PRÁCTICA III: Filtrado frecuencial multirresolución Gaussianos (estudantes de 2ª matrícula)

Xosé R. Fdez-Vidal
xose.vidal@usc.es

EPSE de Lugo — 16 de outubro do 2023

1. Introdución

A información da imaxe prodúcese en moitas escalas espaciais diferentes. Para analizar e manipular esta información nun rango de escalas espaciais as pirámides/bancos de filtros son unha ferramenta de representación axeitada.

Aquí abordaremos dúas opcións diferentes en progresión de complexidade. A primeira é un banco de filtros gaussiano paso-baixa creados en frecuencia, cuxo resultado é a creación de versións da imaxe de entrada en múltiples resolucións. Isto é útil para a análise en diferentes escalas espaciais, pero non separar a imaxe en diferentes bandas de frecuencia. A segunda, un banco de filtros log-gaussiana que nos proporcionará pirámide un nivel adicional de análise, dividindo a imaxe en diferentes bandas de frecuencia espaciais isotrópicas.

As aplicacións das técnicas multirresolución son moi variadas pero neste caso imos poñer un exemplo ilustrativo. Supoñamos que queremos detectar os paxaros na imaxe da Fig. 1 usando o enfoque de correlación normalizado (template matching) descrito na teoría. Se temos un modelo dun paxaro, a correlación normalizada poderá detectar só as aves que teñan un tamaño na imaxe similar ao do modelo. Para introducir a invarianza de escala, unha posible solución é cambiar o tamaño do modelo para cubrir unha ampla gama de tamaños posibles e aplicalos á imaxe. Entón, o conxunto de modelos poderá detectar aves de diferentes tamaños. A desvantaxe deste enfoque é que será computacionalmente custoso xa que precisará detectar grandes aves calculando circunvolucións con núcleos grandes que é moi lento. Outra alternativa é cambiar o tamaño da imaxe como se mostra na Fig. 1 resultando nunha pirámide de imaxes multiescala.

Neste exemplo, a imaxe orixinal ten unha resolución de 848×643 píxeles. Obtense cada imaxe da pirámide reducindo a imaxe desde o nivel anterior reducindo o número de píxeles por factor do 25 %. Esta operación denomínase submostraxe. Agora podemos usar a pirámide para detectar aves de diferentes tamaños usando unha única plantilla. A caixa vermella da figura indica o tamaño do modelo utilizado. A figura mostra como as aves de diferentes tamaños fanse detectables, polo menos, nun dos niveis da pirámide. Este método será máis eficiente xa que o modelo se pode manter pequeno e as circunvolucións permanecerán computacionalmente eficientes.

Outra aplicación divertida é a mestura de imaxes (image blending). O obxectivo é combinar dúas imaxes nunha soa como se amosa na Fig.2. A máscara (función escalón branco-negro) úsase para definir como se combinarán as imaxes.

Facer unha transición brusca dunha imaxe a outra dá unha imaxe cun límite artificioso (ver o bordo recto da mazá/laranxa). Se usamos unha pirámide laplaciana, podemos pasar dunha imaxe a outra a través de moitas diferentes escalas espaciais para facer unha transición gradual entre as dúas imaxes. Primeiro, construímos a pirámide laplaciana para as dúas imaxes de entrada, neste exemplo usamos 7 niveis e tamén mantemos o último residuo paso-baixo:

e a pirámide gaussiana da máscara como se mostra a continuación (ten en conta que usamos 8 niveis, un nivel máis que para a pirámide laplaciana):

Agora combinamos os niveis das tres pirámides (imaxe1, imaxe2 e máscara) como $l_k = l_k^A \times m_k + l_i^B \times (1 - m_k)$ e facemos o mesmos co residuo paso-baixo $g_k = g_k^A \times m_k + g_i^B \times (1 - m_k)$. Cos novos valores de l_k e g_k reconstruímos de novo a sinal. O resultado amosase na seguinte Fig. 5.

Para recapitular brevemente: a transformada de Fourier revela o contido da frecuencia espacial da imaxe perfectamente, pero sofre de non ter localización espacial. Unha pirámide gaussiana proporciona

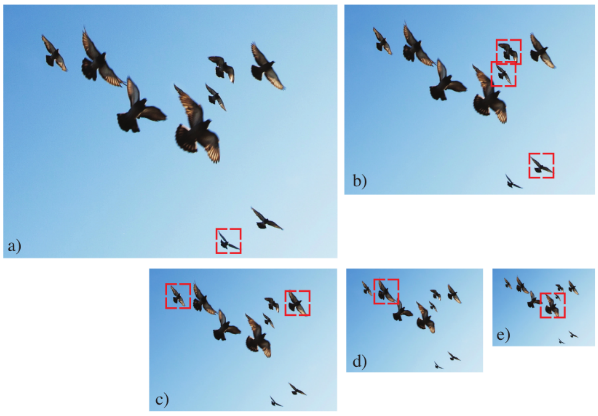

Figura 1: Descomposición multirresolución gaussiana dunha imaxe. Cada imaxe é un 25 % máis pequena que a anterior (ou a gaussina é un 25 % máis ancha). O cadro vermello indica o tamaño dun patrón utilizado para detectar aves voando. Como o tamaño da plantilla está fixado, só poderá detectar as aves que encaixan ben dentro da caixa. As aves que son máis pequenas ou máis grandes non se detectarán nunha única escala. Ao executar o mesmo modelo en moitos niveis descomposición, detéctanse diferentes instancias de aves a diferentes escalas.

