

# Organización del Computador 2

## Preinforme Trabajo Práctico Final

Departamento de Computación  
Facultad de Ciencias Exactas y Naturales  
Universidad de Buenos Aires

## Reconocimiento Óptico de Caracteres.

| Integrante        | LU     | Correo electrónico          |
|-------------------|--------|-----------------------------|
| Izcovich, Sabrina | 550/11 | sizcovich@gmail.com         |
| Vita, Sebastián   | 149/11 | sebastian.vita@yahoo.com.ar |

## Índice

|  |   |
|--|---|
| 1. ¿Qué es el reconocimiento óptico de caracteres?                     | 3 |
| 2. Aplicaciones actuales   | 3 |
| 3. Dificultades encontradas por el reconocimiento óptico de caracteres | 4 |
| 4. Nuestro proyecto  | 4 |
| 5. Tecnologías a utilizar  | 5 |
| 6. Referencias   | 5 |

## 1. ¿Qué es el reconocimiento óptico de caracteres?

El reconocimiento óptico de caracteres (OCR, por sus siglas en inglés) es el proceso por el cual se traducen o convierten imágenes de dígitos o caracteres (sean éstos manuscritos o de alguna tipografía especial) a un formato representable en nuestra computadora (por ejemplo, ASCII). Esta tarea puede ser más sencilla (por ejemplo, cuando tratamos de determinar el texto escrito en una versión escaneada a buena resolución de un libro) o tornarse casi imposible (por ejemplo, cuando intentamos leer la receta de un médico).

## 2. Aplicaciones actuales

La técnica de reconocimiento óptico de caracteres resulta provechosa en diversas situaciones, por ejemplo, para modificar un archivo extenso del cual no poseemos la fuente. Al realizar el reconocimiento de dígitos de forma automática se ahorran grandes gastos en recursos humanos y tiempo por lo que su campo de aplicación resulta muy extenso. A continuación, listaremos algunos de los más comunes:

- **Digitalización de libros:** En la actualidad, existe una gran cantidad de libros de interés general que, ya sea por su antigüedad o por su reducida cantidad de ejemplares, no existen en formato digital. En una era de cultura digital en la que se accede a todo tipo de información a través de internet llevando a la desaparición de los formatos en papel, resulta necesario digitalizar dichos libros con el fin de que no desaparezcan y puedan llegar a la mayor cantidad de personas posible. Del mismo modo, la digitalización de libros facilita la búsqueda de los mismos, como también la de una frase o palabra específica de que puedan ser de interés.
- **Edición de textos:** En muchas ocasiones, se encuentran impresiones que poseen errores y deben ser modificadas. Al digitalizarlas, se pueden reformar de forma simple con cualquier editor de texto para luego volver a imprimirla.
- **Inserción de caracteres en un dispositivo:** Un ejemplo actual de esto puede ser la escritura en una tableta o celular inteligente. Muchos de estos dispositivos proveen un lápiz óptico que permite escribir a mano alzada sobre su pantalla. Para realizar esto, no es necesario el uso de un teclado ya que el dispositivo transforma la letra manuscrita en caracteres reconocibles por el mismo.
- **Formularios:** Muchas veces deben llenarse formularios de forma manual, como por ejemplo para un censo o para una elección. Posteriormente, estos datos deben ser volcados en una base de datos. En este caso, se utilizan las casillas que posee el formulario para separar cada uno de los caracteres y poder obtener los datos de forma eficiente.

**NOMBRE** RAQUEL **D.N.I.** 24834

**PETICIONES QUE SOLICITA POR ORDEN DE**  
 ESCRÍBASE CON LA MAYOR CLARIDAD POSIBLE, DADO QUE CUALQUIER E  
 QUE SE ANULE LA PETICIÓN O SE OBTENGA DESTINO EN UN

| Nº ORDEN | CENTRO O LOCALIDAD | ESPECIALIDAD | CÓDIGO | ESP. | SÍMBOLO | TITULANTE |
|----------|--------------------|--------------|--------|------|---------|-----------|
| 103      | 11022000137        |              |        |      |         |           |
| 104      | 11022000138        |              |        |      |         |           |
| 105      | 11013000237        |              |        |      |         |           |
| 106      | 11013000238        |              |        |      |         |           |
| 107      | 11008000137        |              |        |      |         |           |
| 108      | 11008000138        |              |        |      |         |           |

| Nº ORDEN | CENTRO O LOCALIDAD | ESPECIALIDAD | CÓDIGO | ESP. | SÍMBOLO | TITULANTE |
|----------|--------------------|--------------|--------|------|---------|-----------|
| 142      | 29039000338        |              |        |      |         |           |
| 143      | 29011000237        |              |        |      |         |           |
| 144      | 29011000238        |              |        |      |         |           |
| 145      | 29043000237        |              |        |      |         |           |
| 146      | 29043000238        |              |        |      |         |           |
| 147      | 29091000537        |              |        |      |         |           |

Figura 1: Ejemplo de formulario

### 3. Dificultades encontradas por el reconocimiento óptico de caracteres

Existen varias dificultades a la hora de reconocer caracteres. Entre ellas se encuentran:

- **Caligrafía:** Más allá de que los caracteres manuscritos siguen una forma básica, cada persona les aplica ciertos rasgos personales que los hacen únicos. Por ejemplo, un mismo carácter puede tener distinta firmeza en el trazo, distinta inclinación o distinto tamaño, dificultando su reconocimiento.



