

EVALUACIÓN DE OPERADORES DIFERENCIALES γ -NORMALIZADOS PARA SELECCIÓN AUTOMÁTICA DE ESCALAS

Jorge A. Sánchez

Centro de Investigación en Informática para la Ingeniería (CIII), Universidad Tecnológica Nacional,
Facultad Regional Córdoba, jsanchez@scdt.frc.utn.edu.ar

Palabras clave: scale-space, automatic scale selection, feature detection, normalized derivatives

Resumen. La representación de una imagen en el espacio de escalas presenta una base formal a la extracción de características en las primeras etapas de procesamiento. Dicha representación se plantea en términos generales mediante la convolución de la imagen base con núcleos Gaussianos de longitud variable. De ello resulta una de sus propiedades fundamentales como es la de *no incremento de extremos locales*. Esto es, dada la representación mediante espacios de escala de una imagen determinada, el valor de un máximo local (mínimo local) no puede aumentar (disminuir) con la escala, por lo que, debido a la conmutabilidad de los operadores convolución y diferenciación, esta propiedad se transfiere también a las derivadas espaciales de dicha representación, lo que lleva a que, en general, la magnitud de tales derivadas decrezca con la escala.

Se introduce así el concepto de derivadas γ -normalizadas como una forma de compensar el decrecimiento producido por el suavizado gradual de la imagen. El caso de $\gamma = 1$ se conoce como de *invariancia perfecta de escala* en donde la magnitud de las derivadas permanece constante ante cambios de escala en el dominio de entrada. Si bien la magnitud no se conserva, la ubicación espacial de tales extremos se mantiene aún en el caso de $\gamma \neq 1$.

Una de las primeras operaciones a efectuar en la representación mediante espacios para escala de una imagen dada, es la determinación de las llamadas *escalas características*. Las mismas pueden ser entendidas como valores que reflejan la extensión espacial de las estructuras primitivas que aparecen en la imagen y donde su determinación se emplea en el guiado de etapas posteriores de procesamiento.

Dada una imagen cualquiera, la detección de escalas características se plantea como la búsqueda de extremos locales sobre una función de detección generada a partir de operadores diferenciales aplicados sobre la representación mediante espacios de escalas de la imagen en cuestión.

El trabajo propone la evaluación de diversos operadores para la selección automática de escalas, prestando especial interés a la variación (cualitativa y cuantitativa) de la respuesta de tales operadores ante variaciones de γ . Se realiza una evaluación comparativa de operadores diferenciales para selección automática de escalas en base a un conjunto de imágenes de prueba sobre las que se inducen transformaciones sintéticas.

1. INTRODUCCIÓN

Al observar el mundo mediante imágenes, los objetos y estructuras presentes en él aparecen a diferentes niveles de resolución y donde, en general, no es posible suponer *a priori* la manera (posición y escala) en la que nos serán representados dichos objetos.

Existen diversas aproximaciones, previas a la formulación mediante espacios de escala, que se basan en la representación multi-escala de la información y entre las que podemos mencionar los *quad-tree* (Finkel and Bentley, 1974) y la representación mediante *pirámides de resolución* (Burt, 1981; Crowley, 1981). La primera de ellas corresponde a la representación en forma de árbol de una imagen mediante la subdivisión recursiva de la misma en estructuras (regiones) más simples. La segunda comprende la representación mediante versiones suavizadas y sub-muestreadas de la imagen original, generalmente empleada de forma que la información extraída en niveles bajos de resolución sirve como estimación al cómputo en niveles más finos.

La teoría de los espacios de escala (Witkin, 1983; Koenderink, 1984; Yuille and Poggio, 1986; Lindeberg, 1990; Florack et al., 1994) surge por la necesidad de incorporar la noción de *escala* de forma explícita en la representación de una señal (imagen), mediante la generación de una familia monoparamétrica de señales de ella derivadas. Esta familia debe ser parametrizada en términos de escala y de forma tal que a medida que la misma es incrementada, las estructuras más finas deben ser suprimidas en forma gradual, lo que implica que la imagen a escalas gruesas debe ser una representación simplificada de la misma a escalas más finas (*causalidad*). Además, la ausencia de información *a priori*, implica que no existen ubicaciones espaciales particulares (*homogeneidad*) ni orientaciones preferidas (*isotropía*), como tampoco hay razones para discriminar entre estructuras pequeñas o grandes, con lo que todos los niveles de escalas deben ser tratadas de forma similar (*invarianza de escala*).

La representación en espacios de escala de una imagen no contiene información explícita respecto de cuales estructuras deben ser tratadas como significativas y en que rango de escalas se manifiestan de manera más adecuada.

El trabajo se encuentra organizado como sigue: en la sección 2 se introducen los conceptos básicos de la representación mediante espacios de escala y derivadas γ -normalizadas. en la sección 3 se introducen los conceptos de selección automática de escalas, escalas características y operadores diferenciales rotacionalmente simétricos empleados en su detección. Se efectúa también un análisis respecto de estructuras básicas. Por último, en la sección 4 se realiza una descripción de los métodos empleados en la evaluación y se presentan resultados experimentales.

2. REPRESENTACIÓN EN EL ESPACIO DE ESCALAS

Para cada $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ será definido un $L(f) : \mathbb{R}^2 \times \mathbb{R}_+ \rightarrow \mathbb{R}$, tal que

$$D_3 L(f)(\mathbf{x}, t) = \frac{1}{2} ((D_1^2 L(f))(\mathbf{x}, t) + (D_2^2 L(f))(\mathbf{x}, t)), \quad \forall (\mathbf{x}, t) \in \mathbb{R}^2 \times \mathbb{R}_+ \quad (1)$$

$$L(f)(\mathbf{x}, 0) = f(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^2 \quad (2)$$

donde $\mathbf{x} = (x_1, x_2)^T$ y $t \in \mathbb{R}_+$. Diremos que la representación de f en espacios de escala corresponde a la solución de (1) con condiciones iniciales (2). Se demuestra que dicha solución está dada por

$$L(f)(\mathbf{x}, t) = (g_t * f)(\mathbf{x}) \quad (3)$$

con $g_t : \mathbb{R}^2 \rightarrow \mathbb{R}$ dado por

$$g_t(\mathbf{x}) = \frac{1}{2\pi t} e^{-\frac{1}{2t} \mathbf{x}^T \mathbf{x}} \quad (4)$$

un núcleo Gaussiano parametrizado en términos de $t = \sigma^2$, y con

$$(g_t * f)(\mathbf{x}) = \int_{\mathbb{R}^2} g_t(\mathbf{y})f(\mathbf{x} - \mathbf{y})d\mathbf{y} \quad (5)$$

con $f(\mathbf{x})$ de cuadrado integrable en \mathbb{R}^2 .

Diversos trabajos demuestran la unicidad del núcleo Gaussiano en la generación de espacios de escala (Babaud et al., 1986; Lindeberg, 1994; Florack et al., 1994; Witkin, 1983).

Se observa que los núcleos de la forma (4) conforman un semi-grupo, por lo que dados dos niveles de escala diferentes, t_1 y t_2 (con $t_2 > t_1$), el cómputo de $L(f)(\mathbf{x}, t_2)$ pueda ser realizado mediante convolución de $L(f)(\mathbf{x}, t_1)$ con el núcleo $g_{(t_2-t_1)}(\mathbf{x})$ (propiedad de *suavizado en cascada*).

