

Captura de información impresa por medios ópticos

Enrique Cabrera, Raúl Espejel y Mireya Gally

Instituto de Física, Universidad Nacional Autónoma de México,

Apartado postal 20-364, 01000 México.

(Recibido el 12 de septiembre de 1988; aceptado el 22 de agosto de 1990)

Resumen. En el presente trabajo se discuten los criterios de diseño de los equipos especializados en la digitalización de imágenes gráficas y fotográficas. Se analizan las características del prototipo que se diseñó y construyó con el fin de automatizar la captura de información impresa y que consta fundamentalmente de tres partes: un sistema de movimiento de la cabeza lectora respecto del documento, un sistema electro-óptico para digitalizar la imagen y un conjunto de programas para el procesamiento de la información y el control del flujo de la misma. Este equipo fue diseñado como equipo periférico a una computadora Commodore C-64 y puede ser adaptado en otros equipos. La información así recopilada es almacenada en disco para posteriormente ser procesada.

PACS: 06.50.-x; 06.70.Td

1. Introducción

Una de las principales tareas a la que se enfrenta la mayoría de los usuarios de las computadoras es la captura de información. Debido al vertiginoso ritmo de crecimiento del volumen de material impreso a que se tiene acceso hoy día esta tarea requiere que se le dedique cada vez un mayor tiempo. Con el fin de agilizar este proceso, se ha desarrollado un modelo experimental capaz de alimentar a la computadora a través de un sistema automatizado de lectura.

El material impreso a procesar en una computadora se puede dividir según su presentación de dos formas: la gráfica (blancos y negros) y la fotográfica (tonos de gris). Como ejemplos de la primera se tienen los textos, los listados de programas, las listas de datos y los planos y como ejemplos de la segunda, las fotografías tomadas con los microscopios ópticos o electrónicos, las aéreas, etc. La recopilación de la información se lleva a cabo en función de las propiedades geométricas de la imagen, de la densidad óptica de cada uno de los puntos componentes y del estudio de las variaciones en el contraste entre puntos contiguos.

Los sistemas para la captura de información impresa por medios ópticos a partir del rastreo de la imagen se dividen en tres partes fundamentales:

- i) El dispositivo lector encargado de recibir la información proveniente del impreso, de transducirla en una señal eléctrica y de procesarla,
- ii) El mecanismo que efectúa el barrido del dispositivo lector sobre el área de interés del impreso,

- iii) Un conjunto de programas encargado de controlar al sistema periférico y procesar la información recibida.

Los sistemas especializados para la digitalización de las imágenes se dividen, en función de la información que se quiere capturar, en digitalizadores y densitómetros. Los primeros requieren de una imagen de alto contraste y el proceso de digitalización consiste en almacenar el valor booleano, correspondiente a la señal eléctrica recibida del sensor, en la memoria de la computadora correlacionando dicho valor con la posición de cada punto del impreso. En el caso de los densitómetros el valor almacenado de cada punto es proporcional a la densidad fotográfica relativa de dicho punto. Un sistema para el reconocimiento de caracteres es un caso particular de digitalizador de imágenes gráficas, cuya finalidad es digitalizar una imagen en pequeñas porciones de dimensiones fijas en las que debe quedar incluido un caracter. A estas porciones les corresponde un área de memoria en la que la imagen del caracter ha quedado grabada en forma discreta y se les conoce como mapas de puntos de los caracteres. La imagen así digitalizada posteriormente es interpretada.

En base a los aspectos generales aquí mencionados se desarrolló un modelo experimental para digitalizar textos, con capacidad de interpretarlos.

2. Criterios de diseño

Los criterios de diseño y construcción de un sistema digitalizador de imágenes pueden dividirse para su exposición en tres partes: el movimiento del sensor, la captura óptica y el procesado de la información.

