

Object detection and localization using local and global features

K. Murphy, A. Torralba, D. Eaton, and W. Freeman

Victor Hugo Arroyo Dominguez¹

Instituto Nacional de Astrofísica Óptica y Electrónica

1 de junio de 2009

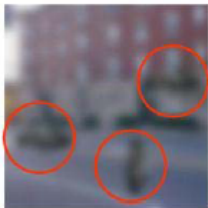
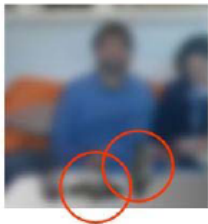
Tabla de Contenido

- 1 Resumen
- 2 Introducción
- 3 Detección de objetos utilizando características locales de la imagen
- 4 Detección de objetos utilizando características globales de la imagen
- 5 Detección de presencia de objetos
- 6 Localización de objetos
- 7 Discusión
- 8 Trabajo a futuro
- 9 Referencias

- Los enfoques tradicionales para detección de objetos solo se basan en piezas pequeñas de la imagen, utilizando una ventana deslizante o regiones alrededor de un punto de interés.
- Sin embargo el utilizar un enfoque local puede ser ambiguo. Esta ambigüedad se puede reducir utilizando características globales de la imagen, al cual se le llama el “*gist*” de la escena. Se muestra que con la combinación de características locales y globales, se obtienen mejores rangos de detección.

- El enfoque más común para la detección/localización genérico de objetos es deslizar una ventana por toda la imagen.
- Una extensión natural de este enfoque se utiliza para detectar partes de un objeto, realizando posteriormente una fusión de las diferentes partes detectadas.
- Otro enfoque es la de extraer puntos de interes de la imagen y clasificar cada región alrededor de dicho punto.

Introducción (2)



Introducción (3)

- Este trabajo es la extensión de un trabajo previo, se realizan las siguientes mejoras:
 - Se utilizan mejores características locales.
 - El conjunto de datos se preparo cuidadosamente.
 - Se corrigieron errores de etiquetado.
 - Se aseguró que los objetos a detectar fueron lo suficientemente grandes para ser detectados
 - Se simplificó el modelo, concentrandose en una sola instancia del objeto.

Introducción (4)

- ¿Que es la detección de la presencia de objetos?
 - Es determinar si una o mas instancias del objeto estan presentes en la imagen.
- ¿Que es la localización de objetos?
 - Es encontrar la localización y la escala de un objeto en la imagen.

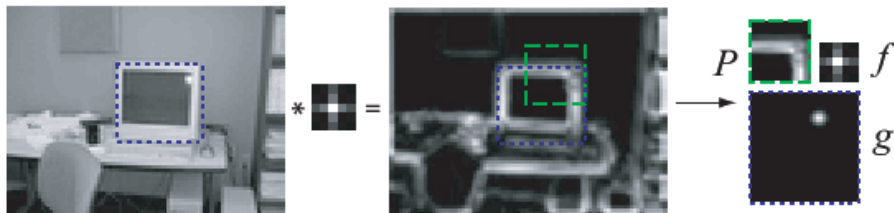
Introducción (5)

	Train +	Train -	Valid +	Valid -	Test +	Test -	Size (hwxw)
Screen	247	421	49	84	199	337	30x30
Keyboard	189	479	37	95	153	384	20x66
CarSide	147	521	29	104	119	417	30x80
Person	102	566	20	113	82	454	60x20

Detección de objetos utilizando características locales de la imagen

- Tipo de características locales de cada ventana.
 - Se trató la imagen con un banco de filtros, los cuales fueron escogidos manualmente.
 - Se obtuvieron fragmentos de la imagen de una de la salida de los filtros.
 - Tamaño y localización de los fragmentos.
 - Se guarda la localización de los objetos para crear una máscara espacial centrada en el objeto.

