

“Espacios de Escala”

Proyecto Final

Visión Computacional II

Yolanda de Lira Domínguez

Contenido

<u>Capítulo 1 Introducción</u>	3
<u>1.1. Representación Multi-escala</u>	4
<u>1.2. Representación Espacio de Escala</u>	5
<u>1.2.1. Espacios de Escala para señales 1-D: Suavizamiento Gaussiano</u>	5
<u>1.2.2. Formulación de Difusión de Espacios de Escala</u>	6
<u>1.3. Definición de Espacios de Escala</u>	6
<u>1.3.1. Unicidad de el kernel Gaussiano</u>	7
<u>1.4. Propiedades de representación de espacios de escala</u>	8
<u>1.4.1. Promedio pesado y apertura finita</u>	9
<u>1.4.2. Propiedad de semi-grupos y suavizamiento en cascada</u>	10
<u>1.4.3. Separabilidad</u>	11
<u>1.4.4. Máximo principio</u>	11
<u>Capítulo 2 Espacios de Escala para señales discretas</u>	12
<u>2.1. Espacios de escala para señales 1-D discretas</u>	12
<u>2.2. Espacios de escala para señales N-D discretas</u>	14
<u>Capítulo 3 Difusión lineal y no lineal</u>	17
<u>3.1. Difusión lineal</u>	17
<u>3.1. Difusión no lineal</u>	18
<u>Capítulo 4 Aplicaciones de Espacios de Escala</u>	22
<u>4.1. Segmentación de imágenes</u>	22
<u>4.2. Suavizamiento preservando bordes</u>	22
<u>4.3. Detección de rasgos</u>	23
<u>Conclusiones</u>	24
<u>Referencias</u>	25

Espacios de Escala

(Scale-Space)

Capítulo 1 Introducción

En nuestra vida diaria usamos la visión como una de nuestras principales fuentes de información acerca del mundo. Comparado con el sentido del oído, el sentido visual da una más rica descripción del mundo. Comparado con el sentido del tacto, la visión nos permite captar información acerca de los objetos a gran distancia y sin afectar los objetos físicamente. Considerando la aparente facilidad con la cual obtenemos información del mundo con la luz que entra a nuestros ojos, se requiere un esfuerzo intelectual para apreciar que no es una tarea trivial [1].

Vistos a distancia, los objetos adquieren cierto encanto porque un *misterio los acecha*. Cuando la distancia se incrementa, los detalles se pierden. La imagen sufre un cambio topológico. Lucretius (Lucretius, 1951) en su "On the nature of the universe" describe cómo las cosas desaparecen a distancia, y además cómo aparentan cambiar. Por ejemplo, torres rectangulares distantes aparentan estar redondeadas. Este efecto es conocido como "emborronamiento" (*blurring*). Él también observó cómo un par de islas distantes aparentan unirse en una sola. Este es otro tipo de cambio topológico. Cuando la distancia se incrementa, algunos detalles pueden unirse o desaparecer [3].

Nosotros percibimos objetos en el mundo teniendo estructuras en escalas finas y gruesas. Un árbol por ejemplo, puede aparentar tener forma cilíndrica cuando es visto a gran distancia, sin embargo, está formado por un gran número de ramas. Visto de cerca, sus hojas son visibles, incluso podemos observar que tienen una textura a una escala fina [1].

El hecho de que los objetos en el mundo aparenten diferentes formas dependiendo de la escala de observación tiene importantes implicaciones en el análisis de datos medibles, tales como imágenes, con métodos automáticos. Estas ideas no son completamente nuevas en Visión Computacional. Las representaciones Multi-escala de imágenes en términos de pirámides fueron desarrolladas alrededor del 1970. La principal motivación era lograr eficiencia

computacional por estrategias de gruesas a finas¹. Sin embargo, pronto se descubrió que relacionar estructuras desde diferentes niveles en la representación multi-escala no era trivial.

En 1983, Witkin propuso que la escala fuera considerada como un parámetro continuo, de ese modo, se generalizaría la existente noción de pirámides Gaussianas. Él notó la relación con la ecuación de difusión, de aquí que encontró una bien fundada manera de relacionar estructuras de imágenes entre diferentes escalas. Pronto Koenderink promovió tal enfoque, el cual ha sido desarrollado en lo que ahora conocemos como Teoría de Espacios de Escala (*Scale-Space Theory*) [1].

1.1. Representación Multi-escala

Aunque algunos aspectos de escala han sido conocidos por mucho tiempo, el concepto de escala ha sido difícil de formalizar en una teoría matemática. Esto sólo durante las últimas décadas en que se ha desarrollado herramientas para manejar el concepto de escala en una manera formal. El interés para este desarrollo viene de la necesidad de desarrollar algoritmos robustos en Procesamiento de Imágenes, Visión Computacional y otros campos relacionados con procesamiento automático de señales.

Una metodología que ha sido propuesta para manejar la noción de escala en datos medibles, es representando señales medibles a múltiples escalas. Dado que, en general, niveles particulares de escalas no pueden ser supuestos previamente sin un fuerte conocimiento a priori, la única solución razonable es que el sistema visual sea capaz de manejar estructuras de imágenes a todas las escalas.

La principal idea de crear una representación multi-escala de una señal es generar una familia uni-paramétrica de señales derivadas, donde la información de escalas finas sea suprimida sucesivamente. Luego, se requiere un mecanismo que simplifica sistemáticamente los datos y remueve detalles de escalas más finas, o información de altas frecuencias. Esta operación, la cual recibe el nombre de suavizamiento de espacios de escala (*scale-space smoothing*) debe estar disponible en algún nivel de escala.

Pero, porqué representar una señal a múltiples escalas cuando toda la información se encuentra en los datos originales? La principal razón de esto, es representar explícitamente el aspecto multi-escala del dato en el mundo real. Otro objetivo es suprimir detalles innecesarios que posteriormente pueden simplificar tareas de procesamiento. Más técnicamente, la más reciente motivación refleja la común necesidad de suavizar como un paso de pre-procesamiento para muchos algoritmos numéricos, como supresión de ruido [1].

¹ En inglés *coarse-to-fine strategies*

1.2. Representación Espacio de Escala

Una metodología propuesta por Witkin (1983) y Koenderink (1984) para obtener tal representación multi-escala de una señal es introducir la señal en una familia uni-paramétrica de señales derivadas, el escala-espacio, donde el parámetro, denotado parámetro de escala $t \in R_+$,² describe el nivel actual de escala.

1.2.1. Espacios de Escala para señales 1-D: Suavizamiento Gaussiano

Dada una señal $f: R \rightarrow R$, la representación espacio de escala $L: R \times R_+ \rightarrow R$ se define tal que la representación a “escala cero”, sea igual³ a la señal original

$$L(\cdot; 0) = f(\cdot), \quad (1.1)$$

y las representaciones a escalas más gruesas están dadas por convoluciones de la señal dada con los kernels Gaussianos de ancho incrementando sucesivamente.

$$L(\cdot; t) = g(\cdot; t) * f(\cdot), \quad (1.2)$$

En términos de integrales explícitas, el resultado de la operación convolución '*' se escribe:

$$L(x; t) = \int_{\xi=-\infty}^{\infty} g(\xi; t) f(x - \xi) d\xi, \quad (1.3)$$

donde $g: R \times R_+ \setminus \{0\} \rightarrow R$ es el kernel Gaussiano (en una dimensión)

$$g(x; t) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t}, \quad (1.4)$$

La Figura 1.1. muestra el resultado de suavizar una señal uni-dimensional a diferentes escalas en la forma anteriormente descrita. Note como este suavizamiento sucesivo captura la noción intuitiva de información de escala fina que va siendo suprimida, y las señales comienzan a ser más suaves gradualmente.