2.1 *El movimiento*

El movimiento de exploración del objeto por el elemento sensor puede realizarse de tres modos equivalentes: desplazando el sensor respecto del objeto, desplazando el objeto respecto del sensor y mediante una combinación de ambas, en cuyo caso el objeto se desplaza en una dirección y el sensor en la dirección ortogonal para cubrir así el total del área en estudio. Dichos movimientos son controlados por la computadora, lo que permite tener una correspondencia biunívoca entre la posición del lector y la dirección de almacenamiento en memoria con lo cual es posible reconstruir la imagen a partir de la información almacenada. El movimiento del sistema en las direcciones ortogonales es en forma discreta lo que a su vez lleva implícita la discretización de la imagen. Los transductores mecánicos más recomendables para realizar dicho movimiento son los motores de pasos (*stepping motors*) donde cada uno de los ejes es controlado por un motor.

Entre las características más relevantes de este tipo de motores se cuenta su precisión angular dada por el número de pasos por revolución, la posibilidad de moverlos con doble precisión angular (medios pasos) y la forma digital en que se alimentan. La lógica de control que se requiere para su funcionamiento los hace especialmente prácticos para ser controlados mediante una computadora; esto se debe a que las señales de alimentación se pueden expresar en forma binaria, modo

natural de trabajo de estas máquinas. De aquí la sencillez de la electrónica asociada y su gran confiabilidad. Para su control, se requiere solamente de una función lógica que determine para cada momento el valor binario de alimentación. Con un solo parámetro de la función queda definido el sentido de giro y el paso que se realice en cada movimiento. Dado que el programa en la computadora es el encargado del movimiento de los motores, es en éste donde se implementó la sincronía entre el movimiento y el sensado de la señal.

Como se mencionó anteriormente, estos motores tienen un número predeterminado de pasos, por ello la resolución lineal del sistema depende de este número y del sistema mecánico. La repetitividad en la posición que presenta este tipo de motores permite agilizar la digitalización de la imagen al implementar en el programa una lectura bi-direccional. El problema medular en este caso es que la precisión del mecanismo de acoplamiento entre el motor de pasos y la cabeza lectora sea tal que la posición de la cabeza sea la misma independientemente de si el movimiento es en una u otra dirección. Estas restricciones hicieron que se seleccionara el mecanismo de una impresora comercial ATI Z-520 de impresión bi-direccional para realizar los movimientos de rastreo de la imagen, dado que en ésta los problemas mecánicos ya han sido resueltos por el fabricante.

Cabe mencionar aquí que la principal desventaja que presenta este tipo de motores es la vibración que producen por su forma discreta de trabajo. A fin de evitar que éstas afecten la resolución del sistema los fabricantes de impresoras utilizan un acoplamiento por bandas dentadas y engranes de reducción, además de una frecuencia óptima de pulsos. Para alcanzar el número de pulsos por segundo al cual vienen optimizadas las impresoras, fue necesario programar la función lógica que los controla en lenguaje de máquina, debido a que el lenguaje Básic resulta demasiado lento para estos fines. En esta rutina se introdujo un parámetro de retardo que permite un ajuste fino de la frecuencia de los pasos.

2.2 La captura

Las condiciones expuestas anteriormente restringen fuertemente las características de los componentes del sistema. La geometría entre el elemento luminoso y el detector debe seleccionarse en función de la distribución espacial del foco, las propiedades de la superficie del objeto (en el caso de transparencias sus propiedades ópticas) y la distribución espacial del elemento sensor. Esto es, finalmente, lo que define la resolución de la imagen digitalizada y por ende las aplicaciones que de él se pueden hacer.

En los prototipos desarrollados se utilizaron foto-darlingtonos como sensores, elementos éstos de estado sólido que entregan a la salida un voltaje proporcional a la cantidad de luz recibida. El factor de amplificación resulta 1000 veces mayor que el de los foto-diodos y su respuesta pico es a longitudes de onda de $0.9 \mu\text{m}$, por lo que como fuente de iluminación se utilizaron IREDS (LEDs infra-rojos). El tiempo de respuesta de estos dispositivos es de aproximadamente $600 \mu\text{seg}$, suficiente para realizar una lectura por cada paso del motor (2 mseg).

Se utilizaron fibras ópticas como medio para acoplar las dimensiones del área

sensible del foto-darlington (1 mm^2) y el área mínima requerida por cada elemento, para construir los mapas de puntos que caracterizan unívocamente a cada uno de los caracteres alfa-numéricos (0.6 mm^2). Esto permite hasta cierto punto controlar la resolución del sistema sensor.

