

## UN METODO DE RECONOCIMIENTO AUTOMATICO DE CARACTERES IMPRESOS

Por Jaime López - Krahe, Universidad Paris VIII - Vincennes.

### Resumen

Descripción de un heurístico de reconocimiento de caracteres impresos mediante el recuento de transiciones del blanco al negro, según una dirección paralela a la línea de texto analizado. Estudio numérico de los parámetros de construcción de un aparato sencillo de lectura óptica para ciegos, alcanzando un porcentaje de reconocimiento del 94 %.

### 1.- Introducción.

El problema tiene un planteamiento diferente al de las lectoras ópticas normales, que al ser periféricos de entrada al ordenador, necesitan un porcentaje de reconocimiento muy elevado.

En nuestro caso, en un 95 % es suficiente ya que el hombre puede completar, mediante el contexto, los caracteres incorrectamente identificados. Veremos que el recuento del nº de transiciones del blanco al negro según una dirección paralela a la línea del texto, nos da un gran número de componentes de un vector del espacio de representación. Haremos un estudio para optimizar la muestra y la elección de los otros parámetros necesarios a la construcción del aparato.

### 2.- Definición del espacio de representación y algoritmo.

Sea  $C$  un conjunto finito de caracteres que hay que reconocer. El sistema conlleva un cierto aprendizaje, ya que tiene que definir un conjunto finito de puntos,  $E \in \mathbb{R}^n$  que representen a  $C$ .

A cada punto  $m \in \mathbb{R}^n$ , se le hace corresponder  $m' \in E$  tal que la distancia  $mm'$  sea mínima. La distancia que nos ha parecido más adecuada es la de tipo euclidiano (1). Puede ser necesario ordenar  $E$ , de tal forma que si dos puntos están a una distancia mínima de  $m$ , se escoge el más pequeño respecto al orden definido en  $E$ . La relación de orden puede ser la frecuencia de aparición de los caracteres en el lenguaje en cuestión.

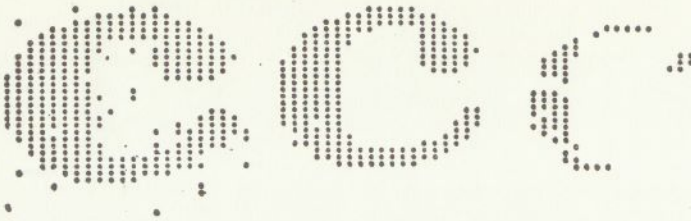


Fig. 1.- Evolución de un mismo carácter (la "C") según tres valores diferentes, en densidad, del umbral blanco/negro: 8, 24 y 125 (la densidad óptica 3D corresponde al valor 255)

Para ello se realiza una digitalización (fig. 1) de una muestra adecuada de caracteres que permiten una simulación numérica del aparato. Se ha buscado una geografía fotosensible y los algoritmos que nos den una buena relación entre el porcentaje de reconocimiento y la economía de los medios empleados para dicho reconocimiento (que dependen fundamentalmente del número de fotoreceptores).

Se proyecta el texto sobre dos columnas de fotoreceptores, colocados transversalmente a la línea impresa, (ésta avanza a mano, en un barrido de izquierda a derecha).

La velocidad de recogida de datos es independiente del procedimiento de reconocimiento: esto se consigue mediante el recuento de las transiciones del blanco al negro según el estado del conjunto de fotoreceptores.

Cada fotoreceptor lleva asociado un contador que acumula el nº de transiciones. Es una codificación de las componentes del vector, - el cual, se encuadra a continuación a partir de la primera componente no nula del vector de representación (fig. 2).

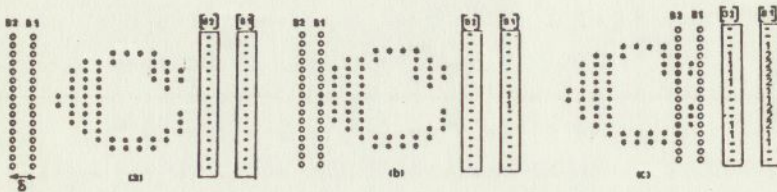


Fig. 2.- Representación esquemática de las tres fases de codificación de un carácter aislado: (a) estado de espera, todos los contadores a cero (viene indicado por un trazo "-" en la figura). (b) comienzo de la detección de un carácter por B1, se empieza el recuento. (c) detección de fin de carácter por B1. Se activa la lectura de B2 que nos da una componente adicional del vector representado.

Vector producido por el contador B1:  $x = (0, 0, 1, 1, 2, 2, 2, 1, 1, 2, 2, 2, 2, 1, 0, 0)$ . Vector encuadrado  $x' = (1, 1, 2, 2, 2, 1, 1, 2, 2, 2, 2, 1, 0, 0, 0, 0)$ .

---

La segunda columna de fotoreceptores B2 da las informaciones indispensables para resolver ciertas ambigüedades de reconocimiento, como las simetrías de codificación (p. ej. los caracteres d, b; p, q ...). Se lee una sola vez, en el momento de la detección de fin de carácter, y es activada por B1.