Si además es f continuamente diferenciable, debido a la propiedad de conmutabilidad de los operadores convolución y diferenciación, las propiedades de la representación de una imagen mediante un espacio de escalas se transfieren también a sus derivadas

$$D_i(L(f))(\mathbf{x}, t) = (D_i(g_t) * f)(\mathbf{x}) = (g_t * D_i(f))(\mathbf{x}), \quad \forall(\mathbf{x}, t) \in \mathbb{R}^2 \times \mathbb{R}_+, \quad i = 1, 2 \quad (6)$$

2.1. Discretización espacial

En forma análoga a la anterior, pero considerando el caso unidimensional por simplicidad, para cada $f : \mathbb{Z} \rightarrow \mathbb{R}$ será definido un $L(f) : \mathbb{Z} \times \mathbb{R}_+ \rightarrow \mathbb{R}$, tal que

$$D_3L(f)(n, t) = L(f)(n + 1, t) - 2L(f)(n, t) + L(f)(n - 1, t), \quad \forall(n, t) \in \mathbb{Z} \times \mathbb{R}_+ \quad (7)$$

$$L(f)(n, 0) = f(n), \quad \forall n \in \mathbb{Z} \quad (8)$$

definido $g : \mathbb{Z} \rightarrow \mathbb{R}$ como

$$g(n) = \begin{cases} 1 & \text{si } u = 1 \text{ ó } u = -1, \\ -2 & \text{si } u = 0, \\ 0 & \text{en caso contrario.} \end{cases}$$

se cumple que

$$D_3L(f)(n, t) = (g * L(f)(\cdot, t))(n) = \sum_{k=-\infty}^{+\infty} g(k)L(f)(n - k, t) \quad (9)$$

Por lo que la versión semi-discretizada de la representación de f en un espacio de escalas, viene determinada por (Lindeberg, 1990)

$$L(f)(n, t) = (P_t * f)(n) \quad (10)$$

con $P_t : \mathbb{Z} \rightarrow \mathbb{R}$ el núcleo simétrico de Poisson

$$P_t(n) = e^{-t}I_n(t) \quad (11)$$

con f cuadrado sumable y $I_n(t)$ es la función modificada de Bessel de orden n (Abramowitz and Stegun, 1964).

Para el caso bidimensional, la representación de $f : \mathbb{Z}^2 \rightarrow \mathbb{R}$ en un espacio de escalas, se obtiene mediante la propiedad de separabilidad del núcleo (11) computado sobre direcciones ortogonales

$$L(f)(\mathbf{u}, t) = ((P_t \oplus P_t) * f)(\mathbf{u}) \quad \forall(\mathbf{u}, t) \in \mathbb{Z}^2 \times \mathbb{R}_+ \quad (12)$$

donde $\mathbf{u} = (m, n)^T$ y con $(P_t \oplus P_t) : \mathbb{Z}^2 \rightarrow \mathbb{Z}$ dada por

$$(P_t \oplus P_t)(\mathbf{u}) = P_t(m)P_t(n) \quad \forall \mathbf{u} \in \mathbb{Z}^2 \quad (13)$$

esto es

$$\begin{aligned} ((P_t \oplus P_t) * f)(\mathbf{u}) &= \sum_{i=-\infty}^{i=\infty} \sum_{j=-\infty}^{j=\infty} P_t(i)P_t(j)f(m-i, n-j) \\ &= \sum_{i=-\infty}^{i=\infty} P_t(i) \sum_{j=-\infty}^{j=\infty} P_t(j)f(m-i, n-j) \\ &= \sum_{j=-\infty}^{j=\infty} P_t(j) \sum_{i=-\infty}^{i=\infty} P_t(i)f(m-i, n-j) \end{aligned} \quad (14)$$

2.2. Comportamiento ante cambios de escala isotrópicos

Para cada $k \in \mathbb{R}_+$, defínase las composiciones $T_k : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ y $S_k : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, los que aplicados respectivamente a $\mathbf{x} \in \mathbb{R}^2$ y $t \in \mathbb{R}_+$ resultan en

$$T_k(\mathbf{x}) = k\mathbf{x} \quad (15)$$

$$S_k(t) = k^2t \quad (16)$$

Defínase así también $Q_k : \mathbb{R}^2 \times \mathbb{R}_+ \rightarrow \mathbb{R}^2 \times \mathbb{R}_+$, de forma tal que

$$Q_k(\mathbf{x}, t) = (T_k(\mathbf{x}), S_k(t)) = (k\mathbf{x}, k^2t) \quad (17)$$

Sean ahora $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ y $h : \mathbb{R}^2 \rightarrow \mathbb{R}$, definidas de forma tal que se cumple

$$f(\mathbf{x}) = (h \circ T_k)(\mathbf{x}) \quad (18)$$

La representación en espacios de escala de h , definidas en términos del dominio transformado, $Q_k(\mathbf{x}, t)$, resulta

$$\begin{aligned} L(h)(T_k(\mathbf{x}), S_k(t)) &= L(h)(k\mathbf{x}, k^2t) \\ &= \int_{\mathbb{R}^2} g_{(k^2t)}(\mathbf{y})h(k\mathbf{x} - \mathbf{y})d\mathbf{y} \\ &= \int_{\mathbb{R}^2} g_{(k^2t)}(k\mu)h(k(\mu - \mathbf{x}))k^2d\mu \\ &= \int_{\mathbb{R}^2} g_t(\mu)f(\mu - \mathbf{x})d\mu \\ L(h)(T_k(\mathbf{x}), S_k(t)) &= L(f)(\mathbf{x}, t) \end{aligned} \quad (19)$$

que resulta de aplicar el cambio de variables $y = k\mu$ y observar que $g_t(\mathbf{x}) = k^2g_{(S_k(t))}(T_k(\mathbf{x}))$.

Así, las derivadas de la representaciones de f respecto del dominio original y de h respecto del compuesto mediante (17), se encuentran relacionadas de la forma

$$(D_1^m D_2^n L(f))(\mathbf{x}, t) = k^{(m+n)}(D_1^m D_2^n L(h))(T_k(\mathbf{x}), S_k(t)) \quad (20)$$

$$(D_3^p L(f))(\mathbf{x}, t) = k^{2p}(D_3^p L(h))(T_k(\mathbf{x}), S_k(t)) \quad (21)$$

Se asume que tanto f como h cumplen con las suficientes condiciones de regularidad para que las expresiones anteriores sean válidas.

2.3. Derivadas γ -normalizadas

A fin de compensar el decrecimiento gradual en la magnitud de las derivadas espaciales con la escala que surge por el suavizado gradual de la imagen original, se introduce el concepto de *derivadas γ -normalizadas* (Lindeberg, 1994).

