

Puntos de interés

Lección 04.1

Dr. Pablo Alvarado Moya

MP6127 Visión por Computadora
Programa de Maestría en Electrónica
Énfasis en Procesamiento Digital de Señales
Escuela de Ingeniería Electrónica
Tecnológico de Costa Rica

I Cuatrimestre 2013

- Varias áreas de la Visión por Computador requieren detección de “**puntos de referencia**” o “**puntos de interés**”
- Preámbulo a búsqueda de correspondencias
- Uso en reconocimiento de objetos, panoramas, reconstrucción. . .
- Los puntos de interés deben ser invariantes ante variaciones como
 - Rotación
 - Escala
 - Traslación
 - Cambio de perspectiva
 - Cambio de iluminación
 - Deterioro de objetos
 - Envejecimiento de personas
 - . . .

Dos métodos principales para encontrar correspondencias

- 1 Rastreo de puntos en secuencias de imágenes (flujo óptico)
- 2 Búsqueda independiente con posterior emparejamiento (*match*)

Pasos con métodos de detección y emparejamiento

Tres pasos:

- 1 Detección de puntos de interés estables
- 2 Generación de descriptores invariantes
- 3 Emparejamiento de descriptores

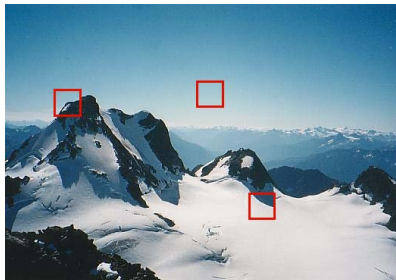
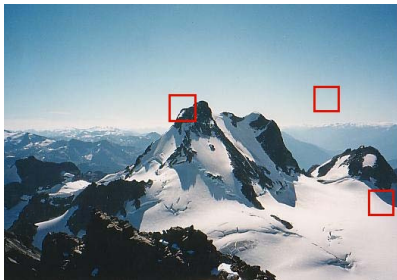
Últimos pasos pueden reemplazarse por

- rastreo de los puntos de interés.

Puntos de interés

¿Qué propiedades son deseables?

(1)



Szelisky, 2011

¿Qué propiedades son deseables?

Obsérvese que:

- Si no hay bordes, incertidumbre es total
- Si hay un borde, incertidumbre es en dirección del borde
- Si hay una esquina, incertidumbre es baja

Necesaria estructura que describa distribución de direcciones de bordes en una región

Recordando el gradiente

- En PDI gradiente se usa como detector de **bordicidad**:

$$\nabla I(x, y) = \begin{bmatrix} \frac{\partial}{\partial x} I(x, y) \\ \frac{\partial}{\partial y} I(x, y) \end{bmatrix} = \begin{bmatrix} I_x(x, y) \\ I_y(x, y) \end{bmatrix}$$



$I(x, y)$



$|\text{Sobel}[I(x, y)]| \approx |\nabla I(x, y)|$

- Aproximaciones: variaciones de diferencias finitas centradas

$$\frac{\partial}{\partial x} I(x, y) \approx \frac{I(x+1, y) - I(x-1, y)}{2}$$

- Varios tipos: Sobel, Roberts, Prewitt, Ando, DoG, etc.

- Idea: realizar en una pequeña región análisis de componentes principales de los gradientes
- Datos son:

$$\mathbf{X} = \begin{bmatrix} l_x(x - \Delta x_1, y - \Delta y_1) & l_y(x - \Delta x_1, y - \Delta y_1) \\ l_x(x - \Delta x_2, y - \Delta y_2) & l_y(x - \Delta x_2, y - \Delta y_2) \\ l_x(x - \Delta x_3, y - \Delta y_3) & l_y(x - \Delta x_3, y - \Delta y_3) \\ \vdots & \vdots \\ l_x(x - \Delta x_n, y - \Delta y_n) & l_y(x - \Delta x_n, y - \Delta y_n) \end{bmatrix}^T$$

donde los desplazamientos $(\Delta x_i, \Delta y_i)$ recorren con $i = 1 \dots n$ la región alrededor de (x, y) .