Detección de objetos utilizando características locales de la imagen (2)



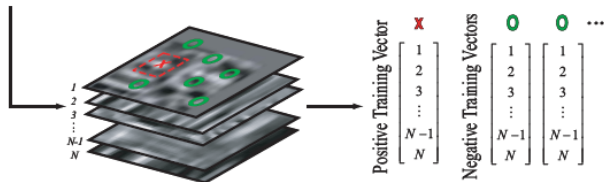
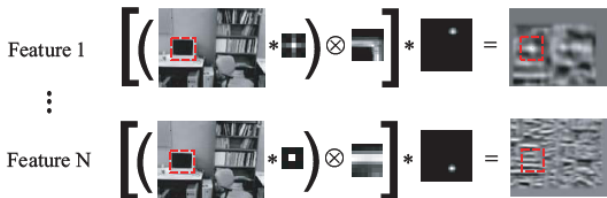
Detección de objetos utilizando características locales de la imagen (3)

- El clasificador que se utilizó fue “boosted decision stumps”, debido a:
 - Su facilidad de implementación.
 - Rapidez para entrenar y para aplicar.
 - Realiza una selección de atributos.

Detección de objetos utilizando características locales de la imagen (3)

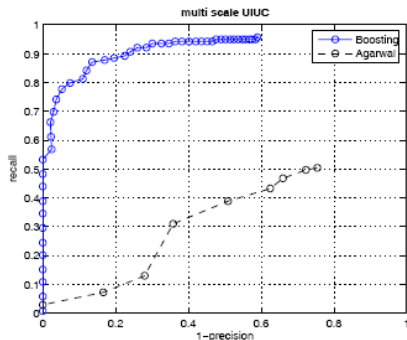
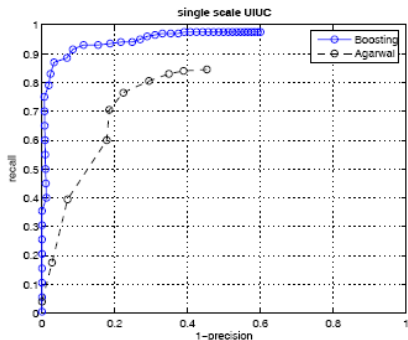
Datos de entrenamiento

$$[(I * f) \otimes P] * g$$



Resultados

Se aplicó el detector a imágenes de carros vistos lateralmente utilizando características locales, en la izquierda a escala simple, a la derecha a multiescala.



Detección de objetos utilizando características globales de la imagen

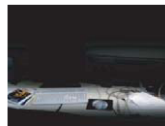
- Se calcula el “gist” de una imagen, realizando lo siguiente:
 - Se calcula una transformación de pirámide dirigida.
 - Utilizando 4 orientaciones y 2 escalas.
 - Se divide la imagen en una rejilla de 4x4.
 - Se calcula el promedio de energía para cada canal en cada celda de la rejilla.
- El “gist” captura de forma burda la textura y la capa espacial de una imagen.

Detección de objetos utilizando características globales de la imagen

El "gist" captura las características de la textura dominante en toda la imagen, además de una capa espacial burda.



Ejemplo de localizacion para pantallas, teclados, carros y gente

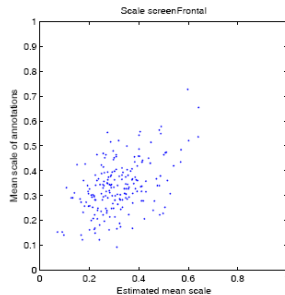
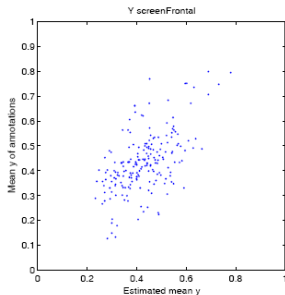


Detección de presencia de objetos

- La detección de la presencia de objetos significa determinar si una o mas instancias de un objeto estan presentes en la imagen.
- Un enfoque exitoso para este problema es el siguiente:
 - Extraer distintas areas alrededor de un punto de interes.
 - Convertirlo a codigos de palabras utilizando un vector de cuantización.
 - Contar el número de ocurrencias de cada posible código de palabra.
 - Clasificar el histograma resultante.

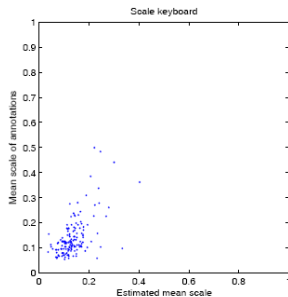
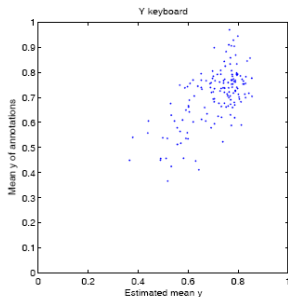
Rendimiento en la localización de pantallas

En la columna de la izquierda se muestra la localización vertical del objeto, en la derecha la escala del objeto. El eje vertical muestra la confianza y el eje horizontal muestra la predicción.