Dados $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ y su representación en espacios de escalas mediante $L(f) : \mathbb{R}^2 \times \mathbb{R}_+ \rightarrow \mathbb{R}$, se define $D_{i,\gamma}^n : (\mathbb{R}^2 \times \mathbb{R}_+) \rightarrow (\mathbb{R}^2 \times \mathbb{R}_+)$, tal que

$$D_{i,\gamma}^n(L(f))(\mathbf{x}, t) = t^{n\gamma/2} D_i^n(L(f))(\mathbf{x}, t), \quad \forall(\mathbf{x}, t) \in \mathbb{R}^2 \times \mathbb{R}_+ \quad (22)$$

con $i = 1, 2$ y $L(f)$ continuamente diferenciable hasta orden n . El comportamiento ante cambios de escala isotrópicos de la forma (17) se puede obtener de (20) y resulta

$$(D_{1,\gamma}^n D_{2,\gamma}^m L(f))(\mathbf{x}, t) = k^{(m+n)(\gamma-1)} (D_1^n D_2^m L(f))(\mathbf{x}, t) \quad (23)$$

el caso de $\gamma = 1$ es conocido como de *invariancia perfecta de escala*. Si se considera el caso de $\gamma \neq 1$ se observa que la magnitud de las derivadas espaciales no se preserva ante cambios de escala del dominio de f , pero sí la ubicación de los puntos extremos sobre t . Ello resulta de observar en (21) que

$$D_3 L(f)(\mathbf{x}, t) = 0 \iff D_3 L(h)(T_k(\mathbf{x}), S_k(t)) = 0 \quad (24)$$

Se demuestra (Lindeberg, 1998), además, que el concepto de derivadas γ -normalizadas equivale a la normalización de la L_p -norma de las derivadas del núcleo Gaussiano (4). Dicho concepto se utiliza en (Lindeberg and Bretzner, 2003) para la normalización de las derivadas espaciales en representaciones mediante espacios de escala computados en forma de pirámides híbridas.

3. SELECCIÓN AUTOMÁTICA DE ESCALAS

Dada una imagen y su representación mediante espacios de escala, la idea de selección automática de escalas para un punto dado, $\mathbf{x}_0 \in \mathbb{Z}^2$, en f , consiste en asignar a dicho punto una escala (*escala característica*) de forma que la misma refleje la extensión espacial de la estructura a la que pertenece.

En (Lindeberg, 1994) se propone un método para la selección automática de escalas basado en la detección de extremos (sobre escalas) en combinaciones, posiblemente no-lineales, de derivadas normalizadas. En la figura 1 se ilustra, para $\gamma = 1$, la evolución con la escala (medida en términos de σ) la magnitud del Laplaciano γ -normalizado (33) computado sobre los centros de cada una de las estructuras que aparecen en la imagen de la izquierda. Obsérvese que la posición del máximo se mueve hacia la derecha a medida que aumenta el diámetro de los círculos blancos respectivos.

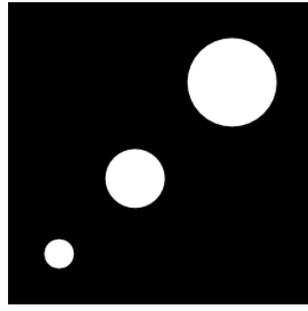
3.1. Simetría rotacional en operadores diferenciales

Dada $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, para cada $\theta \in [0, 2\pi)$ sea $R(\theta)$ una matriz de 2×2 dada por

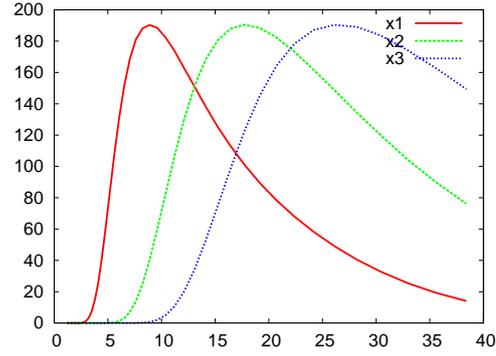
$$R(\theta) = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \quad (25)$$

f es una *función rotacionalmente simétrica* si y solo si

$$f(\mathbf{x}) = f(R(\theta)\mathbf{x}) \quad \forall \mathbf{x} \in \mathbb{R}^2, \forall \theta \in [0, 2\pi) \quad (26)$$



(a)



(b)

Figura 1: Evolución en la magnitud del Laplaciano γ -normalizado con la escala para estructuras (blobs) de diferente tamaño, computada en sus centros respectivos, x_i ($i=1,2,3$).

Sea el operador $\mathcal{O} : (\mathbb{R}^2 \rightarrow \mathbb{R}) \rightarrow (\mathbb{R}^2 \rightarrow \mathbb{R})$. Diremos que \mathcal{O} es un *operador rotacionalmente simétrico* si $\mathcal{O}(f)$ es una función rotacionalmente simétrica, $\forall f \in \mathcal{F}(\mathbb{R}^2, \mathbb{R})$.

En (Brady and Horn, 1983) se define el concepto de invariancia rotacional en términos de las denominadas *formas cuadráticas de operadores*. Así, para el caso de derivadas de primer orden, se dice que una expresión es una *forma cuadrática en D_1 y D_2* si la misma es de la forma

$$[D_1 \quad D_2] \begin{bmatrix} \lambda & \mu \\ \mu & \xi \end{bmatrix} \begin{bmatrix} D_1 \\ D_2 \end{bmatrix} \quad (27)$$

y se demuestra que la única forma cuadrática rotacionalmente simétrica en D_1 y D_2 , es el operador *gradiente cuadrado*, $\mathcal{SG} : (\mathbb{R}^2 \rightarrow \mathbb{R}) \rightarrow (\mathbb{R}^2 \rightarrow \mathbb{R})$, definido de forma tal que

$$\mathcal{SG}(f)(\mathbf{x}) = [D_1 f(\mathbf{x})]^2 + [D_2 f(\mathbf{x})]^2 \quad (28)$$

Además, en el citado trabajo, se extiende el concepto anterior a *formas cuadráticas (de operadores) en D_1^2 , $D_1 D_2$, $D_2 D_1$ y D_2^2* , representadas mediante expresiones de la forma

$$[D_1^2 \quad D_1 D_2 \quad D_2 D_1 \quad D_2^2] \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} D_1^2 \\ D_1 D_2 \\ D_2 D_1 \\ D_2^2 \end{bmatrix} \quad (29)$$

donde A , B y C son matrices de 2×2 con A y C simétricas. Se demuestra que, para $D_1 D_2 = D_2 D_1$ dicho conjunto forma un espacio vectorial de dimensión 2 con bases dadas por los operadores *variación cuadrática*, $\mathcal{O}_{QV} : (\mathbb{R}^2 \rightarrow \mathbb{R}) \rightarrow (\mathbb{R}^2 \rightarrow \mathbb{R})$, y *laplaciano cuadrado*, $\mathcal{O}_{SL} : (\mathbb{R}^2 \rightarrow \mathbb{R}) \rightarrow (\mathbb{R}^2 \rightarrow \mathbb{R})$, definidos de forma tal que

$$\mathcal{O}_{QV}(f)(\mathbf{x}) = [D_1^2 f(\mathbf{x})]^2 + 2[D_1 D_2 f(\mathbf{x})]^2 + [D_2^2 f(\mathbf{x})]^2, \quad (30)$$

$$\mathcal{O}_{SL}(f)(\mathbf{x}) = [D_1^2 f(\mathbf{x}) + D_2^2 f(\mathbf{x})]^2, \quad \forall \mathbf{x} \in \mathbb{R}^2 \quad (31)$$

y con f cumpliendo las condiciones suficientes de regularidad.

3.2. Operadores diferenciales para selección automática de escalas

Una propiedad deseada en los operadores empleados en la selección automática de escalas es la de invariancia rotacional. El Laplaciano es un operador circularmente simétrico empleado usualmente en la detección de *blobs* y selección automática de escalas (Lindeberg, 1998).

