

# Introducción a la difusión en el tratamiento de imágenes

Guillermo Gallego

La idea que motiva el estudio de la difusión anisótropa en el tratamiento de imágenes es la búsqueda de métodos de suavizamiento de imágenes (“filtros”) que atenúen el ruido a la vez que respeten la información de bordes (“señal”) de la imagen.

En adelante, trataremos de pensar en una imagen (en escala de gris) como una función  $I : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ , habitualmente continua. Trataremos de evitar concebir una imagen (digital) como un conjunto de píxeles (muestras de la anterior función en una rejilla/retículo rectangular). Sólo discretizaremos cuando necesitemos implementar alguna ecuación de evolución en un ordenador. En tal caso, lo más habitual es utilizar muestras y diferencias finitas para aproximar derivadas de la imagen.

## I. DIFUSIÓN ISÓTROPICA

Supongamos una imagen dada  $I_0(x, y)$ . Dado que buscamos suavizar la imagen, un posible punto de partida es utilizar nuestros conocimientos sobre procesos de difusión en Física. Por ejemplo, podemos simular el proceso de suavizamiento dictado por la ecuación del calor al hacer evolucionar la temperatura de un cuerpo sólido. Planteamos el siguiente problema con condición de contorno inicial: encontrar la función  $I(x, y, t)$ ,  $t > 0$  tal que resuelve

$$\begin{cases} I_t = \Delta I \\ I(x, y, 0) = I_0(x, y), \end{cases} \quad (1)$$

donde  $\Delta I = I_{xx} + I_{yy}$  es el Laplaciano de  $I$  y los subíndices denotan derivadas parciales:  $I_t = \partial_t I = \frac{\partial I}{\partial t}$ ,  $I_{xy} = \frac{\partial^2 I}{\partial x \partial y}$ , etc. La ecuación (1) modela un proceso de difusión lineal isótropo.

### I-A. Dedución de la ecuación de difusión

Los procesos de difusión están estrechamente relacionados con las leyes de conservación. Sea  $I$  la concentración de una sustancia o temperatura. Si en una región hay diferencias de concentración (o temperatura), se producirá un flujo de sustancia desde las regiones de mayor concentración hacia las regiones de menor concentración, hasta alcanzar un estado de equilibrio. Si se cumple la ley de conservación de la sustancia  $I$  y en cierto dominio (volumen)  $\Omega$  se produce una variación de la cantidad de sustancia  $I$ , esta se deberá a la presencia de un flujo de sustancia  $\mathbf{j}$  a través de dicho dominio:

$$\underbrace{\frac{\partial}{\partial t} \int_{\Omega} I dV}_{\text{Variación de } I \text{ en } \Omega} + \underbrace{\int_{\partial\Omega} \mathbf{j} \cdot \mathbf{n} dS}_{\text{Flujo a través de } \Omega} = 0,$$

siendo  $\mathbf{n}$  el vector normal a  $\partial\Omega$ . Si  $\Omega$  es fijo (respecto al tiempo  $t$ ) y se utiliza el teorema de la divergencia,

$$\int_{\Omega} (I_t + \nabla \cdot \mathbf{j}) dV = 0,$$

por lo que, para regiones arbitrariamente pequeñas, se deduce la ecuación de continuidad<sup>1</sup>

$$I_t + \nabla \cdot \mathbf{j} = 0. \quad (2)$$

Si el proceso de igualación de la concentración sigue la (primera) Ley de Fick (1855), el flujo es proporcional (y en sentido opuesto, como ya sabíamos) al gradiente de la concentración,

$$\mathbf{j} = -c\nabla I, \quad (3)$$

siendo  $c$  el coeficiente de difusión (también llamado difusividad o conductividad). Al sustituir (3) en (2) se obtiene la ecuación de difusión o segunda ley de Fick,

$$I_t = \nabla \cdot (c\nabla I). \quad (4)$$

Cuanto más grande es  $c$ , mayor es el flujo de sustancia, y por tanto, más rápido es el proceso de difusión. Finalmente, si  $c = 1$  se obtiene (1), ya que  $\Delta I := \nabla \cdot (\nabla I)$ , por definición.