- Covarianza dada por

$$\Sigma_{\mathbf{x}}(x, y) = \frac{1}{n} \mathbf{xx}^T = \begin{bmatrix} \langle I_x^2(x, y) \rangle & \langle I_x(x, y) I_y(x, y) \rangle \\ \langle I_x(x, y) I_y(x, y) \rangle & \langle I_y^2(x, y) \rangle \end{bmatrix}$$

donde

$$\langle F(x, y) \rangle = \frac{1}{n} \sum_{i=1}^n F(x + \Delta x_i, y + \Delta y_i)$$

- $\Sigma_{\mathbf{x}}(x, y)$ es el **tensor de estructura**
- La covarianza contiene información de la distribución de los gradientes en la ventana.

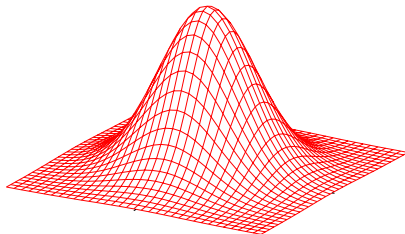
- Sean λ_1 y λ_2 los eigenvalores de $\Sigma_{\mathbf{x}}(x, y)$
- Sean $\underline{\mathbf{e}}_1$ y $\underline{\mathbf{e}}_2$ los eigenvectores correspondientes
- Asíumase $\lambda_1 \geq \lambda_2$
- $\underline{\mathbf{e}}_1$ es la dirección dominante
- λ_i indica la varianza de los gradientes proyectados a $\underline{\mathbf{e}}_i$.
- Si $I(x, y) = \text{cte} \Rightarrow I_x(x, y) = I_y(x, y) = 0 \wedge \lambda_1 = \lambda_2 = 0$
- Si hay borde lineal, **todos** los gradientes o son cero (fuera del borde) o están alineados en la dirección $\underline{\mathbf{e}}_1$, lo que implica $\lambda_2 = 0$
- Si hay esquina, hay varias direcciones y $\lambda_1 \neq 0 \wedge \lambda_2 \neq 0$

Los eigenvalores indican qué tan confiable es (x, y) como punto de interés

- Matriz $\Sigma_{\mathbf{x}}(x, y)$ es siempre 2×2 y simétrica
- Existe forma cerrada de calcular λ_1 y λ_2
- Algunos métodos usan $\underline{\mathbf{e}}_1$, otros no.

Invarianza a la rotación

- Ventana cuadrada: si se rota imagen, cambia $\Sigma_{\mathbf{x}}(x, y)$
- Solución, redefinir $\langle F(x, y) \rangle$
- Usual: usar una ventana gaussiana



- Puntos al centro tienen más peso que en borde

- Shi y Tomasi (1994): ¿Qué tan lejos está λ_2 de cero?
Máximos de $\sqrt{\lambda_2}$ indican menor incertidumbre
- Harris y Stephens (1988), con $\alpha = 0,06$

$$\det(\Sigma_{\mathbf{x}}(x, y)) - \alpha \operatorname{tr}^2(\Sigma_{\mathbf{x}}(x, y)) = \\ \langle I_x^2 \rangle \langle I_y^2 \rangle - \langle I_x I_y \rangle^2 - \alpha (\langle I_x^2 \rangle + \langle I_y^2 \rangle)^2 = \lambda_1 \lambda_2 - \alpha (\lambda_0 + \lambda_1)^2$$

que no requiere resolver el eigensistema ni raíces cuadradas.

- Brown, Szeliski y Winder (2005):

$$\frac{\det \Sigma_{\mathbf{x}}(x, y)}{\operatorname{tr} \Sigma_{\mathbf{x}}(x, y)} = \frac{\lambda_0 \lambda_1}{\lambda_0 + \lambda_1}$$

- **Matriz Hessiana:** alternativa a $\Sigma_{\mathbf{x}}(x, y)$

$$H(x, y; \sigma) = \begin{bmatrix} L_{xx}(x, y; \sigma) & L_{xy}(x, y; \sigma) \\ L_{xy}(x, y; \sigma) & L_{yy}(x, y; \sigma) \end{bmatrix}$$

con $L_{\nu\tau}(x, y; \sigma) = G_{\nu\tau}(x, y; \sigma) * I(x, y)$ y

$$G_{\nu\tau}(x, y; \sigma) = \frac{\partial^2}{\partial \nu \partial \tau} G(x, y; \sigma)$$

y

$$G(x, y; \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/\sigma^2}$$

- σ determina tamaño de ventana
- Grado de interés se mide con $\det[H(x, y; \sigma)]$