La señal eléctrica a la salida del sensor es procesada con el fin de tener una señal compatible con los requerimientos de la siguiente etapa según el modo de operación. La respuesta del sensor a los tonos de gris requiere, en el modo gráfico, de un sistema digital con histéresis para evitar oscilaciones y entregar una señal de tipo TTL [9]. En el modo fotográfico se requiere de una señal proporcional a la densidad óptica de cada punto y de un convertidor analógico-digital que la traduzca a un formato binario. La longitud de la palabra del convertidor determina el número de niveles disponibles; la diferencia en densidad fotográfica entre dos niveles consecutivos está dada por la ganancia del sistema previo al convertidor. Estos parámetros determinan la resolución y el rango dinámico del sistema.

Un aspecto muy importante en el diseño de estos sistemas es la homogeneidad de la iluminación de cada uno de los puntos a ser sensado ya que las inhomogeneidades son interpretadas como información. En ocasiones es preciso determinar el elemento sensor más adecuado a la iluminación disponible, mientras que en otras ocasiones hay que adecuar la iluminación al sensor [7].

La resolución final de los prototipos construídos quedó determinada, en el caso del lector de textos por el área de detección de la fibra óptica (0.035 mm^2), y la resolución del movimiento (0.5 mm en x y 0.2 mm en y); en el caso de la platina micrométrica la resolución en ambos ejes (x, y) es de $2.54 \mu\text{m}$, valor muy superior al área de detección de la fibra óptica, por lo que se utilizó un sistema óptico de microscopio para amplificar la imagen antes de ser capturada por el lector.

2.3 El procesado

Los criterios generales de diseño en el desarrollo de los algoritmos para la digitalización de imágenes y el reconocimiento de caracteres, se basan en la discretización de la imagen. Esto implica mover los motores desde la computadora, capturar y almacenar la información acorde al modo de operación y finalmente procesarla. La información proviene de una serie de puntos del objeto ordenados según un patrón previamente definido y se alimenta a la computadora en una secuencia tal que permita, con un requerimiento mínimo de memoria, generar el mapa de puntos correspondiente. El mínimo número de puntos, así como su distribución espacial, que definen de manera unívoca a todos los caracteres ASCII, dependen de la forma de los caracteres.

En este prototipo se utilizó la forma estándar de caracteres de la computadora Commodore C-64 a la que se acopló el sistema y está formada por una matriz de 64 puntos (8×8). En este caso, mediante un programa especializado, se determinó que con solamente 15 de los 64 puntos se pueden identificar unívocamente los 91 caracteres más utilizados en la escritura de textos (alfa-numéricos), sin dejar margen de error en la lectura, situación poco realista. Admitiendo un error experimental del 10% en la digitalización de los caracteres, se requiere como mínimo de 36 puntos

distribuidos en una matriz de (6×6) , lo que define la posición de las fibras ópticas y la longitud que debe desplazarse el lector para cada paso del motor.

El elevado número de puntos de que se forma la imagen de un impreso impone condiciones restrictivas en cuanto al tiempo de ejecución de los programas de identificación de caracteres, razón por la cual se seleccionó nuevamete el Lenguaje de Máquina como lenguaje de programación [2,3,5].

En el caso particular del reconocimiento de caracteres primero se genera el mapa de puntos de cada caracter a ser identificado y se procede a su comparación con los mapas de puntos previamente almacenados de los caracteres del conjunto ASCII. Para ello se hace uso de dos conceptos fundamentales: el "peso" y la "distancia" de Hamming [4]. El primero se define como el número de puntos "encendidos" en el mapa del caracter en cuestión, mientras que el segundo se define como el número de puntos en que no se tuvo concordancia en el estado lógico al comparar el caracter leído y el caracter en memoria. Esto es: la máxima similitud entre dos caracteres se tiene cuando la distancia de Hamming es mínima (en el caso que ésta sea cero, se tiene la identidad de los mapas de puntos).

La información capturada en la computadora es almacenada en medios magnéticos así como presentada en pantalla o en una impresora gráfica.

