

Sistemas Híbridos para el Reconocimiento de Patrones

Román Katz y Claudio Delrieux

*Universidad Nacional del Sur, Alem 1253, (8000) Bahía Blanca, ARGENTINA.
Voice: (54)(291)5710098 — Fax: (54)(291)4595154 — e-mail: claudio@acm.org*

Palabras Clave: Procesamiento de Imágenes, Reconocimiento de Patrones, Inteligencia Artificial, Programación Lógica, Visión Robótica.

1 Introducción

El Procesamiento Digital de Imágenes (PDI) es un área de permanente importancia tecnológica. Sus objetivos van desde procesar datos adquiridos satelitalmente para mejorar la percepción, detección o interpretación de algún patrón específico [5, 6], o aplicar filtros a imágenes fotográficas para reconstruir o retocar sus características visuales [3, 4], o comprimir información gráfica para facilitar su transporte por las redes de comunicaciones [7, 2]. Entre las áreas de mayor importancia tecnológica del PDI podemos mencionar el reconocimiento de patrones y la visión robótica, en las cuales se requiere el uso combinado de técnicas de todo tipo, desde las puramente numéricas hasta las más cercanas a la inteligencia artificial. Actualmente la mayor parte de las técnicas avanzadas utilizadas en el reconocimiento de patrones se basa en un procesamiento numérico, donde mediante un tratamiento heurístico se busca conjugar los resultados de ciertos estimadores con valores específicos de una distribución [8, 10].

Las técnicas asociadas están adaptadas para cada problema en particular con lo que se logran excelentes rendimientos en tareas específicas, pero su posibilidad de aplicación satisfactoria en otras áreas es muy reducida. En tal sentido el procesamiento simbólico, en donde se busca reconocer la presencia de determinados patrones para disparar luego un proceso de interpretación, se presenta como una alternativa posible. Es indudable que ambos tipos de procesamiento son indispensables en aplicaciones tecnológicas como las mencionadas. En un sistema de visión robótica, por ejemplo, la mayor parte del procesamiento debe realizarse en forma numérica para acomodar las imágenes de entrada de manera de poder facilitar la ejecución de los algoritmos de alto nivel que puedan requerirse (reconocimiento de patrones, identificación, etc.). Sin embargo, no tiene sentido práctico implementar estos últimos desde el bajo nivel, así como tampoco parece adecuado implementar los algoritmos fuertemente numéricos de bajo nivel en un lenguaje de alto nivel.

Por lo tanto, la solución tecnológica parece ser la de construir una arquitectura de software que integre simultáneamente en una misma aplicación el procesamiento numérico de bajo nivel en una plataforma adecuada (C++, Delphi), con la implementación de algoritmos de inteligencia artificial implementados en los lenguajes específicamente desarrollados (Lisp,

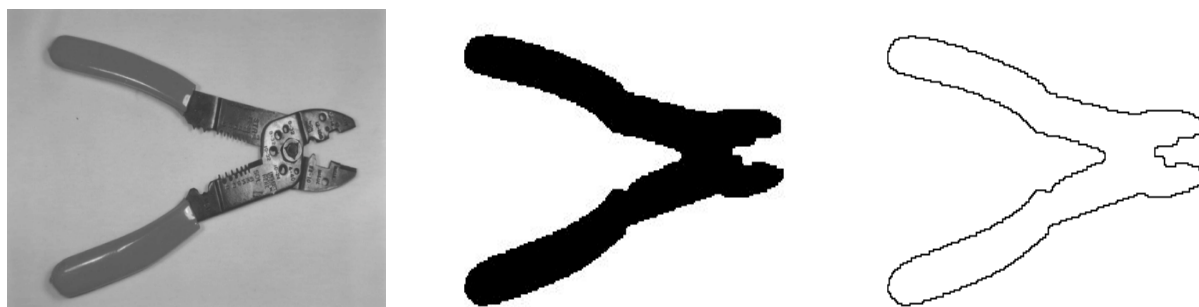


Figura 1: (a) Imagen en escala de grises, (b) umbralizada, y (c) segmentada.

PROLOG). Es este último el enfoque metodológico adoptado para este trabajo, en el cual presentamos una aplicación implementada en Delphi, donde se conjugan las técnicas de procesamiento numérico para el procesamiento de bajo nivel de las imágenes, con las de alto nivel implementadas en el lenguaje PROLOG. El objetivo en esta etapa de nuestro desarrollo es implementar un sistema de reconocimiento de objetos en imágenes digitales.

