + \beta_1)|x - x'| \leq |G_1(x,y) - G_1(x',y')| + |G_2(x,y) - G_2(x',y')| \leq (\lambda_2 + \beta_2)|x - x'|. \quad (2.3.26)$$

Así por (2.3.23),

$$\begin{aligned}\frac{M_1(\lambda_1+\beta_1)}{\sqrt{1+c_5^2}}\|(x,y)-(x',y')\|_2 &= M_1(\lambda_1+\beta_1)|x-x'| \\ &\leq M_1[|G_1(x,y)-G_1(x',y')|+|G_2(x,y)-G_2(x',y')|] \\ &\leq \|G(x,y)-G(x',y')\|_2 \\ &\leq M_2[|G_1(x,y)-G_1(x',y')|+|G_2(x,y)-G_2(x',y')|] \\ &\leq M_2(\lambda_2+\beta_2)|x-x'| \\ &\frac{M_2(\lambda_2+\beta_2)}{\sqrt{1+c_5^2}}\|(x,y)-(x',y')\|_2.\end{aligned}$$

Se concluye que, G es bi-Lipschitz en V_i. ■

Capítulo 3

Modelo de reconocimiento biométrico con etapa de dimensión fractal

En el Capítulo 2 se establecieron las condiciones necesarias para que los filtros de Gabor conserven la dimensión fractal. Ahora, el análisis se centra en el proceso para realizar el reconocimiento biométrico de imágenes de iris de ojos. Se analizará el método que comprende las siguientes etapas:

Normalización. Se extrae el disco del iris y se transforma en una imagen rectangular.

Filtrado. Se aplican los filtros de Gabor para eliminar el ruido e identificar la información relevante.

Clasificación por dimensión fractal. Se realizarán diversos grupos de imágenes basados en el valor de su dimensión.

Entrenamiento y validación de la red neuronal. El problema multiclase con un número relativamente grande de etiquetas (individuos) se reducirá a n problemas multiclase con un número pequeño de etiquetas, incluso llegando a ser clasificación binaria, como consecuencia de la clasificación fractal. Se entrenarán cada uno de los

modelos con sus respectivos conjuntos de entrenamiento y prueba. La validación se realizará contando el número de reconocimientos correctos entre el total (accuracy).

Antes de establecer esta metodología, se incrementará el número de imágenes de la base *Phoenix* utilizando el proceso de *Data Augmentation*, el cual genera diversas imágenes a partir de una sola, esto se realiza mediante rotaciones, reflexiones, acercamientos y alejamientos. Pero debido a que pertenecen todas ellas a un sólo individuo, deberán estar dentro de la misma clasificación fractal, es decir, deben de tener el mismo valor de esta dimensión. Los resultados de la Sección 2.1 muestran que estas transformaciones conservan en general la dimensión, por lo que este problema se encuentra resuelto.

De igual manera, se requiere que la extracción del disco también sea una operación que no altere la dimensión original. Nuevamente, referimos a los resultados de la Sección 2.1, donde se garantiza tal preservación.

Cabe mencionar que, anteriormente, se ha establecido que el filtrado de Gabor no preserva la dimensión en general, por lo que de ahora en adelante se considerarán las condiciones del Teorema 2.3.16 sobre los parámetros para que se garantice que la dimensión no será modificada.

Dicho brevemente, en cada uno de los pasos del método no se debe alterar la dimensión fractal de la imagen original, garantizado esto, posteriormente, se aplicará la metodología propuesta y se concluirá con un análisis de escalabilidad.

3.1. Aumento de datos (Data Augmentation)

La base de fotografías *Phoenix* cuenta con un 128 imágenes correspondientes a los ojos izquierdo y derecho de 64 personas. Sin embargo, las redes neuronales presentan una baja eficiencia al ser entrenada con relativamente pocos datos ([2]). En esta investigación, se plantea contar con al menos 50 imágenes de un mismo ojo. Por esta razón, se incremen-

tará la cantidad de elementos mediante el proceso de Data Augmentation. Este método consiste en realizar transformaciones a la imagen original, en esta investigación se utilizarán reflexiones, rotaciones, acercamientos, alejamientos y cambios de color (definidas en la Sección 2.1). En la Figura 3.1 se muestran algunas imágenes generadas a partir de una sola.