Los mismos algoritmos utilizados para la captura de la información óptica de los caracteres de un impreso son utilizados para el aprendizaje de nuevos conjuntos de caracteres, asociando un mapa de puntos a cada uno de los caracteres ASCII.

3. Prototipos

Los criterios mencionados con anterioridad permitieron el desarrollo de una gran variedad de equipos con aplicaciones muy concretas. En particular se elaboraron tres prototipos: un digitalizador de imágenes, un lector ópto-electrónico y un densitómetro. Las características fundamentales que los diferencian son: el movimiento, la captación de la información y el procesamiento de la misma.

3.1 El movimiento

En lo que al movimiento se refiere, se elaboraron tres sistemas: una mesa digitalizadora, una impresora de matriz y una platina micrométrica. Para cada uno de estos sistemas se utilizaron diferentes medios de acoplamiento entre los motores de pasos y el lector, cuidando en todos ellos que los motores no estuvieran acoplados directamente al soporte del lector y se sujetarán firmemente a la estructura del sistema, con el fin de amortiguar al máximo las vibraciones de los motores y que no se transmitieran éstas a través del mecanismo de acoplamiento.

En la mesa digitalizadora el objeto permanece fijo y el detector realiza un barrido en ambas direcciones. En este caso la transmisión del movimiento se hace básicamente a través de hilos y el detector se desplaza sobre una barra que a su vez le sirve de guía para dar el movimiento vertical. El movimiento horizontal se logra desplazando a la barra sobre dos rieles en sus extremos. El mayor inconveniente de

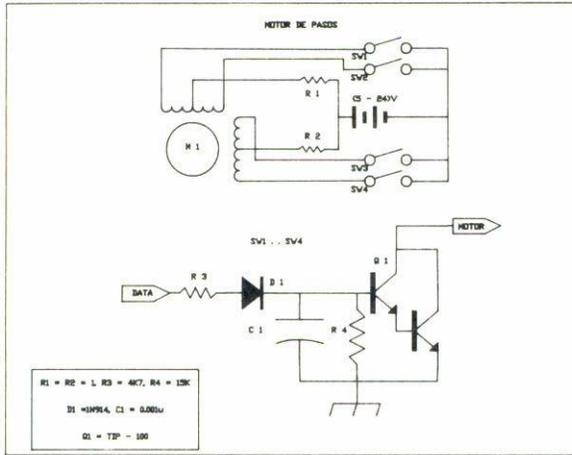


FIGURA 1. Diagrama del circuito electrónico para el movimiento de los motores de pasos desde una computadora.

este mecanismo es la complejidad de su fabricación para garantizar la ortogonalidad de los ejes y la resolución del sistema.

En el segundo sistema se adaptó una impresora de matriz de puntos en la cual se sustituyó la cabeza impresora por el lector óptico. El primero de los motores controla el desplazamiento a lo largo de un renglón mientras que el segundo hace girar un rodillo sobre el cual va firmemente sujeto el impreso [1].

Finalmente se acoplaron los ejes de los motores de pasos directamente a los tornillos micrométricos de una platina de microscopio (uno por cada eje) con una resolución de 1/3840 de pulgada ($6.6 \mu\text{m}$), para cada paso del motor en ambas direcciones. Cabe mencionar que este sistema sirvió de base al densitómetro y como medio para determinar la confiabilidad de los motores de pasos.

En la Figura 1 se muestra el diagrama del sistema de movimiento de los motores de pasos y su interface con la computadora.

3.2 La captación

En lo que al modo de captación se refiere, se diseñaron y construyeron dos sistemas:

- a) Un sensor de imágenes gráficas para los sistemas de digitalización y de reconocimiento de caracteres,
- b) Un sensor de imágenes fotográficas acoplado a un convertidor analógico digital con una resolución de 8 bits (256 tonos de gris) para el densitómetro.