² R_+ denota el conjunto de los números reales no negativos, y $R_+ \setminus \{0\}$ el mismo conjunto excluyendo el cero.

³ La notación $L(\cdot; 0) = f$ significa $L(x; 0) = f(x) \forall x \in R^N$

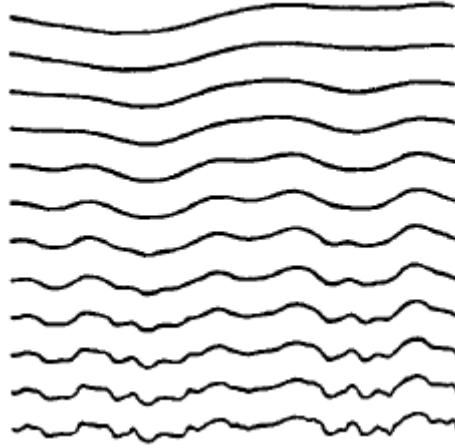


Figura 1.1. Esta figura muestra una señal uni-dimensional que ha sido suavizada sucesivamente por convolución con kernels Gaussianos de ancho incremental. (Adaptado de Witkin (1983))

1.2.2. Formulación de Difusión de Espacios de Escala

En términos de ecuaciones diferenciales, la evolución sobre escalas de la familia L de espacios de escala puede ser descrita por la *ecuación de difusión* (en una dimensión)

$$\partial_t L = \frac{1}{2} \nabla^2 L = \frac{1}{2} \partial_{xx} L, \quad (1.5)$$

De hecho, la representación de espacios de escala puede ser definida equivalentemente como la solución de (1.5) con condición inicial $L(\cdot; 0) = f(\cdot)$.

Esta analogía también da una directa interpretación física de la transformación de suavizamiento. La representación espacios de escala L de una señal f se puede entender como el resultado de dejar una distribución inicial de calor f desarrollarse en el tiempo t en un medio homogéneo. Por lo tanto, es de esperarse que detalles en escalas finas desaparezcan, y las imágenes comienzan a ser más difusas cuando el parámetro de escala incrementa.

1.3. Definición de Espacios de Escala

Cuando se construye una representación de espacios de escala, es de crucial importancia que la transformación de una escala fina a una escala gruesa realmente pueda considerarse como una simplificación, por lo cual los rasgos en escalas finas desaparecen en forma *monótona* con el incremento de escala. Si fuesen creados nuevas estructuras artificiales a escalas gruesas que no correspondan a regiones importantes en la representación de la escala más fina de la señal, sería imposible determinar ya sea que un rasgo una escala gruesa correspondía a una simplificación de alguna estructura de una escala

gruesa de la imagen original, o si fuera un fenómeno accidental, como una amplificación del ruido, creado por el método de suavizamiento –no de los datos. De aquí que, es de suma importancia que no se formen nuevas estructuras artificiales por la transformación de suavizamiento cuando se va de una escala fina a una más gruesa.

Cuando Witkin (1983) introdujo la noción de espacios de escala, estaba interesado en señales de una dimensión. Observó que el número de cruces por cero de la segunda derivada decrecían en forma monótona con la escala, y tomó esta como una característica básica de la representación. De hecho, esta propiedad se cumple para derivadas de orden arbitrario, lo cual implica que el número de extremos locales de cualquier derivada de la señal no puede incrementar con la escala. Desde este punto de vista, la convolución con un kernel Gaussiano posee una fuerte propiedad de suavizamiento.

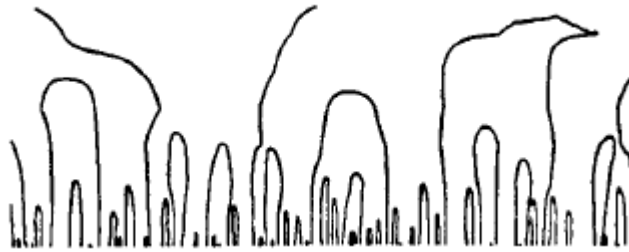


Figura 1.2. Posición de los cruces por cero (cruces por cero del Laplaciano con respecto a x) a través del espacio de escala. (adaptado de Witkin (1983))

1.3.1. Unicidad de el kernel Gaussiano

Koenderink (1984) extendió el concepto de espacios de escala a señales de dos dimensiones, introdujo la noción de *causalidad*, que significa que nuevas superficies de nivel no deben ser creadas cuando el parámetro de escala se incrementa. Equivalentemente, debería ser siempre posible trazar un valor de nivel de gris existente a cierto nivel de escala a un similar nivel de gris a algún nivel de escala más fino. Combinando causalidad con las nociones de *homogeneidad* e *isotropía* que esencialmente significan que todos los puntos espaciales y todos los niveles de escala deben ser tratados de manera similar, mostró que la representación de espacios de escala de una señal bidimensional por necesidad, debía satisfacer la *ecuación de difusión*

$$\partial_t L = \frac{1}{2} \nabla^2 L = \frac{1}{2} (\partial_{xx} + \partial_{yy}) L, \quad (1.6)$$

Dado que la convolución con el kernel Gaussiano

$$g(x, y; t) = \frac{1}{2\pi t} e^{-(x^2 + y^2)/2t}, \quad (1.7)$$

describe la solución de la ecuación de difusión a un dominio infinito, se sigue que el kernel Gaussiano es el único kernel para generar un espacio de escala. Esta formulación se extiende a dimensiones arbitrarias.

1.4. Propiedades de representación de espacios de escala

Existe una gran variedad de maneras posibles para construir una familia uni-paramétrica de señales partiendo de una señal dada. Bajo supuestos generales acerca de la estructura de los cálculos realizados en las primeras etapas de procesamiento, el kernel Gaussiano y sus derivadas son únicos. Las condiciones que especifican esta unicidad son básicamente linealidad e invarianza bajo traslación, combinados con diferentes maneras de formalizar la noción que una representación a escalas gruesas realmente debe corresponder a una simplificación de cualquier representación a escala más fina. No pueden ser creados nuevas estructuras por la transformación de suavizamiento.

Desde el punto de vista de representaciones multi-escala, la representación de espacios de escala definida en la sección 1.2 puede ser vista como un tipo muy especial de representación multi-escala, la cual comprende un parámetro continuo de escala y preserva el mismo muestreo espacial en todas las escalas.

Para señales continuas de dimensión arbitraria N , el concepto de espacios de escala se construye formalmente como sigue: dada una señal, la representación de espacios de escala $L: R^N \times R_+ \rightarrow R$ está definida tal que la representación en escala cero es igual a la señal original

$$L(:,0) = f(\cdot), \quad (1.8)$$

y las representaciones a escalas más gruesas están dadas por la convolución con kernels Gaussianos de ancho incrementando

$$L(:,t) = g(:,t) * f(\cdot), \quad (1.9)$$

En términos de integrales explícitas, el resultado de la operación convolución '*' se escribe:

$$L(x;t) = \int_{\xi \in R^N} g(\xi;t) f(x - \xi) d\xi, \quad (1.10)$$

donde $x = (x_1, \dots, x_N)^T \in R^N$ y $g: R^N \times R_+ \setminus \{0\} \rightarrow R$ es el kernel Gaussiano de dimensión N

$$g(x;t) = \frac{1}{(2\pi t)^{N/2}} e^{-x^T x / 2t}, \quad (1.11)$$

A continuación se mencionan algunas propiedades de la representación de espacios de escala descrita en la sección 1.2.

