

Reconocimiento de Patrones Adaptativo.

Aplicaciones en procesamiento de señales

Laura Lanzarini¹, Franco Chichizola², Waldo Hasperue³ Leonardo Corbalán², César Estrebou⁴,
Juan Maulini⁵, Bernarda Albanesi⁶, Nadia Funes

Instituto de Investigación en Informática LIDI (III-LIDI)⁷
Facultad de Informática. UNLP

CONTEXTO

Esta presentación corresponde al Subproyecto “Sistemas Inteligentes” perteneciente al Proyecto “Algoritmos Distribuidos y Paralelos. Aplicación a Sistemas Inteligentes y Tratamiento Masivo de Datos” del Instituto de Investigación en Informática LIDI.

RESUMEN

Esta línea de investigación se centra en el estudio y desarrollo de estrategias adaptativas aplicables al Reconocimiento de Patrones. Actualmente el énfasis está puesto en el reconocimiento de objetos en video y en la identificación de personas a través de sus medidas biométricas.

Con respecto a la identificación de objetos se trabaja sobre dos aspectos: su representación y el mecanismo de comparación a utilizar. Por tal motivo se estudian y desarrollan métodos detectores y descriptores analizando en cada caso la precisión de los resultados. Con el objetivo de poder realizar el reconocimiento en tiempo real también se analiza la paralelización de los métodos propuestos.

La identificación de personas se efectúa a través de dos medidas biométricas: la imagen del rostro y la señal de voz. En el primer caso se estudian metaheurísticas que ayuden a reducir la tasa de falsos positivos durante el proceso de reconocimiento a través de la selección de los descriptores adecuados. Por otra parte, se investiga en la utilización de redes neuronales dinámicas como herramienta

para identificar a una persona a partir de su señal de voz.

Palabras claves: Procesamiento de señales. Reconocimiento de Patrones. Técnicas de Optimización. Biometría.

1. INTRODUCCIÓN

Las estrategias adaptativas han demostrado ser herramientas sumamente útiles en la resolución de problemas complejos. Su capacidad para responder a los cambios que se producen en el entorno de información les permite brindar buenos resultados en distintas áreas.

En el Instituto de Investigación en Informática LIDI se está trabajando en este tema desde hace varios años. Inicialmente se desarrollaron estrategias basadas en Redes Neuronales y Algoritmos Evolutivos aplicables al Reconocimiento de Patrones en imágenes [Lan00, Lan04, Gar07, Has07]. Luego se profundizó en el estudio de la capacidad de caracterización de este tipo de estrategias a fin de establecer un modelo de la información disponible. Se definieron y desarrollaron varios métodos basados en redes neuronales competitivas con aplicaciones concretas en distintas áreas [Cor06a, Cor06b, Has07, Has07a] y se propusieron nuevos métodos para trabajar con cúmulos de partículas [Lan08, Lop09].

Actualmente, con el objetivo de mejorar la transferencia tecnológica de los resultados obtenidos, se está trabajando en el

¹ Profesor Titular. Facultad de Informática. UNLP

² Profesor Adjunto. Facultad de Informática. UNLP

³ Becario de Postgrado Tipo I - CONICET – Ayudante Diplomado - Facultad de Informática. UNLP

⁴ Jefe de Trabajos Prácticos SD. Facultad de Informática. UNLP

⁵ Becario III-LIDI. Ayudante Diplomado. Facultad de Informática

⁶ Ayudante Diplomado. Facultad de Informática. UNLP.

⁷ Calle 50 y 115 1er Piso, (1900) La Plata, Argentina, TE/Fax +(54) (221) 422-7707. <http://weblidi.info.unlp.edu.ar>

reconocimiento tanto de objetos como de personas. Este es un problema de sumo interés en la actualidad por su aplicabilidad en diversas áreas como seguridad, control industrial y entretenimiento, entre otras.

El reconocimiento de personas a partir en uno o más rasgos físicos o de su conducta es el objetivo central de la biometría. Las tareas que se llevan a cabo en esta área pueden dividirse básicamente en dos categorías: verificación e identificación de individuos. En el primer caso, conociendo la identidad del individuo, se comparan sus rasgos con la información previamente almacenada en una base de datos mientras que en el segundo caso, no se conoce a priori quien es la persona y se comparan sus rasgos contra todos los de la base de datos. Como puede verse, la identificación es un proceso más lento que la verificación. En informática, la mayoría de las soluciones existentes se basan en técnicas matemáticas y estadísticas. Esta línea de investigación tiene por objetivo general el uso de estrategias adaptativas aplicables al reconocimiento de personas poniendo especial interés en su voz y la imagen de su rostro.

A continuación se detallan brevemente los avances realizados últimamente.

