

Ingeniería del conocimiento

C. León, J. Molina, C. Fragoso, A. Algarín
 Dpto. de Tecnología Electrónica
 Universidad de Sevilla

cleon@cica.es

Reconocimiento automático de caracteres empleando técnicas de sistemas expertos

Resumen: El presente artículo presenta la aplicación de técnicas de inteligencia artificial, en este caso sistemas expertos, al reconocimiento automático de caracteres. Concretamente se expone la labor llevada a cabo en la realización de un "Entorno de Reconocimiento" que combina técnicas algorítmicas de tratamiento de imágenes con un "lenguaje de descripción de caracteres" desarrollado para permitir realizar el proceso de reconocimiento en base a reglas de producción.

1. Introducción

Antes de pasar a la descripción de trabajo realizado se hace necesario justificar la necesidad del mismo. Diremos pues que el reconocimiento automático de caracteres, bien sean impresos o manuscritos, forma parte de la creciente demanda de sistemas de tratamiento de documentos digitalizados. El problema del reconocimiento automático de caracteres (OCR) ha sido abordado desde diferentes puntos de vista, distinguiéndose, en general, entre aquella parte del proceso relacionada con la adquisición y tratamiento del texto sobre el que se desea trabajar y el reconocimiento en sí mismo.

En cualquier caso, en las diferentes aproximaciones al reconocimiento existentes (emparejamiento de matrices, análisis estructural, lenguajes descriptivos, redes neuronales, etc.) se establece un balance entre la rapidez de los resultados y la precisión de los mismos [GOM-96].

En el trabajo presentado en este artículo se propone un "Entorno de Reconocimiento", inicialmente aplicado a caracteres numéricos impresos pero fácilmente extensible a cualquier tipo de elemento que se desee reconocer, bien sean caracteres tipográficos o cualquier otro objeto, que combina la flexibilidad de la programación estructurada en Reglas de Producción con los diferentes métodos de reconocimiento contenidos en las mismas, resultando un sistema flexible, con un índice aceptable de aciertos (superior como media al 95%), que es fácil de actualizar y ampliar.

2. Estructura global del sistema

El objetivo del sistema es el reconocimiento de dígitos a partir de una imagen digitalizada que supuestamente los contiene, para ello está dividido en tres módulos de carácter general: Módulo de Adquisición, Módulo de Tratamiento y Descripción y Módulo de Reconocimiento.

El objetivo del primero es capturar la imagen procedente de la digitalización de un documento en la memoria de un ordenador PC, mediante un dispositivo de entrada; el segundo procesará la imagen para obtener y describir la mayor cantidad de información posible, y el tercero aplicará técnicas basadas en sistemas expertos para reconocer esa información y etiquetarla (**Figura 1**). A continuación pasamos a describir detalladamente cada uno de los mencionados componentes.

