INTRODUCCIÓN A LOS FILTROS ANISOTRÓPICOS POR DIFUSIÓN

Los filtros anisotrópicos de difusión fueron introducidos por primera vez por Pietro Perona y Jitendra Malik en su ya histórica publicación [1].

DEFINICIÓN

Perona y Malik formularon el filtro anisotrópico de difusión como un proceso de difusión que favorece el suavizado (*smoothing*) intrarregión frente al suavizado interregión. Analicemos primeramente las palabras que componen nuestra materia de estudio.

- ***Anisotrópico***: significa etimológicamente “desigualdad de movimiento”. Nuestro filtro es un filtro dinámico, se va a comportar de manera diferente dependiendo del gradiente local. El gradiente va a controlar la anisotropía.
- ***Difusión***: porque matemáticamente se basa en la ecuación de difusión del calor térmico.

Previo a un análisis matemático del filtro y al estudio de su implementación, procedemos a adelantar ya las ventajas principales de este tipo de filtros.

VENTAJAS

- Fácil extensión a 3-D
- Validez para todo tipo de imágenes médicas (gran versatilidad)
- Resuelve las desventajas principales de los filtros convencionales:
    - Difuminado (*blurring*) de las regiones frontera
    - Supresión de detalles significativos
- Sencillez algorítmica
- Su naturaleza intrínsecamente adaptativa, al tener en cuenta los estadísticos y las estructuras locales de las imágenes
- A lo largo del proceso (que es iterativo) no se generan puntos espúreos en las discontinuidades de las imágenes
- La localización espacial de los bordes son estacionarios en el tiempo

OBJETIVOS

Los objetivos de los filtros anisotrópicos de difusión son, fundamentalmente:

- Mejorar los bordes de la imagen
- Favorecer el *smoothing* intrarregión frente al interregión
- Minimizar la pérdida de información, preservando discontinuidades y estructuras detalladas, sin permitir la generación de elementos espúreos iteración tras iteración (esto es, que los pasos iterativos sean *causales*)
- Eliminar ruido
- Mejorar la definición morfológica

Básicamente, podemos condensar estos objetivos en dos: ***eficiencia*** y ***fidelidad***.

APLICACIONES

Las aplicaciones de los filtros anisotrópicos de difusión son muchas. En el campo de las imágenes médicas, conviene resaltar:

- Filtran el ruido de una manera muy eficiente
- El filtrado anisotrópico de difusión, previo a un proceso de segmentación, mejora notablemente la eficiencia de este proceso. Cuando la segmentación es semiautomática, por ejemplo, minimiza drásticamente la interacción humana
- Hace innecesarias muchas de las operaciones morfológicas en imágenes: *thinning*, *linking*...
- En imágenes de resonancia magnética, por ejemplo, las adquisiciones de una sola exposición (1 NEX) con nuestro filtrado consiguen mejores resultados que las de exposiciones múltiples (4 NEX,...). Esta ganancia de tiempo conlleva a un tráfico (throughput) de pacientes mucho mayor, evitándoles molestias inherentes a la adquisición de este tipo de imágenes
- La naturaleza intrínseca del proceso de difusión permite *segmentar* imágenes de una manera muy natural.

INTRODUCCIÓN A LA DIFUSIÓN ANISOTRÓPICA

Los sistemas de filtrado convencionales muchas veces no son adecuados para las situaciones no estacionarias a las que están sujetas las imágenes médicas. La rapidez con que aparecen nuevos métodos de adquisición de imágenes médicas (en todas sus modalidades), da un impulso sin igual al desarrollo de nuevos métodos de visualización tridimensional (3-D) en aplicaciones de Medicina.

Los filtros adaptativos, en general, presentan unas ventajas computacionales muy grandes (su implementación matemática discreta les confiere esta ventajosa propiedad).

Los filtros anisotrópicos están basados en la medida de la anisotropía del entorno del elemento considerado de la imagen (píxel).

