

# Estudio de descriptores geométricos para el diseño y optimización de algoritmos de reconocimiento de objetos digitales

María Lorena Bergamini, Jorge Kamlofsky

CAETI – Facultad de Tecnología Informática. Universidad Abierta Interamericana

Av. Montes de Oca 745. Ciudad de Buenos Aires

(+54 11) 43015323; 43015240 ; 43015248

[Maria.Bergamini@uai.edu.ar](mailto:Maria.Bergamini@uai.edu.ar); [Jorge.Kamlofsky@uai.edu.ar](mailto:Jorge.Kamlofsky@uai.edu.ar)

## Resumen

En este proyecto se estudian, diseñan e implementan herramientas de análisis de imágenes orientadas al desarrollo de algoritmos para reconocimiento y clasificación de objetos en imágenes digitales.

Los objetos se pueden reconocer por su forma, y éstas se identifican definiendo sistemas de representación apropiados y métricas acordes, que pretenden ser de fácil y económica generación y manejo, a fin de utilizarse en tiempo real.

Se estudian representaciones compactas de formas que priorizan la reducción en la cantidad de datos a tratar, sin perder información acerca del objeto que se está describiendo. Dichas representaciones deben ser invariantes frente a escalado, rotación y transformaciones afines; ya que las imágenes de objetos no siempre son tomadas desde el mismo ángulo, distancia y posición.

**Palabras clave:** visión artificial, invariantes geométricos, curvatura digital, rectas digitales, reconocimiento de formas

## Contexto

El tema de investigación que abarca este proyecto está siendo estudiado por el

grupo desde 2012. Se desarrolla en el Centro de Altos Estudios en Tecnología Informática (CAETI), dependiente de la Facultad de Tecnología Informática de la Universidad Abierta Interamericana (UAI).

Los proyectos de investigación que se desarrollan en el CAETI se organizan en 5 líneas temáticas identificadas como prioritarias. El proyecto aquí presentado se enmarca en la línea prioritaria de Algoritmos y Software. Esta línea se orienta al diseño de software atendiendo el creciente aumento en la complejidad de las tareas a resolver, manteniendo calidad, tiempos y costos controlados. También se interrelaciona con proyectos de la línea prioritaria Automatización y Robótica. En esta línea se combinan mecánica, electrónica, física, informática, control, e inteligencia artificial para diseñar y manufacturar sistemas robóticos que puedan sustituir al ser humano en ciertas tareas.

## Introducción

El reconocimiento visual de objetos es una capacidad básica con la que cuentan los organismos más desarrollados. La identificación de patrones usando la visión permite realizar tareas como clasificación, vigilancia, manipulación, selección, navegación, etc.

El avance de los desarrollos tecnológicos permitió que esta capacidad se incorpore a dispositivos robóticos diseñados para fines específicos en una amplia variedad de aplicaciones que van desde la asistencia en la industria hasta el desempeño de tareas hogareñas.

Específicamente, entre estas tareas podemos mencionar control de calidad en la línea de fabricación, sistema de monitoreo de un predio para detección de intrusos, sistema automatizado de barreras en un estacionamiento privado, control de tránsito en determinada vía, conteo y seguimiento de objetos en movimiento, digitalización de información proveniente de soporte físico (papel, cartelería, patentes, etc), dactiloscopia, asistencia al diagnóstico médico por imágenes, etc.

La visión artificial en estos dispositivos cuenta con un sistema automatizado de reconocimiento por extracción de rasgos característicos de un objeto, que luego se comparan con una base de datos o con la información aprendida por el sistema incorporado. Una dificultad que surge en la práctica es que el objeto observado puede estar en cualquier posición, con cualquier nivel de brillo y de iluminación, o parcialmente oculto, lo que podría incidir negativamente en la comparación con la base almacenada. Existen investigaciones y avances en el desarrollo de metodologías de reconocimiento que tienen en cuenta estas variantes y responden con buenas tasas de precisión en condiciones controladas.