### I-B. Imágenes como mapas de temperatura o de concentración de una sustancia

La ecuación del calor en el problema (1) lleva a interpretar los niveles de gris de una imagen como si fuera la concentración de una sustancia o temperatura. Entonces, los niveles de gris de la imagen sufren un proceso de difusión igual al que dicta la ecuación del calor para evolucionar la temperatura de un sólido a medida que avanza el tiempo. En realidad, ya no tenemos una imagen  $I_0(x, y)$ , sino una familia de imágenes  $I(x, y, t)$ , una imagen para cada instante de tiempo.

En la práctica, las imágenes tienen una extensión finita. Así que es habitual suponer algún tipo de condición de contorno en la frontera del dominio,  $\partial\Omega$ , para saber cómo evolucionar los niveles de gris en aquellos puntos. Lo habitual es suponer que el flujo es nulo ( $\mathbf{j} = 0$ ) en  $\partial\Omega$ , lo que equivale a condiciones de contorno homogéneas de tipo Neumann puesto que  $\frac{\partial I}{\partial \mathbf{n}} = \nabla I \cdot \mathbf{n} = 0$ , siendo  $\nabla I = (I_x, I_y)^\top$  para  $\Omega \subset \mathbb{R}^2$ . Esto indica una estabilización local de la intensidad de la imagen en la dirección perpendicular a  $\partial\Omega$ ; la intensidad puede variar a lo largo de  $\partial\Omega$ , pero permanece constante al atravesar

<sup>1</sup>La ecuación (2) no debería resultar desconocida, puesto que utilizando un símil electromagnético, si  $I \equiv \rho$  es la densidad de carga eléctrica  $q$  (la “sustancia”) y  $\mathbf{j} \equiv \mathbf{J}$  es la densidad de corriente, (2) es la ecuación de continuidad de la carga que se obtiene de la ley de conservación de la carga.

$\partial\Omega$  perpendicularmente. La condición de flujo nulo a través del dominio implica un aislamiento del mismo respecto del resto de  $\mathbb{R}^2$ : no entra/sale sustancia. En el símil calorífico, el dominio  $\Omega$  está aislado térmicamente del resto de  $\mathbb{R}^2$ .

### I-C. Solución de la ecuación del calor mediante convolución. Linealidad e isotropía

La solución al problema (1) se puede obtener sin necesidad de resolver iterativamente una ecuación en derivadas parciales (EDP), si se conoce la solución fundamental de la ecuación (función de Green). Resulta interesante saber que  $I$  se obtiene a partir de  $I_0$  mediante una convolución con un filtro (paso bajo) Gaussiano cuya desviación típica depende de  $t$ . A medida que avanza el tiempo  $t$ , la difusión es mayor, así que el correspondiente filtro Gaussiano tiene una desviación típica mayor. Por ejemplo, en el caso unidimensional, la solución al problema

$$\begin{cases} I_t = I_{xx} & -\infty < x < \infty, 0 < t < \infty \\ I(x, 0) = I_0(x) & \text{Condición inicial} \end{cases} \quad (5)$$

es la convolución

$$\begin{aligned} I(x, t) &= I_0(x) * \frac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{4t}} \\ &= \frac{1}{\sqrt{4\pi t}} \int_{-\infty}^{\infty} \exp\left(-\frac{(x-\xi)^2}{4t}\right) I_0(\xi) d\xi. \end{aligned} \quad (6)$$

Esto es fácil de comprobar. Tomando Transformadas de Fourier en (5) (respecto de la variable espacial  $x$ ), siendo  $I(x) \xleftrightarrow{\mathcal{F}} \hat{I}(\omega)$ , obtenemos

$$\begin{cases} \hat{I}_t(\omega, t) = (j\omega)^2 \hat{I}(\omega, t) & -\infty < \omega < \infty, 0 < t < \infty \\ \hat{I}(\omega, 0) = \mathcal{F}\{I_0(x)\}, & \text{Condición inicial} \end{cases} \quad (7)$$

que es una ecuación diferencial ordinaria (EDO) de primer orden en la variable  $t$ , del tipo  $y'(t) + \alpha y(t) = 0$ , con condición inicial  $y(0) = A$ , cuya solución es  $y(t) = Ae^{-\alpha t}$  para  $t \geq 0$ . Entonces, la solución de (7) es

$$\hat{I}(\omega, t) = \mathcal{F}\{I_0(x)\} e^{-\omega^2 t},$$

es decir, el producto de dos funciones (en  $\omega$ ), cuya transformada inversa es precisamente la convolución (6). La ecuación del calor es lineal, al igual que la convolución. Los procesos de difusión anisótropa son no lineales.