EVOLUCIÓN HISTÓRICA

Históricamente, se venían empleando técnicas convencionales de filtrado en aplicaciones médicas (filtrados lineales). Los primeros filtros no lineales a emplear fueron los filtros de mediana. Más tarde surgieron los filtros de difusión isotrópicos, en los que el coeficiente de difusión (esto es, el grado de difusión alcanzado en cada región) era constante:

$$c(x, y) = k$$

Entre ellos, los de mayor aceptación fueron los gaussianos. Más adelante, Perona y Malik propusieron su innovador filtro, que ha conocido en pocos años numerosas variaciones. La evolución a grandes rasgos se recoge en los siguientes puntos:

- Los filtros lineales pueden llegar a ser bastante eficaces en la eliminación del ruido, pero degradan detalles significativos (bordes, ángulos cerrados, líneas...)

- Los filtros no lineales consiguen mejorar la preservación de los bordes, no así los detalles estructurales finos.

- Los filtros adaptativos buscan un compromiso entre tres criterios:

  - Eficiencia del suavizado
  - Conservación de discontinuidades

- Generación de artefactos inexistentes

Esto es, pretenden un suavizado óptimo, sin descuidar la conservación de bordes (discontinuidades), y al mismo tiempo preservando la causalidad en las diferentes iteraciones.

BASES DE LA DIFUSIÓN ANISOTRÓPICA

La difusión anisotrópica pertenece al conjunto de métodos "verticales", llamados así por considerar sólo la dimensión vertical (nivel de intensidad) de las imágenes. El tamaño de las estructuras compete a una aproximación "plana" (por contraposición a "vertical") en la que el procesado de las imágenes opera sin tener en cuenta el nivel de intensidad de cada píxel. A este conjunto de técnicas pertenecen, por ejemplo, las operaciones morfológicas de imágenes (erosión, dilación...)

Perona y Malik introdujeron por primera vez la difusión anisotrópica como un proceso que suaviza los píxeles de la imagen dentro de las diferentes regiones (esto es, un suavizado *intrarregión*), pero que paraliza el suavizado a través de los bordes significativos.

Para iniciarnos en el estudio profundo de la difusión anisotrópica, tomemos un ejemplo sencillo, unidimensional. Modelemos un borde como una función escalón, convolucionada con una gaussiana.

En las ecuaciones que manejaremos, vamos a asumir que el coeficiente "c", medida del grado de difusión, no va a ser constante. Veremos cómo una elección apropiada de c(x,y,t) nos permitirá conseguir un sistema de filtrado deseado.

La difusión anisotrópica busca satisfacer tres criterios (enunciados por Witkin y Koenderink):

1. **Causalidad**: la difusión, como proceso, no debe generar detalles espúreos cuando pasamos de una resolución detallada a escalas más toscas.

2. **Localización Inmediata**: a cada grado de resolución deseado, los bordes deberán mantenerse bien visibles, y coincidir con las fronteras semánticamente significativas de esa resolución.

3. **Suavizado dinámico**: a toda escala, el suavizado intrarregión debe prevalecer sobre el suavizado interregión.

Para satisfacer los tres requisitos que exigimos a nuestro filtro, podemos definir un estimador vectorial **E**(x,y,t), que debe cumplir:

- **E**(x,y,t)=0 en el interior de cada región homogénea
- **E**(x,y,t)=*M***e**(x,y,t) en los puntos situados en los bordes, donde **e** es un vector normal al borde en cada punto, y *M* es el contraste local (diferencia de las intensidades de las imágenes a ambos lados del borde)

Si efectivamente existe tal estimador, **E**, el coeficiente de difusión c(x,y,t) puede ser una función de la magnitud de **E**:

$$c = g(\|E\|)$$

Razonando, es intuitivo asignarle una monotonicidad decreciente a la función g( ). Así, para valores bajos de **E** hay difusión, al ser la región homogénea, y a medida que avanzamos en el eje **E**, la difusión va disminuyendo. Efectivamente, g( ) debe ser una función monótona decreciente para que pueda difundir en regiones homogéneas, parando la difusión en regiones que reconoce como bordes.