1.4.1. Promedio pesado y apertura finita

La razón por la que las señales se hacen más suaves se puede entender de diferentes maneras. Una forma simple es ver las gráficas de kernels Gaussianos correspondientes a diferentes valores de parámetros. Considerando primeramente el caso en una dimensión, la raíz del parámetro de escala $\sigma = \sqrt{t}$ es la desviación estándar del kernel Gaussiano, y puede ser interpretado como una característica de la longitud del gráfico. Desde este punto de vista, el valor de la representación de espacio de escala en cierto punto de la señal puede verse como el resultado de promediar la señal original con funciones de pesos simétricas donde el ancho va incrementando.

En dimensiones más altas, un descriptor natural de la longitud del kernel es su matriz de covarianza:

$$C(g(:,t)) = M_2(g(:,t)) - M_1(g(:,t))M_1^T(g(:,t)) = tI \quad (1.12)$$

donde M_2 es la matriz de segundos momentos

$$M_2(g(:,t)) = \int_{x \in R^N} xx^T g(x,t) dx, \quad (1.13)$$

M_1 es el vector de primer momento

$$M_1(g(:,t)) = \int_{x \in R^N} xg(x,t) dx = 0, \quad (1.14)$$

I es la matriz identidad de $N \times N$

Claramente, la transformación (1.10) representa una operación de promedio dado que el kernel Gaussiano es normalizado, es decir:

$$\int_{x \in R^N} g(x,t) dx = 1 \quad (1.15)$$

Filosóficamente, esta operación de promedio pesado puede ser interpretada como una manera de definir una *apertura* de una observación física. Para medir cualquier entidad del mundo real, es necesario integrar la entidad sobre alguna ventana de tamaño finito (no infinitesimal). Las dimensiones de los puntos ideales no pueden ser desarrolladas en realidad, dado que alguna cantidad finita de energía es necesaria para obtener una respuesta no infinitesimal que pueda ser registrada por un detector físico. La representación de espacios de escala de una señal f a cierta escala t corresponde al resultado de medir la señal usando una función de abertura simétrica rotacional con la misma característica de longitud $\sigma = \sqrt{t}$ a lo largo de todas las

direcciones coordenadas. Esto tiene el efecto de que las estructuras con características de longitud menores de σ se eliminarán.

1.4.2. Propiedad de semi-grupos y suavizamiento en cascada

Otra propiedad importante de la representación de espacios de escala radica en la propiedad de semi-grupos de la familia de kernels Gaussianos. El resultado de convolucionar un kernel Gaussiano con un kernel Gaussiano es otro kernel Gaussiano,

$$g(\cdot; t) * g(\cdot; s) = g(\cdot; t + s) \quad (1.19)$$

En términos de representación de espacios de escala, esto significa que la representación a una escala gruesa $L(\cdot; t_2)$ puede ser calculada de la representación a una escala fina $L(\cdot; t_1)$ convolucionando con un kernel Gaussiano con valor de parámetro $t_2 - t_1 > 0$

$$L(\cdot; t_2) = g(\cdot; t_2 - t_1) * L(\cdot; t_1) \quad (1.20)$$

Esta propiedad es la llamada como suavizamiento en cascada de la representación de espacios de escala, la cual significa que las señales a escalas gruesas realmente pueden ser interpretadas como más suaves que las correspondientes señales a escalas más finas, dado que la transformación de un nivel fino a un nivel grueso es del mismo tipo que la transformación de la señal original a un nivel no cero en el espacio de escala.

Estas relaciones se hacen más obvias en el dominio de Fourier. Dada una función $h: R^N \rightarrow R$, se define la transformada de Fourier como: $\hat{h}: R^N \rightarrow C$ por:

$$\hat{h}(w) = \int_{x \in R} h(x) e^{-i w^T x} dx \quad (1.21)$$

Entonces, la propiedad de semi-grupos supone la forma

$$\hat{g}(w, t) \hat{g}(w, s) = \hat{g}(w, t + s) \quad (1.22)$$

Usando:

$$\hat{L}(w, t) = \hat{g}(w, t) \hat{f}(w) \quad (1.23)$$

la propiedad de cascada puede ser escrita como:

$$\begin{aligned}
\hat{L}(w, t_2) &= \hat{g}(w, t_2) \hat{f}(w) \\
&= \hat{g}(w, t_2 - t_1) \hat{g}(w, t_1) \hat{f}(w) \\
&= \hat{g}(w, t_2 - t_1) \hat{L}(w, t_1)
\end{aligned} \tag{1.24}$$

1.4.3. Separabilidad

Una propiedad técnicamente usual del kernel Gaussiano es su separabilidad. El kernel Gaussiano de dimensión N $g: R^N \rightarrow R$ puede ser escrito como el producto de N kerneles de una dimensión $g_1: R \rightarrow R$. Sea $x = (x_1, \dots, x_N)^T \in R^N$. Entonces:

$$g(x, t) = \prod_{i=1}^N g_1(x_i; t) \tag{1.25}$$

En términos de expresiones explícitas para kerneles Gaussianos, tenemos:

$$\frac{1}{(2\pi t)^{N/2}} e^{-x^T x / 2t} = \prod_{i=1}^N \frac{1}{(2\pi t)^{1/2}} e^{-x_i^2 / 2t} \tag{1.26}$$

La separabilidad es importante en términos de eficiencia computacional, al menos cuando se implementa la operación de suavizamiento por convoluciones en el dominio del espacio.

1.4.4. Máximo principio

Una fuerte propiedad de suavizamiento de esta representación es el no realce de extremos locales, que corresponde al *principio máximo* para ecuaciones diferenciales parabólicas:

- Si a una cierta escala $t_0 \in R_+$ un punto $x_0 \in R$ es un máximo local para el mapeo $x \mapsto L(x; t_0)$, entonces el Laplaciano $\nabla^2 L(x_0; t_0)$ en este punto es negativo, es decir, $\partial_t L(x_0; t_0) < 0$.
- Por otro lado, si a una cierta escala $t_0 \in R_+$ un punto $x_0 \in R$ es un mínimo local para el mapeo $x \mapsto L(x; t_0)$, entonces el Laplaciano $\nabla^2 L(x_0; t_0)$ en este punto es positivo, es decir, $\partial_t L(x_0; t_0) > 0$.

Capítulo 2 Espacios de Escala para señales discretas

La teoría de espacios de escala ha sido desarrollada y bien establecida para imágenes y señales continuas. Sin embargo, esto no dice algo acerca de cómo debería implementarse computacionalmente en problemas reales, es decir, para imágenes y señales discretas. En principio, hay dos posibles enfoques:

- Aplicar los resultados de la teoría de espacios de escala continua discretizando las ecuaciones. Por ejemplo, la integral de convolución (1.3) puede ser aproximada por una suma usando métodos numéricos estándar, o la ecuación de difusión (1.5) puede ser discretizada en el espacio. Si los métodos numéricos se eligen con cuidado, se pueden esperar aproximaciones razonables a los valores numéricos continuos. Pero esto no garantiza que las condiciones del original espacio de escala se preserven.
- Formular una teoría genuina en el caso discreto postulando axiomas apropiados.