Figura 2: Ejemplo de cinco caracteres iguales pero con distinta caligrafía.

- **Ruido en la imagen:** Como las imágenes se obtienen a partir de un dispositivo, ya sea escáner o cámara de fotos, éste puede agregarle ruido si es que no es muy preciso, como también disminuir la resolución o calidad de la imagen. Llamamos ruido a la variación del brillo o color en las imágenes, como también a las irregularidades que pueden presentar. De este modo, ciertos dispositivos pueden alterar la escala de grises en el fondo de las mismas o sobre los caracteres que posee dificultando el reconocimiento de los mismos. Por otro lado, muchos textos se borran, manchan o tornan amarillos a medida que pasa el tiempo generando ruido y dificultando enormemente la identificación de sus respectivos caracteres.

Las características mencionadas anteriormente dificultan el reconocimiento óptico de caracteres y no permiten que el resultado del mismo sea óptimo en todos los casos.

### 4. Nuestro proyecto

El objetivo del trabajo práctico consiste en implementar un método de reconocimiento de dígitos manuscritos basado en la descomposición en valores singulares. Para ello, se procederá utilizando la herramienta de SIMD vista en la materia. De este modo, dada una imagen presentando un número entre 0 y 9, se tomará la información relevante de la misma que permita reconocer digitalmente de qué dígito se trata. El procedimiento a realizar es el siguiente: Como instancias de entrenamiento, se tiene un conjunto de  $n$  imágenes de dígitos manuscritos en escala de grises del mismo tamaño y resolución (varias imágenes de cada dígito). Cada una de estas imágenes sabemos a qué dígito se corresponde pues se encuentran etiquetadas. En este trabajo consideraremos la base de datos MNIST<sup>1</sup>, utilizada como referencia en esta área de investigación. Para  $i = 1, \dots, n$ , sea  $x_i \in R^m$  la  $i$ -ésima imagen de nuestra base de datos almacenada por

filas en un vector, y sea  $\mu = (x_1 + \dots + x_n)/n$  el promedio de las imágenes. Definimos  $X \in R_{n \times m}$  como la matriz que contiene en la  $i$ -ésima fila al vector  $(x_i - \mu)^t \div \sqrt{n-1}$  y

$$X = U \Sigma V^t$$

a su descomposición en valores singulares, con  $U \in R^{n \times n}$  y  $V \in R^{m \times m}$  matrices ortogonales, y  $\Sigma \in R^{n \times m}$  la matriz diagonal conteniendo en la posición  $(i, i)$  al  $i$ -ésimo valor singular  $\sigma_i$ . Siendo  $v_i$  la columna  $i$  de  $V$ , definimos para  $i = 1, \dots, n$  la transformación característica del dígito  $x_i$  como el vector  $tc(x_i) = (v_1^t x_i, v_2^t x_i, \dots, v_k^t x_i) \in R^k$ , donde  $k \in \{1, \dots, m\}$  es un parámetro de la implementación. Este proceso corresponde a extraer las  $k$  primeras componentes principales de cada imagen. La intención

<sup>1</sup><http://yann.lecun.com/exdb/mnist/>

es que  $tc(xi)$  resume la información mas relevante de la imagen, descartando los detalles o las zonas que no aportan rasgos distintivos.

Luego, dada una nueva imagen  $x$  de un dígito manuscrito, que no se encuentra en el conjunto inicial de imágenes de entrenamiento, el problema de reconocimiento consiste en determinar a qué dígito corresponde. Para esto, se calcula  $tc(x)$  y se compara con  $tc(x_i)$ , para  $i = 1, \dots, n$ .

Este proyecto nos pareció oportuno como trabajo final ya que para poder realizarlo es necesario utilizar operaciones sobre matrices. Algunas de estas operaciones, como la triangulación de matrices o la obtención de autovalores y autovectores, pueden ser optimizadas mediante SSE. Esto se debe a que aplican las mismas operaciones sencillas sobre un grupo de elementos, luego, las imágenes podrán ser procesadas a gran velocidad.

Nuestra propuesta consiste en programar tanto en *C* como en *Assembler* las funciones principales de los algoritmos que calculan la información más relevante de cada imagen. Una vez hecho esto, analizaremos los tiempos de corrida y evaluaremos la optimización de los mismos.

## 5. Tecnologías a utilizar

Para la realización del trabajo descrito anteriormente, utilizaremos la base de datos MNIST que se encuentra compuesta por imágenes en escala de grises de tamaño 28x28, y presenta un conjunto de entrenamiento de 60000 patrones y uno de testeo de 10000. Aunque ambas bases de datos incluyen casos de dígitos con distorsiones y ambigüedades y que son difíciles de clasificar, la base de datos MNIST tiene en su tratamiento la dificultad de el gran volumen de datos a procesar.

## 6. Referencias

- <http://www.iti.es/media/about/docs/tic/13/articulo2.pdf>