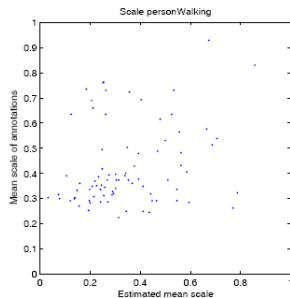
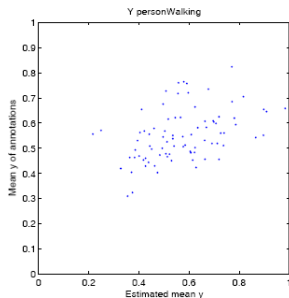
Rendimiento en la localización de teclados

En la columna de la izquierda se muestra la localización vertical del objeto, en la derecha la escala del objeto. El eje vertical muestra la confianza y el eje horizontal muestra la predicción.



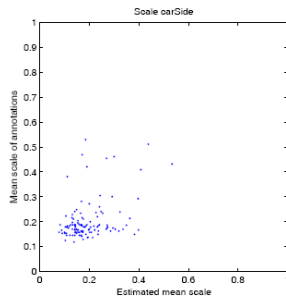
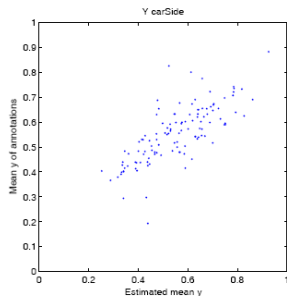
Rendimiento en la localización de peatones

En la columna de la izquierda se muestra la localización vertical del objeto, en la derecha la escala del objeto. El eje vertical muestra la confianza y el eje horizontal muestra la predicción.



Rendimiento en la localización de carros

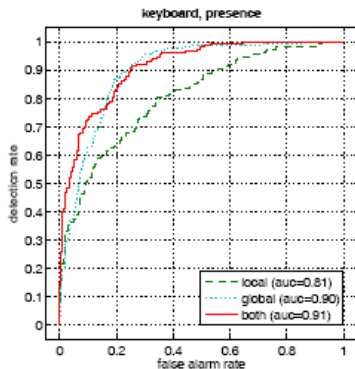
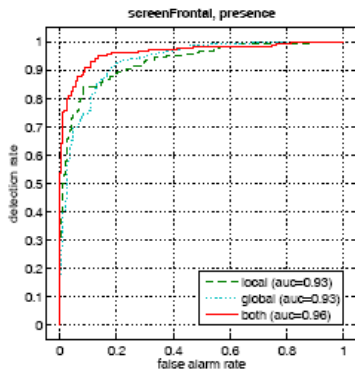
En la columna de la izquierda se muestra la localización vertical del objeto, en la derecha la escala del objeto. El eje vertical muestra la confianza y el eje horizontal muestra la predicción.



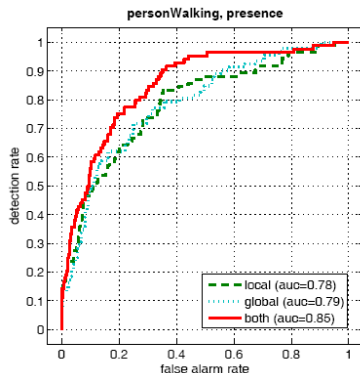
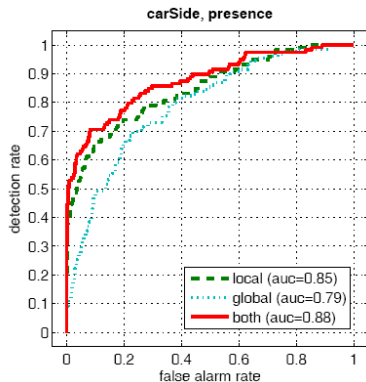
Localización de objetos

- La localización de objetos significa encontrar la localización y escala de un objeto en una imagen.
- Se realiza la comparación de tres metodos:
 - $P(X | L)$
 - $P(X | G)$
 - $P(X | L, G)$

Rendimiento de la detección de la presencia de objetos, para pantallas y teclados