Por tanto, tenemos que

$$g(0) = 1$$
$$g(\infty) = 0$$

De esta manera, el proceso de difusión tendrá lugar principalmente en el interior de las regiones, pero no afectará las fronteras entre regiones, donde la magnitud de **E** es muy grande.

Podemos atisbar ya que el éxito de nuestro proceso de difusión va a depender sobre todo de la calidad de **E** como buen estimador de la posición de los bordes.

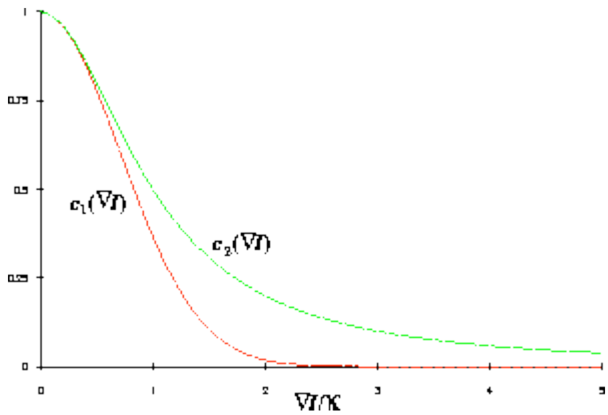
Afortunadamente, un excelente estimador (y el más sencillo de los que podamos imaginar) de la posición de los bordes, es el **gradiente** de la función imagen. Veremos cómo la función gradiente da unos excelentes resultados:

$$E(x,y,t) = \nabla I(x,y,t)$$

Con lo que, finalmente, llegamos a:

$$c(x,y,t) = g\left(\left\|\nabla I(x,y,t)\right\|\right)$$

Existe una infinidad de funciones g( ) que cumplan los requisitos fijados hasta ahora (que sea monótona decreciente en todo su dominio, y cuyo recorrido vaya de uno, en el origen, a cero, en el infinito). Se han adoptado dos funciones (empíricamente) en los últimos años que tienen las mejores prestaciones. Estas dos funciones tienen resultados cuya percepción es similar.



Sus expresiones matemáticas son:

$$c_1(\bar{x},t) = \exp\left(-\left(\frac{\left|\nabla I(\bar{x},t)\right|}{K}\right)^2\right)$$

$$c_2(\bar{x},t) = \frac{1}{1+\left(\frac{\left|\nabla I(\bar{x},t)\right|}{K}\right)^{1+\alpha}}, \alpha > 0$$

En ellas, K es el parámetro espacio-escala: K determina qué estructuras van a preservarse tras el filtrado. Dependiendo de la intensidad del gradiente, nuestro algoritmo puede preservar o difuminar los bordes.

Además del parámetro K, el número de iteraciones es un factor muy importante a tener en cuenta. El proceso de difusión es un proceso iterativo, con una convergencia asintótica que habremos de garantizar. El valor de α suele ser la *unidad*.

Consideremos la siguiente ecuación de difusión anisotrópica, aportación del campo de la Termodinámica:

$$I_t = div(c(x,y,t)\nabla I) = c(x,y,t)\Delta I + \nabla c \cdot \nabla I$$

Donde

- *div* representa el operador divergencia
- ∇ representa el operador gradiente
- Δ representa el operador laplaciano

todos ellos con respecto a las variables espaciales.