2 Procesamiento numérico

Los primeros pasos del tratamiento que recibe una imagen luego de su adquisición se compendian bajo la denominación de preprocesado y son tendientes a producir una mejora en alguna de sus características, de forma de incrementar el rendimiento de las etapas posteriores [5]. En general son operaciones que requieren algoritmos simples pero de gran insumo de operaciones numéricas en su ejecución, entre los que se destacan conversión a escala de grises, filtrado, umbralización, escalamiento, extracción de regiones de interés, etc. [10]. Un paso posterior a la mejora de la imagen es el de la segmentación. Aquí se busca simplificar la imagen reduciéndola a un subconjunto predeterminado de objetos básicos (puntos, segmentos u otras primitivas geométricas sencillas), con el objeto de facilitar luego un procesamiento de un nivel superior.

En la etapa siguiente, la descripción de la imagen, se genera una representación útil para el problema de interpretación de interés a partir de la información proveniente de las operaciones de filtrado y segmentación previas. Existe una gran variedad de descriptores [5] entre los que se pueden encontrar los relacionados con características de forma, de Fourier, momentos estadísticos, códigos de cadena, segmentos del contorno, textura, etc. Bajo estas condiciones una alternativa razonable para nuestro problema en particular es utilizar los segmentos del contorno como descriptor, considerando los determinados por puntos regularmente espaciados y obteniendo de esta forma una representación del perímetro del objeto con una cantidad de información simplificada. Como se observa en la Fig. 2, aún reduciendo la cantidad de vértices para describir el contorno a porcentajes muy bajos (hasta un 25% o menos de los vértices originalmente segmentados) el objeto se presenta altamente distinguible.

3 Reconocimiento

El paso final del proceso de reconocimiento de imágenes es el de la interpretación. En el mismo se busca descubrir, identificar y comprender la información contenida en los descriptores.

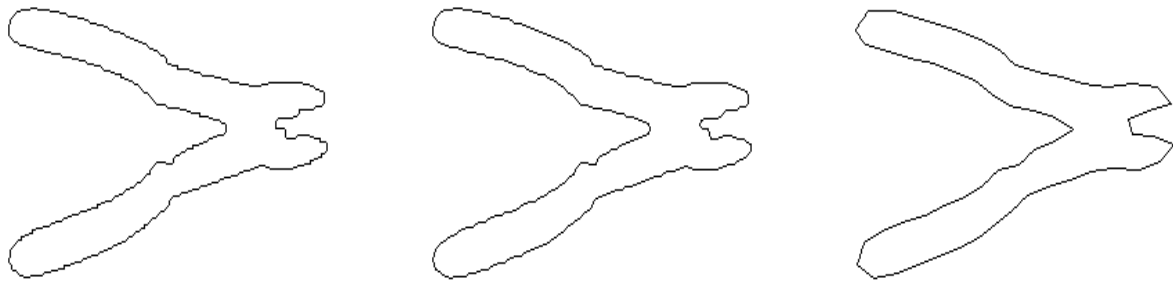


Figura 2: Contorno descrito con diferentes proporciones de los vértices originalmente segmentados: (a) 50%, (b) 25% y (c) 5%.

Se desea además que este grado de inteligencia, representado por una base de conocimientos apropiada, tenga un alto nivel de independencia con respecto a la aplicación, de manera tal que esta experiencia pueda servir en un futuro para otro tipo de reconocimientos. Expresar nuestro conocimiento mediante la lógica de predicados, específicamente utilizando el lenguaje PROLOG [9], nos permitirá verificar la condición de independencia buscada. Su naturaleza declarativa facilitará notablemente la tarea de expresar tanto formas geométricas complejas [1] como relaciones estructurales y posicionales entre las mismas. La representación de conocimiento y de los patrones de inferencia utilizando lógica es lo suficientemente modular como para poder testear y modificar las diferentes partes del sistema independientemente unas de otras, lo cual también garantiza su portabilidad y reuso en futuras aplicaciones. El uso de PROLOG facilita además el desarrollo del sistema de razonamiento o “motor” de inferencia, tanto para implementar los patrones deductivos como para realizar otro tipo de inferencias ampliativas en general (abducción, razonamiento por *default*, razonamiento por intervalos de confianza, etc.).