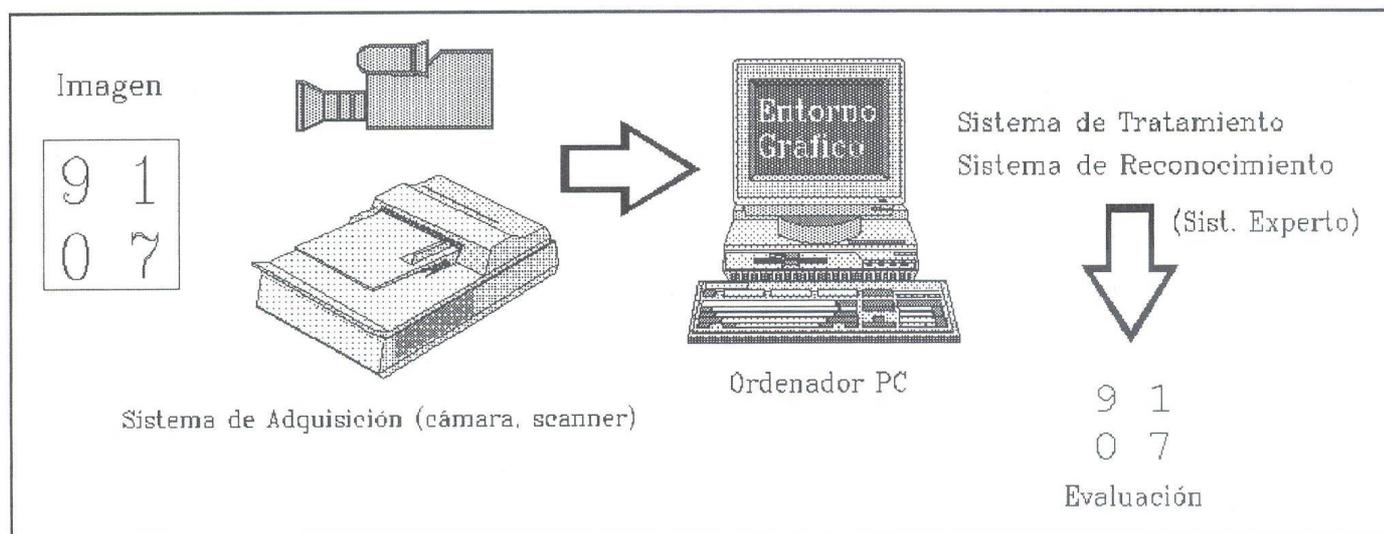


Figura 1

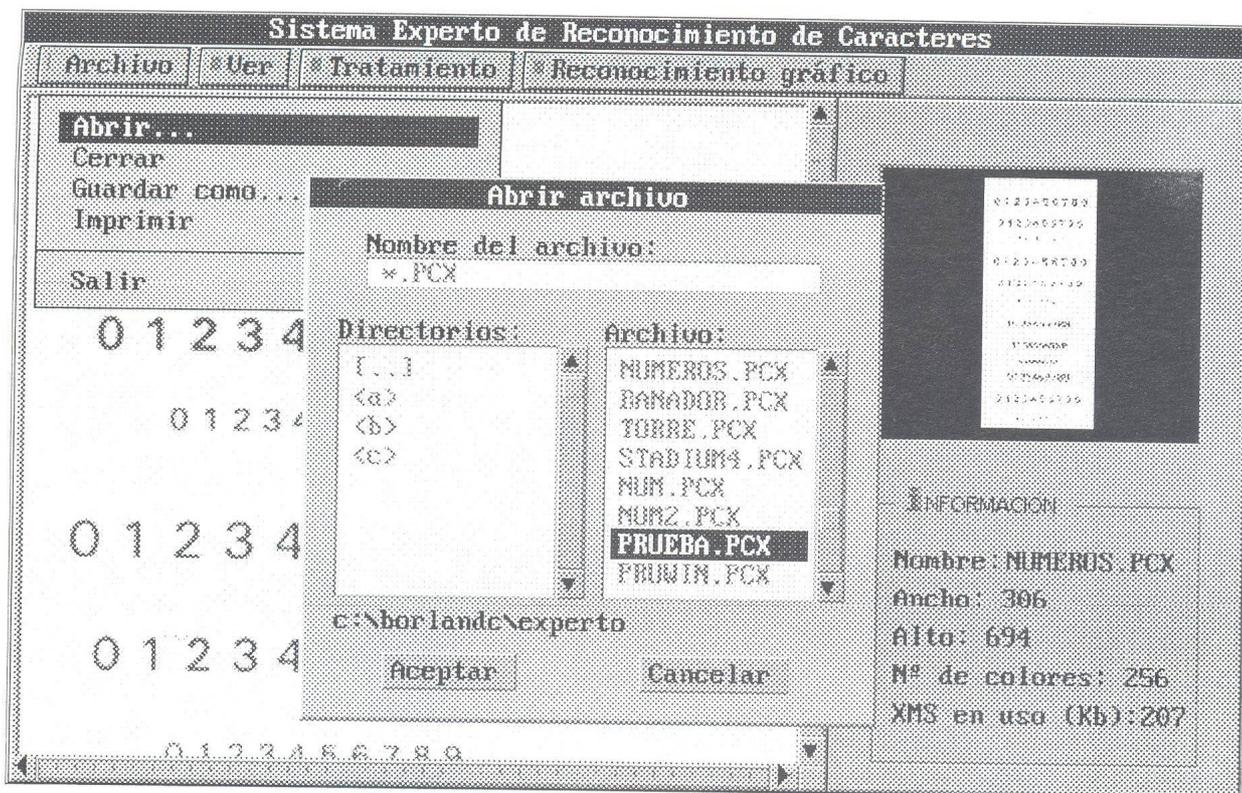


Figura 2

3. Módulo de adquisición

En este módulo se realiza la digitalización de la imagen suministrada por un escáner o por una cámara de TV [GAL-90]. Es capturada en un entorno gráfico instalado en un ordenador PC, cuyas funciones permiten visualizar, editar y modificar la imagen para adecuarla a su tratamiento posterior.

El mencionado módulo de adquisición ha sido programado íntegramente en lenguaje C y diseñado para trabajar estrechamente con las posibilidades gráficas del modo SVGA [TIS-93], para ello se han desarrollado diversas librerías de funciones y objetos gráficos, que permiten al usuario situarse ante un entorno amigable (Figura 2), distribuido en ventanas, de fácil manejo gracias a su menú multinivel.

4. Módulo de tratamiento y descripción

En este módulo se procesa la imagen con el fin de aislar y describir los elementos gráficos contenidos en ella con objeto de completar los términos de un lenguaje descriptivo, que enlaza directamente con la fase de reconocimiento. Para el correcto tratamiento de la imagen ésta sufre un proceso en el que se eliminan ruidos y desperfectos ocasionados en la fase de adquisición, todo ello de manera opcional y configurable a partir de las funciones de edición y tratamiento incorporadas en el entorno (Figura 3).

4.1. Algoritmos y funciones de tratamiento

El bloque de tratamiento está constituido por tres grupos de algoritmos: Aislamiento o División parcial de la imagen, Detección de los bordes o contornos y Esqueletización [NIB-86],[FU-88].

El Aislamiento consiste en separar los grupos de puntos susceptibles de ser reconocidos como caracteres en la fase posterior. Para ello se barre la imagen en toda su extensión y se almacena la localización de estos grupos de información; por otro lado la División Parcial [VER-91] se realiza gráficamente sobre la imagen, obteniendo el sistema, una vez concluido el proceso, la ubicación de todos los objetos gráficos contenidos en ella.

Las funciones de Detección de bordes y los algoritmos de Esqueletización permiten describir los objetos gráficos; así el contorno y el esqueleto aportan información útil en la generación de un código direccional, en tanto que el esqueleto permite además calcular puntos significativos del objeto.

4.2. Descriptores empleados

La descripción del objeto es la fase que sigue al tratamiento, consistiendo en extraer características de éste para recono-