Atributos de un punto de interés

- Lo anterior permite encontrar posición de puntos de interés
- Invarianza a escala: requiere además radio
- Invarianza a rotación: requiere orientación
- Invarianza afín: requiere aspecto

- Mikolajczyk (2002) demostró empíricamente que método más estable busca máximos y mínimos del laplaciano de gaussianos normalizado en escala:

$$\begin{aligned}\sigma^2 \nabla^2 G(x, y; \sigma) &= \sigma^2 \nabla \cdot \nabla G(x, y; \sigma) \\ &= \sigma^2 \left[\frac{\partial^2}{\partial x^2} G(x, y; \sigma) + \frac{\partial^2}{\partial y^2} G(x, y; \sigma) \right] \\ \sigma^2 \nabla^2 G(x, y; \sigma) * I(x, y) &= \sigma^2 \operatorname{tr} [H(x, y; \sigma)]\end{aligned}$$

¿Cómo alcanzar la invarianza en escala?

- Si imagen se escala pero ventana es constante $\Rightarrow \Sigma_{\mathbf{x}}(x, y)$ o $H(x, y; \sigma)$ varían
- La idea básica: detectar máximos de mediciones de interés no solo en $I(x, y)$, sino en $L(x, y; \sigma)$ donde

$$L(x, y; \sigma) = G(x, y; \sigma) * I(x, y)$$

¿Cómo alcanzar la invarianza en escala?

- Si imagen se escala pero ventana es constante $\Rightarrow \Sigma \mathbf{x}(x, y)$ o $H(x, y; \sigma)$ varían
- La idea básica: detectar máximos de mediciones de interés no solo en $I(x, y)$, sino en $L(x, y; \sigma)$ donde

$$L(x, y; \sigma) = G(x, y; \sigma) * I(x, y)$$



Carlos Moya, 2004

¿Cómo alcanzar la invarianza en escala?

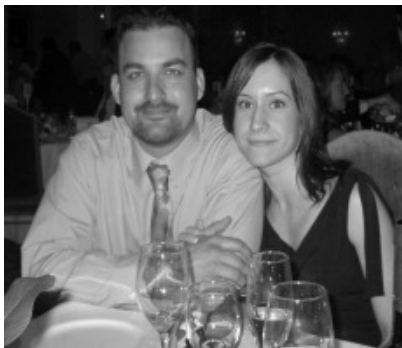
- Si imagen se escala pero ventana es constante $\Rightarrow \Sigma \mathbf{x}(x, y)$ o $H(x, y; \sigma)$ varían
- La idea básica: detectar máximos de mediciones de interés no solo en $I(x, y)$, sino en $L(x, y; \sigma)$ donde

$$L(x, y; \sigma) = G(x, y; \sigma) * I(x, y)$$



Carlos Moya, 2004

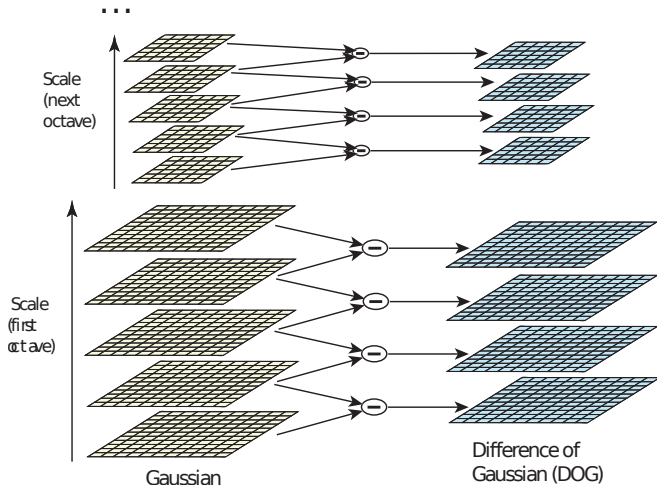
Espacio de escalas



- Método aproxima laplaciano de gaussianos normalizado en escala con diferencias de gaussianas

$$\begin{aligned} D(x, y; \sigma) &= (G(x, y; k\sigma) - G(x, y; \sigma)) * I(x, y) \\ &= L(x, y; k\sigma) - L(x, y; \sigma) \end{aligned}$$