2.1. Espacios de escala para señales 1-D discretas

Una representación multi-escala es una familia de señales derivadas que pretenden representar la señal original a varios niveles de escala. Cada miembro de la familia está asociado con un valor de parámetro de escala que describe el nivel actual de escala. El parámetro de escala, denotado por t , puede ser discreto $t \in \mathbb{Z}_+$ o continuo $t \in \mathbb{R}_+$, lo cual nos da dos tipos diferentes de espacios de escala discretos: señales discretas con un parámetro de escala discreto y señales discretas con un parámetro de escala continuo. En ambos casos, se tiene como supuestos básicos:

- Cada representación debería ser generada por una transformación lineal e invariante bajo traslación de la señal original. De donde, el operador de suavizamiento puede ser expresado como un operador de convolución.
- Un valor incremental del parámetro de escala t debería corresponder a niveles más gruesos de la escala y señales con menos estructura. En particular $t = 0$ debería representar la señal original.
- Todas las señales deberían ser funciones de valores reales $\mathbb{Z} \rightarrow \mathbb{R}$ definidas en el mismo gris infinito; en otras palabras, no deberán usarse representaciones en pirámides.

El requerimiento esencial es que una señal a un nivel más grueso de escala debería contener menor estructura que una señal a niveles más finos de escala. Si consideramos el número de extremos locales como una medida de la cantidad de estructura, es necesario que el número de extremos locales en el espacio no incremente cuando se pasa de niveles de escalas finas a gruesas. Se puede demostrar que la familia de funciones generadas por convolución con

el kernel Gaussiano cumple esta propiedad en el caso continuo. A continuación se darán algunos axiomas básicos del análisis en una dimensión.

Definición 2.1. (Kernel de espacios de escala discretos 1-D)

Un kernel discreto de una dimensión $K : Z \rightarrow R$ se dice que es un kernel de espacios de escala si para todas las señales $f_{in} : Z \rightarrow R$ el número de extremos locales en la señal convolucionada $f_{out} : h * f_{in}$ no excede el número de extremos locales de la señal original.

Aquí, se puede hacer una importante observación, que esta definición puede expresarse equivalentemente en términos de cruces por cero, sólo reemplazando “extremos locales” por “cruces por cero”. El resultado se sigue del hecho de que un extremo local en una función discreta f es equivalente a un cruce por cero de su primera diferencia Δf , definida por $(\Delta f)(x) = f(x+1) - f(x)$, y este operador de diferencia conmuta con el operador de convolución.

Sin embargo, la definición mencionada tiene algunas consecuencias, significa que el número de extremos locales (cruces por cero) en cualquier diferencia de n de la señal convolucionada no puede ser más grande que el número de extremos locales (cruces por cero) en la diferencia de orden n de la señal original. Actualmente, el resultado puede generalizarse a operadores lineales arbitrarios.

Proposición 2.2. (Propiedad general de suavizamiento de kernels discretos de espacios de escala)

Sea $K : Z \rightarrow R$ un kernel discreto de espacios de escala y sea L un operador (del espacio de funciones discretas reales) que conmuta con K . Entonces para cualquier $f : Z \rightarrow R$ el número de extremos locales en $L(h * f)$ no excede el número de extremos locales de $L(f)$.

Los siguientes son las propiedades más importantes que deben tener los kernels discretos de espacios de escala en una dimensión.

- Positividad y unimodalidad en el dominio espacial. La positividad significa que todos los coeficientes de un kernel de espacios de escala deben tener el mismo signo. La unimodalidad significa que la secuencia de los coeficientes de un kernel de espacios de escala $\{K(n)\}_{n=-\infty}^{\infty}$ debe ser unimodal.
- Positividad y unimodalidad en el dominio de la frecuencia. Estas dos propiedades deben cumplirse también en el dominio de la frecuencia.

Una secuencia real se dice que es unimodal si es primero ascendente (descendente) y luego descendente (ascendente).

En términos de ecuaciones diferenciales, podemos interpretar la representación de espacios de escala discretos de acuerdo al siguiente:

Teorema 2.3. (Formulación de difusión del espacio de escala discreto)

Dada una señal discreta $f: Z \rightarrow R$ en l_1 sea $L: Z \times R_+ \rightarrow R$ la representación de espacios de escala dada por:

$$L(x; t) = \sum_{n=-\infty}^{\infty} T(n; t) f(x - n), \quad (2.1)$$

donde $T: Z \times R_+ \rightarrow R$ es el análogo discreto del kernel Gaussiano. Entonces, L satisface la solución del sistema de ecuaciones diferenciales ordinarias ($x \in Z$)

$$\partial_t L(x; t) = \frac{1}{2} (L(x + 1; t) - 2L(x; t) + L(x - 1; t)), \quad (2.2)$$

con condiciones iniciales $L(x; 0) = f(x)$, es decir, el sistema de ecuaciones diferenciales obtenidas si la ecuación de difusión (1.5) es discretizada en el espacio, pero resuelta analíticamente en tiempo.

2.2. Espacios de escala para señales N-D discretas

La extensión de la teoría de espacios de escala en una dimensión a dos o más dimensiones no es trivial. Koenderink (1984) deriva el espacio de escala para imágenes continuas de dos dimensiones de tres supuestos: causalidad, homogeneidad e isotropía. La idea principal es que debería ser posible trazar cualquier nivel de gris en a escalas gruesas a un correspondiente nivel de gris a una escala más fina. En otras palabras, no deben crearse nuevas curvas de nivel cuando se incrementa el parámetro de escala.

Por supuesto, es imposible aplicar estas ideas directamente en el caso discreto dado que no hay correspondencias directas a curvas de nivel o geometría diferencial para señales discretas. Una forma alternativa de expresar el requerimiento de causalidad, sin embargo, es requerir que si para algún nivel de escala t_0 un punto x_0 es un máximo local para la representación de espacios de escala a ese nivel (considerado como una función de el espacio de coordenadas solamente), entonces su valor no debe incrementarse cuando el parámetro de escala se incrementa. Análogamente, si un punto es un mínimo local entonces su valor no debe decrecer cuando el parámetro de escala incrementa.

La representación de espacios de escala para señales de más altas dimensiones se puede construir análogamente al caso uni-dimensional. Primero, estableceremos que el espacio de escala debe ser generado por

convolución con una familia uni-paramétrica de kerneles, es decir:
 $L(x;0) = f(x)$ y para $t > 0$

$$L(x;t) = \sum_{\xi \in Z^N} T(\xi;t) f(x - \xi), \quad (2.3)$$

Ahora se hará un resumen de las propiedades básicas que una familia debe satisfacer para ser familia candidata para generar un espacio de escala (lineal).