Además, dicho operador puede ser aproximado mediante *Diferencias de Gaussianas (DoG)*, computadas desde niveles de escala adyacentes en la representación en espacio de escalas de una imagen, construido de forma tal que dados dos niveles sucesivos, σ_n y σ_{n-1} , se cumple que $\sigma_n = k\sigma_{n-1}$, con $k > 1$ (Lowe, 1999), resultando en una importante reducción del costo computacional.

En (Lindeberg, 1996) se emplea el determinante de la matriz Hessiana sobre $L(f)(\mathbf{x}, t)$ para la selección de escalas, donde se observa que puede ser representado, para $(D_1D_2L(f)(\cdot, t))(\mathbf{x}) = (D_2D_1L(f)(\cdot, t))(\mathbf{x})$, de la forma

$$\begin{aligned} \det(\mathcal{H}(\mathbf{x}, t)) &= \begin{vmatrix} D_1^2L(f)(\mathbf{x}, t) & D_1D_2L(f)(\mathbf{x}, t) \\ D_2D_1L(f)(\mathbf{x}, t) & D_2^2L(f)(\mathbf{x}, t) \end{vmatrix} \\ &= \frac{1}{2} [(\mathcal{O}_{SL}L(f)(\cdot, t))(\mathbf{x}) - (\mathcal{O}_{QV}L(f)(\cdot, t))(\mathbf{x})] \end{aligned} \quad (32)$$

con \mathcal{O}_{QV} y \mathcal{O}_{SL} dados mediante (30) y (31).

En (Mikolajczyk, 2002) se emplea la detección simultánea de extremos tanto en el determinante como de la traza de la matriz Hessiana a fin de favorecer la selección de puntos que presentan autovalores de magnitudes similares y elevadas. Obsérvese que la traza de dicha matriz corresponde al laplaciano.

Dados $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ y su representación en un espacio de escalas $L(f) : \mathbb{R}^2 \times \mathbb{R}_+ \rightarrow \mathbb{R}$, se definirá, a fines de evaluación, el siguiente conjunto de operadores diferenciales γ -normalizados, $\mathcal{O}_i : (\mathbb{R}^2 \times \mathbb{R}_+ \rightarrow \mathbb{R}) \rightarrow (\mathbb{R}^2 \times \mathbb{R}_+ \rightarrow \mathbb{R})$, tales que

Laplaciano

$$(\mathcal{O}_1L(f))(\mathbf{x}, t) = t^\gamma \{ (D_1^2L(f))(\mathbf{x}, t) + (D_2^2L(f))(\mathbf{x}, t) \} \quad (33)$$

Variación cuadrática

$$\begin{aligned} (\mathcal{O}_2L(f))(\mathbf{x}, t) &= t^{2\gamma} \{ [(D_1^2L(f))(\mathbf{x}, t)]^2 + \\ &\quad + 2[(D_1D_2L(f))(\mathbf{x}, t)]^2 + [(D_2^2L(f))(\mathbf{x}, t)]^2 \} \end{aligned} \quad (34)$$

Magnitud del gradiente (cuadrado)

$$(\mathcal{O}_3L(f))(\mathbf{x}, t) = t^\gamma \{ [(D_1L(f))(\mathbf{x}, t)]^2 + [(D_2L(f))(\mathbf{x}, t)]^2 \} \quad (35)$$

Determinante de la matriz Hessiana

$$\begin{aligned} (\mathcal{O}_4L(f))(\mathbf{x}, t) &= t^{2\gamma} \{ [(D_1^2L(f))(\mathbf{x}, t)] [(D_2^2L(f))(\mathbf{x}, t)] - \\ &\quad [(D_1D_2L(f))(\mathbf{x}, t)]^2 \} \end{aligned} \quad (36)$$

La aproximación al Laplaciano mediante *Diferencia de Gaussianas* resulta de expresar la *ecuación de difusión* (1) en términos de $\sigma = \sqrt{t}$ y expresar $D_3L(f)(\mathbf{x}, t)$ mediante diferencias. Dicha aproximación incluye el factor de normalización, σ^2 , requerido para invariancia perfecta de escalas, $\gamma = 1$ (Lowe, 1999). A fin de compensar el efecto ante valores de $\gamma \neq 1$ introduciremos el concepto de operador *DoG γ -normalizado*, $\mathcal{O}_5 : (\mathbb{R}^2 \times \mathbb{R}_+ \rightarrow \mathbb{R}) \rightarrow (\mathbb{R}^2 \times \mathbb{R}_+ \rightarrow \mathbb{R})$, tal que

$$(\mathcal{O}_5L(f))(\mathbf{x}, \sigma) = \frac{\sigma^{2(\gamma-1)}}{(k-1)} [(L(f)(\cdot, k\sigma))(\mathbf{x}) - (L(f)(\cdot, \sigma))(\mathbf{x})] \quad (37)$$

con $\sigma_n = k\sigma_{n-1}$, $k > 1$ y $n = 1, 2, \dots, N$, donde N es el número de niveles de escala computados en la representación de f . El factor $(k-1)^{-1}$ es constante para todo valor de σ y puede ser obviado en la implementación.

3.3. Escala característica

Dada una $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ y su representación en un espacio de escalas $L : \mathbb{R}^2 \times \mathbb{R}_+ \rightarrow \mathbb{R}$. Sea $h_i : \mathbb{R}^2 \times \mathbb{R}_+ \rightarrow \mathbb{R}$ una función para selección de escalas, definida de forma tal que tal que

$$h_i(\mathbf{x}, t) = (\mathcal{O}_i L(f))(\mathbf{x}, t) \quad (38)$$

para alguno de los operadores $\mathcal{O}_i : (\mathbb{R}^2 \times \mathbb{R}_+ \rightarrow \mathbb{R}) \rightarrow (\mathbb{R}^2 \times \mathbb{R}_+ \rightarrow \mathbb{R})$ presentados en la sección 3.2. Se definirá como *escala característica* para un punto $\mathbf{x}_0 \in \mathbb{R}^2$ en f , a los valores de t que (en caso de existir) sean soluciones de

$$(D_3 h_i)(\mathbf{x}_0, t) = 0 \quad (39)$$

3.4. Análisis de respuesta para modelos simples de señales

Los operadores presentados en la sección anterior presentarán respuestas diferentes de acuerdo al tipo de características (estructuras) presentes en la imagen. Considérese los siguientes modelos Gaussianos de señales (Lindeberg, 1996, 1998) y sus respectivas representaciones mediante espacios de escala (centradas en $(x_1, x_2)^T = (0, 0)^T$)

Borde,

$$\begin{aligned} f_{t_0}^e(\mathbf{x}) &= \int_{u=-\infty}^{x_1} \frac{1}{t_0 \sqrt{2\pi}} e^{-\frac{1}{2} \frac{u^2}{t_0}} du = \Phi(x_1, t_0) \\ L(f_{t_0}^e)(\mathbf{x}, t) &= (g_t * f_{t_0}^e)(\mathbf{x}) \\ &= \int_{\mathbb{R}^2} g_t(\mathbf{y}) f_{t_0}^e(\mathbf{x} - \mathbf{y}) d\mathbf{y} \\ &= \int_{\mathbb{R}^2} g_t(\mathbf{y}) \Phi(x_1 - y_1, t_0) d\mathbf{y} \\ L(f_{t_0}^e)(\mathbf{x}, t) &= \Phi(x_1, t_0 + t) \end{aligned} \quad (40)$$

Esquina (unión L),

$$\begin{aligned} f_{t_0}^c(\mathbf{x}) &= \Phi(x_1, t_0) \Phi(x_2, t_0) \\ L(f_{t_0}^c)(\mathbf{x}, t) &= \Phi(x_1, t_0 + t) \Phi(x_2, t_0 + t) \end{aligned} \quad (41)$$

Blob,

$$\begin{aligned} f_{t_0}^b(\mathbf{x}) &= g_{t_0}(\mathbf{x}) \\ L(f_{t_0}^b)(\mathbf{x}, t) &= g_{(t_0+t)}(\mathbf{x}) \end{aligned} \quad (42)$$

En cada caso, el parámetro t_0 refleja la *extensión* de la estructura considerada y puede considerarse como la escala característica de la misma. La Tabla 1 muestra, para cada uno de los operadores diferenciales γ -normalizados considerados y para cada una de las estructuras arriba mencionadas, la escala característica correspondiente, $t_{caract.}$.

