

# Transformaciones geométricas para facilitar la identificación de objetos en imágenes digitales

Jorge Kamlofsky, María Lorena Bergamini

CAETI – Facultad de Tecnología Informática. Universidad Abierta Interamericana

Av. Montes de Oca 745. Ciudad de Buenos Aires

(+54 11) 43015323; 43015240 ; 43015248

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

## Resumen

La visión artificial por reconocimiento de patrones es un concepto que permite dar visión de mayor calidad, ver en tiempo real y seguir objetos. Dependiendo de su posición, orientación o tamaño, un objeto puede generar millones de imágenes diferentes, lo que dificulta su identificación.

Las características topológicas de un objeto pueden estudiarse por medio de un análisis de los bordes. Convexidad, concavidad, curvatura, puntos extremos pueden detectarse siguiendo los puntos frontera del objeto, y en su conjunto permiten identificar el mismo.

Estos patrones del borde tienen relación con su ubicación dentro del borde del objeto, y también con la inclinación y escala del mismo. Dotando a una imagen digital de una topología y aprovechando las propiedades invariantes, se estudian estrategias de reconocimiento de formas.

## Palabras claves

Topología digital, visión artificial, invariantes geométricos, reconocimiento de patrones.

## Contexto

El proyecto presentado comenzó en agosto de 2011. 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. Con mejores herramientas, se alcanza mayor complejidad del software que se construye y son más demandantes los requerimientos que debe cumplir. Es esperable que la tendencia en cuanto al ritmo de crecimiento de la complejidad del software y la dependencia que las empresas, el estado y la sociedad en general tienen sobre los servicios provistos por software continuará en las próximas décadas.

En definitiva, la investigación, transferencia, innovación y capacitación en el área de algoritmos y software es fundamental para poder acompañar las necesidades de producción, operación y mantenimiento de software en los próximos años.

## Introducción

El reconocimiento de patrones específicos para la identificación de un objeto dentro de

una imagen hace que el proceso de visión artificial sea mucho más eficiente, permitiendo incrementar la calidad de visión, trabajar en tiempo real o bien reducir los requisitos tecnológicos para crear un sistema de visión. El desarrollo de esta tecnología abre nuevas ventanas en los procesos de automatización: identificación y conteo de vehículos, seguimiento de objetos en movimiento, detección de fuego sin utilización de sensores de temperatura, diferenciación de formas humanas y animales en alarmas perimetrales, lectura de chapas patentes, dactiloscopia, diagnóstico médico por imágenes, 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 el tiempo real. Es fundamental el diseño de algoritmos eficientes que puedan realizar tareas de reconocimiento de objetos manejando sólo la información necesaria de la imagen. Tratar a ciertos píxeles que conforman un objeto dentro de una imagen como un conjunto, y detectar sus propiedades topológicas es un enfoque simple que ha demostrado ser altamente efectivo cuando se pretende interpretar automáticamente el contenido de una imagen.

Las aplicaciones de visión artificial se sostienen en el concepto matemático de topología digital (Eckhardt, Latecki, 2003). La topología digital ha demostrado ser un marco teórico adecuado para el análisis de objetos dentro de imágenes digitales, desde su concepción a inicios de la década de 1970. En este contexto se aplica el *Teorema de curvas de Jordan* en el plano digital, que afirma que toda curva simple cerrada contenida en un conjunto conexo separa al conjunto en dos subconjuntos: su interior o agujero y su complemento o fondo. Este teorema requiere las nociones topológicas de vecindad, adyacencia, camino, conectividad, arco y curva. El teorema de curvas de Jordan permite separar un objeto del resto de la imagen digital en la que se encuentra inmerso. Los trabajos en el área de topología digital y geometría digital de Eckhardt y Latecki (1994) y Latecki

et al. (2000) han marcado una clara guía en esta línea.

En la práctica, se aplica un algoritmo para demarcar la curva que representa el borde o frontera de un objeto dentro de una imagen digital. El algoritmo permite localizar los puntos de un conjunto conexo que limitan con su complemento y los del complemento que limitan con el conjunto, conformando así la frontera del mismo (Kamlofsky, 2011).