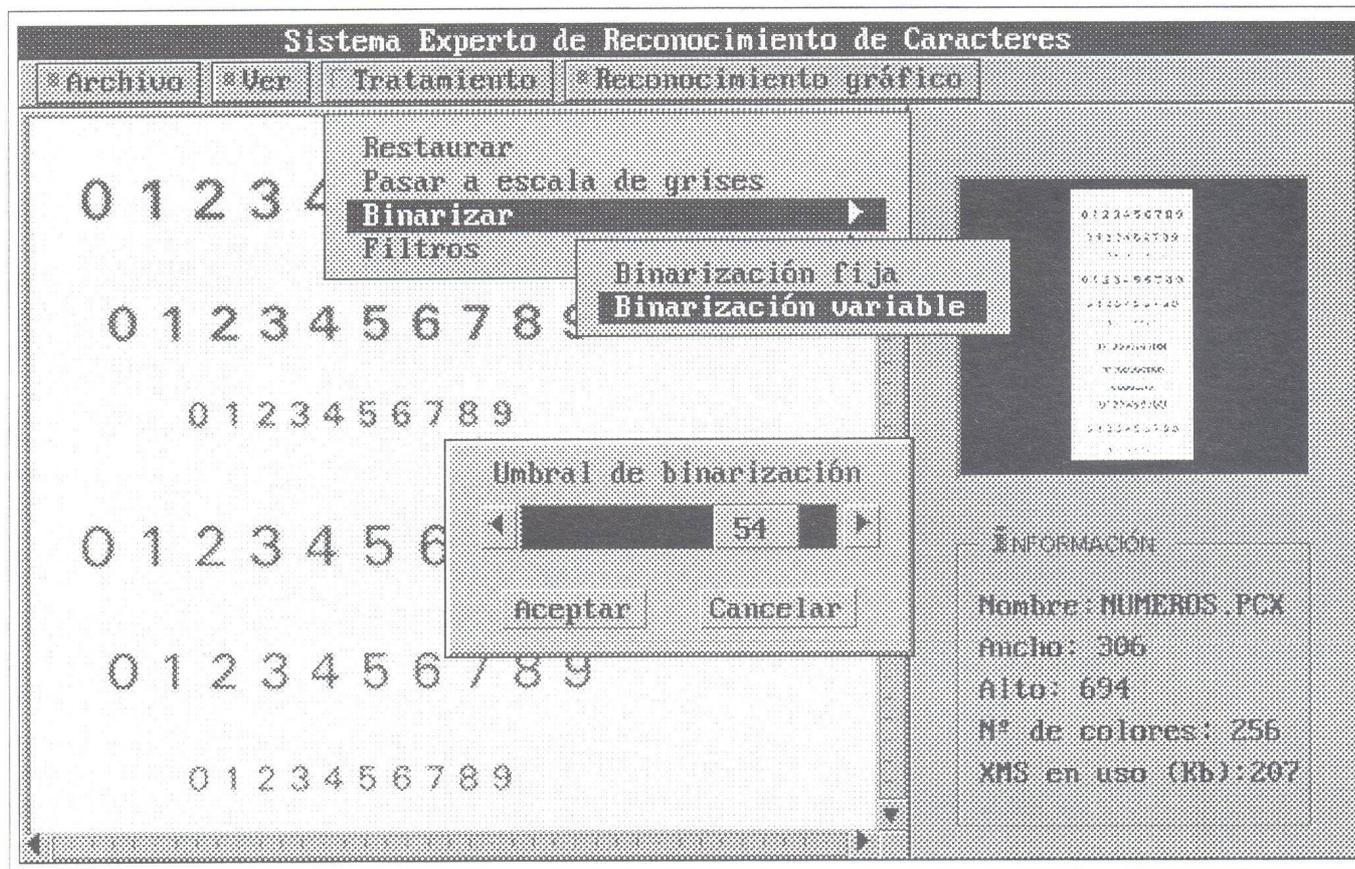


Figura 3. Pantalla del Módulo de Tratamiento

cerlo. Los descriptores deben ser independientes del tamaño, la localización y orientación del objeto y deben contener la suficiente información para distinguir un objeto de otro. La descripción es un tema central en el diseño del sistema, ya que no sólo afecta a la complejidad de las reglas de reconocimiento, sino también a su rendimiento. Los descriptores utilizados han sido: Descriptores de frontera, Descriptores de región y Puntos significativos.

Como descriptor de frontera se ha creado el código de cadena, empleado para representar una frontera como un conjunto de segmentos con longitud y dirección especificadas, de modo que a partir de la detección de bordes del objeto, se codifica la frontera empleando los códigos de dirección (N, S, E, O, NE, etc.).

Como descriptores de región se han utilizado el área, que se define como el número de puntos contenidos en su frontera, los ejes mayor y menor del objeto, así como el cociente de ambos (excentricidad), el perímetro o longitud de su frontera, la compacidad, que se define como $\text{perímetro}^2/\text{área}$, además de las áreas parciales y algunos más.