La forma general en los extremos detectados (para los casos considerados) corresponde a

$$t_{caract.} = \frac{\gamma}{a - \gamma} t_0 \quad (43)$$

de forma que si $\gamma = a/2$, la misma corresponderá a t_0 , lo que para el caso de *blobs* lleva a la elección de $\gamma = 1$ y corresponde al caso de invariancia perfecta de escalas, mientras que para *bordes* y *esquinas* conduce a $\gamma = 0,5$.

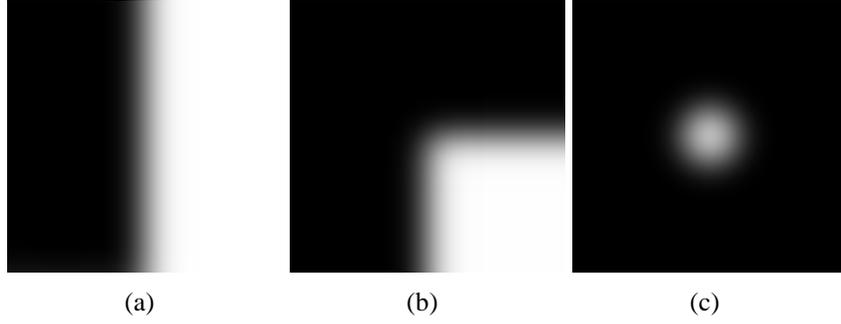


Figura 2: Ejemplo de estructuras Gaussianas: (a) borde, (b) esquina y (c) blob.

	Laplaciano	Gradiente	Var. Cuadrática	Det. Hessiano
borde	-	$t_{caract.}^e = \frac{\gamma}{1-\gamma}t_0$	-	-
esquina	-	$t_{caract.}^e = \frac{\gamma}{1-\gamma}t_0$	$t_{caract.}^c = \frac{\gamma}{1-\gamma}t_0$	$t_{caract.}^c = \frac{\gamma}{1-\gamma}t_0$
blob	$t_{caract.}^b = \frac{\gamma}{2-\gamma}t_0$	-	$t_{caract.}^b = \frac{\gamma}{2-\gamma}t_0$	$t_{caract.}^b = \frac{\gamma}{2-\gamma}t_0$

Tabla 1: Escalas características para el punto $\mathbf{x}_0 = (0, 0)^T$ detectadas mediante extremos en $L_t(\mathbf{x}_0)$

La selección de valores de $\gamma \neq a/2$ lleva a la selección de escalas características diferentes, pero proporcionales, a t_0 . Como se señala en (Mikolajczyk, 2002), esto puede resultar en una ventaja a nivel de implementación, ya que mientras menor sea t_0 para una estructura determinada, menores serán los niveles de escala necesarios a computar en la representación de f .

Un caso interesante de analizar es la respuesta en la selección de escalas características mediante el Laplaciano. Se ha señalado que una de las desventajas de dicho operador es que presenta extremos en torno a bordes y líneas rectas, estructuras que se caracterizan por cambios de intensidad en una sola dirección, llevando esto a la detección de puntos poco estables en el sentido de que su localización es muy sensible al ruido y a pequeños cambios en la textura (Mikolajczyk, 2002).

Considérese el operador *Laplaciano cuadrado γ -normalizado*, $\mathcal{O}_{SL,\gamma} : (\mathbb{R}^2 \times \mathbb{R}_+ \rightarrow \mathbb{R}) \rightarrow (\mathbb{R}^2 \times \mathbb{R}_+ \rightarrow \mathbb{R})$, tal que

$$(\mathcal{O}_{SL,\gamma}L(f))(\mathbf{x}, t) = t^{2\gamma}[D_1^2L(f)(\mathbf{x}, t) + D_2^2L(f)(\mathbf{x}, t)]^2, \quad \forall(\mathbf{x}, t) \in \mathbb{R}^2 \times \mathbb{R}_+ \quad (44)$$

Dicho operador, aplicado sobre el modelo para un borde Gaussiano (40) centrado en $\mathbf{x}_0 = (0, 0)^T$ y $\gamma > 0$,

$$\begin{aligned} (\mathcal{O}_{SL,\gamma}L(f_{t_0}^e))(\mathbf{x}, t) &= t^{2\gamma}[D_1^2L(f_{t_0}^e)(\mathbf{x}, t) + D_2^2L(f_{t_0}^e)(\mathbf{x}, t)]^2 \\ &= t^{2\gamma}[D_1^2\Phi(x_1, t + t_0)]^2 \end{aligned} \quad (45)$$

los extremos (escalas características), son las soluciones de

$$(D_3\mathcal{O}_{SL,\gamma}L(f_{t_0}^e))(\mathbf{x}, t) = 0 \quad (46)$$

y corresponden a

$$(2\gamma - 3)t^{2\gamma+1} + (x_1^2 + t_0(4\gamma - 3))t^{2\gamma} + (2\gamma t_0^2)t^{2\gamma-1} = 0 \quad (47)$$

dentro del espacio de soluciones de la expresión anterior se encuentra $t = 0$, que corresponde al nivel base (imagen original) y no reviste mayor interés. Así, resulta una expresión cuadrática

de la forma

$$(2\gamma - 3)t^2 + (x_1^2 + t_0(4\gamma - 3))t + (2\gamma t_0^2) = 0 \quad (48)$$

cuyas soluciones corresponden a

$$t = \pm \frac{\sqrt{x_1^4 + 2(4\gamma - 3)t_0x_1^2 + 9t_0^2} \mp (x_1^2 + (4\gamma - 3)t_0)}{4\gamma - 6} \quad (49)$$

en donde se observa que no aparecerán extremos si se elige $\gamma = 3/2$, lo que corresponde a la normalización de L_2 -norma en las derivadas espaciales del núcleo Gaussiano bidimensional (Lindeberg, 1996).

Para el caso del Laplaciano, las soluciones de $(D_3\mathcal{O}_1L(f_{t_0}^e))(\mathbf{x}, t) = 0$ corresponden a las de una expresión cúbica en t , cuyas raíces son a las dos dadas en (49) mas una en el punto $t = -t_0$, valor que no pertenece al dominio de $L(f)$.