La formulación lógica de los patrones a reconocer se expresa mediante un *modelo de detección* genérico. Dicho modelo es una estructuración de reglas que reduce el costo de explicitación en predicado de nuevos patrones. De esta forma posibilita insertar el predicado que define a una estructura geométrica dada en una “plantilla” conocida, impidiendo que se repitan resultados y que se ejecuten *backtracking* innecesarios. Dicha plantilla tiene una estructura esquemática como se observa en la Tabla 1. Manteniendo esta estructura, sólo se deben elaborar las “relaciones geométricas” del patrón, es decir la formulación en forma de predicado del elemento particular a detectar.

```

patron geometrico (patron):-
    lista segmentos que verifican patron
    cut[!] (para que la lista sea única)
    patrón está en lista

lista segmentos que verifican patron:-
    relaciones geométricas
    verifica que no sea repetido
    agrega a lista
  
```

Tabla 1: Modelo de Detección.

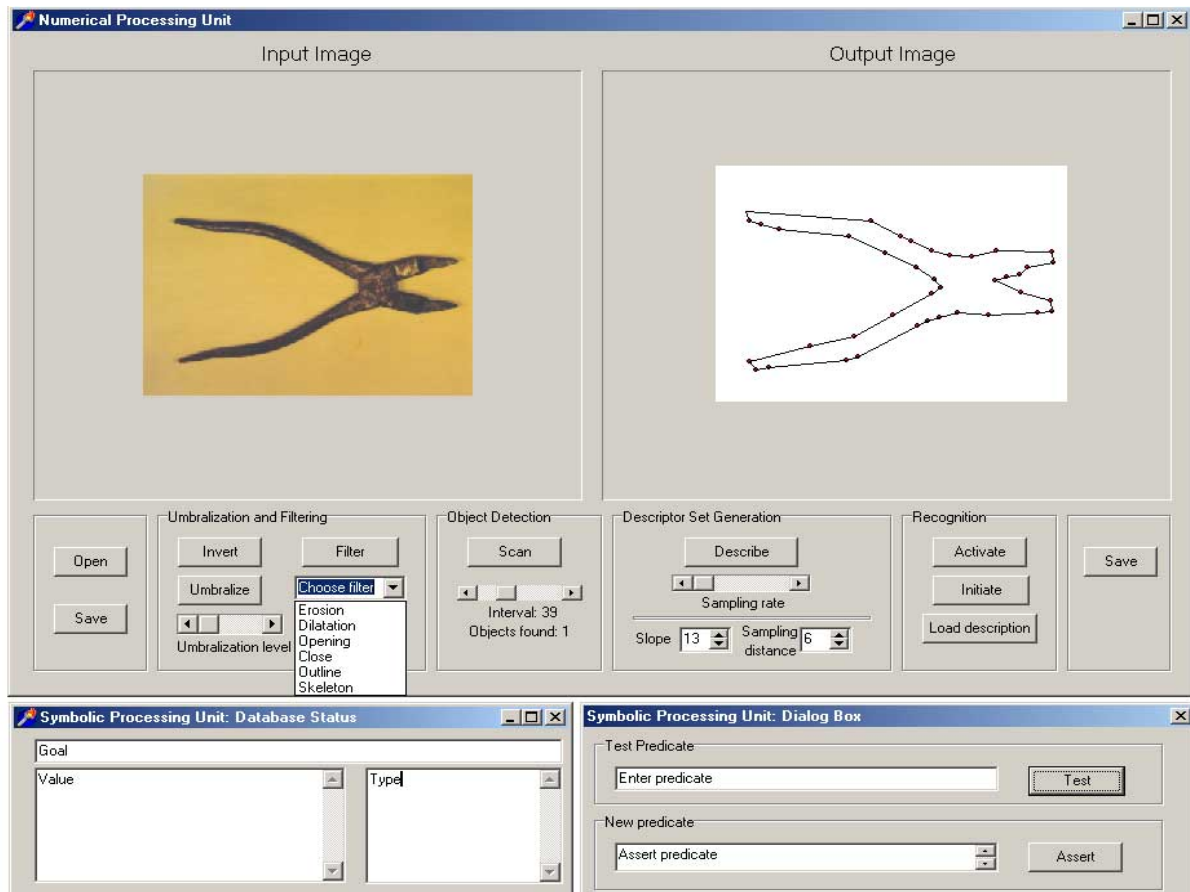


Figura 3: Interfase gráfica del sistema de reconocimiento.