Para imágenes digitales que representan objetos tridimensionales, la tarea de reconocimiento es desafiante, ya que se cuenta con una proyección planar del mismo, perdiendo una dimensión. Esta representación parcial del objeto se puede enriquecer con proyecciones en distintas direcciones (multicámara), lo

que aporta más información (sin llegar a reproducir completamente el objeto real). Más desafiante aún es la tarea cuando hay problemas de iluminación, oclusión, ruido, etc.

Dado que las imágenes digitales, sobre todo en secuencias de tiempo real, involucran una alta cantidad de información digital, no es razonable aplicar métodos basados en análisis pixel a pixel, que requerirían tiempos computacionales incompatibles con la generación de imágenes en tiempo real. Es fundamental el diseño de algoritmos optimizados que puedan realizar tareas de reconocimiento de objetos manejando eficientemente la mínima información necesaria de la imagen.

El desarrollo de herramientas de soporte para la detección automática de objetos en tiempo real es un área que se ve favorecida por el crecimiento de la capacidad computacional, que permite alta velocidad de procesamiento y el aumento del poder de almacenamiento. En general, los algoritmos se estructuran en dos fases: aprendizaje y testeo. Se han desarrollado diversas estrategias para cada una de estas fases. Sin embargo, ningún algoritmo para reconocimiento de objetos es 100% preciso y óptimo [1].

El reto es mejorar los procedimientos involucrados en visión artificial para lograr desempeños correctos en las distintas aplicaciones.

La *forma* es una característica distintiva de los objetos, y es uno de los principales elementos estudiados para reconocerlos.

Existe un amplio desarrollo en descriptores de forma, que se pueden dividir en aquellos basados en regiones y los basados en el contorno, y a su vez pueden ser estructurales o globales [2]. Para evaluarlos, se aplican en una base de imágenes. Se evalúa, además de la precisión al reconocer formas similares,

su aplicabilidad al caso general, complejidad computacional, performance de respuesta, etc.

Obviamente, los descriptores deben ser invariantes por rotación, escala, transformaciones afines.

Una noción importante en visión artificial es similaridad de forma. Se han definido medidas de similaridad entre formas para aplicar en algoritmos basados en aprendizaje por entrenamiento. Muchos investigadores han propuesto medidas de similaridad, o distancias entre formas, basados en una representación compacta de la misma [3,4,5].

En una imagen binaria, los puntos del objeto que son adyacentes al complemento constituyen el borde del mismo. En imágenes en escala de grises o en color, el borde se identifica con los puntos de máximo gradiente. Este subconjunto es de gran interés en el análisis de objetos dentro de imágenes, ya que muchas características geométricas del objeto (conexión, convexidad, dimensiones, agujeros, puntos angulosos, circularidad, partes rectas, etc.) pueden estudiarse por medio de su borde. El reconocimiento por análisis de bordes requiere eficientes procedimientos de extracción de características geométricas: curvatura, segmentos rectos y zonas de convexidad y concavidad. Estos procedimientos deben ser suficientemente insensibles al ruido presente en imágenes reales, a fin de obtener resultados confiables.

Una tarea central en geometría digital es el reconocimiento de rectas digitales [6]. Varios enfoques del análisis diferencial en geometría digital se basan en reconocimiento de segmentos de líneas rectas [7]. Relacionado con esto, un concepto clave en el estudio geométrico de imágenes digitales es el de curvatura.

Varias generalizaciones de la definición de curvatura proveniente de

geometría diferencial a la geometría digital han sido sugeridas. La definición de curvatura para curvas digitales ha dado lugar a diversos estimadores [8,9,10]. Varios de estos estimadores de curvatura se basan en medir segmentos digitales, lo cual es muy sensible a la presencia de ruido en la imagen. En [11] se trata el ruido en imágenes digitales mediante el concepto de segmentos borrosos. Este concepto se fundamenta en la noción de barra digital [12], que es una generalización de la definición de segmentos digitales.