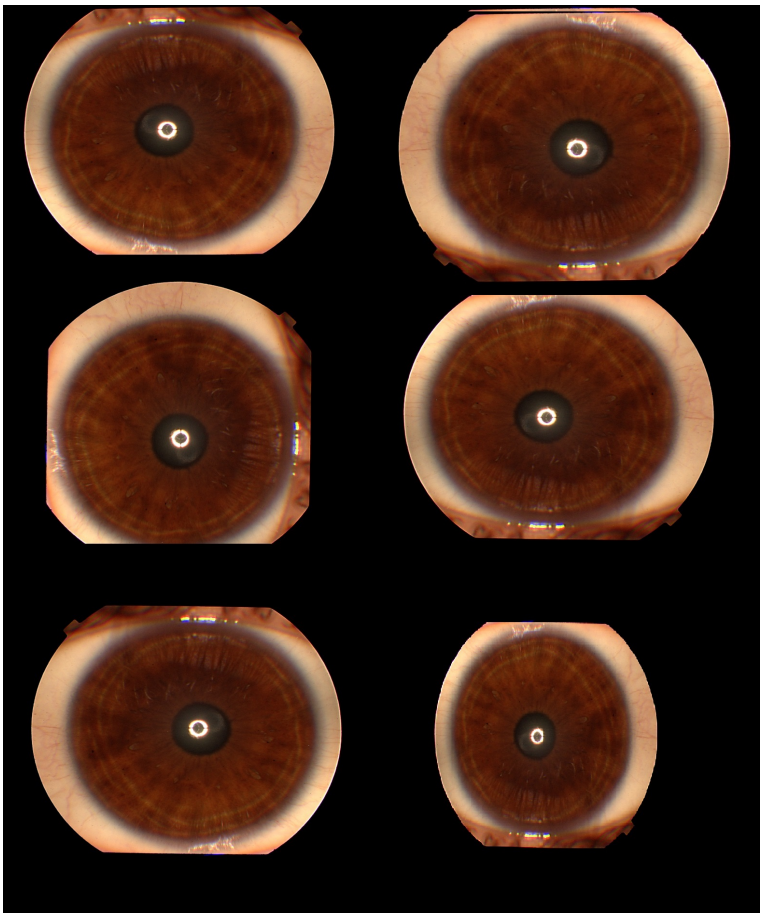

Figura 3.1: Imágenes generadas a través de reflexiones, rotaciones, acercamientos y alejamientos de la original.

Con esta técnica, se extiende la base hasta 6,400 imágenes. El siguiente paso es tratar a cada una de ellas, en la sección siguiente se exhibirán los resultados del proceso de normalización.

3.2. Normalización de las imágenes

El primer paso en la clasificación es extraer el iris y transformar la figura en una imagen única y estándar. La extracción se realiza determinando el centro, estimando el radio interior y exterior y recortando hasta obtener sólo el disco; posteriormente, mediante una transformación se mapea la forma circular en una rectangular.

Daugman en [13] ideó un modelo que reasignaba cada punto dentro de la región del iris en un par de coordenadas polares (r, θ), donde r oscila entre $[0, 1]$ y θ oscila entre $[0, 2\pi]$, a una región rectangular (x, y) (véase la Figura 2.3). Dicha transformada se escribe como $T(r, \theta) = (x(r, \theta), y(r, \theta))$ con

$$x(r, \theta) = (1 - r)x_p + rx_l, \tag{3.2.1}$$

$$y(r, \theta) = (1 - r)y_p + ry_l, \tag{3.2.2}$$

donde

$$\begin{aligned} x_p(\theta) &= x_c + r_{int}\cos(\theta), \\ y_p(\theta) &= y_c + r_{int}\,\mathrm{sen}(\theta), \\ x_l(\theta) &= x_c + r_{ext}\cos(\theta), \\ y_l(\theta) &= y_c + r_{ext}\,\mathrm{sen}(\theta), \end{aligned}$$

con (x_p, y_p) las coordenadas de la pupila, (x_l, y_l) son las coordenadas del disco entre la pupila y la esclerota (denominado limbo), (x_c, y_c) son las coordenadas del centro de la pupila, r_{int} y r_{ext} son los radios de la pupila y el limbo, respectivamente (véase la Figura 3.2). En el Teorema 2.1.7 se ha demostrado que esta transformación preserva la dimensión fractal.

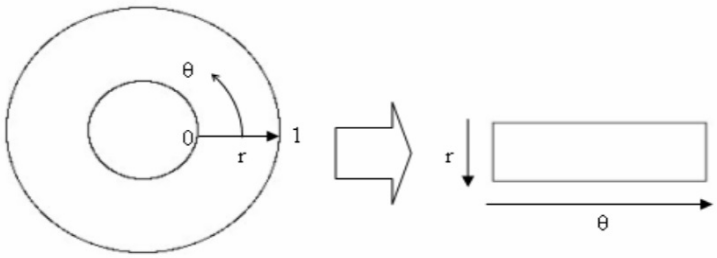

Figura 3.2: Normalización del disco del iris. Fuente: [11].

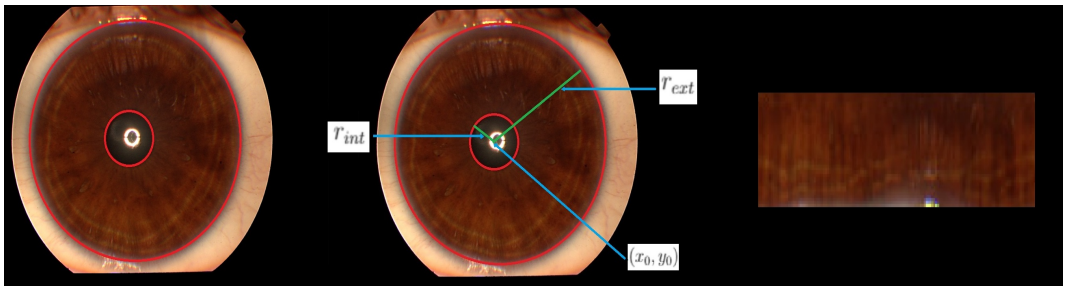

Figura 3.3: Normalización del disco del iris.