Una propiedad importante del operador Laplaciano es que es isótropo (o invariante por rotaciones). El valor  $\Delta I = I_{xx} + I_{yy}$  es el mismo que el obtenido tras una rotación del plano en la que los nuevos ejes estén alineados con los vectores de una base ortonormal  $\{\xi, \eta\}$ ,  $\Delta I = I_{\xi\xi} + I_{\eta\eta}$ . Esto es fácil de comprobar ya que el Laplaciano es la traza de la matriz Hessiana (matriz de segundas derivadas parciales  $(\nabla^2 I)_{ij} = I_{x_i x_j}$ ),

$$\Delta I = \text{trace}(\nabla^2 I), \quad (8)$$

y la traza es invariante por rotaciones puesto que para una matriz cuadrada  $A$  y una base ortonormal cualquiera  $\{\mathbf{u}_i\} \in \mathbb{R}^n$ , se cumple

$$\text{trace}(A) = \sum_{i=1}^n \mathbf{u}_i^\top A \mathbf{u}_i. \quad (9)$$

Esta propiedad hace que la difusión mediante la ecuación del calor no tenga direcciones preferentes, sino que sea igualmente importante en todas las direcciones del plano. Esto se aprecia claramente en la figura 1.

## II. DIFUSIÓN ANISÓTROPA

En un proceso de difusión anisótropa hay direcciones preferentes y es deseable que dichas direcciones coincidan con las de los bordes de los objetos para así poder suavizar la imagen a lo largo de los bordes, evitando suavizar en la dirección perpendicular a los mismos (para no producir un esparcimiento espacial del borde, lo cual lo destruiría). De esta forma se preservan los bordes mientras se suaviza / reduce el ruido en el resto de zonas de la imagen. La figura 2 muestra un ejemplo de difusión anisótropa.

Una posible forma de introducir anisotropía en el problema (1) es utilizando la segunda ley de Fick (4) con una difusividad anisótropa  $c(x, y, t)$  que controle la tasa de difusión en cada punto del dominio  $\Omega$ . El resultado es la familia de imágenes que resuelven el problema

$$\begin{cases} I_t = \text{div}(c(x, y, t)\text{grad}(I)) \\ I(x, y, 0) = I_0(x, y), \end{cases} \quad (10)$$

donde  $\text{div}(c(x, y, t)\text{grad}(I)) = c\Delta I + \nabla c \cdot \nabla I$ . Normalmente, se diseña  $c$  como una función del gradiente de la imagen para preservar los bordes de los objetos. En el conocido modelo de Perona y Malik (1990) se proponen dos posibles valores del coeficiente de difusividad,

$$c(\|\nabla I\|) = e^{-(\|\nabla I\|/K)^2} \quad \text{y} \quad c(\|\nabla I\|) = \frac{1}{1 + \left(\frac{\|\nabla I\|}{K}\right)^2},$$

donde la constante  $K$  controla la sensibilidad a los bordes de los objetos y se suele determinar experimentalmente o en función de la cantidad de ruido en la imagen. Ambas funciones son decrecientes, de tal forma que la difusión disminuya a medida que  $\|\nabla I\|$  aumente, es decir, a medida que se detecten bordes de objetos en la imagen.

Un modelo más general sustituye  $c$  en (3) por un tensor de difusión, esto es, una matriz de  $2 \times 2$ ,  $\mathbf{C} = (c_{ij})$ ,

$$\begin{cases} I_t = \text{div}(\mathbf{C}(x, y, t)\text{grad}(I)) \\ I(x, y, 0) = I_0(x, y). \end{cases}$$

Para que el proceso de difusión esté bien definido y realice su función de suavizamiento, la matriz simétrica  $\mathbf{C}$  debe ser definida positiva.

