

Máster en Ingeniería Industrial
2021-2022

Trabajo Fin de Máster

Detección y reconocimiento de imagen con aplicación ferroviaria para la validación de equipos industriales y operación automática

Antonio Rodríguez Alhambra

Tutor/es

M. Isabel Herreros Cid

Miguel López Hernández

Lugar y fecha de presentación prevista

DETECCIÓN DEL PLAGIO

La Universidad utiliza el programa **Turnitin Feedback Studio** para comparar la originalidad del trabajo entregado por cada estudiante con millones de recursos electrónicos y detecta aquellas partes del texto copiadas y pegadas. Copiar o plagiar en un TFM es considerado una **Falta Grave**, y puede conllevar la expulsión definitiva de la Universidad.



[Incluir en el caso del interés en su publicación en el archivo abierto]

Esta obra se encuentra sujeta a la licencia Creative Commons **Reconocimiento - No Comercial - Sin Obra Derivada**

RESUMEN

Palabras clave:

DEDICATORIA

ÍNDICE GENERAL

1. ESTADO DEL ARTE	1
2. MARCO TEÓRICO.	2
2.1. Extracción de características	2
2.1.1. Momentos de imagen.	2
2.2. Histograma de gradientes	4
BIBLIOGRAFÍA	6

ÍNDICE DE FIGURAS

ÍNDICE DE TABLAS

2.1	4
-----	-------	---

1. ESTADO DEL ARTE

2. MARCO TEÓRICO.

Todo proceso de clasificación de imágenes puede dividirse en los siguientes pasos:

- Tratado inicial de las imágenes
- Extracción de características
- Tratamiento de datos
- Entrenamiento de modelo de clasificación
- Evaluación modelo de clasificación

2.1. Extracción de características

En este paso el objetivo es obtener información cuantitativa (numérica) de las muestras a partir de diversos métodos.

Por ejemplo, de una clasificación de números escritos a mano se pueden obtener características como el número de trazos, anchura del contorno, color, etc. Es decir, se obtienen una serie de propiedades informativas de la muestra para, después, entrenar el modelo de clasificación.

Para este trabajo, se han utilizado los siguientes métodos:

- Descriptores de Fourier.
- Momentos de imagen Hu.
- Extracción de contornos.
- Descriptores locales binarios (*Local Binary Patterns*)
- Histograma de gradientes.
- *Método propio.

2.1.1. Momentos de imagen.

Los momentos de imagen son un promedio de las intensidades de una imagen binaria. La convención habitual define un momento M para una imagen binaria B de la siguiente forma:

$$M_{ij} = \sum_x \sum_y x^i y^j B(x, y) \quad (2.1)$$

Utilizando la ecuación 2.1, se pueden obtener características la imagen como los centroides:

$$\bar{x} = \frac{M_{10}}{M_{00}} \quad (2.2)$$

$$\bar{y} = \frac{M_{01}}{M_{00}} \quad (2.3)$$

Sin embargo, aplicando una transformación a la ecuación 2.1, se pueden obtener momentos invariantes a la traslación:

$$\mu_{i,j} = \sum_x \sum_y (x - \bar{x})^i (y - \bar{y})^j I(x, y) \quad (2.4)$$

A la ecuación 2.4 se la conoce como **momentos centrales**.

Además, aplicando otra transformación se pueden obtener momentos invariantes al escalado también:

$$\eta_{i,j} = \frac{\mu_{i,j}}{\frac{i+j}{2} + 1} \quad (2.5)$$

La fórmula 2.5 se conoce como **momento centralizado**.

Momentos de Hu

Los momentos de Hu [1] son un conjunto de siete fórmulas obtenidas a partir de los momentos centralizados que permiten obtener siete momentos invariantes tanto a traslación, rotación, escalado y volteado (el séptimo momento es el invariante a volteado, cambiando de signo cuando la imagen es reflejada).

$$\begin{aligned} M_1 &= \eta_{20} + \eta_{02} \\ M_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\ M_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\ M_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\ M_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ &\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ M_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ &\quad + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ M_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ &\quad - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (2.6)$$

A partir de 2.6, se pueden obtener siete características numéricas.

2.2. Histograma de gradientes

El histograma de gradientes [2] es una técnica de extracción de características muy utilizada en la detección de objetos.

A partir de una imagen binaria de dimensiones $n \times m$, se calculan los gradientes de intensidad, así como la magnitud y el ángulo:

$$\begin{aligned} G_x &= B(x+1, y) - B(x, y) \\ G_y &= B(x, y+1) - B(x, y) \end{aligned} \quad (2.7)$$

La ecuación 2.7 determina los gradientes de la imagen binaria B . Tanto G_x como G_y son dos imágenes binarias con la información de gradientes en ejes X e Y, respectivamente. A partir de la ecuación 2.7, se obtienen las magnitudes y ángulos:

$$Mag_{(x,y)}(\mu) = \sqrt{G_x^2 + G_y^2} \quad (2.8)$$

$$Ang_{(x,y)}(\theta) = |\tan^{-1}(\frac{G_y}{G_x})| \quad (2.9)$$

Tanto la ecuación 2.8 como 2.9 representan imágenes binarias.

El siguiente paso consiste en dividir las imágenes de magnitud y ángulos en N cuadrículas, pudiendo ser $N = 1$ (una única cuadrícula).

Para cada cuadrícula se representa un histograma de 9 posiciones, con cada posición en el rango $[\theta, \theta + \delta\theta)$. Convencionalmente, se suele utilizar $\delta\theta = 20^\circ$. De esta forma se obtiene el siguiente histograma H_N :

Magnitud									
Ángulo	0	20	40	60	80	100	120	140	160

TABLA 2.1

A cada intervalo angular y de magnitud, se le denomina θ_j y μ_j , respectivamente, para $j \in [0, 8]$.

De tal forma, si en la cuadrícula N_i existe un $\theta_{x,y}$ que pertenece al rango $[\theta, \theta + \delta\theta j)$, para $j \in [0, 8]$, se determina que $\mu_j = \mu_{x,y} + \mu_j$.

$$\mu_j = \sum \mu_{x,y} \iff \theta_{x,y} \in [\theta, \theta + \delta\theta \cdot j) \quad (2.10)$$

Existen casos en los que $\theta_{x,y} > 160^\circ$, entonces $\mu_{x,y}$ contribuye tanto a 0° como a 160° .

$$\begin{aligned} \frac{180 - \theta_{x,y}}{20} \cdot \mu_{x,y} &\implies [160, 180) \\ \left(1 - \frac{180 - \theta_{x,y}}{20}\right) \cdot \mu_{x,y} &\implies [0, 20) \end{aligned} \tag{2.11}$$

El último paso consiste en normalizar el histograma:

$$\mu_j = \frac{\mu_j}{\max(\mu_j)} \tag{2.12}$$

Tras la aplicación de 2.12, se tienen $9 \cdot N$ características, siendo N el número de cuadrículas.

BIBLIOGRAFÍA

- [1] M.-K. Hu, “Visual pattern recognition by moment invariants,” *IRE Transactions on Information Theory*, vol. 8, n.º 2, pp. 179-187, 1962. DOI: [10.1109/TIT.1962.1057692](https://doi.org/10.1109/TIT.1962.1057692).
- [2] R. K. McConnell, “Method of and apparatus for pattern recognition,” ene. de 1986. [En línea]. Disponible en: <https://www.osti.gov/biblio/6007283>.