Como consecuencia de todo lo anterior, se aplica esta transformación a cada una de las imágenes de la base y se obtienen sus respectivas transformaciones rectangulares como las mostradas en la Figura 3.3.

3.3. Filtrado de imágenes y su dimensión

Posterior a la normalización, se realiza el filtrado a cada una de las cincuenta imágenes mediante la aplicación de la transformada de Gabor:

$$G(x,y) = Ke^{-\pi(a^2(x-x_0)_r^2+b^2(y-y_0)_r^2))}e^{i(2\pi(u_0x+v_0y)+P)},$$

con valores de los parámetros:

- $(x_0, y_0) = (27, 27)$,

- $a = \sqrt{\frac{2}{\pi}}$,
- $b = 0$,
- $u_0 = \frac{\cos(\pi/4)}{8}$,
- $v_0 = \frac{\text{sen}(\pi/4)}{8}$,
- $P = 10$.

De esta manera, por el Teorema 2.3.16, se garantiza que la dimensión no se encuentra alterada hasta este punto.

Luego, se realiza el cálculo de la dimensión de conteo de cajas a cada imagen normalizada mediante el límite:

$$\dim_B = \lim_{\delta \to 0} \frac{\log(N_\delta)}{\log(\frac{1}{\delta})}.$$

A manera de ejemplo, consideremos la imagen 001L (Figura 3.4). A través de la metodología se generan cincuenta imágenes en total, se procesan y se obtienen los valores de la dimensión fractal mostrados en la Cuadro 3.3.1.

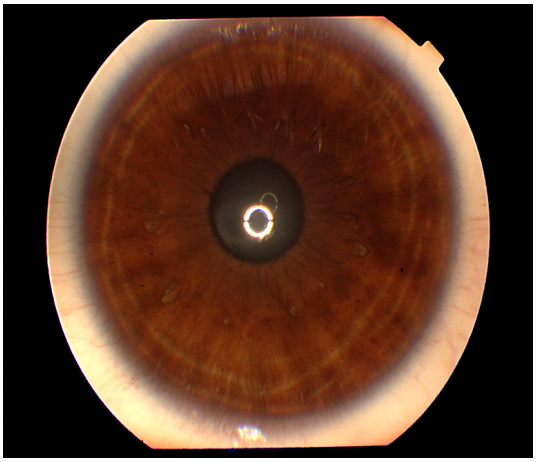

Figura 3.4: Imagen 001L.

1.83678689	1.83045492	1.83044097	1.83042878	1.83237551
1.83679563	1.83046672	1.83678834	1.8323431	1.83678689
1.83239831	1.83045492	1.83241435	1.83242502	1.83045492
1.83044097	1.83239356	1.8323884	1.83243951	1.83676358
1.83238847	1.83045492	1.836809	1.83678834	1.83679617
1.83045263	1.83239356	1.83241179	1.83047961	1.83678834
1.83045587	1.83678689	1.83678086	1.83237213	1.83677161
1.83678834	1.83245164	1.83681162	1.83674396	1.83047961
1.83235512	1.83681162	1.83681162	1.83043365	1.83676358
1.83048206	1.83678834	1.83243951	1.83238141	1.83236908

Cuadro 3.3.1: Dimensiones de las imágenes generadas a partir de la Figura 3.4.

Este proceso genera cincuenta valores de la dimensión fractal para una sola imagen que difieren, esto se debe a pequeños errores en la detección de los círculos del disco del iris generados por pequeñas variaciones en la toma de imágenes. Para solventar este problema, calcularemos el promedio de estos valores, por ejemplo, para la Figura 3.4 es de $\bar{x} = 1.8335216$, con una desviación estándar de $\sigma = 0.002694611$. Esto conduce a diseñar una prueba de hipótesis para garantizar que cada nueva captura de iris tenga un valor dimensional en el intervalo, con una confianza del 95 %. La prueba de hipótesis se construye como:

$$\begin{cases} H_0 : \mu = 1.8335216, \\ H_1 : \mu \neq 1.8335216, \end{cases} \tag{3.3.1}$$

donde μ es la media poblacional, H_0 es la hipótesis nula y H_1 es la alternativa. De esta manera, el intervalo de confianza del 95 % está dado por el estadístico

$$\bar{x} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}},$$

Donde $Z_{\alpha/2}$ es el estadístico de la prueba de hipótesis para la media, como se quiere una

confianza del 95 %, se tiene $\alpha = 0.05$, de modo que, se tienen los siguientes valores de los parámetros:

- $\bar{x} = 1.8335216$,
- $\sigma = 0.002694611$,
- $\alpha = 0.05$,
- $n = 50$,
- $Z_{\alpha/2} = 1.96$.