Además B2, permite detectar el espacio en blanco que separa dos palabras. La elección de la distancia , que separa B1 de B2 tiene que cumplir las condiciones siguientes:

- Debe ser inferior a la suma de la dimensión del carácter más pequeño y de la separación mínima entre caracteres (para evitar recoger información que proviene de otros caracteres).
- Debe ser inferior al menor espacio en blanco que separe dos palabras.
- Debe ser superior al mayor espacio entre caracteres que exista (para poder distinguir el espacio entre palabras).

De las experiencias de simulación llevadas a cabo con imágenes numéricas podemos concluir que la distancia más discriminatoria corresponde al límite inferior permitido por las condiciones precedentes. Esto es debido, a que las nubes de representación en el espacio  $R^n$  de un mismo carácter, se encuentran menos diseminadas al tener una codificación más fiable.

### 3.- Elección de parámetros y resultados.

Para evitar que se produzcan codificaciones diferentes para un mismo carácter, la proyección del texto se normaliza mediante una óptica de grosor variable (zoom). Se ha determinado el número de receptores por columnas que nos da el porcentaje de reconocimiento más elevado por el algoritmo siguiente:

- Un número de fotoreceptores demasiado pequeño, no da suficiente información para producir una "firma" discriminatoria de los caracteres (2), sino que atenúa el ruido parásito de alta frecuencia espacial por reducción de la banda (fig. 3).



Fig. 3.- Legibilidad en función del número de fotoreceptores.

Al contrario, un número demasiado elevado no es ventajoso, ya que la calidad de la información se ve degradada por los ruidos parásitos, como p. ej. los "agujeros" blancos en los caracteres negros.

Además el crecimiento de información en la entrada produce el aumento correspondiente en el tratamiento de la fase de reconocimiento propiamente dicha.

Para este estudio hemos limitado voluntariamente el aprendizaje a un UNICO REPRESENTANTE de cada carácter que hay que reconocer: el aparato en proyecto podrá incluir, por simple conmutación esta función simplificada de aprendizaje para adaptarse a los caracteres más comunes. En la evaluación del porcentaje de reconocimiento global obtenido con las experiencias realizadas sobre las líneas de un libro de texto de difusión corriente, no se han considerado los caracteres utilizados en el aprendizaje.

La figura 4 presenta los resultados obtenidos en función de dos parámetros importantes en la definición del aparato.

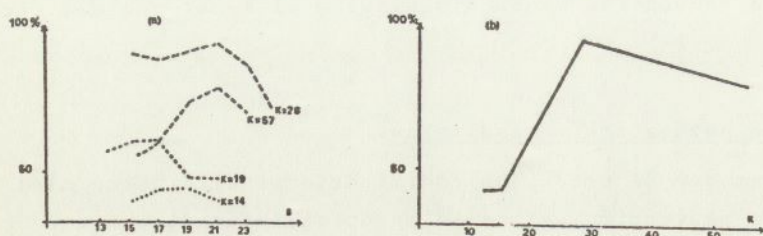


Fig. 4.- Evolución del porcentaje de reconocimiento en función de los parámetros de muestreo: (a) de S, umbral de separación blanco/negro (la figura 1 da una indicación subjetiva); (b) de K, número de fotoreceptores por columna en una elección optimizada del umbral blanco/negro.

Se puede observar que el número óptimo se encuentra entre 20 y 30 fotoreceptores. Un análisis de los resultados ha mostrado que el 62 por ciento de los errores se refieren a ciertas confusiones sistemáticas entre ciertos caracteres ("u" y "n", "a" y "s", ...)

Se puede prever que el usuario se acostumbre a estos fallos y los corrija fácilmente dado su carácter sistemático.

#### 4.- Conclusión.

Podemos observar que las características más importantes de este método son las siguientes:

- Porcentaje de reconocimiento del orden del 94 por ciento asociado a unos medios simples.
- Independencia de la codificación respecto a la velocidad de barrido.
- Independencia respecto a la dimensión horizontal del carácter.
- Tratamiento de los casos en que no exista una alineación vertical (gracias al encuadre)
- Velocidad de reconocimiento elevada respecto a la economía de los medios utilizados.
- Errores sistemáticos (sin duda susceptibles de ser reducidos).

A la vista de estos resultados preliminares, parece que son grandes las posibilidades de construir una máquina barata de lectura óptica de cadencia elevada, a partir de microprocesadores, teniendo en cuenta la buena relación que existe entre el porcentaje de éxito y la simplicidad en el algoritmo descrito. (3), (4) .

#### Referencias

- (<sup>1</sup>) E. DIDAY et G. GORAERT, *Apprentissage et mesures de ressemblances adaptatives (I.R.I.A., Le Chesnay, 1974)*.
- (<sup>2</sup>) C. E. SHANNON, *Proc. Inst. Radio Eng.*, 37, 1949, p. 10-21.
- (<sup>3</sup>) J. LOPEZ-KRAHE, *Simulation d'une machine de lecture optique à l'usage des aveugles (Maîtrise d'Informatique, Université Paris VIII, 1976)*.
- (<sup>4</sup>) UNESCO, *Congrès International pour le 150<sup>e</sup> anniversaire du système Braille [(Rapports), Paris, mai 1975]*.