

Identificación de edificios y monumentos a partir de fotografías tomadas con dispositivos móviles

Esteban C. Fornal, Christian N. Pfarher, Mauro J. Torrez
Trabajo práctico final de “Procesamiento Digital de Imágenes”, II-FICH-UNL.

Resumen—Se presenta un método para la identificación de edificios y monumentos, a partir de fotografías tomadas con la cámara de un dispositivo móvil. Se extrae un vector de características **globales y locales?** de cada imagen. Se presentan dos métodos para la extracción de características, uno basado en la transformada de Hough y otro que utiliza estadísticas de los histogramas. Se evalúa el desempeño utilizando **error cuadrático medio?** para ambos... ambos métodos por separado y en conjunto, para una base de datos de prueba de imágenes.

Palabras clave—Identificación de edificios, *building recognition*, histograma, transformada de Hough, clasificación.

I. INTRODUCCIÓN

EN la actualidad, la gran difusión de los dispositivos móviles nos permite llevar la información importante siempre con nosotros. Además es muy común utilizar nuestros celulares, PDAs, GPS para *obtener* información desde cualquier lugar donde estemos, gracias a la proliferación de redes de datos inalámbricas. Dado que una gran parte de estos dispositivos poseen una cámara digital, surge la idea de utilizarla junto con las conexiones de datos para obtener información acerca del lugar donde nos encontramos. Se presenta aquí un método que combina técnicas de Procesamiento Digital de Imágenes orientadas en este sentido, que permite la identificación de edificios, monumentos y otros puntos de interés, a partir de una fotografía capturada con un dispositivo móvil.

II. MÉTODO PROPUESTO

En el método propuesto se ha puesto énfasis en la etapa de extracción de características de la imagen, considerando un método trivial de clasificación. La base de datos se generó tomando **X ??** imágenes representativas de cada monumento. Se extrajeron sus características y se promediaron para obtener un prototipo “generalizado” de cada monumento/edificio en particular.

Extracción de características

La extracción de características se realiza en este trabajo mediante dos técnicas diferentes:

- 1) por transformada de Hough, y
- 2) por estadísticas del histograma.

Extracción de características mediante Transformada de Hough

La transformada de Hough nos permite visualizar, a partir de una imagen de bordes, los parámetros de aquellas rectas¹ que son principales en la imagen.

¹El procedimiento es general, sirve para cualquier geometría que se pueda expresar en términos de sus parámetros. En este trabajo, se utiliza el espacio de los parámetros de las rectas, el más sencillo.

-1	-2	-1	-1	0	1
0	0	0	-2	0	2
1	2	1	-1	0	1

Fig. 1. Máscaras de filtrado (operadores gradiente) de Sobel.

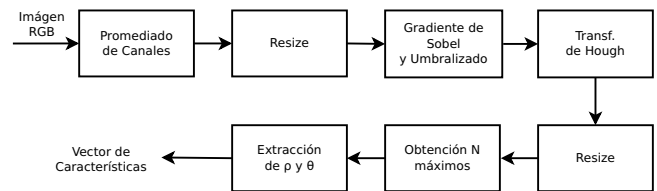


Fig. 2. Proceso de extracción de características por Transformada de Hough

Para la extracción de características con esta técnica se siguen los siguientes pasos:

- 1) A partir de la imagen original, se obtiene su versión en escala de grises promediando los tres canales RGB, y se la escala a un tamaño normalizado.
- 2) Se obtiene una imagen de sólo bordes, aproximando la magnitud del gradiente según

$$\nabla f \approx |G_x| + |G_y|, \quad (1)$$

donde G_x , G_y son el resultado de aplicar los operadores gradiente de Sobel (fig. 1) a la imagen. Finalmente se umbraliza esta imagen de bordes utilizando un parámetro U :

$$f(x) = \begin{cases} 0, & x \leq U \\ 255, & x > U \end{cases} \quad \text{f(x,y) ? en realidad es f(intensidad) sería} \quad (2)$$