### II-A. Ecuación geométrica del calor

Otra ecuación de difusión anisótropa bien conocida en el tratamiento de imágenes es la ecuación geométrica del calor (geometric heat equation - GHE),

$$\begin{cases} I_t = \frac{I_y^2 I_{xx} - 2I_x I_y I_{xy} + I_x^2 I_{yy}}{I_x^2 + I_y^2} \\ I(x, y, 0) = I_0(x, y), \end{cases} \quad (11)$$

la cual trataremos en más detalle. Esta ecuación se puede interpretar como un proceso de difusión local unidimensional

en la dirección de los bordes de los objetos en la imagen. Si interpretamos la imagen de escala de grises como un mapa de elevación,  $z = I(x, y)$ , por cada punto  $\mathbf{x} = (x, y)^\top$  pasa una curva equinivel (iso-intensidad). Los vectores tangente  $\mathbf{T}(\mathbf{x})$  y normal  $\mathbf{N}(\mathbf{x})$  a dicha curva en  $\mathbf{x}$  establecen una base ortonormal  $\{\mathbf{T}, \mathbf{N}\}$  (la base de Frenet), siendo  $\mathbf{N} = \frac{\nabla I}{\|\nabla I\|}$  y  $\mathbf{T}$  una rotación de 90 grados de  $\mathbf{N}$ . Utilizando (8) y (9), si escribimos el Laplaciano en dicha base,  $\Delta I(\mathbf{x}) = I_{TT} + I_{NN}$ , y sustraemos la difusión en la dirección normal, obtendremos el operador de difusión de la ecuación geométrica del calor,

$$I_{TT} = \Delta I - I_{NN},$$

siendo  $I_{NN} = \mathbf{N}^\top (\nabla^2 I) \mathbf{N}$ , donde

$$\nabla^2 I = \begin{pmatrix} I_{xx} & I_{xy} \\ I_{xy} & I_{yy} \end{pmatrix}$$

es la matriz Hessiana de  $I$ . Suponemos que  $I$  es suficientemente suave como para que se cumpla la igualdad de derivadas cruzadas  $I_{xy} = I_{yx}$  ( $I \in C^2(\Omega)$ , según el teorema de Schwarz) y así  $\nabla^2 I$  sea simétrica. Entonces,

$$\begin{aligned} I_{TT} &= \Delta I - \mathbf{N}^\top (\nabla^2 I) \mathbf{N} \\ &= \Delta I - \frac{1}{\|\nabla I\|^2} (\nabla I)^\top (\nabla^2 I) \nabla I \\ &= \frac{I_y^2 I_{xx} - 2I_x I_y I_{xy} + I_x^2 I_{yy}}{I_x^2 + I_y^2}. \end{aligned}$$

Una forma más compacta de escribir la GHE que además es válida para cualquier dimensión (no sólo para imágenes en  $\mathbb{R}^2$ ) e independiente del sistema de coordenadas utilizado es

$$I_t = \|\nabla I\| \operatorname{div} \left( \frac{\nabla I}{\|\nabla I\|} \right). \quad (12)$$

Comparando esta fórmula con anteriores modelos de difusión anisótropa, vemos que no coinciden, si embargo, (12) se asemeja a (10) si la difusividad térmica es  $c = \|\nabla I\|^{-1}$ , función decreciente en la misma línea que las propuestas por Perona y Malik.

### III. EVOLUCIÓN DE CURVAS PLANAS

Los procesos de difusión de imágenes guardan estrecha relación con los procesos de evolución de curvas planas (contornos activos, etc.). En dicho contexto, si la curva  $\mathbf{C}(p) = (x(p), y(p))^\top$  se representa mediante una ecuación implícita  $\phi(\mathbf{C}) - k_0 = 0$  se dice que la curva  $\mathbf{C}$  es el corte de nivel  $k_0$  (“the  $k_0$  level set”) de la función de nivel  $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ . Es decir,  $\mathbf{C} = \{(x, y) \in \mathbb{R}^2 \text{ tales que } \phi(x, y) - k_0 = 0\} \equiv \phi^{-1}(k_0)$ . El vector normal a la curva es  $\mathbf{N} = -\frac{\nabla \phi}{\|\nabla \phi\|}$  y la curvatura de la curva está dada por  $\kappa = \operatorname{div}(-\mathbf{N}) = \operatorname{div} \left( \frac{\nabla \phi}{\|\nabla \phi\|} \right)$ . Para que la curva evolucione en el tiempo según la ley