Definición 2.4. (Familia de Kerneles de espacios de pre-escala)

Una familia uni-paramétrica de kerneles $T: Z^N \times R_+ \rightarrow R$ se dice que es una familia de kerneles de espacios de pre-escala si satisface:

- $T(\cdot;0) = \delta(\cdot)$
- La propiedad de semi-grupos $T(\cdot;s) * T(\cdot;t) = T(\cdot;s+t)$
- La propiedad de simetría $T(-x_1, x_2, \dots, x_N; t) = T(x_1, x_2, \dots, x_N; t)$ y $T(P_k^N(x_1, x_2, \dots, x_N); t) = T(x_1, x_2, \dots, x_N; t)$ para toda $x = (x_1, x_2, \dots, x_N) \in Z^N$, toda $t \in R_+$, u todas las posibles permutaciones P_k^N de N elementos, y
- El requerimiento de continuidad $\|T(\cdot;t) - \delta(\cdot)\|_1 \rightarrow 0$ cuando $t \downarrow 0$

Definición 2.5. (Representaciones pre-escala)

Sea $f: Z^N \rightarrow R$ una señal discreta y sea $T: Z^N \times R_+ \rightarrow R$ una familia de kerneles de espacios de pre-escala. Entonces la familia uni-paramétrica de señales $L: Z^N \times R_+ \rightarrow R$ dada por (2.3) se dice que es una representación pre-escala de f generada por t

Definición 2.6. (Propiedad de espacios pre-escala: no realce)

Una familia uni-paramétrica de señales $L: Z^N \times R_+ \rightarrow R$ se dice que posee propiedades de pre-escala, o equivalentemente no realza extremos locales, si para todo valor del parámetro de escala $t_0 \in R_+$ se cumple que si $x_0 \in Z^N$ es un punto extremo local para el mapeo $x \mapsto L(x; t_0)$ entonces la derivada de L con respecto a t en este punto satisface:

$$\partial_t L(x_0; t_0) \leq 0 \quad \text{si } x_0 \text{ es un punto máximo local} \quad (2.4)$$

$$\partial_t L(x_0; t_0) \geq 0 \quad \text{si } x_0 \text{ es un punto mínimo local} \quad (2.5)$$

Con esto, podemos decir que una familia de kerneles de espacios de pre-escala es una familia de kerneles de espacios de escala si esta propiedad se satisface para cualquier señal de entrada.

Definición 2.7. (Familia de kerneles de espacios de escala)

Una familia uni-paramétrica de kernels de espacios de pre-escala $T: Z^N \times R_+ \rightarrow R$ se dice que es una familia de kernels de espacios de escala si para cualquier señal $f: Z^N \rightarrow R$ en l_1 la representación de espacios de pre-escala de f generada por T posee las propiedades de espacios de pre-escala.

Definición 2.8. (Representación de espacios de escala)

Una representación de espacios de pre-escala $L: Z^N \times R_+ \rightarrow R$ de una señal $f: Z^N \rightarrow R$ generada por una familia de kernels $T: Z^N \times R_+ \rightarrow R$ que son kernels de espacios de escala, se dice que es una representación de espacios de escala de f .

Con esto, podemos formar representaciones de espacios de escala para señales discretas de dimensión N .

Capítulo 3 Difusión lineal y no lineal

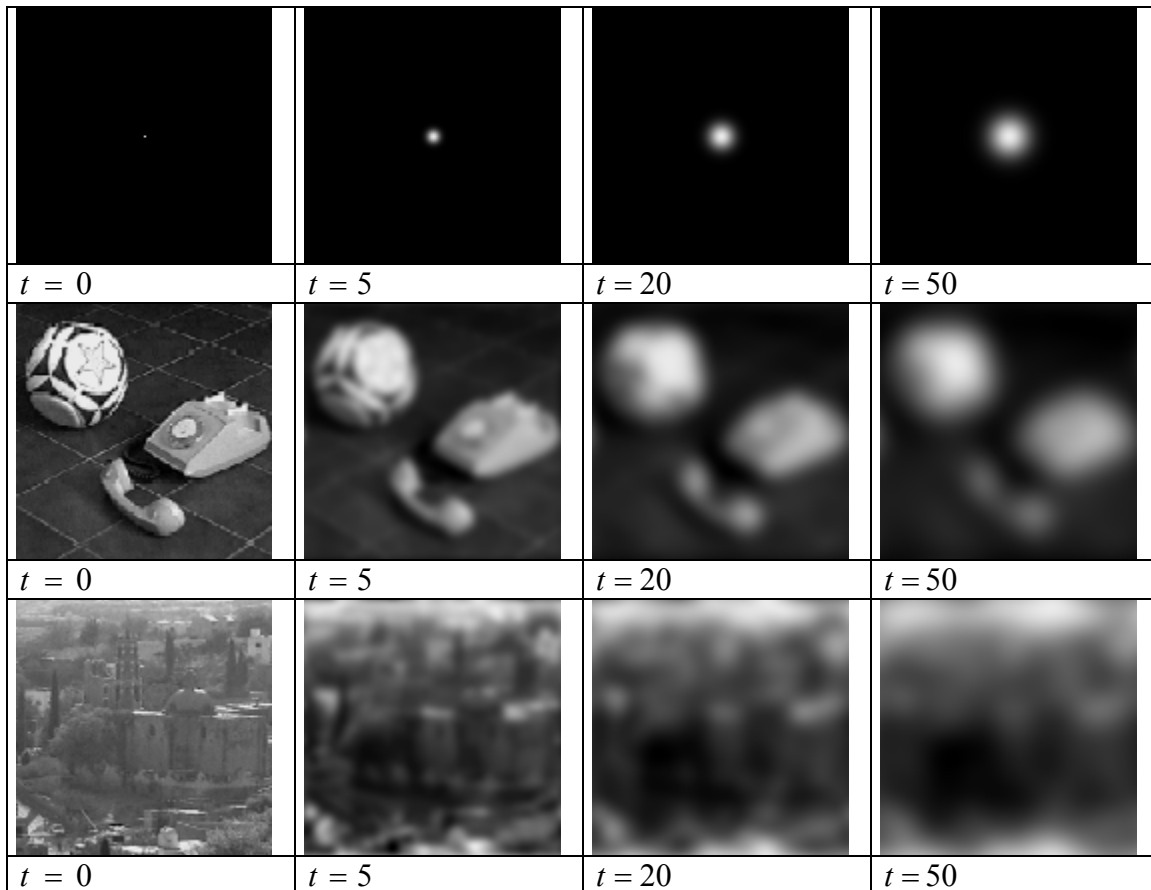
En los capítulos anteriores se ha visto la teoría básica para las representaciones de espacios de escala, donde hemos visto las condiciones que se deben cumplir para tener una representación de espacios de escala. Algunos autores han estudiado estas condiciones, llegando a la conclusión de que las representaciones de espacios de escala pueden clasificarse como lineales y no lineales, que corresponden al tipo de difusión realizada en la señal original. En las siguientes secciones se explica cada tipo de difusión y se muestran algunos ejemplos de los resultados que se obtienen en cada caso.

3.1. Difusión lineal

La representación de espacios de escala lineal es una familia uni-paramétrica de señales derivadas de la ecuación lineal de difusión o ecuación de calor (1.5). Esta representación tiene las propiedades mencionadas en el capítulo 1, que son básicamente: causalidad, homogeneidad e isotropía.

Este tipo de representación consiste básicamente en convolucionar la señal original con un kernel Gaussiano de parámetro t (varianza) sucesivamente, incrementando t en cada nivel de escala.