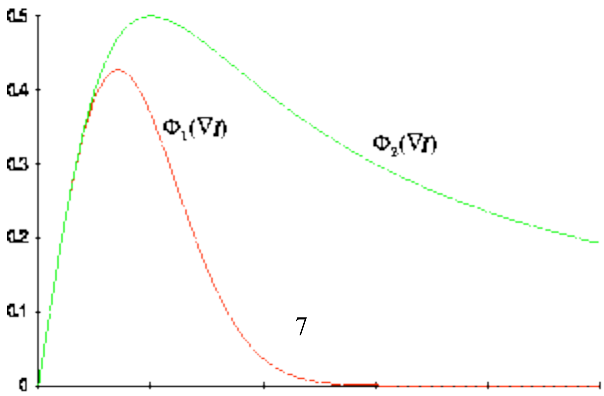
I es la función de intensidad de la imagen (I_t, su derivada respecto de la variable temporal, *t*). Si c(*x,y,t*) es constante, se reduce a

$$I_t = c\Delta I$$

Definimos la función flujo ϕ, como:

$$\phi(I_x) = g(I_x) \cdot I_x = c \cdot I_x$$

Y que tendrá el siguiente aspecto (para las funciones c que hemos establecido):



Dada una función de flujo determinada, podemos un encontrar un valor de K, por encima del cual la función es monótona decreciente, y por debajo del cual la función es monótona creciente. De esta manera conseguimos perfilar aún más los bordes de interés y suavizar en regiones homogéneas sin elementos representativos.

Vemos que el flujo máximo se genera en las regiones donde el gradiente ∇I es igual a K. Para regiones donde los valores del gradiente son inferiores a K, el flujo se reduce a cero, como debe ser en este tipo de regiones homogéneas (donde el gradiente es muy pequeño). Pero también para las regiones donde los valores del gradiente son superiores a K, el flujo vuelve a reducirse a cero, como era de esperar: paramos el proceso de difusión en regiones de gradientes altos.

Como ya adelantamos, elegimos "c" como función del gradiente:

$$c(x,y,t) = g(I_x(x,y,t))$$

Nuestra expresión se simplifica al trabajar en una sola dimensión: el operador *divergencia* se reduce a una derivada (asumimos, sin ninguna pérdida de generalidad, que operamos sobre el eje *x*):

$$div(c(x,y,t\nabla I)) = \frac{\partial}{\partial x}(c(x,y,t)I_x)$$

Nuestra primera versión unidimensional de la ecuación de difusión tiene este aspecto:

$$I_t = \frac{\partial}{\partial x}\phi(I_x) = \phi'(I_x)\cdot I_{xx}$$

Conviene reseñar que las derivadas primera, segunda, y tercera (con respecto a la variable espacial) las representaremos así, respectivamente:

$$I_x, I_{xx}, I_{xxx}$$

Estamos interesados en analizar la variación del gradiente ("la pendiente del borde") con respecto al tiempo, esto es:

$$\frac{\partial}{\partial t}(I_x) = \frac{\partial}{\partial x}(I_t) = \frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}\phi(I_x)\right) = \phi''\cdot I_{xx}^2 + \phi'\cdot I_{xxx}$$

O lo que es la misma expresión, menos compacta:

$$\frac{\partial}{\partial t}\frac{\partial I}{\partial x} = \frac{\partial^2 c}{\partial x^2}\frac{\partial I}{\partial x} + 2\frac{\partial c}{\partial x}\frac{\partial^2 I}{\partial x^2} + c\frac{\partial^3 I}{\partial x^3}$$

Donde hemos intercambiado el orden de diferenciación por ser la función de difusión c(·) positiva y continua en todas sus variables.

Supongamos que el borde está orientado de tal manera que $I_x>0$. Entonces, en el punto de inflexión, $I_{xx}=0$, e $I_{xxx}<<0$.

De acuerdo a la expresión [], en la proximidad del punto de inflexión, la variación con el tiempo de I_x tiene signo contrario al de $\phi'(I_x)$:

- Si $\phi'(I_x)>0$, entonces la pendiente del borde *disminuirá* con el tiempo
- Si $\phi'(I_x)<0$, entonces la pendiente del borde *se incrementará* con el tiempo

Hemos conseguido una mejora selectiva del borde, que se opone, por ejemplo, al efecto de emborronado que sufren los bordes al ser sometidos a filtrados lineales.

