

# Reconocimiento de instancias y categorías

## Lección 12.1

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

# Contenido

- 1 Introducción
- 2 Reconocimiento de instancias
  - Entrenamiento
  - Reconocimiento
- 3 Reconocimiento de categorías

# Tipos de reconocimiento

- En la literatura se distingue entre reconocimiento de *instancias* y de *categorías*
- En el primero se buscan etiquetas de objetos entrenados
- En el segundo, métodos infieren la categoría o clase de objetos

# Reconocimiento de categorías

# Reconocimiento de instancias

- Técnicas recientes usan descriptores 2D locales
- Dos fases:
  - 1 entrenamiento
  - 2 reconocimiento



Lowe, 2004

# Fase de entrenamiento

Entrenamiento:

- 1 Detección de puntos de interés en imágenes de entrenamiento
- 2 Normalización de regiones (rotación, transformación afín, etc.)
- 3 Extracción de descriptores para cada punto de interés
- 4 Almacenamiento en estructuras de rápido acceso de los descriptores, parámetros de puntos de interés, y clase correspondiente

# Fase de reconocimiento

Reconocimiento:

- 1 Detección de puntos de interés
- 2 Normalización de regiones
- 3 Extracción de descriptores
- 4 Votación por clases presentes
- 5 Validación geométrica

# Fase de entrenamiento

Entrenamiento:

- 1 **Detección de puntos de interés**
- 2 Normalización de regiones
- 3 Extracción de descriptores para cada punto de interés
- 4 Almacenamiento en estructuras de rápido acceso de los descriptores, parámetros de puntos de interés, y clase correspondiente



# Detección de puntos de interés

- Se utilizan métodos descritos en lección 4:
  - Matriz hessiana
  - Tensor estructural
  - MSER
  - ...



Gordon y Lowe, 2006

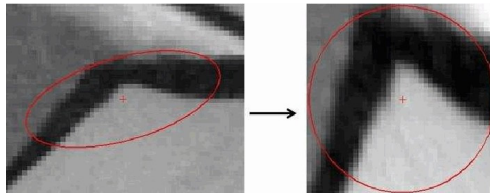
# Fase de entrenamiento

Entrenamiento:

- 1 Detección de puntos de interés
- 2 **Normalización de regiones**
- 3 Extracción de descriptores para cada punto de interés
- 4 Almacenamiento en estructuras de rápido acceso de los descriptores, parámetros de puntos de interés, y clase correspondiente

# Normalización de regiones

- Normalización implica eliminar información irrelevante:
  - Rotación
  - Escalado
  - Distorsión afín
- Usualmente región de interés se deforma a referencia canónica



Fan, Wu y Hu, IEEE PAMI, 34(10), 2012

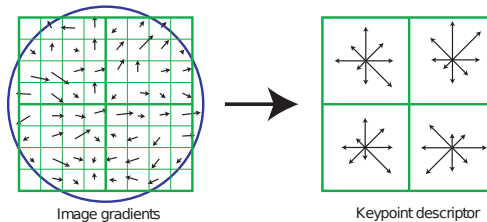
# Fase de entrenamiento

Entrenamiento:

- 1 Detección de puntos de interés
- 2 Normalización de regiones
- 3 **Extracción de descriptores para cada punto de interés**
- 4 Almacenamiento en estructuras de rápido acceso de los descriptores, parámetros de puntos de interés, y clase correspondiente

# Extracción de descriptores

- Mayoría de literatura usa SIFT (Lección 4)
- SURF, MOPS, PCA-SIFT, GLOH, etc.



Lowe, 2004

# Fase de entrenamiento

Entrenamiento:

- 1 Detección de puntos de interés
- 2 Normalización de regiones
- 3 Extracción de descriptores para cada punto de interés
- 4 **Almacenamiento en estructuras de rápido acceso**

# Almacenamiento en estructuras de rápido acceso

- Es posible usar clasificadores neuronales y estadísticos (SVM, MLP, RBF, boosting, random forests, etc.)
- Por cantidades de datos, es mejor usar estructuras de datos para búsqueda
- Se almacena también información de puntos de interés (orientación, posición, escala) y clase
- Hash, *kd-tree*, ...
- Idea: reducir número de comparaciones
- Algoritmos buscan datos almacenados, semejantes a presentados
- Se han adaptando técnicas de búsquedas en texto
- ¿Equivalente de “palabras”? : idea aglomerar descriptores y así discretizar espacio

# Fase de reconocimiento

Reconocimiento:

- ① **Detección de puntos de interés**
- ② **Normalización de regiones**
- ③ **Extracción de descriptores**  
⇒ **Como en el entrenamiento**
- ④ Votación por clases presentes
- ⑤ Validación geométrica



# Fase de reconocimiento

Reconocimiento:

- 1 Detección de puntos de interés
- 2 Normalización de regiones
- 3 Extracción de descriptores
- 4 **Votación por clases presentes**
- 5 Validación geométrica

# Votación

- En estructuras de datos se buscan patrones más cercanos
- Patrones encontrados votan por la clase equivalente
- Se puede votar considerando distancia a descriptores más cercanos
- Se consideran solo clases que superen número mínimo de votos

# Fase de reconocimiento

Reconocimiento:

- 1 Detección de puntos de interés
- 2 Normalización de regiones
- 3 Extracción de descriptores
- 4 Votación por clases presentes
- 5 **Validación geométrica**

# Validación geométrica

- Aquellos objetos detectados se validan geoméricamente
- Con puntos de interés del objeto presentado se selecciona mejor referencia en base de datos
- Se buscan correspondencias de puntos a esa referencia
- Se estima pose
- Estimación basada en transformada de Hough
- Si error estimado de pose supera umbral, se descarta clase.



Lowe, 2004

# Reconocimiento de categorías

# Bolsa de palabras

- Bolsa de palabras (*bag of words*) método más sencillo para reconocer categorías
- Equivalente a método anterior, con técnica de palabras visuales, pero se elimina fase de validación geométrica
- Se basan en histogramas de palabras

# Modelos de partes

- Otra familia de reconocimiento de categorías se basan en reconocimiento de partes
- Cada parte de un objeto es identificada por separado
- Interrelaciones entre partes conocidas a priori
- Se evalúa factibilidad de arreglos geométricos

# Enlaces

- Prof. DiCarlos Lab, MIT. Reconocimiento de objetos (neurociencia+visión por computador)
- Prof. Hoff. SIFT Based Object Recognition



# Resumen

- 1 Introducción
- 2 Reconocimiento de instancias
  - Entrenamiento
  - Reconocimiento
- 3 Reconocimiento de categorías

*Este documento ha sido elaborado con software libre incluyendo L<sup>A</sup>T<sub>E</sub>X, 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