Espacio de escalas de Lowe y Brown



Lowe, 2004

- La escala la determina el σ del nivel con el máximo
- La orientación se puede determinar con eigenvectores
- Lowe utiliza histograma de direcciones y selecciona máximo con interpolación
- Lowe también replica puntos de interés si hay varios máximos (el segundo mayor a 80 % del máximo)

- Transformaciones proyectivas se aproximan con afines
- Necesaria cuando hay cambios de perspectiva en 3D
- Opción: usar ambos eigenvectores del tensor de estructura o de matriz Hessiana
- MSER

- Imagen: I mapeo $I : D \in \mathbb{N}^2 \rightarrow S$
- S es ordenado, y existe predicado de adyacencia A (4-vecinos, p.ej.)
- Región: $Q \subset D$
- Borde externo de region ∂Q

$$\partial Q = \{q \in D \setminus Q : \exists p \in Q : A(p, q)\}$$

- Región extrema: $Q \in D$ region que para todo $p \in Q$:
 - $q \in \partial Q : I(p) > I(q)$ (región de intensidad máxima), o
 - $q \in \partial Q : I(p) < I(q)$ (región de intensidad mínima)

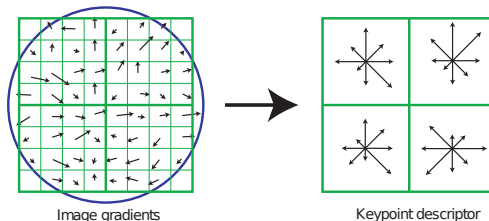
- Región extrema de máxima estabilidad (MSER):
 - Sea $Q_1, \dots, Q_{i-1}, Q_i \dots$ una sucesión de regiones extremas anidadas ($Q_i \subset Q_{i+1}$).
 - Q_{i^*} es de máxima estabilidad sii $q(i) = |Q_{i+\Delta} \setminus Q_{i-\Delta}| / |Q_i|$ tiene un mínimo local en i^* .
 - $|\cdot|$ denota cardinalidad.
 - $\Delta \in S$ es un parámetro del método.
- Existen algoritmos muy eficientes para su cálculo (quasi $\mathcal{O}(n)$), similares al algoritmo de divisorias (*watersheds*)



Matas et al., 2004

- ¿Cómo describir las regiones alrededor del punto de interés?
- Imagen: sensible a iluminación: intensidad, sombras, distorsiones perspectivas, etc.
- Bordes deseables
- Muchas propuestas: MOPS, SIFT, PCA-SIFT, SURF, GLOH, etc.

- SIFT: Scale Invariant Feature Transform
- Proposed by Lowe, 1999 (U British Columbia)



- Histogramas de orientación, con entradas proporcionales a $|\nabla I|$
- Ejemplo: descriptor 2×2 a partir de 8×8 muestras
- Lowe: descriptor 4×4 a partir de 16×16 muestras con 8 bins por histograma (128 dimensiones)

- SURF: Speeded Up Robust Features
- Propuesto por Bay, Tuytelaars, Van Gool, 2006 (ETHZ, KU Leuven)
- Interés principal: aceleración del cálculo
- Descriptor se basa en wavelets de Haar
- Se describen subregiones con sumas de componentes de la transformada de Haar en las direcciones x y y:

$$\underline{\mathbf{v}} = (\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|)$$

Búsqueda de correspondencias

- Objetivo: encontrar descriptores más cercanos en otra imagen
- Suposición: distancia entre descriptores inversamente proporcional a similitud
- Métodos: dependen de la aplicación
- Panoramas: buscar en regiones cercanas (se supone traslape)
- Reconocimiento: estructuras de datos (*kd-tree*)

- 1 Revisar artículo de Lowe sobre SIFT
- 2 Trabajar en el proyecto

Este documento ha sido elaborado con software libre incluyendo L^AT_EX, Beamer, GNUPlot, GNU/Octave, XFig, Inkscape, LTI-Lib-2, GNU-Make, Kazam, Xournal y Subversion en GNU/Linux



Este trabajo se encuentra bajo una Licencia Creative Commons Atribución-NoComercial-LicenciarIgual 3.0 Unported. Para ver una copia de esta Licencia, visite <http://creativecommons.org/licenses/by-nc-sa/3.0/> o envíe una carta a Creative Commons, 444 Castro Street, Suite 900, Mountain View, California, 94041, USA.

© 2013 Pablo Alvarado-Moya Escuela de Ingeniería Electrónica Instituto Tecnológico de Costa Rica