El esqueleto de una región es el resultado de aplicar sobre ella un algoritmo de adelgazamiento (esqueletización) que permita reducirla a un grafo o eje intermedio continuo; sobre

el esqueleto se calculan los puntos significativos del objeto, como son los puntos finales, puntos de corte y puntos de ruptura o de cambio de dirección, así como la posición relativa de ellos.

Cuanto más completa es la descripción del objeto, más posibilidades tiene el sistema de reconocimiento de discriminarlo, si bien el uso de descriptores debe limitarse a aquellos que contengan suficiente información sobre la particularidad del mismo. En la **figura 4** se observa la información que el sistema obtiene tras la descripción de un objeto.

4.3. El lenguaje de descripción de caracteres

A partir de la información recopilada en la descripción se definen los términos de un lenguaje descriptivo que el sistema de reconocimiento utilizará para su decisión. El mencionado lenguaje es próximo a un lenguaje natural [GON-93], de tal forma que completar o modificar las reglas de reconocimiento sea tarea fácil para el usuario.

Términos como *CompruebaCurva*, *CompruebaRecta*, *PosiciónInicialRecta*, *NúmeroCurvasDerecha* o *AreaCerrada*, forman parte de este lenguaje de conexión con el módulo de reconocimiento. Estos términos permiten canalizar toda la información acerca del objeto, de tal forma