4. EVALUACIÓN EXPERIMENTAL

La evaluación de los operadores para selección automática de escalas sigue el método propuesto en (Mikolajczyk, 2002), pero incluyendo variaciones de γ . Dada una imagen, se computó su representación en espacios de escala para valores de σ tales que $\sigma_n = \sigma_0^n$, con $\sigma_0 = 1,2$ y $n = 0, 1, \dots, 16$ y se aplicó a tal representación cada uno de los operadores definidos en la sección 3.2, para valores de γ comprendidos en el intervalo $[0,2; 1,9]$ y a incrementos $\Delta\gamma = 0,1$. Se consideraron factores de escala (original/escalada) de hasta 5,15, los que fueron inducidos en forma sintética mediante interpolación bilineal. La figura 3 muestra el conjunto de imágenes de prueba empleado¹.

A fines de evaluación, se considerarán: (a) puntos para los que se ha detectado al menos una escala característica, (b) puntos a los que no se les ha detectado una escala característica y (c) puntos a los que se les ha detectado una escala característica y la misma es correcta. Aquí, la detección (no detección) de escalas características resulta de la existencia (no existencia) de soluciones en (39), mientras que la escala detectada para un punto se considera correcta si, dada la imagen original, $f_b(\mathbf{x})$, y una versión escalada de la misma, $f_s(\mathbf{x}) = (f_b \circ T_k)(\mathbf{x})$, la relación entre escalas características detectadas para puntos correspondientes es igual a la relación de escala entre las imágenes. Aquí, y a diferencia de (Mikolajczyk, 2002), no es aplicado ningún umbral durante la detección de escalas características para un punto determinado.

Para valores de γ en el intervalo $[\gamma_{min}, \gamma_{max}]$, se consideran las siguientes evaluaciones cuantitativas: (a) porcentaje de puntos para los que se ha detectado una escala característica, (b) porcentaje de puntos para los que se ha detectado correctamente una escala característica y (c) porcentaje de puntos en la imagen para los que se ha detectado correctamente una escala característica, respecto de puntos con escalas detectadas.

La figura 4 ilustra, para cada operador considerado, la variación de dichos parámetros respecto de variaciones de γ para una relación de escalas entre imágenes de 2,48. Los resultados corresponden al promedio de las evaluaciones individuales sobre cada imagen de prueba.

Si se considera el porcentaje máximo de puntos con escalas correctamente detectadas, se observa que, sin aplicación de umbrales en la detección, el mejor comportamiento lo presenta el esquema basado en el *determinante de la matriz Hessiana*. Dicho máximo lo alcanza para $\gamma = 1$, lo que corresponde al caso de invariancia perfecta de escalas. Esto se observa a su vez en los operadores *gradiente cuadrado* y *variación cuadrática*, en tanto que para los operadores

¹<http://www.robots.ox.ac.uk/~vgg/research/affine/>



Figura 3: Imágenes de prueba

laplaciano y *DoG* el máximo en el porcentaje de puntos con escalas correctas respecto de las detectadas, se alcanza para valores de γ próximos a 1,5.

Como se ha señalado, puntos que presentan cambios de intensidad importantes solo en una dirección (bordes) presentan inconvenientes debido a la pobre localización. Debido a ello, surge la necesidad de incluir una medida cualitativa que caracterice en cierta forma la *calidad* de los puntos que presentan soluciones a (39), aplicado un operador determinado.

El detector de esquinas de Harris (Harris and Stephens, 1988) se basa en asignar a cada punto de una imagen dada, una medida de la distribución local del gradiente en un entorno del mismo.

Brevemente, dada una $f : \mathbb{Z}^2 \rightarrow \mathbb{R}$, se define la *medida de Harris*, $\mathcal{C}(f) : (\mathbb{Z}^2 \rightarrow \mathbb{R}) \rightarrow (\mathbb{Z}^2 \rightarrow \mathbb{R})$, tal que

$$(\mathcal{C}(f))(\mathbf{u}) = |\mathbf{H}(\mathbf{u})| - k[\text{tr}(\mathbf{H}(\mathbf{u}))]^2 \quad \forall \mathbf{u} \in \mathbb{Z}^2, \forall k \in \mathbb{R} \quad (50)$$

con $\mathbf{H}(\mathbf{u})$ una matriz de 2×2 dada por

$$\mathbf{H}(\mathbf{u}) = \begin{bmatrix} \langle [(D_1 f)(\mathbf{u})]^2 \rangle & \langle [(D_1 f)(\mathbf{u})][(D_2 f)(\mathbf{u})] \rangle \\ \langle [(D_1 f)(\mathbf{u})][(D_2 f)(\mathbf{u})] \rangle & \langle [(D_2 f)(\mathbf{u})]^2 \rangle \end{bmatrix} \quad (51)$$

donde la notación $\langle \cdot \rangle$ denota promediación en un entorno reducido de \mathbf{u} .

Puntos con variaciones de intensidad elevada en direcciones ortogonales, corresponderán a valores positivos y elevados de la *medida de Harris*, mientras que bordes y regiones homogéneas presentarán valores negativos y próximos a cero. El parámetro k representa el grado de rechazo a bordes y en las pruebas efectuadas fue fijado en un valor de 0,04.

Se incluyen las siguientes evaluaciones cualitativas basadas en la medida de Harris: (a) valor medio para puntos con escalas detectadas y cuya medida de Harris es positiva y mayor a un umbral predefinido; (b) valor medio en detecciones correctas con medida de Harris positiva y mayor a un umbral predefinido. La figura 5 ilustra dichas evaluaciones para una relación de escala de 2,48 y un umbral de 1,0 aplicado sobre la medida de Harris.

Se puede observar que, en general, existe una relación aproximadamente lineal entre el incremento de γ y el valor medio en la medida de Harris para puntos con escalas correctamente

detectadas. En el caso de los operadores *magnitud de gradiente* y *variación cuadrática* el mismo se estabiliza en forma aproximada a valores de 1,3 y 1,6 respectivamente.

En la figura 6 se ilustra, respecto de variaciones en el factor de escala y para cada uno de los operadores considerados, los valores de γ que corresponden a los máximos en los porcentajes de puntos con escalas detectadas, puntos con escalas correctamente detectadas y puntos con escalas correctas respecto de las detectadas.

Allí, para el caso del *determinante de Hessiano* y *variación cuadrática*, se observa que el mayor porcentaje de puntos con escalas correctamente detectadas se ubica en torno a valores de $\gamma = 1$. Para los operadores *laplaciano*, *DoG* y *gradiente cuadrado*, se observa que el porcentaje de puntos con escalas correctamente detectadas se produce para valores de $\gamma > 1$, siendo mayor cuanto mayor es el factor de escala considerado.

5. CONCLUSIONES

Fueron evaluados diferentes operadores diferenciales γ -normalizados para la selección automática de escala prestando especial interés en el comportamiento de los mismos ante variaciones de γ .

Se ha observado (Mikolajczyk, 2002) que el empleo de valores de $\gamma < 1$ permite reducir el número de niveles computados en la representación mediante espacios de escala de una imagen dada. Esto resulta de que los máximos en la función detección para una estructura determinada, aparecen a valores de escala inferiores respecto de los correspondientes para $\gamma = 1$. En el citado trabajo se propone la detección de extremos, en forma simultánea, tanto en el *laplaciano* como en el *determinante del Hessiano*, a fin de mejorar la calidad de los puntos detectados.

Los resultados aquí presentados demuestran que el empleo de valores de $\gamma > 1$ presenta ventajas en términos cualitativos. Ello resulta muy interesante si se considera la detección de escalas características como una de las primeras etapas de procesamiento y en donde una mejora en tales términos repercute en el rendimiento de procesos posteriores.