Gross y Latecki (1995) proponen un conjunto de patrones de frontera que no se deforman en el proceso de digitalización, y demuestran que, bajo ciertas condiciones, son los únicos patrones locales que pueden ocurrir. Esto permite estudiar propiedades geométricas del objeto digitalizado (convexidad, curvatura) y detectar ruido.

Un concepto clave en el estudio geométrico de imágenes digitales es el de curvatura. Se han sugerido varias generalizaciones de la definición de curvatura proveniente de geometría diferencial a la geometría digital. En particular, en Liu et al. (2008) se plantea medir la curvatura en imágenes digitales basándose en una función altura medida con una rotación de ejes. Este nuevo concepto coincide con la definición estándar de curvatura en curvas suaves cuando el número de alturas es suficientemente alto, y coincide con el concepto de ángulo de giro en una curva poligonal.

La tarea siguiente consiste en definir qué tipo de objeto tiene a esa curva por borde, analizando las propiedades geométricas del mismo. Éstas se pueden comparar con ciertos patrones previamente recopilados, identificando de este modo al objeto en cuestión.

Pero estos patrones están relacionados con la posición, ángulo de inclinación y tamaño del objeto. Pueden tenerse patrones almacenados en memoria, en diferentes posiciones, tamaños o inclinaciones, o bien, más lógica y eficientemente, puede guardárselos en una única posición, tamaño e inclinación.

La tecnología de visión por reconocimiento de patrones utiliza el concepto de *voting* para decidir acerca de a qué objeto corresponde la imagen analizada (van Erp et al., 2002).

Mediante voting se compara la imagen con patrones y se decide con cuál se identifica, midiendo la similitud entre el objeto y los patrones candidatos.

Para esta comparación, se aplican transformaciones geométricas en los objetos.

Wu et al. (2010) proponen un modelo de base activa para representar clases de objetos, junto con un algoritmo de aprendizaje y reconocimiento de patrones. Sugieren representar el borde de una clase de objetos por una base de elementos en determinada posición y orientación. En la etapa de aprendizaje, los elementos de la base son seleccionados y aplicados a plantillas de aprendizaje, perturbando levemente su posición y orientación, ajustando estos parámetros a la imagen. Finalmente, se propone un esquema de detección y reconocimiento.

La identificación de formas en general se basa en una medida de similaridad con los patrones existentes en una base de datos. Muchos investigadores han propuesto medidas de similaridad, o distancias entre formas, que tienen en cuenta la distancia interna (Ling y Jacobs, 2007), partes visualmente significativas (Latecki y Lakämper, 2000), entre otros.

Enfoques actuales proponen técnicas de aprendizaje de similaridad, para mejorar la performance de reconocimiento. Éstos se basan en propagación en vecindades lineales – LNP- (Wang et al. 2006), aprendizaje multi-visual (Brefeld y Buscher, 2005), modelos generativos (Lawrence, Jordan, 2004), entre otros.

En otra línea de investigación, se estudian transformaciones de imágenes digitales que preservan las propiedades topológicas (conexidad, agujeros, etc). Estas transformaciones abarcan afinamiento mediante la eliminación de puntos simples, o conjuntos simples, segmentación basado en atlas, etc. (Faisan et al., 2008; Passat et al., 2009)

## **Líneas de Investigación y Desarrollo**

Dos estrategias para identificar y seguir objetos en movimiento son usadas: análisis de los bordes del objeto y análisis de sus esqueletos. En ambos casos, sólo se analiza una pequeña cantidad de información (píxeles), pero ese conjunto de puntos obtenidos mantiene propiedades topológicas respecto del conjunto original.

Los puntos del objeto que son adyacentes al complemento constituyen el borde del objeto. Dicho borde puede verse como una curva digital simple cerrada. Este subconjunto es de gran interés en el análisis de objetos dentro de imágenes, ya que muchas características (convexidad, dimensiones, agujeros, etc.) pueden estudiarse por medio de ese subconjunto.