En la Figura 2 se presenta el diagrama esquemático del digitalizador de imágenes y en la Figura 3 el del foto-densitómetro.

diferencia entre ellos el modo de almacenamiento y si ésta va o no a ser procesada para la identificación de los caracteres:

- a) en el caso de imágenes digitalizadas se requiere de un vector para el almacenamiento de valores booleanos (bits) y de la forma en que éste debe ser interpretado para el reestablecimiento de la imagen,
- b) en el caso del densitómetro, se requiere también de un solo vector para el almacenamiento, en este caso de valores enteros (bytes) y la forma en que este debe ser interpretada es análoga a la anterior.
- c) en el caso del sistema para el reconocimiento óptico se requiere de un vector para el almacenamiento de los mapas de puntos y otro para el valor ASCII de cada uno de los caracteres identificados, con lo que la información queda muy compacta.

4. Conclusiones

El desarrollo de este trabajo dió como resultado un sistema periférico a una computadora Commodore C-64, capaz de capturar información a partir de material impreso. La eficiencia de este dispositivo viene dada en función de su rapidez de ejecución, la cual está limitada por la velocidad angular de los motores de pasos (300 rev/min = 480 pulsos/seg) ya que el tiempo de interpretación de cada caracter (40 caracteres/seg) es equivalentemente inferior. La información adquirida por el sistema puede reproducirse a través de los periféricos asociados a la computadora: la impresora, el monitor o el disco flexible.

La relevancia del lector ópto-electrónico se basa en la capacidad que tiene de “ver”, “interpretar” y “aprender” diferentes tipos de caracteres que constituyen a los impresos. El tiempo de interpretación del lector elaborado es comparable con el del sistema comercial OMNI READER para PC [6].

La relevancia de la mesa digitalizadora estriba en el bajo costo que implica la construcción de este sistema y la posibilidad que se tiene de convertirla fácilmente en una graficadora al sustituir el elemento sensor por una plumilla, aplicación que para el presente trabajo está fuera de los objetivos.

La precisión del convertidor ADC0804 [8] de 8 bits limita a 256 el número de tonos de gris en los cuales se divide la señal eléctrica recibida del sensor. Debido a que la mayoría de los monitores solo son capaces de reproducir un número limitado de tonos de gris, entre el negro y el blanco, la señal es procesada internamente y se recurre al concepto de falso color para ampliar la gama de grises [10]. En la actualidad existen en el mercado diversos dispositivos que realizan una función equivalente de adquisición de datos por medios ópticos. Estos dispositivos están enfocados para su uso en máquinas de tipo PC, con una mayor capacidad de memoria que la usada en este trabajo y también una mayor velocidad de ejecución por la frecuencia del cristal asociado y por el tamaño de la palabra, sin embargo su costo es muy elevado.

Referencias

1. ATI de México S.A. de C.V., *Manual de usuario, Impresora ATI Z-520 C.I.* (1986).
2. Bredon, Glen, *Merlin 64 (A macro assembler for the Commodore 64)*, Southwestern Data System (1983).
3. Commodore Business Machines, *The Commodore 64 (Macro Assembler Development System)* (1982).
4. Macwilliams, F.F. y Sloane N.F.A., *The Theory of Error Correcting Codes*, North-Holland (1977).
5. Mansfield, Richard, *Machine Language for Beginners*, Compute! Publication Inc. (1983).
6. Oberon International, Ltd., *OMNI READER*, Software design AIM, Cambridge (1986).
7. Fairchild *Semiconductor Data Book, Linear Integrated Circuits* (1976).
8. National Semiconductor Corporation, *Linear Data Book* (1980).
9. Sahm, W.H. *General Electric Optoelectronics Manual*, Semiconductor Products Dept. (1976).
10. Gonzalez, Rafael C. and Wintz, Paul, *Digital Image Processing*, Addison-Wesley Pb. Cmp. (1987).

Abstract. This prototype of image scanner digitizes graphics or photographics images. In the case when the digitized image is a text, the software developed is capable to recognize the characters. The scanned information or the ASCII value of the recognized characters are saved in disk for future use. In this paper are discussed the different criteria followed for the construction of this prototype, which can be divided in three fundamental parts: the scanning element or reading head, the mechanical system that moves the reading head and the paper, and the software needed for the character recognition and to control the whole system. This scanner was designed to work with the Commodore C-64 but can easily be implemented in any kind of personal computer.