- 3) Con la imagen de bordes umbralizada se calcula la transformada de Hough para rectas.
- 4) Se aplica un escalado (submuestreo) a la transformada obtenida, llevándola a un tamaño pequeño para obtener mayor tolerancia tanto en el parámetro angular θ como en el de distancia ρ .
- 5) Se toman los N máximos de esta transformada y se guardan en el vector de características sus coordenadas (ρ, θ) , mapeadas al rango $[-1, 1]$, obteniendo así un vector de $2N$ valores.

El proceso completo puede verse en la Fig. 2.

Extracción de características por estadísticas del histograma

A partir de la imagen original, se normaliza su tamaño y se toman 2 “perfiles de intensidad” **promedio o acumulado?**: uno horizontal, calculado promediando **promediando?** cada columna de la imagen,

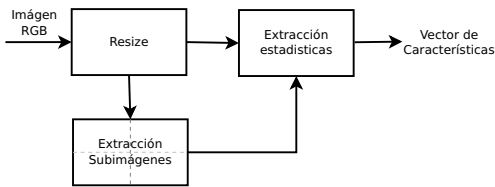


Fig. 3. Proceso de extracción de características utilizando estadísticas del histograma

y otro vertical obtenido al promediar cada fila. También Se obtiene un histograma de la imagen entera. De estos 3 vectores, se calculan y se guardan en el vector de características la media aritmética m , la mediana M (posición del percentil 50), y la desviación absoluta D_{abs} respecto de la mediana:

$$D_{\text{abs}} = \sum_i |x_i - M|$$

Así, se ha obtenido un vector de 9 valores que caracterizan la imagen entera.

Con la idea de incorporar características locales, se subdivide **o subdividió -no se entiende la corrección** la imagen en cuatro cuadrantes y se obtienen, para cada uno, las mismas medidas que se calcularon para la imagen entera. Como resultado, se obtiene un vector de 45 características asociado a cada imagen.

A. Clasificación

revisar si es subseccion* o subsection El entrenamiento de la base de datos se realiza obteniendo las características para cada imagen, junto a su etiqueta. Se extraen las características de todas las imágenes con la misma etiqueta, y se guarda un “prototipo” para esa clase, obtenido de promediar estas características.

La clasificación de las imágenes consiste en obtener la etiqueta del prototipo, cuyas características minimicen el error cuadrático medio con las de la imagen a identificar.

III. BASE DE DATOS

ver título de la sección.. Las imágenes de prueba fueron tomadas con teléfonos celulares corrientes, a una resolución VGA estándar de 640 píxeles de ancho por 480 de alto, en diferentes condiciones de iluminación: día a pleno sol, día nublado, noche, e interior, a monumentos/estatuas y edificios en diferentes lugares de la ciudad de Santa Fe (Argentina). Se muestran en la figura 4 dos ejemplos de las mismas.

IV. EXPERIMENTOS Y RESULTADOS

A. Experimentación

Para analizar el método se utilizó validación cruzada, entrenando la base de datos con 10 imágenes por clase y dejando 3 de prueba por cada clase.

Se utilizaron 3 conjuntos de imágenes diferentes, con 5 clases representativos de edificios basados en nuestras previas clasificaciones de imágenes.

Se clasificó utilizando extracción de características mediante transformada de Hough y mediante histogramas por separado. Para el método de Hough se probaron diferentes valores de U (el umbral aplicado a la imagen de bordes) y N (el número de máximos en el espacio de la transformada de Hough considerado). Mientras que, para el método de histograma se consideró el histograma y los perfiles del canal de intensidad en el espacio de colores HSI.



Fig. 4. Ejemplo de dos imágenes utilizadas para probar el método.