Figura 2: Fusión de dúas imaxes empregando a máscara binaria indicada na figura.

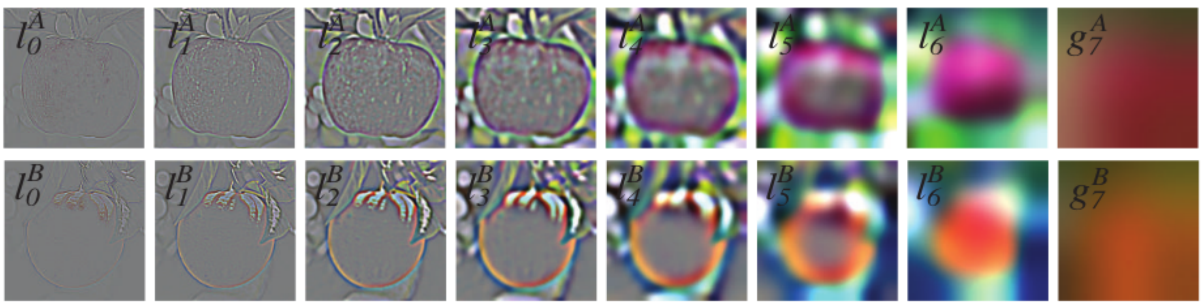

Figura 3: Pirámides laplacianas das dúas imaxes a fusionar.

unha representación multiescala da imaxe, útil para aplicar un algoritmo a escala fixa a unha imaxe nun rango de escalas espaciais. Pero non rompe a imaxe en bandas de frecuencias senón que é simplemente unha serie de versións filtradas paso baixo. Esta é representación sobre-completa é dicir, hai máis píxeles na representación da pirámide gaussiana dunha imaxe que os que hai na propia imaxe.

Outras representacións capturan en cada nivel a información dunha banda de frecuencias (Laplaciana, Gaussiana paso-banda) o que permite revelar estruturas da imaxe para un conxunto de frecuencias. Igual

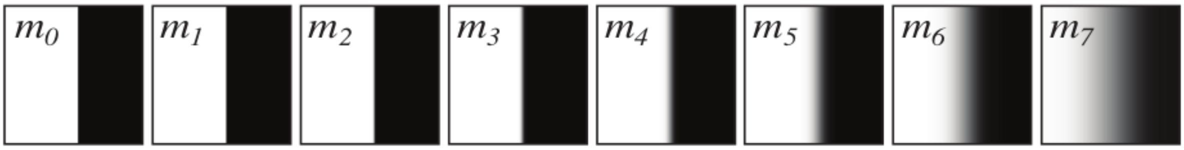

Figura 4: Pirámide de gaussianas sobre a máscara empregada para a fusión das imaxes.

Figura 5: Fusión da imaxe empregando unha pirámide Laplaciana.

que a pirámide de Gaussiana, soen ser sobre-completas.

2. Filtro gaussiano no dominio espacial e frecuencial

Un filtro gaussiano discreto bidimensional (con $\mu = 0$) defínese no espazo do seguinte xeito:

$$G_\sigma(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-(\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2})\right), \tag{1}$$

onde σ_x e σ_y denotan as desviacións típicas na dimensión x e y, respectivamente. Se as sigmas σ_x e σ_y son iguais, a gaussiana é isotrópica como vemos na Fig. 6.

A transformada de Fourier dunha gaussiana pode ser achada analiticamente e obtemos a seguinte expresión:

$$G_\sigma(u, v) = \exp\left(-2\pi^2(u^2\sigma_u^2 + v^2\sigma_v^2)\right), \tag{2}$$

3. Filtro log-Gauss paso-banda no dominio frecuencial

Para construír un banco de filtros gaussiano paso-banda 2D podemos facelo mediante unha función log-Gauss, que é radialmente simétrica sobre a orixe (par) tanto no dominio da frecuencia como no dominio da espacial. A función de transferencia dun filtro logGauss 2D radial ven dada por:

$$G(f) = \exp\left(\frac{-(\log(f/f_o))}{2(\log(\sigma_f/f_o))^2}\right). \tag{3}$$