Rendimiento de la detección de la presencia de objetos, para carros y peatones



Combinando características locales y globales

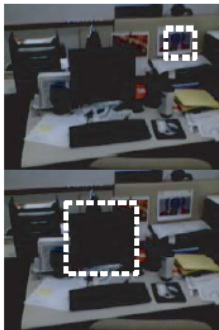
- Se hace la combinación de características locales y globales utilizando un modelo de “producto de expertos”

$$P(X = i|L, G) = \frac{1}{Z}P(X = i|L)^\gamma P(X = i|G)$$

- El exponente γ se obtiene realizando validacion-cruzada y es utilizado para balancear la confianza de los dos detectores, dado que fueron balanceados de forma independiente. Se utiliza $\gamma = 0.5$

- La desventaja del modelo “producto de expertos”, comparado con la regla Bayes, es que los pesos de la combinación son fijos. Sin embargo se cree que las ventajas de un framework discriminante lo compensa.

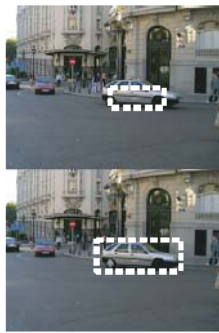
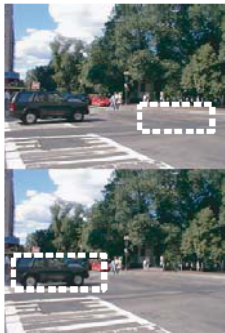
Resultados



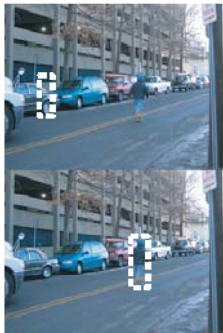
Resultados (2)



Resultados (3)

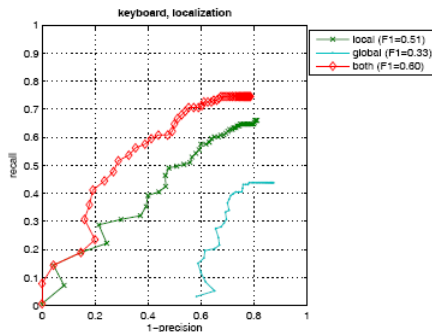
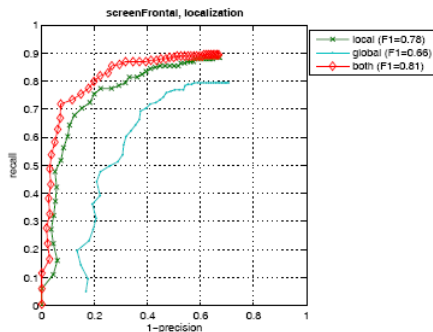


Resultados (4)



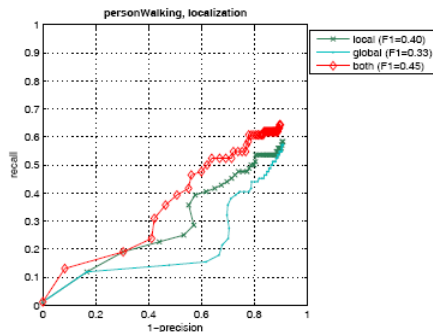
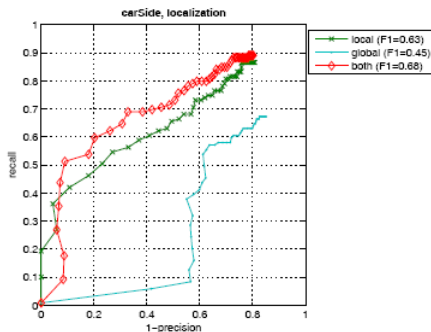
Resultados (5)

Rendimiento de la localización para pantallas y teclados. Se muestran las curvas de precisión para $P(X | G)$, $P(X | L)$ y $P(X | L, G)$



Resultados (6)

Rendimiento de la localización para carros y peatones.



- Utilizando características globales se puede eliminar la ambigüedad mostrada en los métodos de detección locales.
- Debido a que las características globales son compartidas en todas las clases y localizaciones, proporciona un método más barato computacionalmente.

- Una extensión natural de este trabajo es modelar las relaciones espaciales entre objetos.



K. Murphy, A. Torralba, D. Eaton, and W. Freeman. Object detection and localization using local and global features. Towards Category-Level Object Recognition, 1, 2005.