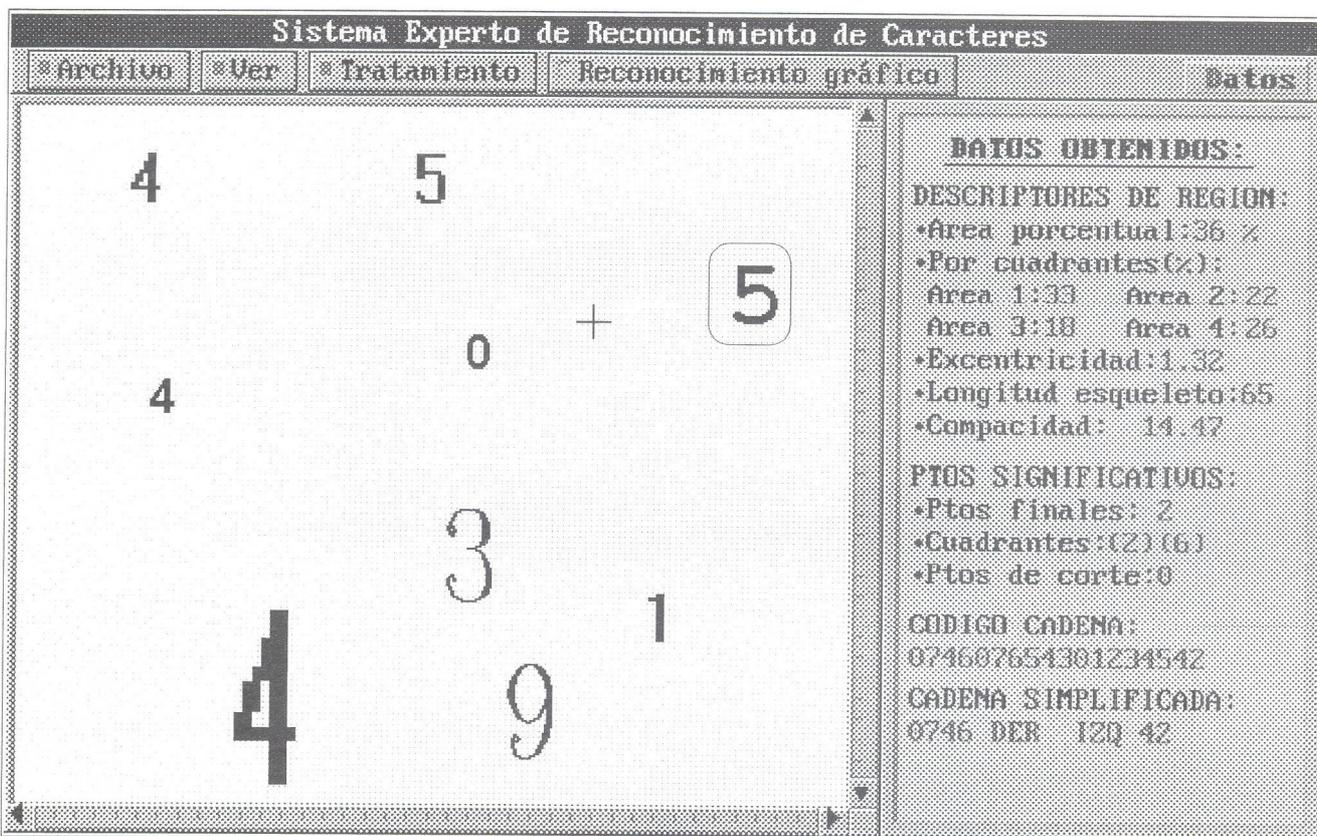


Figura 4

que imiten la descripción que una persona haría del mismo empleando los mismos descriptores.

En la programación del código de reconocimiento se hace uso de este lenguaje para, en el caso que nos ocupa, reconocer los caracteres; se pueden comprobar puntos de corte y puntos finales por cuadrantes en los esqueletos de los objetos, posiciones iniciales o finales de rectas y curvas, verificar que existen áreas cerradas, contar puntos característicos, evaluar áreas totales y parciales, comprobar y comparar la composición de cadenas, sus tamaños y sus tipos, etcétera; en definitiva, permite al sistema de reconocimiento acceder a toda la información descrita en los módulos anteriores.

5. Sistema de reconocimiento

El reconocimiento se plantea como un proceso de etiquetado, esto es, la función del Sistema Experto es identificar cada objeto segmentado de una imagen y asignarle una etiqueta. En la mayoría de los sistemas de visión, al igual que en éste, los niveles de reconocimiento trabajan sobre objetos que suponen han sido segmentados como unidades individuales.

Para llevar a cabo el reconocimiento se ha elaborado una base de conocimiento introduciendo reglas de experiencia,

en la que cada objeto (en este caso caracteres numéricos) queda definido por un conjunto de hechos en términos del lenguaje que fue descrito en la sección anterior. Como se comentó en la introducción, la combinación de técnicas de tratamiento algorítmico con un módulo de reconocimiento basado substancialmente en reglas de producción permite disponer de un mecanismo de programación de nuevas reglas modular y flexible, reduciendo además el código necesario (en torno a un factor de 10 a 1).

Utilizando los resultados ofrecidos por el sistema de descripción se han elaborado las reglas que permiten reconocer los caracteres numéricos objeto del trabajo; cada uno de los números ha sido definido por un conjunto de reglas que sopesan la información que el sistema va obteniendo del objeto gráfico en los módulos anteriores. Así, se asimilan una serie de hechos experimentales, que determinarán la actuación en los distintos estados del proceso. Las reglas emplean términos del lenguaje definido en el módulo de descripción, lo que, como se comentó con anterioridad, permite hacer una aproximación lingüística para el estudio de los objetos gráficos de la imagen. Cuando el sistema ha tomado una decisión, ésta queda registrada, al igual que todos los datos descriptivos del objeto, de esta forma podrán incorporarse nuevas reglas que ayuden al sistema al reconocimiento del mismo.