$$\mathbf{C}_t = \kappa \mathbf{N}, \quad (13)$$

la función  $\phi$  ha de evolucionar según la ley  $\phi_t = \|\nabla \phi\| \kappa$ . Obsérvese que ésta es la ecuación (12) si la curva  $\mathbf{C}$  es cualquier curva de intensidad constante de la imagen  $I$ . Así pues, la GHE hace evolucionar cada curva de iso-intensidad de la imagen según su curvatura  $\kappa$ , en la dirección normal a

la curva. La ley de evolución (13) es la que más rápidamente decrece la longitud Euclídea de la curva  $\mathbf{C}$ ,

$$\mathcal{L} = \int_C \|\mathbf{C}_p\| dp = \int_{p_0}^{p_1} \sqrt{x_p^2 + y_p^2} dp.$$

Es decir, (13) es la ecuación del método de máximo descenso aplicado al funcional “longitud Euclídea de la curva”. Esta reducción de la longitud se aprecia en la figura 3. Las curvas de iso-intensidad de la imagen se contraen (y suavizan) cada vez más a medida que avanza el tiempo, comenzando por los trozos de curva con mayor curvatura. Se aprecia que curvas cerradas pequeñas (provenientes de parte de las componentes de alta frecuencia espacial de la imagen) pronto desaparecen debido a este proceso de suavizamiento.

### IV. IMPLEMENTACIÓN

Para implementar numéricamente las EDPs anteriores, es típico utilizar diferencias finitas para aproximar las derivadas de la imagen  $I$ . Los operadores de difusión son de tipo elíptico y se suelen aproximar por diferencias centrales (en el espacio). La derivada temporal se suele aproximar por diferencias finitas progresivas, lo cual lleva a una esquema de actualización de tipo explícito, en el que la solución  $I(t + \Delta t)$  se puede resolver fácilmente dada la solución en el instante  $t$ ,  $I(t)$ . Los esquemas explícitos llevan asociada una condición de estabilidad en el paso de actualización para resolver la EDP.

Por ejemplo, la ecuación del calor (1) se puede implementar numéricamente de la siguiente forma:

$$\frac{I_{ij}^{n+1} - I_{ij}^n}{\Delta t} \approx \frac{I_{i+1,j}^n - 2I_{ij}^n + I_{i-1,j}^n}{(\Delta x)^2} + \frac{I_{i,j+1}^n - 2I_{ij}^n + I_{i,j-1}^n}{(\Delta y)^2},$$

donde  $I_{ij}^n = I(i\Delta x, j\Delta y, n\Delta t)$  coincide con el valor de intensidad de la imagen  $I$  en el píxel  $(i, j)$  para el instante  $t$ . Al trabajar con imágenes, se suele tomar  $\Delta x = \Delta y = 1$  píxel, por lo que

$$I_{ij}^{n+1} \approx (1 - 4r)I_{ij}^n + r(I_{i+1,j}^n + I_{i-1,j}^n + I_{i,j+1}^n + I_{i,j-1}^n), \quad (14)$$

siendo  $r = \frac{\Delta t}{(\Delta x)^2} = \Delta t$ . La condición de estabilidad exige que el paso no exceda de  $r < \frac{1}{2}$  en cada iteración. La ecuación (14) se inicializa (en el instante  $n = 0$ ) mediante la imagen  $I_0$  y se evoluciona hasta el instante  $T = N\Delta t$  deseado. Cuanto mayor sea  $T$ , mayor será la difusión efectuada en la imagen.