Luego de recorrer la frontera de un objeto en una imagen digital, la tarea siguiente consiste en identificar qué tipo de objeto tiene a esa curva por borde, analizando las propiedades geométricas del mismo. Esta identificación puede llevarse a cabo por comparación con ciertos patrones previamente recopilados y estandarizados. Dado que cada patrón se guarda en una determinada posición y escala, se requiere modificar la imagen analizada (o su frontera) mediante transformaciones geométricas, de modo que mantenga las propiedades de su geometría intrínseca (conexión, agujeros, convexidad, concavidad, puntos de curvatura máxima y mínima, puntos de inflexión, etc).

Se pretende profundizar el estudio del comportamiento de curvas digitales mediante transformaciones geométricas, el efecto de éstas sobre las características topológicas y geométricas, a fin de determinar cuáles son factibles de ser aplicadas. Los resultados teóricos obtenidos fundamentarán un algoritmo de reconocimiento de imágenes, que será diseñado e implementado para tal fin.

El proceso de digitalización, y el sistema de adquisición de la imagen generan ruido, que tiene una presencia notoria en las fronteras de los objetos donde aparecen formas de dientes de serrucho. Esto dificulta su estudio, ya que el análisis de bordes por geometría diferencial requiere el análisis de longitud, rectitud, curvatura, convexidad, características que se

ven afectadas por las irregularidades del borde, producto del proceso de digitalización y adquisición de imagen. Es necesario eludir la influencia del mencionado ruido, a fin de, por un lado, reconocer eficientemente el objeto, y por otro, optimizar el uso de memoria, y no ocuparla con datos superfluos.

## Resultados y objetivos

El proyecto pretende avanzar en el desarrollo teórico de topología digital, específicamente estudiando procesos de discretización y su efecto sobre las características geométricas del objeto, y cómo estas características se mantienen o no en un proceso de transformación geométrica (rotación, escala, perspectiva). Y desde el punto de vista práctico, enfocando hacia las aplicaciones concretas, se pretende diseñar e implementar un algoritmo que, dado un objeto en una imagen digital, lo transforme en un objeto equivalente (aprovechando propiedades invariantes de la topología) a fin de facilitar la identificación del objeto mediante comparación con patrones.

A fin de suavizar y simplificar la curva de borde, se diseñó un algoritmo para aproximar una curva digital por un polígono, con tolerancia de error variable.

Así, el borde se reduce ahora a un conjunto de puntos que son los vértices del polígono aproximante, lo que brinda una enorme ganancia en capacidad de procesamiento.

Con tolerancias muy pequeñas, se generan polígonos de muchos vértices y una mejor aproximación. En el caso límite de tolerancia cero, los vértices del polígono son casi todos los puntos de la curva. Con tolerancia más grandes se obtienen polígonos de menos lados pero una aproximación menos ajustada.

Como preprocesamiento de este algoritmo, se aplica un método para disminuir la cantidad de puntos de la curva, sin perder información. Este preprocesamiento consiste en detectar sucesiones de puntos de borde que están sobre una recta (conjunto de ancho cero) y eliminar los todos esos puntos, excepto los extremos. Luego, con la tolerancia deseada, se elimina puntos que están casi alineados.

Teniendo un polígono que aproxima una curva digital, se puede aproximar la curvatura en cada parte de la curva mediante los ángulos de giro en los vértices del polígono. La sucesión de ángulos de giro muestra un patrón específico de la forma analizada, y es invariante a la escala y rotación. Además, ciertas características que se pueden detectar con dicho patrón son invariantes frente a movimientos de perspectivas (puntos de inflexión, puntos de curvatura máxima, bordes rectos, etc.). El patrón normalizado luego se compara con patrones de objetos conocidos, se mide la similaridad, y se reconoce cuál es el objeto en la imagen.