TABLA I
TASAS DE ERROR PARA LOS MÉTODOS

Método	5 clases	15 clases
Histogramas	0%	0%
Hough	35.5%	60.43%
Ambos	2.22%	4.17%

Luego se evaluó el rendimiento del método utilizando todas las características en forma conjunta, para los parámetros de Hough U y N óptimos encontrados.

Se realizó además una prueba para las 15 clases de imágenes del conjunto, con el objetivo de tener una estimación de cómo responde el método para un mayor número de imágenes.

Se considera la tasa de error (%) del método según:

$$E_{\%} = 100 \cdot \frac{\text{número de errores}}{\text{número de pruebas}}, \quad (3)$$

considerando como error a cada prueba en que la imagen es mal etiquetada.

Cabe destacar que no se considera “error” cuando el método identifica una imagen como el monumento acertado, pero en otras condiciones de iluminación (cuando por ejemplo, la imagen de un monumento tomada de día es clasificada como la del mismo monumento, pero tomada de noche). Esta consideración se hace debido a que la detección de bordes elimina la información de iluminación de la imagen; y está claro que no influye en la clasificación por histograma; ya que éste varía significativamente entre las versiones de día y noche, en particular, la media y mediana tendrán valores bastante menores en la imagen nocturna que en aquella tomada de día. Esto es además consistente con el objetivo del método, que es la correcta identificación independientemente de las condiciones en que se toma la imagen.

B. Resultados

Los resultados de las pruebas con el método de Hough se muestran en la figura 5. Se ha tomado un rango de valores **representativos de edificios basados en nuestras previas clasificaciones de imágenes.** Hemos identificado regiones de mínimo error para $U \in (60, 120)$ y $N \in (20, 60)$. En la misma figura, se puede apreciar que la tasa de error mínima se obtuvo para un umbral $U = 100$ y unos $N = 30$ máximos (en promedio).

En la tabla I se presentan los resultados para las pruebas con conjuntos de 5 clasesel resultado con 15 clases. Se puede observar que para el método de histogramas, la tasa de error fue cero en ambas pruebas. En tanto que para

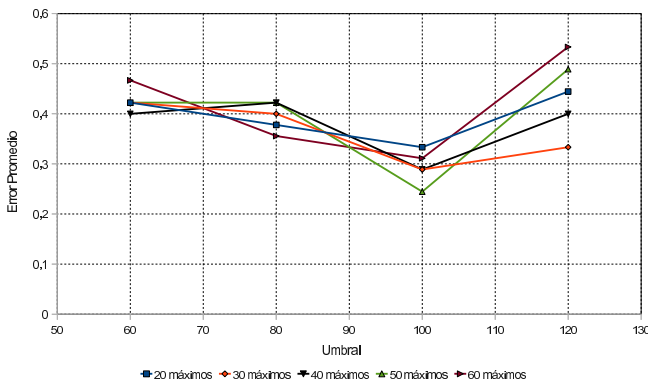


Fig. 5. Error de clasificación promedio en función de distintos U y N para el método de Hough.

el método de Hough, se obtiene menor tasa de error con el conjunto de 5 clases que con el de 15, de 35.5% y 60.43% respectivamente. En la última fila, se muestran los resultados de considerar todas las características en conjunto, asignándole un peso equivalente a cada una.

C. Discusión

A pesar de la muy buena performance del método de histograma, debemos tener presente que los conjuntos de imágenes de entrenamiento/prueba para el método, han sido obtenidas en todos los casos a partir de una secuencia tomada en el lapso de unos minutos, lo cual implica que las condiciones de iluminación en cada conjunto son prácticamente iguales. /resaltaclarar antes lode los conjuntos sino aca no se entiende Los histogramas, miden la distribución estadística de los diferentes niveles de intensidad presentes en la imagen, luego están estrechamente relacionados con las condiciones de iluminación de la escena. Por este motivo, tenemos que los histogramas entre las imágenes de entrenamiento/prueba son, a los efectos prácticos, similares, lo cual explica el buen funcionamiento de este método.