De forma similar se puede implementar la GHE (11), utilizando diferencias centrales para aproximar las derivadas espaciales y diferencias progresivas para aproximar la derivada temporal. La condición de estabilidad del esquema explícito resultante es  $r = \frac{\Delta t}{(\Delta x)^2} < \frac{1}{2}$ . Se debe evitar la división por cero en aquellas zonas de gradiente nulo de la imagen (zonas homogéneas), en las cuales, por convenio no hace falta suavizar más la imagen ( $I_t = 0$ ).

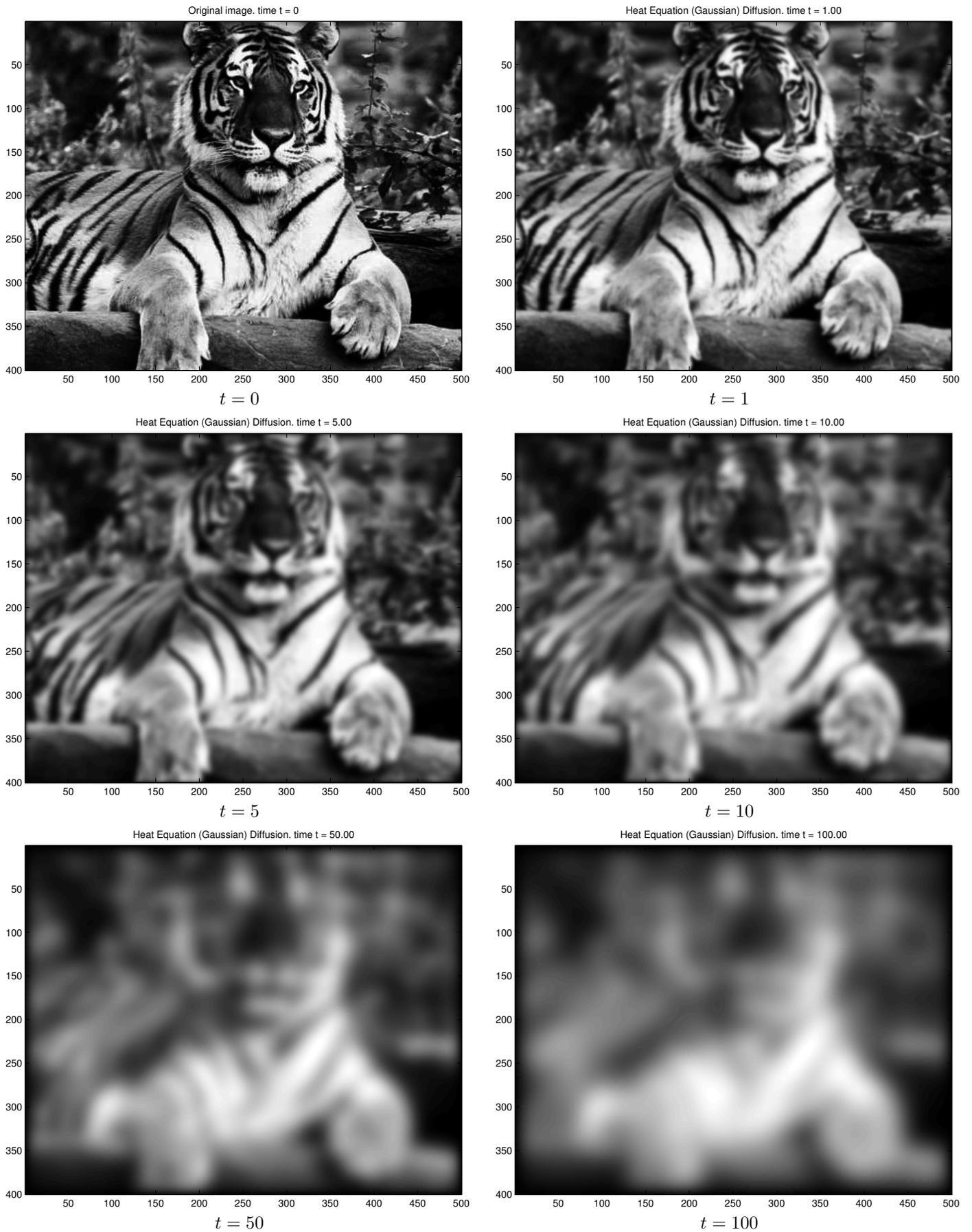


Figura 1. Difusión isotrópica mediante la ecuación del calor (1). Los bordes de los objetos en la imagen se difuminan sin importar su dirección.