1.1. Reconocimiento de Objetos en Video

Si bien el procesamiento digital de imágenes ha avanzado mucho, este tema aun presenta distintos aspectos a considerar dado que la apariencia visual de un mismo objeto puede variar en función de la distancia, del punto de vista y de la iluminación; incluso, podría ocurrir que se encuentre parcialmente oculto en algunas escenas.

Generalmente, un objeto se representa a través de un conjunto de regiones superpuestas, cada una de las cuales, lleva asociado un vector de características calculado a partir de su apariencia. Los descriptores involucrados son construidos con un cierto grado de invariancia al punto de vista y a la iluminación. Se han analizados distintos métodos detectores y descriptores analizando en cada caso la precisión de los resultados [Har88, Smi97, Low04, Ste97, Mat02] concluyendo que SIFT es el más adecuado.

El algoritmo de reconocimiento que se ha propuesto utilizar a partir de esta representación, posee un costo computacional importante. Este aspecto dificulta la obtención de resultados en tiempo real al ser aplicado sobre un video (o secuencias de frames). Por tal motivo se está trabajando en su paralelización sobre una arquitectura de Memoria Compartida utilizando para su implementación las librerías de OpenMP [Gra03, Lei07, Zha08].

1.2. Reconocimiento de Rostros

El reconocimiento de rostros es una técnica biométrica muy utilizada. Se trata de identificar si la imagen del rostro de una persona se corresponde o no con alguna de las imágenes existentes en una base de datos. Este problema es difícil de resolver automáticamente debido a los cambios que distintos factores, como la expresión facial, el envejecimiento e incluso la iluminación, producen en la imagen.

En general, puede observarse que las representaciones utilizadas presentan una tasa elevada de falsos positivos. Este aspecto dificulta el reconocimiento a la hora de establecer un umbral de corte. Como forma de resolver este problema se propone utilizar sólo los descriptores SIFT más representativos [Low04]. Esta selección se realiza a través de una metaheurística poblacional [Shi99, Moj07]. Los resultados obtenidos demuestran que puede lograrse un buen reconocimiento resolviendo a la vez los dos grandes problemas que posee esta caracterización: la detección de falsos positivos y el tiempo requerido para realizar el reconocimiento.

1.3. Reconocimiento de voz

La identificación de la voz para determinar que persona es la que habla es una herramienta sumamente útil en el área de seguridad ya que permite validar que la persona es quien dice ser o bien identificar por algún motivo específico que persona es la que está hablando. El tratamiento digital de señales de audio es una de las técnicas

biométricas que puede utilizarse para realizar las tareas de identificación y verificación [Che07, Mez05, Ras08].

Se han desarrollado varios métodos para lograr un modelado que resuelva el problema de identificación [Has05, Mez05, New08]. Los mapas auto-organizativos dinámicos [Ala00] han demostrado ser herramientas muy eficaces en el armado de modelos. Su aprendizaje les permite organizar su estructura interna respetando la topología de los datos de entrada.

En esta línea de investigación se ha propuesto y desarrollado un algoritmo de tres etapas que genera un único modelo para identificar varias personas por medio del análisis de voz. La primera etapa consiste en el pre-tratamiento de la señal y extracción de componentes característicos; la segunda consiste en el entrenamiento de una red SOM dinámica usando [Has07] y la última etapa establece la correspondencia entre cada elemento de la red y la persona a la cual representa. Una vez armado el modelo es posible realizar las tareas de identificación y validación de personas.

2. TEMAS DE INVESTIGACIÓN Y DESARROLLO

- Desarrollo e implementación, a partir de los métodos existentes, de estrategias adaptativas capaces de construir y mantener modelos adecuados en entornos de información dinámicos.
- Técnicas de procesamiento de imágenes poniendo especial atención en: Filtro Gaussianos, derivadas discretas 3D, representación escala-espacio, técnicas de agrupamiento, transformaciones afines y eliminación de ruido en las imágenes.
- Estudio de arquitecturas multicore, modelos de comunicación, alternativas de lenguajes/bibliotecas para programación paralela (OpenMP, MPI). Métricas de performance en arquitecturas de memoria compartida.
- Análisis de los distintos tipos de Redes Neuronales competitivas dinámicas. Estudio de las estrategias existentes que permiten determinar, durante la adaptación, el tamaño de la arquitectura y forma de conexión de los elementos que componen la red neuronal.
- Estudio y aplicación de diferentes métricas que permitan analizar la preservación de la topología de los datos tanto en el espacio de los patrones de entrada como en el espacio de salida de la red.
- Estudios de performance de los algoritmos desarrollados. Análisis de eficiencia en la resolución de problemas concretos.