Desde los orígenes de la visión artificial, la importancia de disponer de ***descripciones multiescala*** de las imágenes ha ido creciendo. Actualmente, el campo de las *wavelets* (u ondículas) es puntero en este tipo de descripción de imágenes.

El enlace entre los procesos de difusión y el análisis de imágenes comenzó con las descripciones multiescala de las imágenes. El objetivo del análisis multiescala es la computación de una familia de descripciones que dependen de un parámetro, llamado *parámetro espacio-escala*.

Inicialmente se utilizaron funciones gaussianas. Se trataba de combinar (convolucionar) la imagen original con un *kernel* (núcleo) gaussiano G(*x,y;t*) de varianza *t*, para formar una familia de imágenes derivadas (*I,x,t*):

$$I(x,y,t) = I_0(x,y) * G(x,y;t)$$

Es fácil entender que el parámetro espacio-escala en cada resolución, es precisamente la varianza, *t*. Valores altos de *t* corresponden a resoluciones más altas (peores). La variable *t* es el parámetro de ordenación del proceso. En nuestra implementación discreta se utilizará para enumerar los pasos iterativos.
En las difusiones isotrópicas, otro parámetro espacio-escala es el número de iteraciones.

(Hemos visto, en cambio, que en las difusiones anisotrópicas, el número de iteraciones es un factor, y no el parámetro espacio-escala).

PROPIEDADES DE LA DIFUSIÓN ANISOTRÓPICA

La propiedad básica de la difusión anisotrópica es su capacidad para no sólo preservar, sino incluso mejorar (hacer más intensos) los bordes, siempre que la función

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

esté elegida adecuadamente. Es decir, podemos enunciar este principio:

Un proceso de difusión en el que el coeficiente de conducción (o conductividad) se elige localmente, como función de la magnitud del gradiente de la función intensidad de la imagen, es capaz de mejorar el contraste de los bordes.

Discutamos dos de las propiedades de este tipo de difusión:

A. EL PRINCIPIO MÁXIMO

El criterio de causalidad (ya comentado), exige que no aparezcan nuevos elementos en las imágenes cuando pasamos de representaciones de resolución fina a otras de peor resolución.

Bastará probar que todos los máximos y mínimos en toda la batería de imágenes de diferentes resoluciones, pertenecen a la imagen original (y por tanto, la no aparición de estructuras espúreas).

Las ecuaciones de difusión presentadas hasta ahora constituyen un caso especial de una clase de funciones matemáticas más general, llamadas *ecuaciones elípticas*, que satisfacen todas ellas el principio máximo. Este principio se enuncia así:

Si el coeficiente de difusión es positivo, se garantiza que todos los máximos de la solución de la ecuación (en el espacio y en el tiempo) pertenecen a la condición inicial (imagen original).

B. MEJORA DE LOS BORDES DE LA IMAGEN

Ya hemos comentado la limitación de los procesos de filtrado convencionales (filtrados paso bajo, difusión lineal). El precio que pagamos al eliminar el ruido con estos filtros es el emborronado de los bordes. Esto hace que la localización de éstos sea una tarea difícil y costosa.

Podemos lograr la mejora de los bordes y la reconstrucción por tanto de imágenes borrosas, mediante técnicas de filtrado paso alto, o ejecutando la ecuación de difusión reversiblemente (“hacia atrás”). Estas técnicas dan pie a métodos computacional y numéricamente inestables.

Veremos cómo, tal y como hemos elegido las funciones de difusión, podemos hacer que la difusión anisotrópica mejore los bordes, ejecutando la ecuación “hacia delante” en el tiempo, garantizando la estabilidad del proceso.