La figura 3.1. muestra algunos ejemplos de este tipo de representación



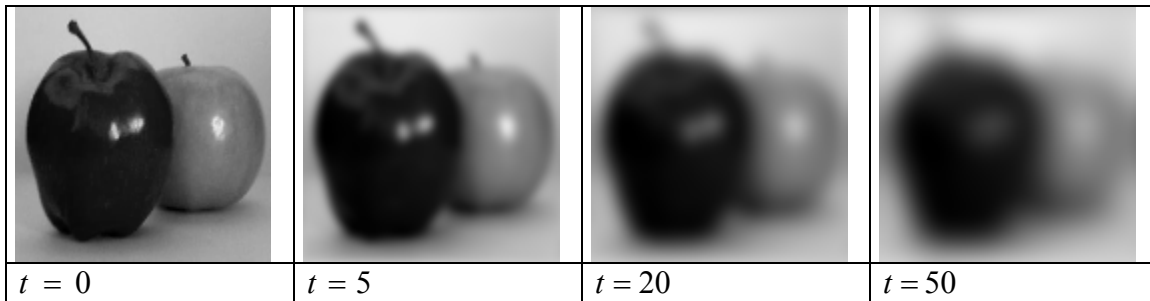


Figura 3.1. Ejemplos de representaciones de espacios de escala en imágenes.

3.1. Difusión no lineal

En la difusión lineal, muchas veces se pierden rasgos importantes, entre los que podemos mencionar los bordes de la imagen. La verdadera ubicación de un borde a una escala gruesa no está directamente disponible en la imagen a nivel grueso, lo cual puede verse en el ejemplo en 1-D de la figura 1.2, donde la ubicación de los bordes se traslada de su posición real. Algo similar sucede con señales 2-D.

La razón de esto es obvia, el suavizamiento Gaussiano no respeta los bordes naturales de los objetos. Esto puede verse en las imágenes de ejemplo: los bordes naturales de las imágenes se tornan difíciles de localizar conforme se incrementa el valor del parámetro de escala.

Con esta motivación, Perona y Malik [4] enunciaron el criterio candidato que debe satisfacer el generar descripciones multi-escala, “significativas”.

- *Causalidad*: Como lo señalaron Witkin (1983) y Koenderink (1983), una representación de espacios de escala debe tener la propiedad de que detalles espurios no deben ser generados cuando se pasa de una escala fina a una gruesa.
- *Localización inmediata*: En cada resolución, las regiones de frontera, deben coincidir con las regiones de frontera significativas semánticamente a esta resolución.
- *Suavizamiento por piezas adecuado*: En todas las escalas, el suavizamiento intrarregión debe ocurrir preferentemente sobre suavizamiento interregión.

Hay una manera simple de modificar el paradigma lineal de espacios de escala, para lograr los objetivos mencionados. En la ecuación de difusión, el coeficiente de difusión c se asume constante independientemente de la ubicación en el espacio. No hay una razón específica por lo cual esto deba ser así.

Perona y Malik [4] encontraron una conveniente elección de $c(x, y, t)$ para satisfacer el segundo y tercer criterios, sacrificando el primero. Consideraron la ecuación de difusión anisotrópica

$$I_t = \text{div}(c(x, y, t)\nabla I) = c(x, y, t)\Delta I + \nabla c \nabla I \quad (3.1)$$

donde div indica el operador divergencia, Δ y ∇ el Laplaciano y el gradiente con respecto a las variables espaciales. Si $c(x, y, t)$ es constante, ésta se reduce a la ecuación de difusión de calor $I_t = c\Delta I$. Supongamos que conocemos la ubicación de los bordes al tiempo (escala) t . Sería deseable suavizar entre una región preferentemente que a través de los bordes. Esto podría lograrse fijando un coeficiente de conducción 1 en el interior de cada región y 0 en los bordes. El suavizamiento se hace entonces separadamente en cada región sin interactuar entre ellas. La región de los bordes quedaría bien marcada.

Por supuesto, no conocemos las regiones de frontera en cada escala, por lo que hace una estimación de ésta. Sea $E(x, y, t)$ tal estimación: una función definida sobre la imagen que debe satisfacer las propiedades:

- $E(x, y, t) = 0$ en el interior de cada región
- $E(x, y, t) = Ke(x, y, t)$ en un borde, donde e es un vector unitario normal al punto que es borde y K es el contraste local del borde.

Si se dispone de una estimación de $E(x, y, t)$, el coeficiente $c(x, y, t)$ se puede elegir como función de la magnitud de E : $c = g(\|E\|)$. El gradiente de la función de brillo es un estimador que arroja buenos resultados.

La figura 3.2. muestra la forma de la función g , la cual se puede elegir por ejemplo como $g(\|\nabla I\|) = e^{-(\|\nabla I\|/K)^2}$ o como $g(\|\nabla I\|) = \frac{1}{1 + \left(\frac{\|\nabla I\|}{K}\right)^2}$

Donde K es una constante.

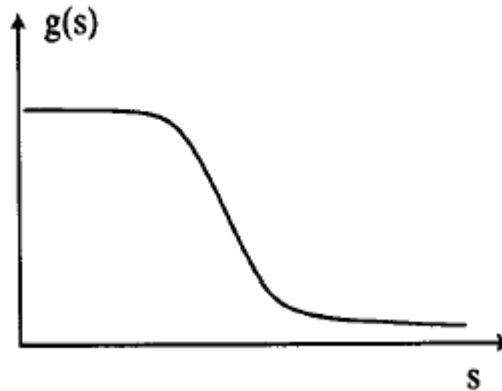


Figura 3.2. La forma de la función g que representa el coeficiente de difusión.

Así pues, se tiene que:

$$c(x, y, t) = g(\|\nabla I(x, y, t)\|) \quad (3.2)$$

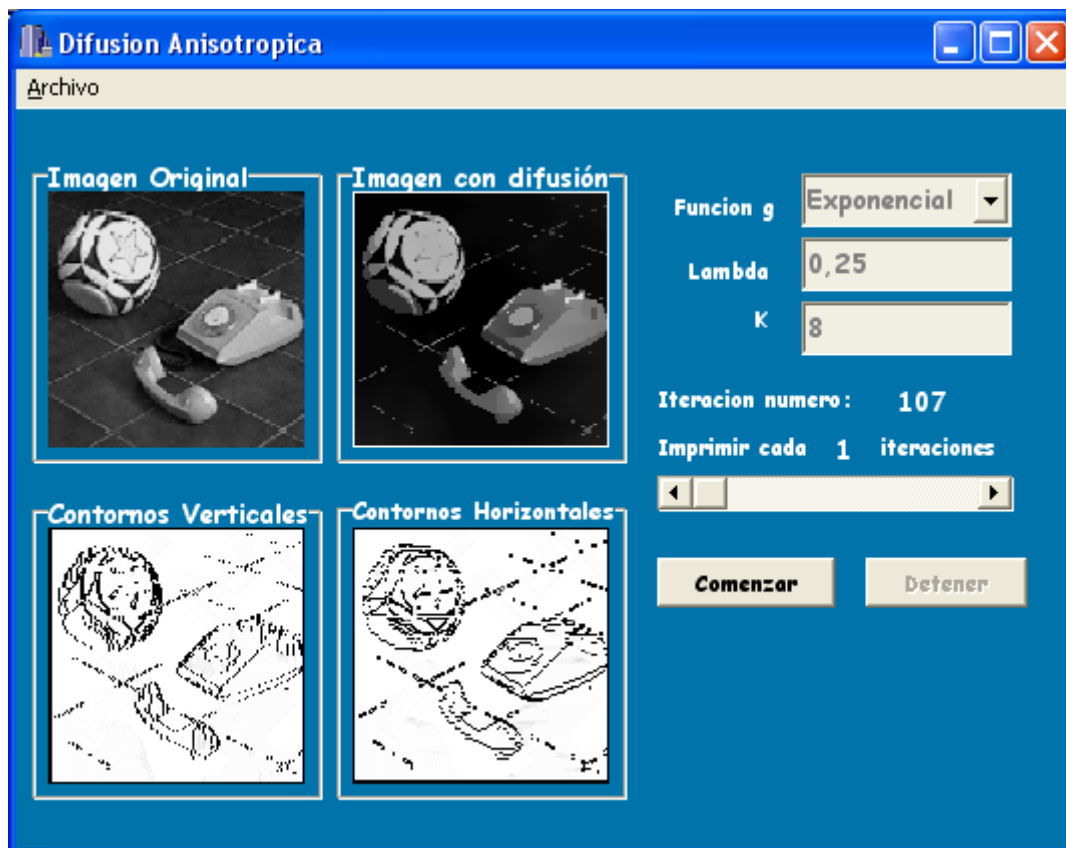


Figura 3.3. Ejemplo de difusión anisotrópica y detección de bordes usando espacios de escala con bola.bmp