## Formación de recursos humanos

Este proyecto bianual se desarrolla bajo la dirección de la Dra. María Lorena Bergamini. Actualmente el equipo se completa con el investigador Licenciado Jorge Kamlofsky y Luciana Klacherian (estudiante de Licenciatura en Matemática - UAI).

La tesis de grado de Kamlofsky con la que accedió al título de Licenciado en Matemática, y dirigida por la Dra. Samira Abdel Masih y el Ing. Néstor Balich, sirvió de investigación alentadora del proyecto aquí presentado. Luciana Klacherian, con la dirección de la Dra. Bergamini, se encuentra desarrollando su tesis de Licenciatura, en la que estudia caracterizaciones de rectas digitales y segmentos de rectas digitales. Se estima que en mayo de este año finalizará dicha tesis.

## Referencias

1. U. Brefeld, C. Buscher, T. Scheffer. Multiview Discriminative Sequential Learning, Proc. European Conf. Machine Learning, 2005.
2. U. Eckhardt, L. Latecki. Topologies for the digital spaces  $Z^2$  and  $Z^3$ . Computer Vision Image Understanding, 90, pp. 295-312, 2003.
3. U. Eckhardt, L. Latecki. Digital Topology. In Current Topics in Pattern Recognition Research, Research Trends,

- Council of Scientific Information, Vilayil Gardens, Trivandrum, India, 1994.
4. S. Faisan, N. Passat, V. Noblet, R. Chabrier, C. Meyer. Topology Preserving Warping of Binary Images: Application to Atlas-Based Skull Segmentation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008 Lecture Notes in Computer Science*, 5241, pp. 211-218, 2008.
  5. A. Gross, L. Latecki. Digital geometric invariance and shape representation, IEEE, 1995.
  6. J. Kamlofsky. Topología digital: base para la visión artificial. Tesis de grado. Facultad de Tecnología Informática. UAI. Buenos Aires, 2011.
  7. J. Kamlofsky, M. L. Bergamini. Aproximación de formas de objetos digitales por polígonos, VII Jornadas Argentinas de Robótica, 2012.
  8. L. J. Latecki, R.-R. Ghadially, R. Lakaemper, U. Eckhardt. Continuity of the discrete curve evolution. *Journal of Electronic Imaging* 9 (3), pp. 317-326, 2000.
  9. L. J. Latecki, R. Lakämper. Shape Similarity Measure Based on Correspondence of Visual Parts, *IEEE Trans. Pattern Analysis and Machine Intelligence* 22(10), pp. 1185-1190, 2000.
  10. N. D. Lawrence, M.I. Jordan. Semi-Supervised Learning via Gaussian Processes, *Advances in Neural Information Processing Systems*, MIT Press, 2004
  11. H. Ling, D. Jacobs. Shape Classification Using the Inner-Distance, *IEEE Trans. Pattern Analysis and Machine Intelligence* 29(2), pp. 286-299, 2007.
  12. H. R. Liu , L. J. Latecki, W. Liu. A Unified Curvature Definition for Regular, Polygonal, and Digital Planar Curves. *Int J Comput Vis* 80, pp 104–124, 2008.
  13. N. Passat, M. Couprie, L. Mazo, G. Bertrand. Topology-Preserving Thinning in 2-D Pseudomanifolds. *Discrete Geometry for Computer Imagery Lecture Notes in Computer Science*, 5810, pp. 217-228, 2009.
  14. M. van Erp, L. Vuurpijl, L. Schomaker. An overview and comparison of voting methods for pattern recognition. *Proceedings of the Eighth International Workshop on Frontiers in Handwriting Recognition (IWFHR'02)*, pp 195-200, 2002.
  15. F. Wang, J. Wang, C. Zhang, H. Shen. Semi-Supervised Classification Using Linear Neighborhood Propagation, *Proc. IEEE CS Conf. Computer Vision and Pattern Recognition*, 2006.
  16. Y. N. Wu, Z. Si, H. Gong, S. C. Zhu. Learning active basis model for object detection and recognition, *Int J Comput Vis* 90, pp 198–235, 2010.