Resultando que el intervalo con una confianza del 95 % es

$$(1.8327747433174355, 1.8342685596799655). \tag{3.3.2}$$

Todo el proceso anterior se realiza para cada una de las 128 imágenes y se obtiene 6,400 imágenes correspondientes a 50 para un sólo ojo, cada una de ellas con un intervalo de confianza para la dimensión fractal, que denotaremos por $\mu_i \pm \sigma_i$, $i = 1, ..., 128$. Cabe mencionar que si se toman los valores con los dos primeros decimales en el intervalo (3.3.2) se garantiza al 100 % que es su valor real. Esto mismo ocurre con el resto de las imágenes procesadas.

En la siguiente Sección 3.4 se analizará la manera en la que se van a distribuir estas imágenes de tal forma que la red neuronal se entrene en un tiempo optimizado.

3.4. Modelo de redes neuronales para el reconocimiento

3.4.1. Redes neuronales convolucionales

La red neuronal convolucional se puede definir como una red neuronal creada de manera artificial [43] para que las computadoras puedan procesar imágenes. Es decir, que la red neuronal convolucional ha sido diseñada para imitar la estructura de cómo trabajan las redes neuronales del cerebro humano para llevar a cabo el proceso de la visión. Son muy utilizadas en la inteligencia artificial y el aprendizaje profundo. Especialmente, en el campo del procesamiento de imágenes. También, son conocidas como CNN (Convolutional Neural Network).

Básicamente, las redes de neuronas convolucionales pueden distinguir por medio de una arquitectura de capas tres dimensiones [43]. Estas dimensiones son el ancho, la altura y la profundidad. Además de las capas de color para distinguir el rojo, el verde y el azul. En otras palabras, son capaces de aprender a distinguir estas características por medio del cálculo de convoluciones en una serie de datos.

Una CNN consta de una capa de entrada, una capa de salida y varias capas ocultas entre ambas.

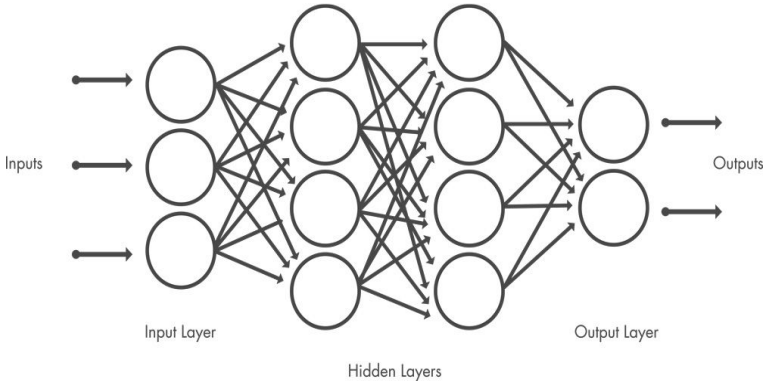

Figura 3.5: Red neuronal de cuatro capas. Obtenida de https://acortar.link/xSKLoO.

Estas capas realizan operaciones que modifican los datos, con el propósito de compren-

der sus características particulares. Las 3 capas más comunes son: convolución, activación o ReLU, y agrupación.

- **Convolución**: Aplica un conjunto de filtros convolucionales a las imágenes de entrada; cada filtro activa diferentes características de las imágenes.
- **Unidad lineal rectificada (ReLU)**: Mantiene los valores positivos y establece los valores negativos en cero, lo que permite un entrenamiento más rápido y eficaz. También se lo conoce como activación, ya que sólo las características activadas prosiguen a la siguiente capa.
- **Agrupación**: Simplifica la salida mediante reducción no lineal de la tasa de muestreo, lo que disminuye el número de parámetros que la red debe aprender.

Las CNN pueden tener decenas o cientos de capas, y cada una de ellas aprende a detectar diferentes características de una imagen. Se aplican filtros a las imágenes de entrenamiento con distintas resoluciones, y la salida resultante de convolucionar cada imagen se emplea como entrada para la siguiente capa. Los filtros pueden comenzar como características muy simples, tales como brillo y bordes, e ir creciendo en complejidad hasta convertirse en características que definen el objeto de forma singular.

A diferencia de una red neuronal tradicional, una CNN tiene pesos y valores de sesgos compartidos, que son los mismos para todas las neuronas ocultas de una capa determinada.

Esto significa que todas las neuronas ocultas detectan las mismas características, tales como bordes o formas, en diferentes regiones de la imagen. Esto hace que la red sea tolerante a la traducción de objetos de una imagen. Por ejemplo, una red entrenada para reconocer automóviles podrá hacerlo independientemente de dónde aparezca el automóvil en la imagen.

Capas de clasificación. Después de aprender características en muchas capas, la arquitectura de una CNN pasa a la clasificación.