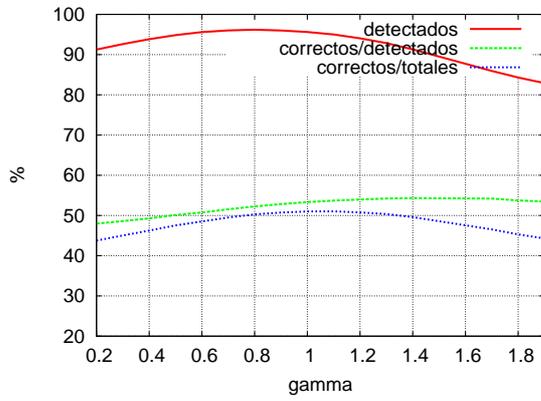
AGRADECIMIENTOS

A los integrantes del Centro de Investigación en Informática para la Ingeniería, CIII, por su constante apoyo. Al Dr. Oscar Bustos, por sus comentarios y recomendaciones.

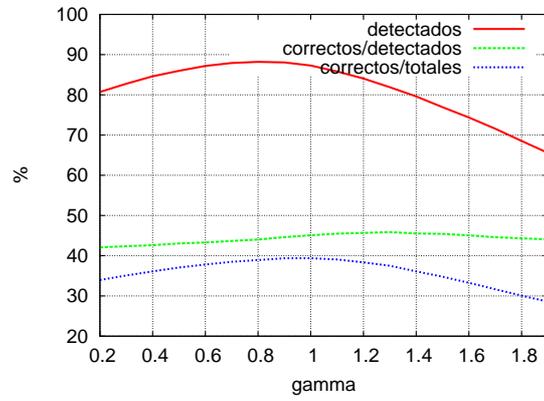
REFERENCIAS

- Abramowitz M. and Stegun I.A. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York, 1964.
- Babaud J., Witkin A.P., Baudin M., and Duda R.O. Uniqueness of the gaussian kernel for scale-space filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:26–33, 1986.
- Brady M. and Horn B.K.P. Rotationally symmetric operators for surface interpolation. *Computer Vision, Graphics, and Image Processing*, 22(1):70–94, 1983.
- Burt P. Fast filter transforms for image processing. *CGIP*, 16(1):20–51, 1981.
- Crowley J.L. A representation for visual information. Technical Report CMU-RI-TR-82-07, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, 1981.
- Finkel R.A. and Bentley J.L. Quad trees: A data structure for retrieval on composite keys. *Acta Inf.*, 4:1–9, 1974.
- Florack L.M.J., Haar Romeny B.M.t., Koenderink J.J., and Viergever M.A. Linear scale-space. *Journal of Mathematical Imaging and Vision*, 4(4):325–351, 1994.

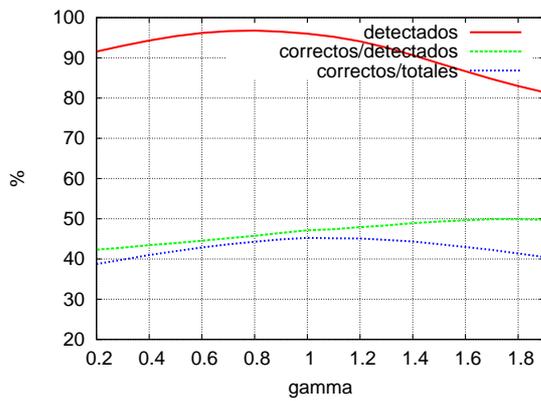
- Harris C. and Stephens M. A combined corner and edge detection. In *Proceedings of The Fourth Alvey Vision Conference*, pages 147–151. 1988.
- Koenderink J.J. The structure of images. *Biological Cybernetics*, 50:363–370, 1984.
- Lindeberg T. Scale-space for discrete signals. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-12(3):234–254, 1990.
- Lindeberg T. *Scale-Space Theory in Computer Vision*. Kluwer Academic Publishers, Norwell, MA, USA, 1994. ISBN 0792394186.
- Lindeberg T. Edge detection and ridge detection with automatic scale selection. *International Journal of Computer Vision*, 30(2):77–116, 1996.
- Lindeberg T. Feature detection with automatic scale selection. *International Journal of Computer Vision*, 30(2):77–116, 1998.
- Lindeberg T. and Bretzner L. Real-time scale selection in hybrid multi-scale representations. In *Proc. Scale-Space'03*, volume 2695 of *Springer Lecture Notes in Computer Science*, pages 148–163. 2003.
- Lowe D.G. Object recognition from local scale-invariant features. In *Proc. of the International Conference on Computer Vision ICCV, Corfu*, pages 1150–1157. 1999.
- Mikolajczyk K. *Detection of local features invariant to affines transformations*. Ph.D. thesis, INPG, Grenoble, 2002.
- Witkin A.P. Scale-space filtering. In *International Joint Conference on Artificial Intelligence*, pages 1019–1022. Karlsruhe, West Germany, 1983.
- Yuille A.L. and Poggio T.A. Scaling theorems for zero crossings. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8(1):15–25, 1986. ISSN 0162-8828.



