

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. Para la identificación se extrae un vector de características de la imagen, que es almacenado en una base de datos para su consulta. Se presentan dos métodos para la extracción de características en la imagen, uno basado en la transformada de Hough y otro que utiliza estadísticas de los histogramas. Se evalúa el desempeño utilizando ambos métodos por separado y en conjunto, para una base de datos de prueba de unas pocas imágenes.

Palabras clave—Identificación/reconocimiento de edificios, *building recognition*, histograma, extracción de características, transformada de Hough, clasificación.

I. INTRODUCCIÓN

LA presencia de gran cantidad de dispositivos tecnológicos móviles como celulares, PDAs, y otros, poseen mayormente cámaras digitales incorporadas y han pasado a ser de uso común debido a su facilidad de uso y portabilidad.

Muchas personas toman fotografías de paisajes, monumentos o edificios históricos que por diferentes motivos les resulta de interés conservar. Surge así la idea de desarrollar una técnica de procesamiento de imágenes, que posibilite la identificación de edificios, monumentos u otros puntos de interés, para poder obtener información pertinente acerca de los mismos. Con este artículo se trata de hacer un aporte en vías a resolver dicha inquietud.

II. MÉTODO PROPUESTO

El método que proponemos se basa en la extracción de características de la imagen y la comparación de éstas con las de una base de datos. Esta base de datos se genera tomando X imágenes representativas del monumento, extrayendo sus características y promediándolas para obtener un prototipo “generalizado” del monumento/edificio a detectar.

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

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

El entrenamiento de la base de datos se realiza obteniendo las características para cada imagen, junto con su etiqueta. Para cada etiqueta, se extraen las características de las X imágenes etiquetadas con la misma, y se guarda un “prototipo” obtenido de promediar estas características, junto con la etiqueta.

La clasificación de las imágenes consiste en encontrar la etiqueta del prototipo cuyas características minimicen el error cuadrático medio.

-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.

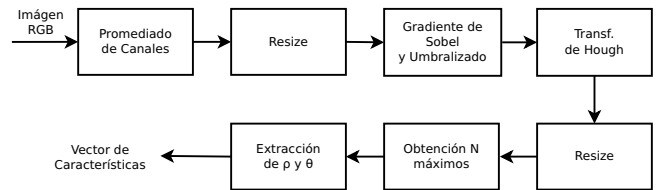


Fig. 2. Proceso de la imagen mediante el método por T. de Hough

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.

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

- 1) A partir de la imagen original, se obtiene una 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 en U :

$$f(x) = \begin{cases} 0, & x \leq U \\ 255, & x > U \end{cases} \quad (2)$$

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

El proceso completo puede verse en la Fig. 2.

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

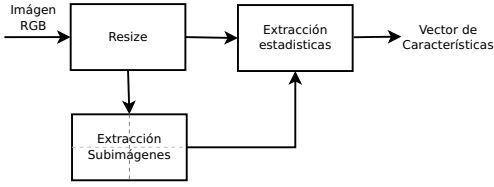


Fig. 3. Proceso de la imagen mediante el método de estadísticas del histograma



Fig. 4. Imágenes de pruebas utilizadas con los algoritmos

Extracción de características por estadísticas del histograma

A partir de la imagen original, se normaliza su tamaño y se toman 2 “perfiles de intensidad”: uno horizontal, calculado promediando cada columna de la imagen, y otro vertical obtenido al promediar cada fila. Se obtiene un histograma para la imagen entera y uno para cada perfil, y se guardan en el vector de características la media aritmética m , mediana M (posición del percentil 50), y desviación absoluta D_{abs} respecto de la mediana:

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

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

Luego se subdivide 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 a partir de histogramas que se guardarán en la base de datos para comparación.

III. EXPERIMENTOS Y RESULTADOS

A. Descripción de las pruebas

Para la prueba del método se utilizó validación cruzada, entrenando la base de datos con 10 imágenes por clase y 3 de prueba por cada clase. Se utilizaron 3 conjuntos de imágenes diferentes, con 5 clases cada uno.

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 algunos ejemplos. Los conjuntos se formaron sorteando las 15 clases de imágenes disponibles en 3 de 5 clases cada uno.

TABLA I
TASAS DE ERROR PARA MÉTODOS

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

Se probaron, en primer lugar, la 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 a tener en cuenta para armar el vector de $2N$ características.

Para el método de histograma se consideró el histograma del canal de intensidad en el espacio de colores HSI.

Luego se consideró el rendimiento del método utilizando ambas técnicas en forma conjunta, para los parámetros de Hough U y N óptimos encontrados. Se realizó además una prueba para las 15 clases de imágenes en conjunto, con el objetivo de tener una estimación de cómo escala el método para mayor número de imágenes.

Tasa de error: 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 cada prueba en que la imagen es mal etiquetada.

Cabe observar que no se considera como error la vez que 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 **luminosidad** 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 para el método de Hough se muestran en la figura 5. Se ha tomado un rango de valores representativos de U y N basados en pruebas previas, donde 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 iniciales con conjuntos de 5 clases, y la estimación del 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 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

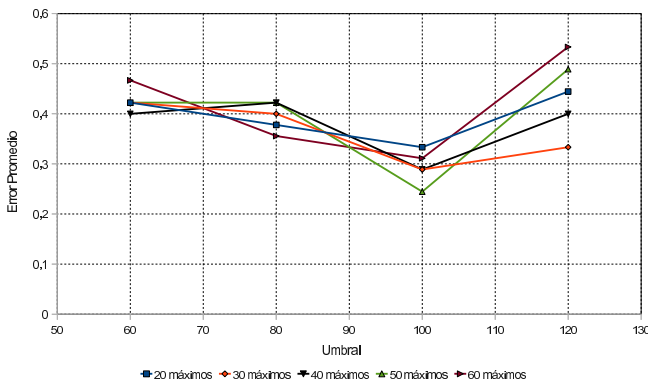


Fig. 5. Error de clasificación promedio en función de distintos U y N para el método de Hough.

los resultados de considerar ambos criterios en conjunto, asignándole un peso equivalente a cada uno.

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.

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, idénticos, 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), el cual es un resultado alentador.

No podemos 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 detector de bordes como el propuesto por Canny [4] **que tiene de bueno el de canny?**, así como también técnicas de preprocesamiento 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.

IV. 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 deberá optimizar la implementación para portarlo a dispositivos móviles con capacidad de procesamiento limitada.

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

V. TRABAJOS FUTUROS

A partir del diseño aquí presentado, nos proponemos continuar investigando esta técnica considerando las siguientes posibilidades:

- 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 una implementación óptima para dispositivos móviles con poder de procesamiento limitado.
- Desarrollo de técnicas de extracción de características más robustas que las aquí presentadas.

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,” en *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.