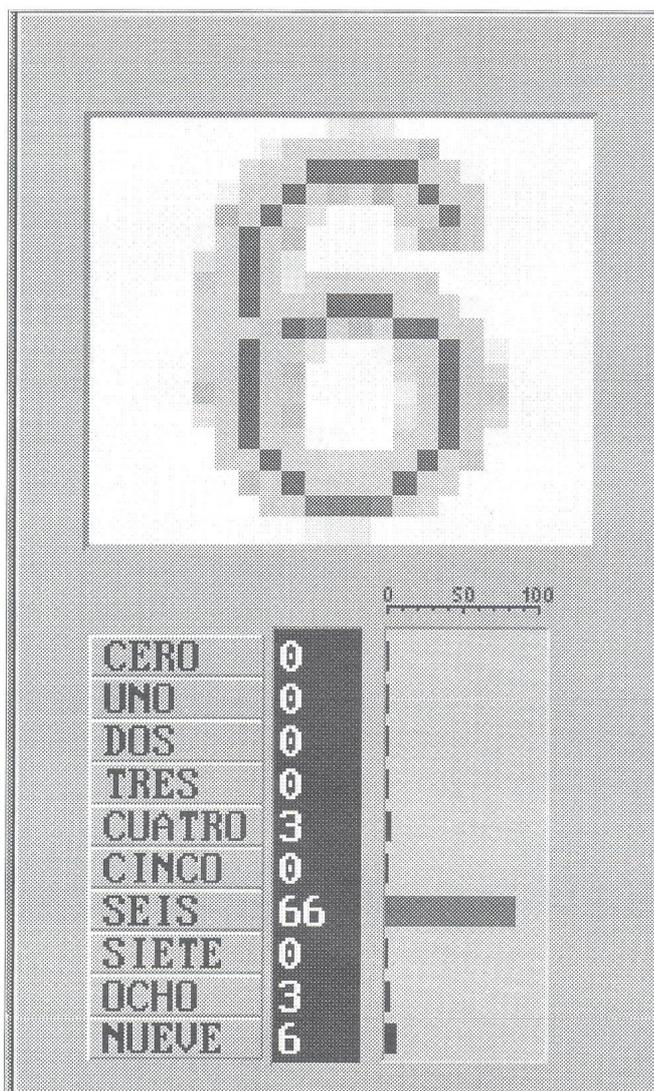


Figura 5

5.1. Las Reglas para el Reconocimiento de Caracteres

Para la programación de las reglas del sistema se ha empleado como soporte un lenguaje denominado DSL [RTE-94]; un ejemplo de las reglas programadas en él puede verse en la **figura 6**.

La sintaxis básica de una regla de producción en este lenguaje es del tipo "IF...THEN...ELSE" [DYM-91]. La parte izquierda de la regla (IF) se denomina antecedente y puede consistir en una o más condiciones; la parte derecha (THEN) es el consecuente, donde se ejecutan las acciones cuando son verificadas las condiciones. Como ejemplo puede verse la siguiente regla, extraída del código DSL del programa de reconocimiento, en concreto del grupo de reglas que definen al dígito cero. En ella se aprecia el uso del lenguaje descriptivo en el antecedente, una función que calcula el número de curvas a derechas que tiene el objeto, otra que comprueba si contiene áreas cerradas y dos variables que contienen el tamaño del código cadena simplificado y

la cuenta de puntos de corte del esqueleto. Si se cumple la combinación de estas condiciones la regla provoca el aumento en un factor decisivo del peso que el sistema daría al número cero. La programación de las reglas correspondientes a cada carácter se ha realizado de forma genérica para que abarque el mayor número de formas posibles del mismo; las excepciones son fácilmente agregadas a la base de conocimiento en forma de nuevas reglas.

6. Resultados obtenidos

Los resultados obtenidos en las diferentes pruebas a las que se ha sometido el sistema son prometedores, los porcentajes de acierto en algunos documentos tipográficos alcanzan 100%, sin embargo, tratándose de caracteres numéricos en los que hay diversidad de formas para un mismo número según la fuente de texto empleada, estos porcentajes fluctúan en torno al 90%.