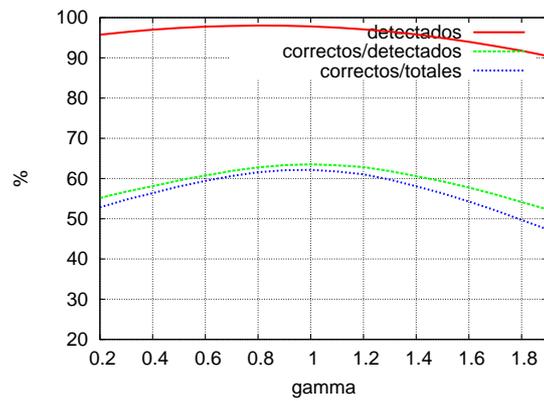
(a) Laplaciano



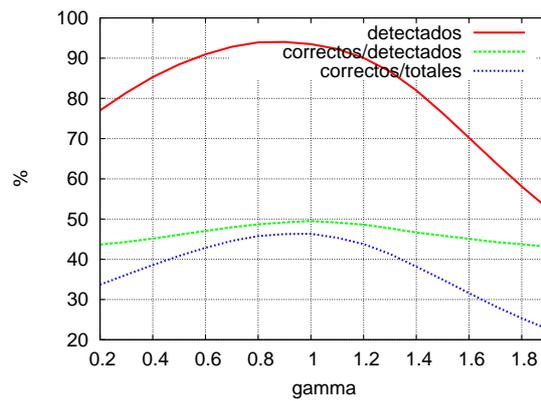
(b) Magnitud del Gradiente



(c) DoG

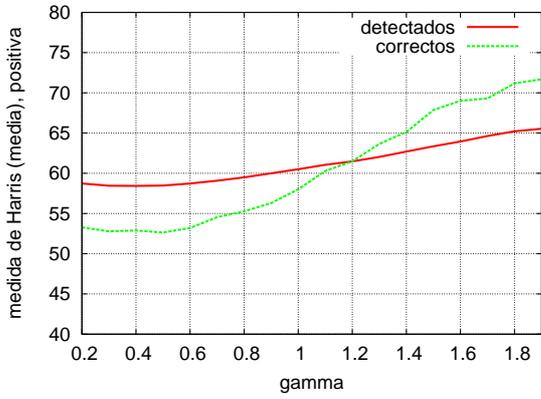


(d) Determinante del Hessiano

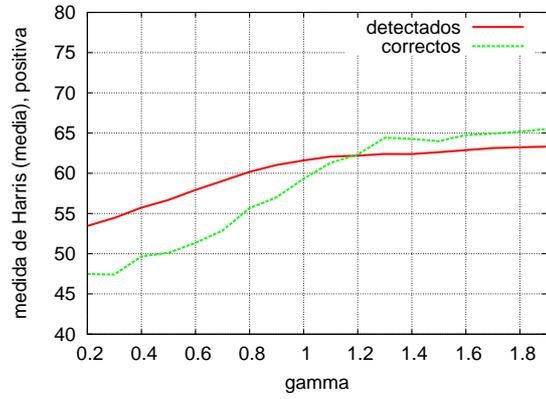


(e) Variación cuadrática

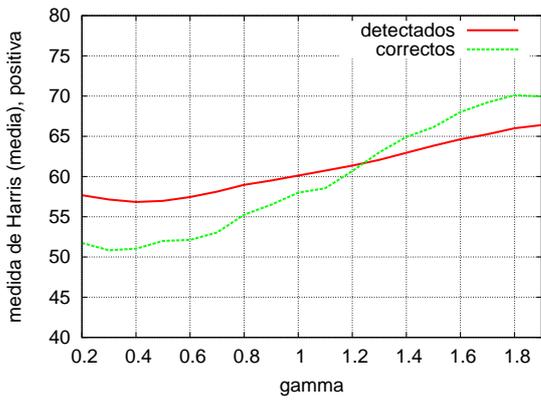
Figura 4: Variación respecto a γ en el porcentaje de puntos con escalas detectadas (rojo), correctas/detectadas (verde), correctas/totales (azul), para un cambio de escala efectivo de 2,48.



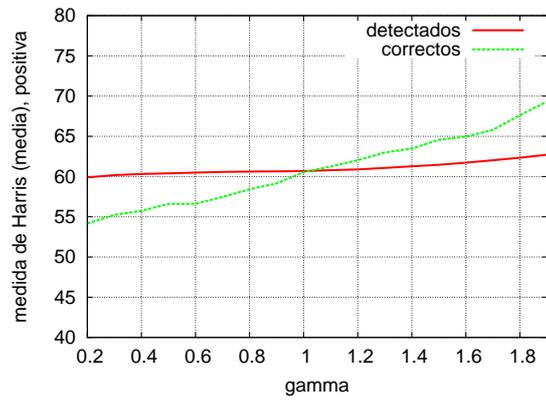
(a) Laplaciano



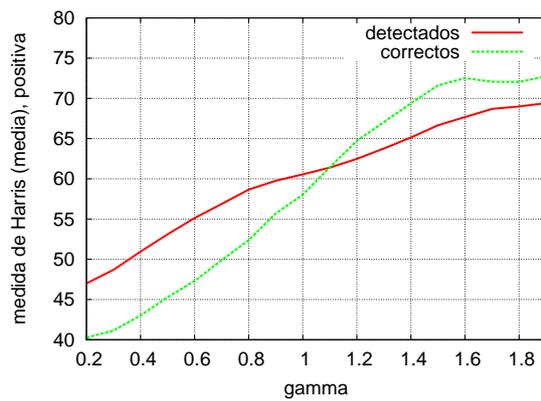
(b) Magnitud del Gradiente



(c) DoG

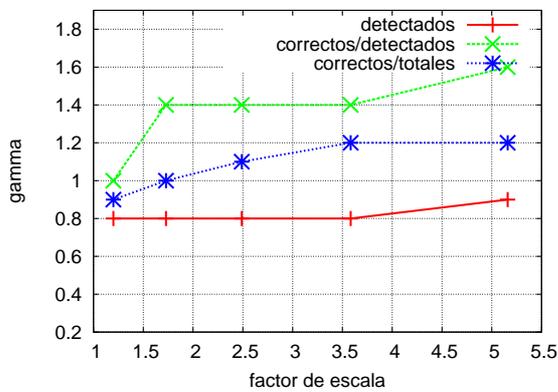


(d) Determinante del Hessiano

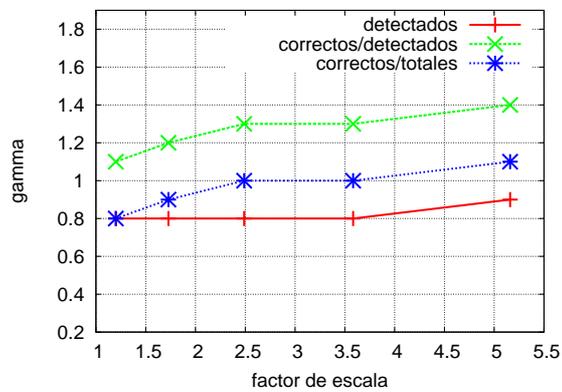


(e) Variación cuadrática

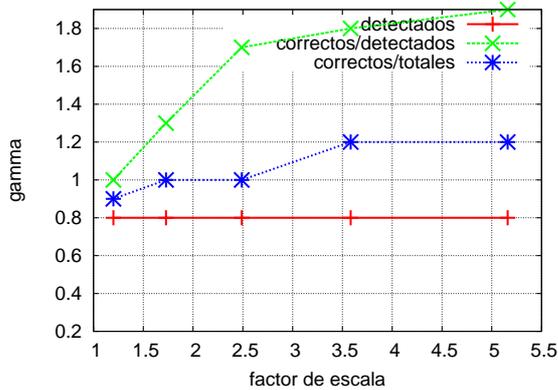
Figura 5: Valor medio de la medida de Harris (positiva) para puntos detectados (rojo) y correctamente detectados (verde), para un cambio de escala de 2,48.



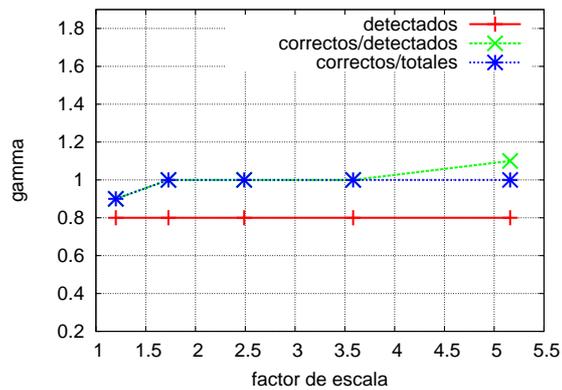
(a) Laplaciano



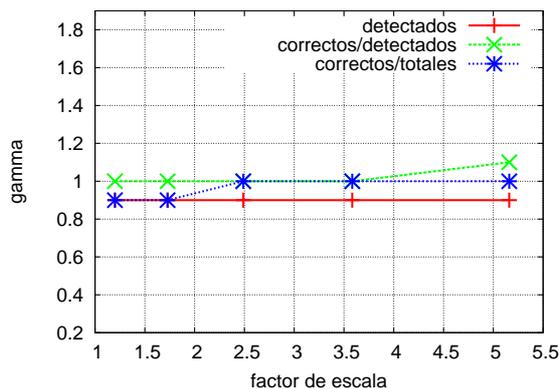
(b) Magnitud del Gradiente



(c) DoG



(d) Determinante del Hessiano



(e) Variación cuadrática

Figura 6: Valores de γ correspondientes a los máximos detectados en los porcentajes de puntos con escalas detectadas (rojo), porcentajes de correctas/detectadas (verde) y porcentajes de correctas/totales (azul), en función del factor del escala