

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 de diferente índole ha abierto un sin número de nuevas aplicaciones para satisfacer las necesidades diarias de seres humanos. Los dispositivos móviles como celulares, PDAs, etc. han pasado a formar parte del común de nuestras vidas brindando nuevas posibilidades de interacción. Es aquí donde surge la idea de la realización de este trabajo.

Día a día, las personas toman fotografías de diferentes objetos ya sean monumentos públicos, edificios históricos, etc. sin saber si quiera que se está fotografiando. Con este artículo se trata de hacer un aporte en vías hacia dicho problema, de manera que mediante el procesamiento de imágenes se tenga dicha información en el instante mismo de la adquisición de la foto. Cabe aclarar, que se hará una implementación en C++, dejando como trabajo futuro el desarrollo de una aplicación para dispositivos móviles.

En principio y ya que no es el objetivo de este trabajo hacer un análisis profundo, sino tan solo dar una aproximación inicial a la resolución del problema, se considerará la aplicación de los métodos en condiciones ideales o semi-ideales. ver...

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.

-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.

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.

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.

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.

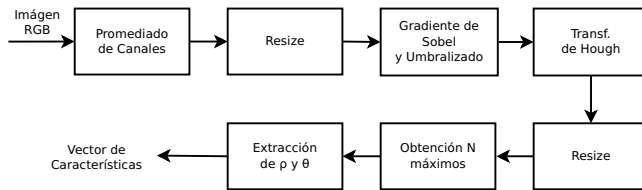


Fig. 2. "Proceso de la imagen mediante el método por T. de Hough"

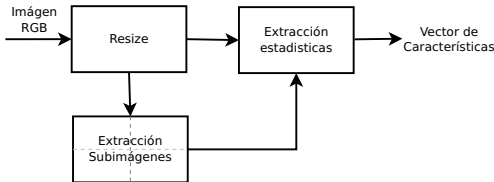


Fig. 3. "Proceso de la imagen mediante el método de estadísticas del histograma"

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 han 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.

Así 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

Para las pruebas se procedió a evaluar el desempeño de ambos métodos, en primera instancia por separados y luego, conjuntamente.

A. Descripción de la Base de datos de imágenes

Se trabajó con imágenes que fueron adquiridas en la ciudad de Santa Fe, mediante un Dispositivo Móvil con una resolución de 640x480 pixels. Se construyó una base de datos de las mismas sobre un total de cinco edificios, tomando trece realizaciones de cada uno de ellos (diez con el propósito de usarlas como prototipo y tres para las prueba con los algoritmos). Cabe aclarar que cada conjunto de imágenes de los cinco edificios fueron tomadas en condiciones ambientales similares y con la cámara en la misma posición respecto del objetivo. **arreglar esto**.

B. Vectores característicos y prototipos

Sobre diez de las trece imágenes de cada edificio, se aplicó el método de la Transformada de Hough citado en este artículo, obteniéndose diez vectores representativos (uno por cada imagen) y se procedió a promediar dichos vectores dando como resultado un prototipo por cada edificio. De la misma manera se realizó el mismo proceso, con el método estadístico descrito también en este documento.

C. Descripción de las pruebas

Por cada método se calculó el MSE o error cuadrático medio definido en 3 entre los cinco prototipos y las quince imágenes de pruebas, para finalmente obtener la tasa de reconocimiento mediante 4

$$MSE = \dots \quad (3)$$

$$\%reconocimiento = \frac{\#aciertos}{15} \quad (4)$$

D. Tablas

columna1	columna2
----------	----------

E. Discusión

blablabla

IV. CONCLUSIONES

blablabla

V. TRABAJOS FUTUROS

A partir del diseño aquí presentado, seguiremos investigando esta técnica con las siguientes posibilidades:

- Considerar la aplicación de un filtrado homomórfico en imágenes que lo requieran.
- Independizarse de la posición en que se tomó la fotografía con alguna técnica de warping.
- Arreglar el metodo de Hough que es una X onga jaja

REFERENCIAS

- [1] R. C.Gonzalez y R. E.Woods, *Digital Image Processing*. Prentice-Hall, 2001.