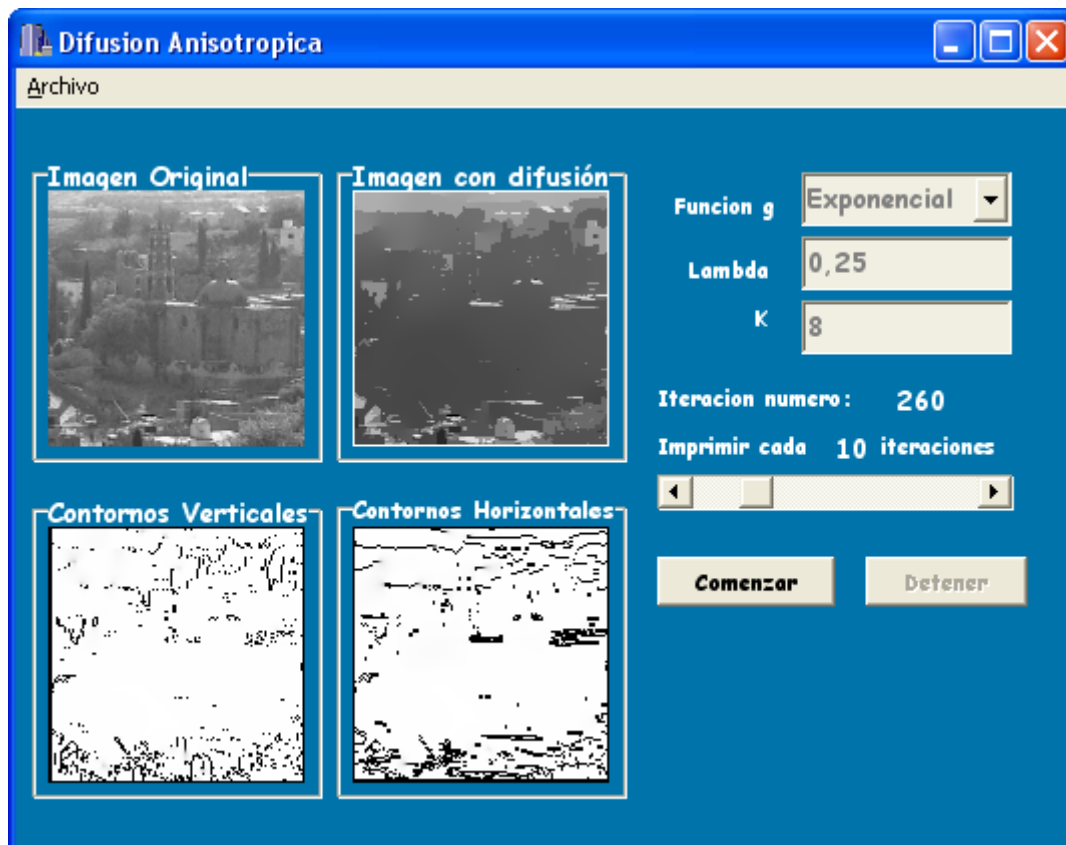


Figura 3.4. Ejemplo de difusión anisotrópica y detección de bordes usando espacios de escala con iglesia.bmp

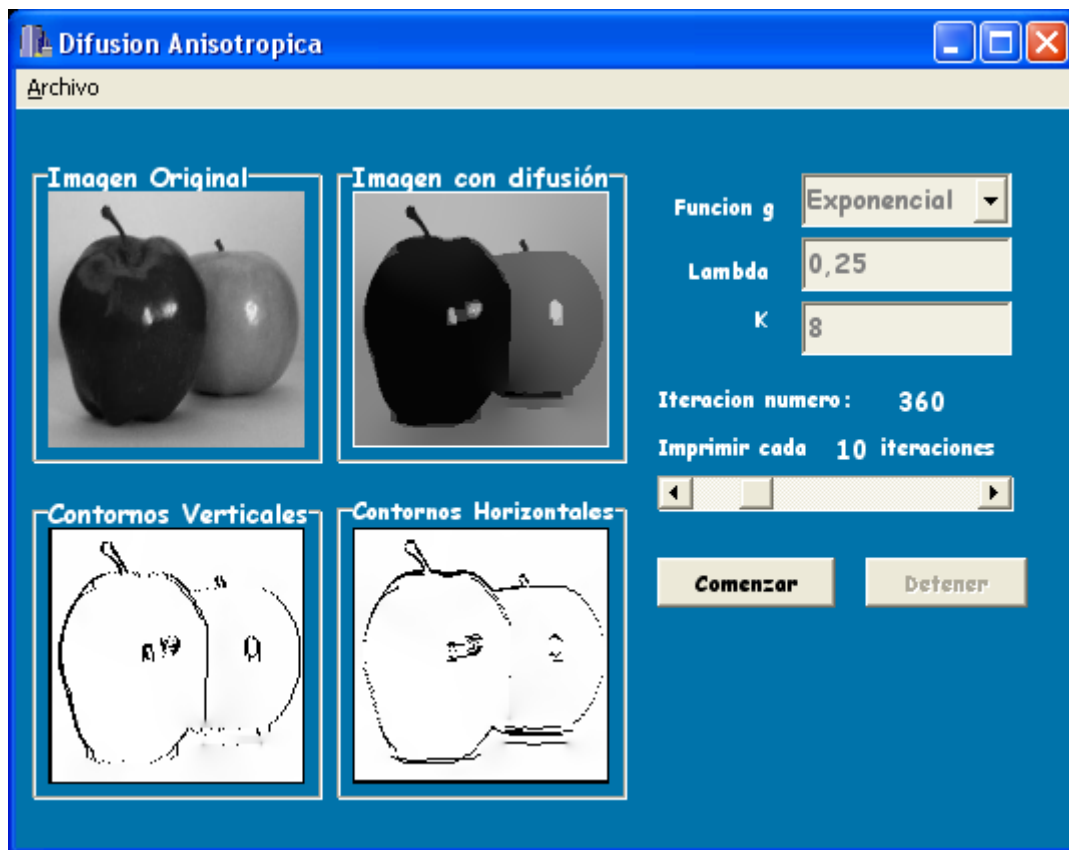


Figura 3.5. Ejemplo de difusión anisotrópica y detección de bordes usando espacios de escala con manzana.bmp