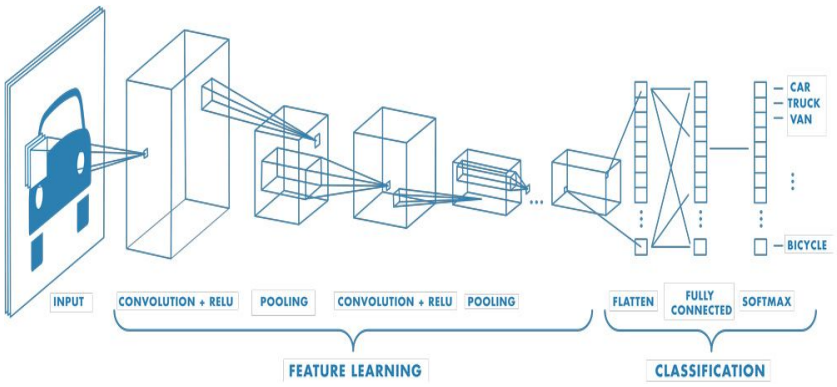

Figura 3.6: Ejemplo de red con múltiples capas convolucionales. Se aplican filtros a las imágenes de entrenamiento con distintas resoluciones, y la salida resultante de convolucionar cada imagen se emplea como entrada para la siguiente capa. Imagen obtenida de https://acortar.link/QdQT3l.

La penúltima capa está completamente conectada y genera un vector de K dimensiones, donde K es el número de clases que se pueden predecir, y contiene las probabilidades para cada clase de una imagen que se está clasificando.

La capa final de la arquitectura de la CNN utiliza una capa de clasificación para proporcionar la salida de clasificación final.

3.4.2. Entrenamiento por subclases

En el artículo [24], se propone dividir tal intervalo en cuatro grupos, por lo que se obtienen los siguientes subintervalos:

$$\text{Grupo A}=[1.30,1.45],$$
$$\text{Grupo B}=[1.45,1.60],$$
$$\text{Grupo C}=[1.60,1.75],$$
$$\text{Grupo D}=[1.75,1.90].$$

Sin embargo, el problema de la escalabilidad de los métodos de identificación aún no

Intervalo	No. de imágenes
A	24
B	28
C	52
D	24
Total	128

Cuadro 3.4.1: Frecuencia de imágenes por intervalo.

se resuelve con esta división, debido a que los intervalos son arbitrarios y pueden contener cantidades no uniformes de elementos, por ejemplo, en el Cuadro 3.4.1 se muestra la cantidad de imágenes con dimensión fractal dentro de cada categoría, obsérvese que si se multiplica por 50 y suponiendo que se sigue la misma tendencia, entonces en el intervalo C existirán alrededor de 2,600 imágenes para entrenar la red. De manera experimental se midió el tiempo de entrenamiento de la red neuronal con sólo dos usuarios (clases), esto es, con 100 imágenes, y se demoró alrededor de 10 minutos, esto con los recursos proporcionados por *Google Colab* (a lo más 12.7 GB en memoria RAM y 107.7 GB en disco). Este tiempo crece al incrementar el número de usuarios alrededor de 7 minutos. Por lo que se debe optimizar el tiempo de entrenamiento. En esta investigación, se propone dividir las imágenes en clases que contengan a lo más tres iris diferentes, de esta forma el entrenamiento se haría con 150 imágenes. Para lograr este entrenamiento, se divide el intervalo $[1, 2]$ de la siguiente manera:

- Se ordenan las dimensiones promedio de mayor a menor, $d_1, d_2, d_2, ..., d_{64}$, posteriormente se calculan los valores

$$\rho_{n,n+1} = d_{n+1} - d_n.$$

- Consideremos n fijo y $r_{n,n+1} = \frac{d_n + d_{n+1}}{2}$. Si se toman dos cifras decimales, se define

el intervalo como

$$[r_{n,n+1} - \rho_{n,n+1}, r_{n,n+1} + \rho_{n,n+1}]\,.$$

Si se consideran más de tres cifras decimales, tómese σ_n y σ_{n+1}, las correspondientes desviaciones estándar, en este caso, se define el intervalo como

$$[r_{n,n+1} - \rho_{n,n+1} - \sigma_n - \sigma_{n+1}, r_{n,n+1} + \rho_{n,n+1} + \sigma_n + \sigma_{n+1}]\,. \tag{3.4.1}$$

De esta manera se minimiza el numero de clases (usuario) en el problema. La diferencia entre tomar dos cifras o más es que en el primer caso se garantiza al 100 % que los futuros cálculos de la dimensión estarán en el intervalo, en el segundo, el nivel de confianza es del 95 %.

- Cada nuevo usuario que es ingresado entra en un intervalo de este tipo, el entrenamiento de la red es realizado sólo sobre este.
- Ya entrenada la red, se calcula la dimensión fractal a dos dígitos y se dirige hacia los usuarios dentro del intervalo donde se encuentre este valor, por lo que se optimiza el tiempo de reconocimiento.

A manera de ejemplo, considérese la imagen 001L, cuya dimensión en promedio a siete cifras decimales es 1.8335216 y la imagen 061L, con dimensión promedio 1.82250371, por lo que $d_{061L} < d_{001L}$.

Si se concideran dos cifras decimales, el valor de d_{061L} es 1.82 y el de d_{001L} es de 1.83. De aquí, $\rho_{061L,001L} = 0.01$ y $r_{061L,001L} = 1.825$, con esto, el intervalo propuesto es

$$[1.815, 1.835],$$

de esta manera, se realiza en entrenamiento para estas dos clases, luego, para hacer el reconocimiento, el algoritmo busca sólo sobre el intervalo propuesto. En el caso de que existiera intersección de este intervalo con otro, el mismo usuario se agregaría, y la búsqueda se realiza en cualquiera de ellos.