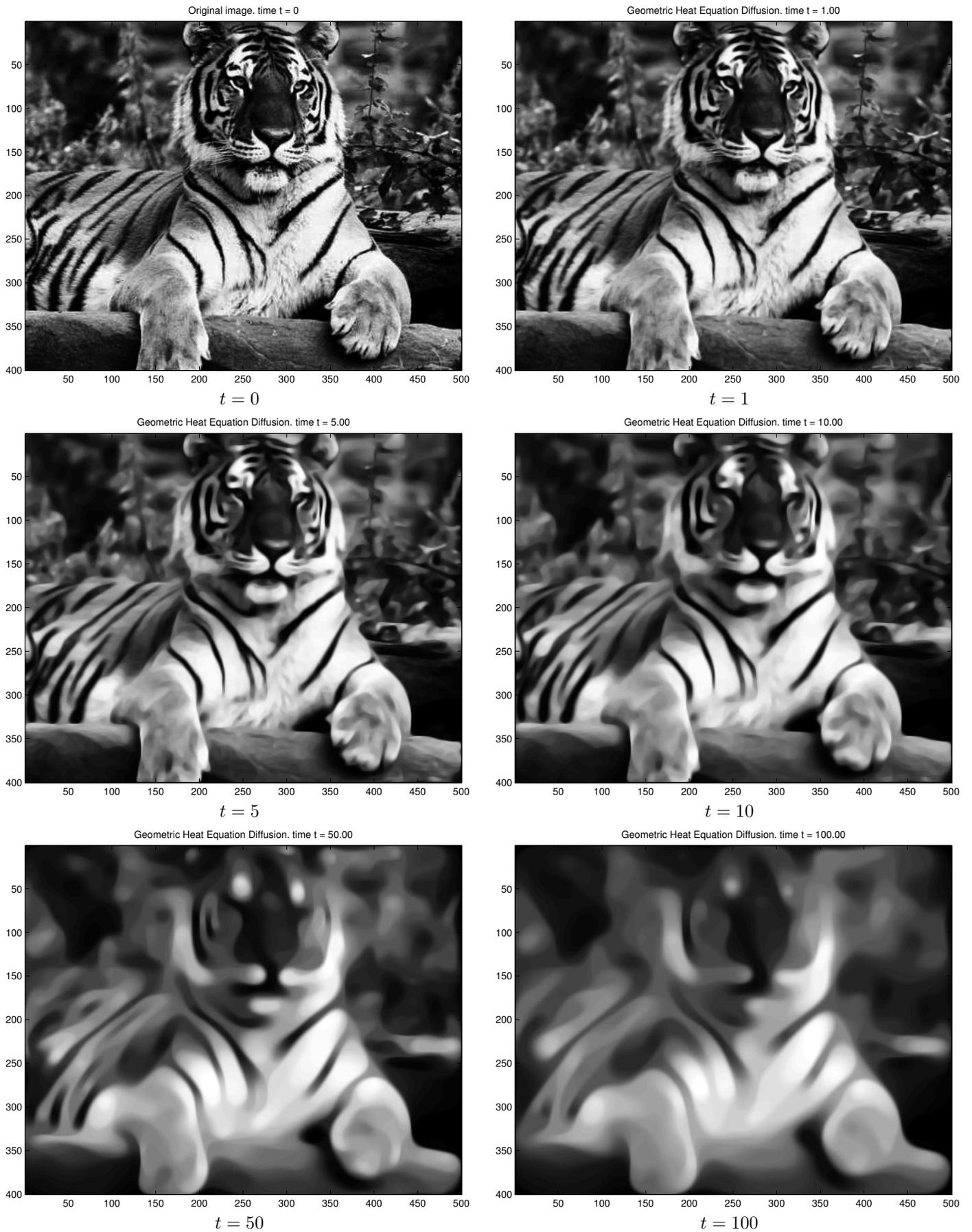


Figura 2. Difusión anisótropa mediante (11). Se preservan mejor los bordes de los objetos en la imagen que utilizando difusión isotrópica (figura 1).