En aplicaciones en tiempo real (por ejemplo, reconocimiento de intrusos en una cámara de seguridad, control de ingreso y egreso vehicular en rutas) el uso de los recursos de procesamiento y almacenamiento debe ser eficiente. Un enfoque que motiva muchas investigaciones actuales es reducir la cantidad de puntos del contorno que deben ser tenidos en cuenta, quedándose sólo con un conjunto de puntos distintivos. Esto, en general, se efectiviza localizando arcos cuyos puntos están aproximadamente alineados.

Con estos conceptos, en [3] se propone un método de aproximación de curvas digitales por polígonos, y se define una representación compacta de la forma, a través del patrón de curvatura acumulada, lo que disminuye notablemente la cantidad de información a tratar al hacer la comparación de formas para su reconocimiento. Además, se propone una medida de similaridad de formas, basada en la curvatura acumulada.

## **Líneas de Investigación, Desarrollo e Innovación**

El objetivo de este proyecto es generar conocimiento teórico y práctico en el área de análisis de imágenes. En el aspecto

teórico se estudian cuestiones de geometría digital, definición, extracción y cálculo de características geométricas de objetos en una imagen digital. Desde el punto de vista práctico se pretende desarrollar algoritmos que permitan llevar a cabo reconocimiento automático y eficiente de objetos, para ser aplicados a sistemas de visión artificial.

En términos específicos, se profundiza el estudio de las características invariantes de una imagen digital frente a ciertas transformaciones geométricas. Estas invariantes se destinan a definir métodos de representación simplificada y medidas de similaridad.

Se estudian descriptores invariantes frente a transformaciones afines, es decir, descriptores independientes de la proporción.

Se definen medidas de similaridad relacionados con cada uno, o con ciertos conjuntos pertinentes de descriptores.

Los resultados teóricos obtenidos fundamentan un algoritmo de reconocimiento de objetos, que será diseñado e implementado para tal fin.

Finalmente, los algoritmos se implementarán en un sistema de visión artificial en tiempo real, para reconocimiento y clasificación de objetos.

## **Resultados y Objetivos**

Se ha desarrollado un método de representación de formas para el reconocimiento eficiente de objetos. El método se basa en una representación de la evolución de la curvatura a lo largo del borde del objeto. Dicha representación simplificada permite realizar comparaciones entre curvas digitales, de modo de poder obtener una medida de similaridad entre formas, como una distancia entre patrones de giro.

Un concepto central en la aproximación de formas por polígonos es

el de segmento recto digital (SRD). Se han estudiado caracterizaciones algebraicas y sintácticas de SRD [6] y de barras digitales (generalización del concepto de SRD). Este tema está siendo estudiado por un tesista de Licenciatura en Matemática.

Para lograr una simplificación de la curva se aplica el concepto de segmento borroso, obteniéndose una aproximación por polígonos. Este concepto es menos sensible al ruido generado en la captura de la imagen que el de SRD. La aproximación sustituye conjuntos de puntos por segmentos rectos, cuando tales conjuntos tienen un ancho menor que cierta tolerancia.

A partir del polígono aproximante, se genera el patrón de giro, que involucra la idea de la curvatura acumulada. Este patrón contiene toda la información necesaria para extraer características geométricas de la forma analizada, ya que resulta ser invariante frente a traslaciones, simetrías y escalado uniforme.

La invariancia frente a rotaciones se logra identificando puntos de curvatura extrema (sharp points). Esto permite dividir el contorno en partes significativas. Así, el patrón de giro se puede aplicar a tramos parciales para focalizar la comparación en ciertas áreas de las formas. Esta comparación por tramos permite determinar diferencias entre un conjunto de siluetas de objetos de la misma especie, permitiendo así la clasificación de objetos del mismo tipo [13].