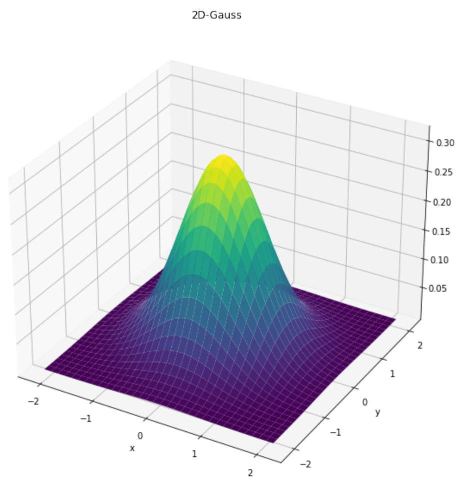

Figura 6: Función gaussiana isotrópica no dominio espacial.

onde f_0 e σ_f son os parámetros do filtro. O valor de f_0 dará a frecuencia central do filtro e o de σ_f afecta o seu ancho de banda. É útil manter a mesma forma mentres se varía o parámetro de frecuencia. Para iso, a razón σ_f/f_o debe permanecer constante. O ancho de banda na frecuencia vén dado por:

$$B = 2\sqrt{2/\log(2)}\,(\|\log(\sigma_f/f_0)\|)\,. \tag{4}$$

Ten en conta que o ancho de banda resultante está en unidades de oitavas.

A posición central dos filtros do banco (conxunto de $\{f_{o_{s_i}}\}$) pódese poñer en función da escala do filtro, da distancia que queremos entre filtros consecutivos (fixámola nós a $d = 2$ que equivale a 1 oitava) e da frecuencia máis alta elixida para o primeiro filtro do banco de filtros (f_{Max}). Polo tanto, se collemos a frecuencia máxima do banco de filtros como $f_{Max} = 1/8$ que está próxima a frecuencia de Nyquist ($f_{Max} = 1/2$) pero non no borde para evitar un truncamento excesivo do filtro, a relación buscada é a seguinte para un conxunto de escalas $s = \{0, 1, 2, \dots\}$

$$f_{o_{s_i}} = \frac{f_{Max}}{d^s}. \tag{5}$$

Con estas relacións, podes contruír facilmente un banco de filtros paso-banda log_gaussiano en función dos parámetros de deseño que fixe o usuario: nº de escalas, ancho de banda e frecuencia máxima do filtro próximo a frecuencia de Nyquist (filtro coa primeira escala).

4. Tarefa a realizar co banco de filtros Gaussiano paso-baixa

Tarefa 1 : Actividades cos filtros Gaussiano paso-baixa

As tarefas que os estudantes teñen que levar a cabo na parte de filtros de Gaussianos paso-baixa, son as seguintes (empregar cadernos Jupyter para introducir os comentarios, explicacións e visualizar os resultados).

(a) Estuda o código aportado (LogGabor2D.py) aportado na práctica II que implementa un filtrado mediante un banco de log-Gabors 2D ata que teñas claro como funciona. É moi importante posto que cando teñas que implementar os filtros gaussinas (paso-baixa ou paso banda) nos vai a servir como base. O profesor resolverá todas as dúbidas necesarias.

(b) Implementa un banco de filtros gaussianos paso-baixa. O banco de filtros debe construirase tendo en conta parámetros fixados polo usuario: número de escalas, factor de forma, ancho de banda. Debe ter métodos para poder visualizar graficamente a cobertura espectral do banco, a forma espacial dos filtros, o resultado das convolucións da imaxe co banco de filtros en 2D.

(c) Un método que me permita reconstruír a imaxe orixinal a partir das respostas dos filtros.

(d) Implementa o mesmo banco pero agora para videos (3D) e repite o mesmo que para 2D.

5. Tarefa a realizar co banco de filtros Gaussiano paso-banda

Neste apartado os estudantes teñen que implementar dende cero un banco de filtros gaussiano paso-banda multiescala (a especificar polo usuario) para imaxes 2D e tamén estendelo a 3D para vídeo.

Tarefa 2 : Actividades cos filtros Gaussianos paso-banda

As tarefas que os estudantes teñen que levar a cabo na parte de filtros de log-Gauss, son as seguintes (en cadernos Jupyter).

(a) Estuda o código aportado (LogGabor2D.py) na parte radial e a teoría previa deste documento para ter claro o que se pretende. O profesor resolverá todas as dúbidas necesarias.

(b) Implementar unha clase que constrúa un banco de filtros log-Gauss paso-banda (segundo parámetros aportados polo usuario) para posteriormente aplicalo a distintas imaxes. Debe ter métodos para poder visualizar graficamente a cobertura espectral do banco, a forma espacial dos filtros, o resultado das convolucións da imaxe co banco de filtros en 2D.

(c) Un método que me permita reconstruír a imaxe orixinal a partir das respostas dos filtros .

(d) Proba a facer fusión de imaxes con este banco de filtros como fixemos na teoría empregando a Laplaciana e a ver cal é o resultados acadado.

(e) Implementa o mesmo banco pero agora para videos (3D) e repite o mesmo que para 2D.

Referencias

[1] J. Smith, *Mathematics of discrete Fourier transform (DFT) with audio applications*.

[2] C. P. Bridge, “An introduction to the monogenic signal,” *Draft*, pp. 1–21, 2016.