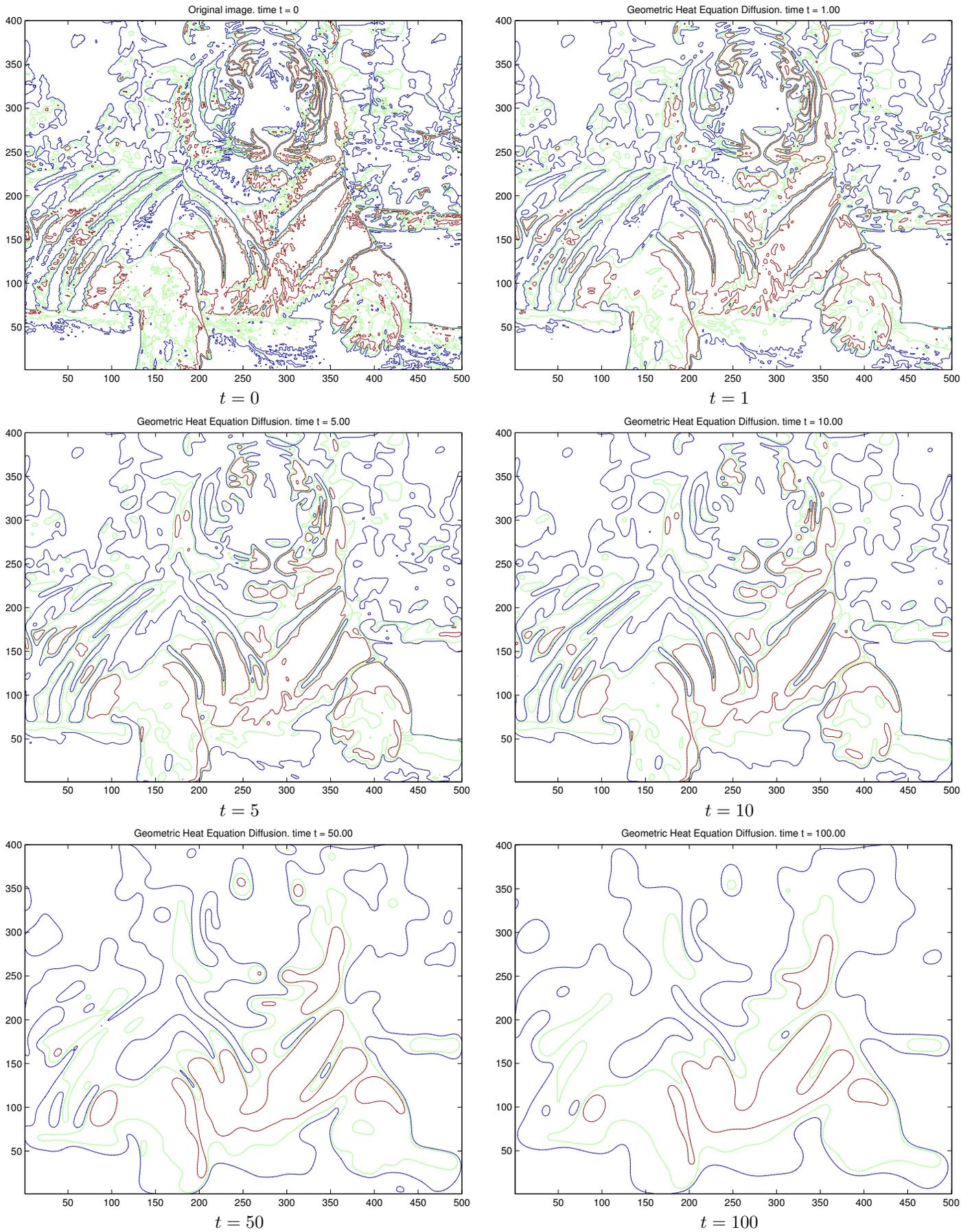


Figura 3. Difusión anisotrópica mediante la ecuación geométrica del calor (11). Evolución de las curvas de iso-intensidad para valores de  $I = \{64, 128, 196\}$ .