Las pruebas han sido realizadas a partir de los juegos de

```

--uno *****
IF area3=0
    THEN uno:=uno+DECISIVO;
IF area1<20 AND area2>38 AND area4>35
    THEN uno:=uno+NORMAL;
IF CompruebaCorte(8)=1 AND
    (CompruebaCorte(3)+CompruebaCorte(4)+CompruebaCorte(5)=0)
    AND (CompruebaCorte(0)>0 OR CompruebaFinal(0)>0)
    OR CompruebaCorte(1)>0 OR CompruebaFinal(1)>0
    OR CompruebaCorte(2)>0 OR CompruebaFinal(2)>0)
    THEN uno:=uno+IMPORTANTE;

IF (CompruebaRecta(6,6,6)+CompruebaRecta(2,3,3))>tamano_cadena/2
    AND tamaño_codigo > 1
    THEN uno:=uno+DECISIVO;

--cinco *****
IF (area1-area2)>6 AND (area4-area3)>4 THEN cinco:=cinco+IMPORTANTE;

IF contadorpc=0 THEN cinco:=cinco+NORMAL;
IF contadorpf=2 THEN cinco:=cinco+NORMAL;
IF CompruebaFinal(2)=1 AND CompruebaFinal(6)=1
    THEN cinco:=cinco+IMPORTANTE;

IF CompruebaRecta(2,2,2)>0 AND PosicionFinalRecta(2,2,2)=tamano_cadena
    AND (CompruebaCurva(DER,0)=3 OR CompruebaCurva(DER,0)=2)
    AND PosicionInicialCurva(DER,0)<tamano_cadena_simple/2
    AND PosicionFinalCurva(DER,0)>tamano_cadena_simple/2
    AND contadorpc=0 AND contadorpf=2
    THEN cinco:=cinco + DECISIVO;

-- IF CompruebaCurva(DER,0)=3
    THEN cinco:=cinco+IMPORTANTE;
-- IF CompruebaCurva(ZQ,7)=5
    THEN cinco:=cinco+IMPORTANTE;

```

Figura 6

caracteres más extendidos en el mercado, generándose tanto documentos con los números alineados en varias filas, como otros donde estaban salteados sin orden alguno, la división parcial se realizó sin problemas y el reconocimiento fue rápido en un ordenador PC 386 (un carácter se reconoce de forma prácticamente instantánea). En el caso de documentos caligráficos la respuesta del sistema sigue siendo más que aceptable, pues los porcentajes se sitúan en torno al 70%, dependiendo, sobre todo, de la calidad del número caligráfico digitalizado.

7. Conclusiones

El objeto del trabajo ha sido proporcionar una solución más al reconocimiento automático de caracteres. El proceso de reconocimiento se ha basado en técnicas de sistemas expertos, desarrollando patrones de cada objeto mediante reglas de experiencia y a través del apoyo de un lenguaje descriptivo que posibilita una comunicación eficiente y adecuada entre el usuario y el sistema de reconocimiento.

Los resultados obtenidos estimamos que resultan satisfactorios, planteándonos en un futuro la extensión de las reglas de reconocimiento para abarcar todo el conjunto de caracteres ASCII.

En cualquier caso es preciso destacar que el sistema resulta lo suficientemente flexible como para permitir su aplicación a cualquier tipo de objeto individual sin más que codificar las reglas adecuadas.

8. Referencias

- [DYM-91] C. L. Dym y R. E. Levitt. *Knowledge Based Systems in Engineering*. McGraw-Hill, 1991.
- [FU-88] K.S. Fu, R.C. González, C.S. Lee. *Robótica: Control, Detección, Visión e Inteligencia*. McGraw-Hill, 1988.
- [GAL-90] L. J. Galbiati. *Machine Vision and Digital Image Processing Fundamentals*. Prentice-Hall International, 1990.
- [GOM-96] P. E. Gómez. "Proceso de Formularios". *Novática*, pp 8-13, Sep.-Oct. 1996.
- [GON-93] A. J. González y D. D. Dankel. *The Engineering of Knowledge-Based Systems. Theory and Practice*. Prentice-Hall International, 1993.
- [NIB-86] W. Niblack. *An Introduction to Digital Image Processing*. Prentice-Hall International, 1986.
- [TIS-88] M. Tischer. *PC Interno*. Marcombo, 1993.
- [RTE-94] *RT-Expert Manual*. Real Time Intelligent System Corporation, 1994.
- [VER-91] D. Vernon. *Machine Vision*. Prentice-Hall International, 1991.