La implementación de estos conceptos se realizó mediante el desarrollo de un sistema para la identificación de patrones mediante su contorno o su esqueleto. En el sistema se conjuga el procesamiento de imágenes de forma numérica con la utilización de un mecanismo de inferencia adecuado para el reconocimiento. El procesamiento numérico genera una representación adecuada de la imagen mientras que el “motor” de inferencia busca cotejar la información contenida en dicha representación con una base de conocimientos estipulada. Se observa en la Fig. 3 la interfase gráfica del sistema, con los comandos asociados a las etapas antes descritas contenidos en la ventana de **Procesamiento Numérico** y con una imagen de entrada arbitraria, junto con la de su descripción final asociada. En el proceso de reconocimiento, el mecanismo de inferencia se obtiene mediante el manejo de una componente de software que contiene una librería DLL del lenguaje PROLOG. La librería de PROLOG requiere una inicialización apropiada, luego de la cual el sistema queda preparado para recibir consultas. En nuestro caso se realiza, junto con la inicialización, la carga de una base de conocimientos estipulada a través de archivos. La misma contiene la formulación lógica de las relaciones geométricas que conforman el patrón que se desea detectar y los predicados apropiados para generar una base de datos a partir de los segmentos que provienen de la descripción del objeto. Las consultas se realizan a través de la ventana **Consola de consultas**, en la que también se pueden generar nuevos predicados para ser testeados y posiblemente agregados a la base de conocimiento. La ventana **Consola base de conocimientos** muestra los resultados obtenidos a partir de las consultas, incluyendo los

tipos y valores finales de las variables involucradas en la misma. Ambas consolas se pueden ver en la Fig. 3.

4 Conclusiones y trabajo futuro

Se presentó un sistema integrado de reconocimiento de patrones que utiliza simultáneamente algoritmos básicos de tratamiento numérico de imágenes y técnicas de reconocimiento simbólico en lenguaje PROLOG. La complementación de procesamiento de bajo y alto nivel en una plataforma de programación común rescata las capacidades salientes de ambos enfoques. Nos permite conjugar la excelente respuesta en velocidad de cálculo, la disponibilidad de librerías gráficas y la versatilidad de implementación de los sistemas en entornos *RAD (Rapid Application Development)* con las facilidades de interpretación del lenguaje PROLOG. La posibilidad de inferir resultados a partir de información incompleta, de “explicitar” el conocimiento de una forma simple, natural y altamente portable y reciclable y la robustez frente a variaciones posicionales y de orientación hacen que dicho mecanismo de identificación resulte ideal para el problema de detección patrones en multiles aplicaciones. El sistema mostrado permite la adaptación de diversas técnicas numéricas de procesamiento adicionales y la ampliación y reemplazo de la base de conocimientos para problemas de detección particulares.

Entre los trabajos futuros se considera la posibilidad de realizar una extensión del sistema para contemplar esquemas de descripción alternativos para los objetos. Se contempla también la incorporación de *meta-interpretés* en el proceso de reconocimiento, realizando un seguimiento en el mecanismo de inferencia de forma de obtener información particular del objeto y “niveles” lógicos que cuantifiquen el éxito obtenido en la tarea. Esto permitiría, entre otras cosas, darle un manejo operacional a las situaciones de reconocimiento parcial, de manera de poder elaborar razonamientos tentativos, explorarlos modificando el procesamiento numérico, y proveyendo respuestas basadas en intervalos de confianza.

Referencias

- [1] B. Batchelor. *Intelligent Image Processing in Prolog*. Springer-Verlag, London, 1991.
- [2] K. Castleman. *Digital Image Processing*. Prentice-Hall, New York, 1989.
- [3] A. Glassner. *Principles of Digital Image Synthesis*. Morgan Kaufman, 1995.
- [4] J. Gomes y L. Velho. *Image Processing for Computer Graphics*. Springer, 1997.
- [5] R. González y R. Woods. *Digital Image Processing*. Addison-Wesley, 1996.
- [6] Anil Jain. *Fundamentals of Digital Image Processing*. Prentice-Hall, Cambridge, 1996.
- [7] Jae Lin. *2D Signal and Image Processing*. Prentice-Hall, Cambridge, 1991.
- [8] Arie Roszenfeld. *Digital Picture Processing*. MIT Press, Cambridge, MA, 1995.
- [9] L. Sterling y E. Shapiro. *The Art of Prolog*. MIT Press, 1994.
- [10] S. Umbaugh. *Computer Vision and Image Processing*. Prentice-Hall, 1998.