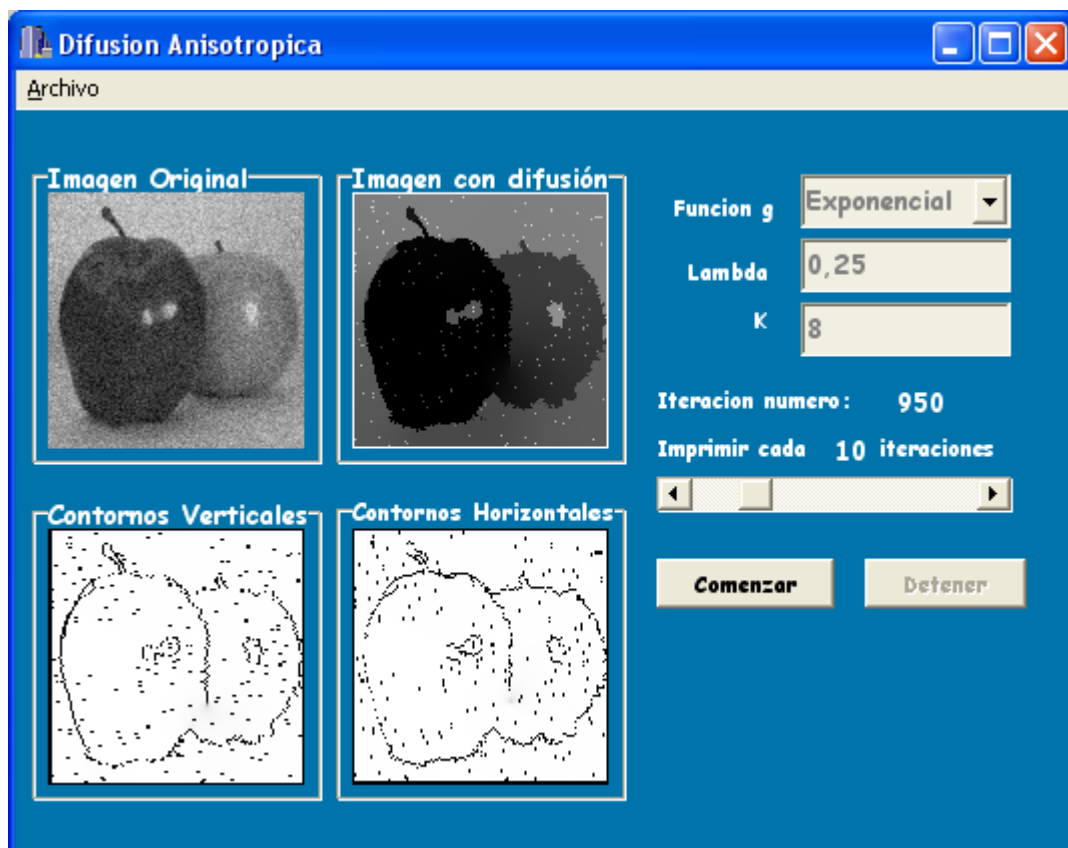


Figura 3.5. Ejemplo de difusión anisotrópica y detección de bordes usando espacios de escala con manzana.bmp con ruido.

Capítulo 4 Aplicaciones de Espacios de Escala

Existen muchas aplicaciones del concepto de representaciones multi-escala. La representación de espacios de escala tiene como aspecto atractivo sus bases matemáticas, lo cual ha motivado a enfatizar el desarrollo de sus fundamentos teóricos.

Actualmente, la teoría de espacios de escala tiene aplicaciones en áreas como:

- Segmentación
- Emborronamiento
- Suavizamiento preservando bordes
- Detección de rasgos

entre otras. En las siguientes secciones se hablará brevemente de algunas de estas aplicaciones.

4.1. Segmentación de imágenes

Oscar Escoda, et. Al. [8], presenta un trabajo de segmentación con dos diferentes algoritmos, uno que emplea espacios de escala lineales y el otro no lineales. La idea del algoritmo es básicamente ligar las imágenes de cada nivel en una especie de estructuras de árboles.

En el caso lineal, se obtuvieron resultados que mejoraban cuando se realizaba una supervisión en detección de bordes, ya que en este caso, no se toman en cuenta los bordes a la hora de suavizar. En el caso no lineal, los resultados fueron mejores, ya que el hecho de preservar los bordes les trajo varias ventajas.

4.2. Suavizamiento preservando bordes

En la sección 3.2. expliqué la difusión anisotrópica propuesta por Perona y Malik [4]. En su trabajo, ellos presentan un algoritmo para suavizar imágenes preservando bordes.

Esto tiene aplicaciones por ejemplo si se desea eliminar ruido, o simplemente como pre-procesamiento para otros algoritmos de procesamiento de imágenes.

El método además de suavizar preservando bordes, los detecta en forma simultánea, partiendo de una estimación de los bordes (en este caso la magnitud del gradiente), y mejorando esta estimación conforme se incrementa la escala.

4.3. Detección de rasgos

Existe un gran número de aplicaciones para detección de rasgos publicadas. Como un ejemplo, se puede mencionar que la teoría de espacios de escala se ha empleado en la detección de esquinas en imágenes.

Para la detección de esquinas den imágenes 2D, se consideran aquellos puntos donde la curvatura de isofotos (curvas de igual brillo) es relativamente grande. Para la medición de tal curvatura, se usan representaciones de espacios de escala.

Conclusiones

La teoría de espacios de escala es muy extensa, y en lo particular me pareció interesante, más que nada por toda la teoría matemática que hay detrás. Las siguientes son algunas conclusiones generales que obtuve en la realización de esta investigación.

-La representación de espacios de escala es una familia uni-paramétrica de señales donde la información de escalas finas se va eliminando conforme se incrementa la escala.

-Existen dos tipos de difusión, o representaciones de espacios de escala, la lineal y la no lineal.

-La difusión lineal consiste en convolucionar una señal con kernels Gaussianos a diferentes escalas.

-La difusión no lineal tiene las propiedades de representación de espacios de escala, y conserva rasgos importantes en la imagen al incrementar la escala.

-Existen muchas aplicaciones para los espacios de escala, y dependiendo de la finalidad de la aplicación, se usa lineal, o no lineal.

Referencias

- [1] T. Lindeberg, *Scale-Space Theory in Computer Vision*, Kluwer Academic Publishers, Netherlands, 1994.
- [2] J. Koenderink, et.al. (eds), *Scale-Space in Computer Vision*, Springer-Verlag, Germany, 1997.
- [3] J. Sporring, et.al. (eds), *Gaussian Scale-Space Theory*, Kluwer Academic Publishers, Netherlands, 1997
- [4] P. Perona and J. Malik, *Scale-Space and edge detection using anisotropic diffusion*, IEEE Trans, Pattern Analysis Machine Intelligence, vol. 12, pp. 629-639, 1990
- [5] P. Jackway, *Scale-Space Properties of the Multiscale Morphological Dilation-Erosion*
- [6] E. Trucco and A. Verri, *Introductory Techniques for 3-D Computer Vision*, Prentice Hall, USA 1998, pp. 71-85
- [7] R. Jain, et.al., *Machine Vision*, Mc. Graw Hill, USA, 1995 pp. 96-97, 104, 198.
- [8] O. Divorra, et.al. *Segmentation of natural images using scale-Space representations : a linear and a non-linear approach* , In Proceedings of EUSIPCO, Toulouse, France, September 2002.