Si se consideran siete cifras decimales, tenemos que, $\sigma_{061L} = 0.001203542$, $\sigma_{001L} = 0.002694611$, $\rho_{061L,001L} = 0.01101789$ y $r_{061L,001L} = 1.828012655$, y el intervalo propuesto es

$$[1.813796458, 1.837694611].$$

Finalmente, siguiendo esta metodología, se encontró que a lo más en cada intervalo se encontraban tres usuarios, el tiempo de entrenamiento fue de alrededor de 7 minutos por intervalo y el reconocimiento posterior es realizado casi inmediatamente con una eficiencia del 99.37 %. Resta realizar la experimentación con equipos más potentes.

Conclusiones

El objetivo general de este trabajo de tesis se divide en dos temas principales: el primero es el análisis de la preservación de la dimensión fractal bajo transformaciones y filtros de Gabor, y el segundo es el de proporcionar un modelo de reconocimiento biométrico con etapa de dimensión fractal para optimizar el tiempo de entrenamiento. Para lograrlo, el Capítulo 1 está dedicado a introducir algunos conceptos y propiedades que son utilizados en el desarrollo del trabajo, principalmente, la definición de dimensión de conteo de cajas (Sección 1.3), transformaciones bi-Lipschitz (Sección 1.4) y la definición de transformaciones que preservan la dimensión fractal (Sección 1.5).

En el Capítulo 2 se mostraron diferentes transformaciones de interés en nuestro estudio: reflexión, rotación, normalización y filtros de Gabor; luego, se analizó si tales transformaciones y filtros preservan la dimensión fractal, se encontró que las primeras tres sí conservan la dimensión, sin embargo, el filtro de Gabor no preserva la dimensión fractal en general, la función del Ejemplo 2.2.2 es muestra de ello, este es el motivo principal de la búsqueda de las condiciones para que lo haga. Sin embargo, dada la complejidad de la función de Gabor, se comienza con un caso particular de la misma: las funciones campanas gaussianas, que son un ejemplo de funciones que no son bi-Lipschitz (Teorema 2.3.5) sino que son DP sobre intervalos acotados. También se discute el caso si el producto de una función de campana por alguna otra función resulta ser bi-Lipschitz, sin embargo,

esto no es cierto en general, el Ejemplo 2.3.7 muestra el caso cuando la función que multiplica es una exponencial , y el Teorema 2.3.6 muestra cuándo es una función diferenciable arbitraria, por lo que la condición de bi-Lipschiz no es suficiente para garantizar que se preserve la dimensión fractal. En el Teorema 2.3.8 se muestra que una clase especial de funciones de campana sí conserva la dimensión. Dado que el interés inicial es aplicar los filtros a funciones en la dimensión 2 y siguiendo con esta idea, en la Sección 2.3.1 se amplía el caso a los filtros de Gabor desde $\mathbb{R}^2$ hasta $\mathbb{R}^2$, que es la función de interés, ya que normalmente las imágenes que se filtran son bidimensionales y se generan imágenes de esa misma dimensión, obteniendo las condiciones sobre los parámetros para que la dimensión fractal se conserve bajo el filtrado de Gabor, tales como se encuentran en Teorema 2.3.16.

En el Capitulo 3 se muestra el proceso para el reconocimiento biométrico con etapa de dimensión fractal. En la Sección 3.1 se mostró la manera en que se puede aumentar la base *Phoenix* utilizando diversas técnicas de data augmentation, entre las cuales se encuentran las rotaciones y reflexiones. En la Sección 3.2 se mostró la forma de procesar una imagen del iris del ojo, la manera de extraerlo y la función con la que se transforma en una imagen rectangular. A continuación, para cada imagen en la base se generan cincuenta nuevas imágenes, se procesan y se obtienen los valores de la dimensión fractal para dichas imágenes generadas, esto se muestra en la Sección 3.3. Este proceso genera cincuenta valores de la dimensión fractal para una sola imagen, los cuales difieren entre ellos, se encontró que esto se debe a pequeños errores en la detección de los círculos del disco del iris generados por pequeñas variaciones en la toma de imágenes. Para solventar este problema, se diseña una prueba de hipótesis 3.3.1 sobre la media de las dimensiones fractales para garantizar que cada nueva captura de iris tenga un valor dimensional en el intervalo 3.3.2, con una confianza del 95 %. Finalmente, se propone dividir las imágenes en clases que contengan a los más tres iris diferentes 3.4.1 para hacer un entrenamiento con a lo más 3 usuarios. En es esta etapa donde se realiza el entrenamiento de la red neuronal, se descubrió que con los recursos proporcionados por Google Colab el tiempo es relativamente grande, alrededor de 10 minutos por usuario, sin embargo, el aprendizaje en

subclases reduce a sólo 7 minutos por subintervalo, posterior a ello, de forma experimental se encontró que el reconocimiento se ejecuta de manera inmediata con una eficiencia del 99.37 %. Al parecer, tal propuesta sí mejora los tiempos de entrenamiento, aunque el problema de escalabilidad aún persiste debido a que sólo se garantiza al 95 % que el calculo de la dimensión es el correcto si se consideran más de dos cifras, lo que es lo ideal cuando se considera una base con una cardinalidad relativamente grande.