Se ha probado la técnica de histogramas sobre el canal I del espacio de color HSI; por lo antes expuesto, debemos tener en cuenta consideraciones similares para el resto de los canales de la imagen.

En general, se deberá poner énfasis en definir técnicas de “normalización” orientadas a mejorar los resultados en condiciones más realistas. Se deberá considerar la utilización de métodos de histograma mejorados, por ejemplo el *histograma de combinación espacial de DCT ponderada* [1], *vectores de coherencia de color* [2], o el procesamiento de histogramas borrosos presentado en [3].

Respecto de la extracción de características mediante la transformada de Hough, vemos que el rendimiento no es tan bueno como en la técnica de histogramas, sin embargo en varios casos ha sido capaz de identificar correctamente edificios a pesar de las diferentes condiciones de iluminación (día y noche), lo cual es un resultado alentador.

No se puede evitar la mención al costo computacional del proceso, que aunque no es tan elevado como para considerar impráctico el método, sí será una limitante al considerar implementaciones en tiempo real o de alta velocidad de respuesta: en una computadora promedio el cálculo se realiza en 1-2 segundos, luego es esperable que este tiempo se triplique en un dispositivo móvil.

Se deberá considerar la utilización de un mejor detector de bordes, como el propuesto por Canny [4], así como también técnicas de pre-procesamiento de la imagen, como puede ser el filtrado homomórfico, en pos de mejorar los resultados obtenidos en nuestras pruebas.

En lo que respecta a la escalabilidad de la base de datos, los resultados no son particularmente alentadores, lo que nos obliga a considerar otra vez la utilización de métodos más refinados de histograma y transformada de Hough. **esto no lo entiende marcelo**

V. CONCLUSIONES

Se ha presentado una técnica para la identificación de edificios, monumentos, esculturas con extracción de características mediante medidas de histograma y transformada de Hough.

El rendimiento ha sido satisfactorio considerando las restricciones a las que se han sometido las pruebas.

Se debe optimizar la implementación para portarlo a dispositivos móviles con capacidad de procesamiento limitada.

Se hace necesario un preprocesamiento de la imagen, así como la incorporación de métodos más refinados de extracción de características, para mejorar los resultados obtenidos y así poder usar el método con una base de datos de mayor magnitud.

VI. TRABAJOS FUTUROS

A partir del diseño aquí presentado, han surgido líneas para continuar investigando, y lograr un método más robusto:

- Aplicación de filtrado homomórfico y otros tipos de pre-procesamiento en las imágenes.
- Aplicación de técnicas de *warping* y otras transformaciones en busca de lograr invarianza respecto de rotación y escalado de la imagen.
- Desarrollo de técnicas de extracción de características más robustas.
- Desarrollo de una implementación óptima para dispositivos móviles con poder de procesamiento limitado.

REFERENCIAS

- [1] L. S. Yonghua Xie y H. Burkhardt, “Color image retrieval based on weighted dct spatial combination histogram,” *International Journal of Information Technology*, vol. 14, no. 2, pp. 12–23.
- [2] G. Pass y R. Zabih, “Histogram refinement for content-based image retrieval,” in *IEEE Workshop on Applications of Computer Vision*, pp. 96–102, 1996.
- [3] K. Konstantinidis, A. Gasteratos, y I. Andreadis, “Image retrieval based on fuzzy color histogram processing,” *Optics Communications*, vol. 248, no. 4–6, pp. 375 – 386, 2005. [Online]. Disponible: <http://www.sciencedirect.com/science/article/B6TVF-4F2MDYT-3/2/074827181182cd309f59f57ee820d26f>
- [4] J. F. Canny, “A computational approach to edge detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 679–698, 1986.
- [5] R. C. Gonzalez y R. E. Woods, *Digital Image Processing*. Prentice-Hall, 2001.