Otro enfoque para la poligonalización de una curva digital es mediante la cobertura tangente. Esto implica determinar arcos de curva que son tangentes de todos sus puntos. Estas ideas constituyen el tema de estudio de una tesista de Licenciatura en Matemática.

## **Formación de Recursos Humanos**

Este proyecto se desarrolla bajo la dirección de la Dra. María Lorena Bergamini. Actualmente el equipo se completa con el investigador Licenciado Jorge Kamlofsky, la Licenciada Nora Sarti, la Profesora Carina Hamilton, la Licenciada Luciana Klacherian.

La tesis de grado de Kamlofsky (2011) y la de Klacherian (2014), ambas para alcanzar el título de Licenciado/a en Matemática, se enmarcaron en esta línea de investigación. Además, se están desarrollando dos tesis más para obtener el grado de Licenciado en Matemática (los alumnos Mónica Real y Darío Falasca).

Actualmente Kamlofsky está finalizando su tesis de maestría en Tecnología Informática, en UAI; y comenzará el doctorado en Ingeniería en la Universidad Nacional de Lomas de Zamora.

## Referencias

- [1] Prasad D. *Survey of The Problem of Object Detection In Real Images*. Int. J. of Im. Proc. 6 (6), pp. 441- 466, 2012.
- [2] Zhang D., Lu G. *Review of shape representation and description techniques*. Patt. Recogn. 37, pp 1-19, 2004.
- [3] Latecki L. J., Lakämper R. *Shape Similarity Measure Based on Correspondence of Visual Parts*, IEEE Trans. Pattern Anal. Mach. Intell 22(10), pp. 1185-1190, 2000.
- [4] Ling H., Jacobs D. *Shape Classification Using the Inner-Distance*, IEEE Trans. Pattern Anal. Mach. Intell. 29(2), pp. 286-299, 2007.
- [5] Kamlofsky J., Bergamini M.L. *Patrón de Evolución de Angulo de Giro para el Reconocimiento de Objetos en Imágenes Digitales*, JAIIO, 2013.
- [6] Ucska-Wehlou H. *Digital lines, Sturmian words and continued fraction*. Uppsala Dissert. in Math. 65, 2009.
- [7] Lachaud J-O., Vialard A., Vieilleville F. *Fast, accurate and convergent tangent estimation on digital contours*. Im. Vision Comput. 25(10), pp. 1572–1587, 2007.
- [8] Kerautret B., Lachaud J.-O. *Curvature Estimation along Noisy Digital Contours by Approximate Global Optimization*. Patt. Recogn. 42(10), pp. 2265 – 2278, 2008.
- [9] Liu H., Latecki L., Liu W. *A Unified Curvature Definition for Regular, Polygonal, and Digital Planar Curves*. Int. J. Comput. Vis. 80, pp. 104–124, 2008.
- [10] Roussillon T., Lachaud J-O. *Accurate Curvature Estimation Along Digital Contours With Maximal Digital Circular Arcs*. In: Aggarwal J., Barneva R., Brimkov V. Koroutchev K., Korutcheva E. (eds.) *Combinatorial Image Analysis*. LNCS, 6636, pp. 43-55, Springer, Heidelberg, 2011.
- [11] Nguyen T.P., Debled-Renneson I. *Curvature Estimation in Noisy Curves*. In: Kropatsch W., Kampel M., Hanbury A. (eds.) *Computer Analysis of Images and Patterns*, LNCS, 4673, pp. 474-481, Springer, Heidelberg, 2007.
- [12] Klette R., Rosenfeld A. *Digital Straightness – a review*. Discrete Appl. Math. 139, pp. 197-230, 2004.
- [13] Kamlofsky J.; Bergamini M. L. *Herramientas de Análisis de Bordes para Reconocimiento y Clasificación de Objetos en Imágenes Digitales*. CoNaIISI 2014.