Trabajo futuro

Durante el desarrollo de esta tesis han surgido algunas líneas futuras que se han dejado abiertas, y que se esperan abordar en un futuro, son el resultado de cuestiones que han ido surgiendo durante la realización de la misma. A continuación, se presentan algunas líneas que pueden desarrollarse como resultado de esta investigación.

- En el proceso de extracción del iris se detecta el círculo que esta entre el iris y la esclerótica, para luego extraer la información del iris, sin embargo, este círculo puede verse afectado por pequeñas variaciones en la toma de imágenes, se sugiere optimizar la detención de dicho círculo.
- Se ha demostrado que, en general, las transformaciones a las que son sometidas las imágenes no conservan el valor de la dimensión fractal. Existen diversos filtrados de imágenes que no se ha estudiado su preservación de la dimensión fractal, por ejemplo, la transformada de Fourier del filtro de Gabor (Fourier-Gabor) que es utilizado en diversos filtrados de señales.
- En el entrenamiento de la red neuronal se utilizaron los recursos proporcionados por *Google Colab*, sin embargo, se tiene la hipótesis que usando un equipo más potente el tiempo se minimiza.

Bibliografía

[1] Alberverio, S., Pratsiovytyi, M. y Torbin, G., *Transformations preserving the Hausdorff-Besicovitch dimension*, Central European Journal of Mathematics, Vol. 6, no. 1, pp. 119-128, 2008.

[2] Alvaro Artola, M., *Clasificación de imágenes usando redes neuronales convolucionales en Python*, Trabajo final de grado en Ingeniería de las Tecnologías de Telecomunicación, Universidad de Sevilla, España, 2019.

[3] Boles, W. y Boashash, B., *A human identification technique using images of the iris and wavelet transform*, IEEE Transactions Signal Processing, Vol. 46, no. 4, pp. 1185-1188, Abril 1998.

[4] Buldyrev, S. V., *Fractals in Biology*, Springer, New York, NY., pp. 3779-3802, 2009.

[5] Caicedo Ortiz, H. E., Santiago Cortes, E., López Bonilla, J. y Castañeda, H. O., *Fractal dimension and turbulence in Giant HII Regions*, IOP Publishing, Vol. 582, no. 1, `https://dx.doi.org/10.1088/1742-6596/582/1/012049`, 2015.

[6] Chien Chang Chen y D. C. Chen, *Multi-resolutional Gabor filter in texture analysis*, Pattern Recognition Letters, Vol. 17, no. 10, pp. 1069-1076, https://www.sciencedirect.com/science/article/pii/0167865596000657, 1996.

[7] Cuevas Juárez, B. L., *Funciones Lipschitz sobre Espacios Métricos*, Tesis de licenciatura, Facultad de Ciencias Físico Matemáticas, Benemérita Universidad Autónoma de Puebla, 2016.

[8] Daugman, J., *Advantage and disvantage of iris recognition*, http://www.cl.cam.ac.uk/~gd1000/addisadvans.html.

[9] Daugman, J., *Comparative biometric system accuracy*, http://www.cl.cam.ac.uk/jgdlOOOINPLsummary.pdf.

[10] Daugman, J., *Uncertainty relation for resolution in space, spatial-frequency and oriented optimized by two-dimensional visual cortical filters*, J. Opt. Sae. Am., Vol. 2, pp. 1160-1169, 1985.

[11] Daugman, J., *High confidence visual recognition of persons by a test of statistical independence*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 15, no. 11, pp. 1148-1161, Noviembre 1993.

[12] Daugman, J., *High confidence visual recognition of persons by a test of statistical independence*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 15, no. 11, pp. 1148-1161, Noviembre 1993.

[13] Daugman. J., *Biometric personal identification system based on iris analysis*, United States Patent, Patent Number: 5,291,560, 1994.

[14] Daugman, J., *How iris recognition works*, IEEE, Vol. 14, no. 1, pp. 21-30, 2004.

[15] Daugman, J., *How iris recognition works*, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 14, no. 1, pp. 21-30, 2004.

[16] Daugman, J., *Probing the uniqueness and randomness of irisCodes: Results from 200 billion iris pair comparisons*, Proceedings of IEEE, Vol. 94, no. 11, pp. 1927-1935, Noviembre 2006.

[17] Dobeš, M. y Machala, L., Iris Database, http://phoenix.inf.upol.cz/iris/

[18] Dunker, M., *Don't blink: Iris recognition for biometric identification*, SANS Institute, EUA, 2021.

[19] Eftekhari, A., *Fractal Dimension of Electrochemical Reactions*, The Electrochemical Society, Inc., Vol. 151, no. 9, pp. E291, Agosto 2004.

[20] Falconer K.J., *Fractal geometry*, University of Michigan, Chichester, EUA, 1990.

[21] Gabor D., *Theory of communications*, Inst. Elect. Eng., Vol. 93, pp. 429-457, 1946.

[22] Jampour, M., Estilayee, M., Naserasadi, A. y Ashourzadeh, M., *Extract and classification of iris images by fractal dimension and efficient color of iris*, International Journal of Computer Applications (975-8887), Vol. 18, no. 1, pp. 11-14, Marzo 2011.

[23] Jampour, M., y Naserasadi, A., *Extract and classification of iris images by fractal dimension and efficient color of iris*, International Journal of Computer Applications (975-8887), Vol. 18, no. 1, pp. 11-14, Marzo 2011.

[24] Jampour, M., y Naserasadi, A., *Extract and classification of iris images by fractal dimension and efficient color of iris*, International Journal of Computer Applications (975-8887), Vol. 18, no. 1, pp. 11-14, Marzo 2011.

[25] Jones J. y Palmer L., *An evaluation of the two-dimensional Cabor filters model of simple receptive fields in cat striate cortex*, Neurophisíalagy, Vol. 58, pp. 538-539, 1987.

[26] Kaufman. R., *On Hausdorff dimension of projections*, Cambridge University, Vol. 15, pp. 153–155, 1968.

[27] Khaw, P., *Iris recognition technology for improved authentication*, SANS Institute, EUA, 2002.

[28] Liu-Jimenez, J., Sanchez-Reillo, R. y Sanchez-Avila, C., *Biometric coprocessor for an authentication system using iris biometrics*, IEEE International Conference on Security Technology, pp. 131-135, 2004.

[29] Losa, G. A, Merlini, D., Nonnenmacher, T. F. y Weibel, E. R., *Fractals in Biology and Medicine*, Birkhäuser Basel, Vol. 4, Marzo 2006.

[30] López Aguilar, F. y Brambila Paz, F., *Antropología fractal*, CIMAT, México, 2007.

[31] Li Ma, Tieniu Tan, Yunhong Wang y Dexin Zhang, *Personal identification based on iris texture analysis*, IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1519–1533, 2003.

[32] Mandelbrot, B., *How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension*, Science, Vol. 156, pp. 636-638, 1967.

[33] Mandelbrot, B., The fractal geometry of nature, *W.H. Freeman*, San Francisco, 1982.

[34] Marcelja S. , *Mathematical description of the response of simple cortical cells* , J Opt Soc Am., Vol. 70, pp. 1297-1300 , Noviembre 1980.

[35] Martin Roche, D., Sanchez Avila, C. y Sanchez Reillo, R., *Iris recognition for biometric identification using dyadic wavelet transform zero-crossing*, IEEE International Carnahan Conference on Security Technology, pp. 272-277, 2001.

[36] Miranda Martínez, M., Oleschko, K., Parrot, J. F., Castrejón Vacio, F., Taud, H., Brambila Paz, F., et al., *Porosity in naturally fractured media: A fractal classification*, Instituto de Geología, UNAM, Vol. 23, pp. 199-214, 2006.

[37] Montesinos Ibáñez, J. F., *Dimensiones de Hausdorff y de Conteo de Cajas y aplicación en identificación biométrica del iris del ojo*, Tesis de licenciatura en matemáticas aplicadas, Universidad Tecnológica de la Mixteca, Huajuapan de León, Oaxaca, México, 2022.

[38] Montesinos Ibáñez, J. F., Pérez Becerra, T. y Sánchez Perales, S., *Fractal dimension preservation by Gabor filters and applications in eye iris extraction*. En proceso de publicación.

[39] Oleschko, K., Fuentes, C., Brambila, F. y Alvarez, R., *Linear fractal analysis of three Mexican soils in different management systems*, Soil Technology, Vol. 10, pp. 207-223, https://www.sciencedirect.com/science/article/pii/S0933363096001262 1997.

[40] NSTC Subcommittee on Biometrics and Identity Management Room, *Biometrics history*, http://www.biometrics.gov/Documents/BioHistory.pdf.

[41] Pollen D.A. y Ronner S.F., *Visual cortical neurons as localized spatial filters* , IEEE Transactions on Systems, Man, and Cybernetics, Vol. 13, pp. 907-916, Septiembre 1983.

[42] Santiago Cortés, E. y Martínez Ledezma, J. L., *Dimensión fractal en retinas humanas*, Journal de Ciencia e Ingeniería, Vol. 8, no. 1, pp. 59-65, Agosto 2016.

[43] Santillana Quesada, S., *Introducción a las redes neuronales para eltratamiento de imágenes*, Trabajo final de grado en Matemáticas, Universidad de Granada, España, 2022.

[44] Valdez Cepeda, R. D. y Olivares Sáenz, E., *Geometría fractal en la ciencia del suelo*, Terra Latinoamericana, https://www.redalyc.org/articulo.oa?id=57316311, 1998.

[45] Wildes, R. P., *Iris recognition: An emerging biometric technology*, IEEE, Vol. 85, no. 9, pp. 1348-1363, Septiembre 1997.

[46] Wildes, R. P., Asmuth, J. C., Green, G. L., Hsu, S.C., Kolczynski, R. J., Matey, J.R. y McBride, S.E., *A system for automated iris recognition*, Second IEEE Workshop on Applications of Computer Vision, pp. 121-128, 1994.