3. RESULTADOS OBTENIDOS/ ESPERADOS.

- Desarrollo e implementación de una red neuronal entrenada a partir del método AVGSOM como herramienta para reconocer patrones de señales de voz.
- Desarrollo e implementación de un sistema de votación a partir de varias redes neuronales competitivas dinámicas que permite reducir el error en la etapa de reconocimiento.
- Desarrollo e implementación de una estrategia paralela que, utilizando memoria compartida, permita extraer las características SIFT de una imagen y compararla con la base de datos existentes de una manera eficiente.
- Desarrollo e implementación de un mecanismo para reconocimiento de rostros basado en características SIFT que permite reducir el tamaño de la base de datos utilizando una variante de PSO binario.
- Resolución de problemas concretos, tanto en ambientes simulados como en el mundo real. En este último caso, resulta de fundamental importancia la optimización del algoritmo propuesto.

4. FORMACIÓN DE RECURSOS HUMANOS

Dentro de los temas involucrados en esta línea de investigación se están desarrollando actualmente 1 tesis de doctorado, 1 de maestría y al menos 3 tesinas de grado de Licenciatura. También participan en el desarrollo de las tareas becarios y pasantes del III-LIDI.

5. REFERENCIAS

- [Ala00] Alahakoon, D., Halgamuge, S.K., Srinivasan, B: Dynamic Self-Organizing Maps with Controlled Growth for Knowledge Discovery. IEEE Transactions On Neural Networks 11, 601--614 (2000)
- [Che07] Chen, W., Peng, C., Zhu, X., Wan, B., Wei, D.: SVM-based Identification of Pathological Voices. In: 29th Annual International Conference of the IEEE, Engineering in Medicine and Biology Society, EMBS, pp. 3786--3789 (2007).
- [Cor06a] Corbalán, Osella Massa, Russo, Lanzarini. Image recovery using a new nonlinear Adaptive Filter based on Neural Networks. CIT - Journal of Computing and Information Technology. Vol. 14, No.4. December 2006 - Pág. 315-320.
- [Cor06b] Corbalán, Hasperue, Osella Massa, Lanzarini. BPNn-CPN. Nuevo método para segmentación de Imágenes basado en Redes Neuronales Artificiales. IV Workshop de Computación Gráfica, Imágenes y Visualización (WCGIV). CACIC 2006. San Luís. Argentina. Octubre de 2006.
- [Gar07] Garbi, Mercado, Lanzarini, Russo. Reconocimiento de números manuscritos".XIII Congreso Argentino de Ciencias de la Computación CACIC 2007.

Corrientes y Resistencia, Argentina. Octubre de 2007

- [Gra03] "An Introduction to Parallel Computing. Design and Analysis of Algorithms. 2nd Edition". Grama, Gupta, Karypis, Kumar. Pearson Addison Wesley, 2003.
- [Har88] Chris Harris & Mike Stephens. A combined corner and edge detector. Proceedings of The Fourth Alvey Vision Conference (1988), pp. 147-151.
- [Has05] Hasperué, Lanzarini. Dynamic Self-Organizing Maps. A new strategy to enhance topology preservation. XXXI Conferencia Latinoamericana de Informática. CLEI 2005.
- [Has07] Hasperue, Corbalán, Lanzarini, Bría. Skeletonization of Sparse Shapes using Dynamic Competitive Neural Networks. Journal Inteligencia Artificial. Revista Iberoamericana de Inteligencia Artificial. Vol.11. Nro.35. pp.33-42 2007.
- [Lan00] Lanzarini. Reconocimiento de Patrones en Imágenes Médicas utilizando Redes Neuronales. Journal of Computer Science and Technology. Vol.4 . Dic 2000.
- [Lan04] Lanzarini, Yanivello. Reconocimiento de Comandos Gestuales utilizando GesRN. V Workshop de Agentes y Sistemas Inteligentes. CACIC 2004. Bs.As. Argentina. 2004. ISBN 987 9495 58 6
- [Lan08] Lanzarini, Leza, De Giusti. Particle Swarm Optimization with Variable Population Size. Lecture Notes in Computer Science. Vol 5097/2008. Artificial Intelligence and Soft Computing. Pags. 438-449. Junio de 2008. ISBN 987-3-540-69572-1.
- [Lei07] Lei Chai, Qi Gao, Dhabaleswar K. Panda. Understanding the Impact of

- Multi-Core Architecture in Cluster Computing: A Case Study with Intel Dual-Core System. IEEE International Symposium on Cluster Computing and the Grid 2007, pages 471-478, 2007
- [Lop09] Lopez, Lanzarini, De Giusti. Particle Swarm Optimization with Oscillation Control. Genetic and Evolutionary Computation Conference. ACM GEECCO Proceeding. Montréal, Canada. July 2009.
- [Low04] David Lowe. Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision. Springer Netherlands. ISSN 0920-5691. Vol. 60, nro. 2/2004. Pages 91-110.
- [Mat02] J. Matas, O. Chum, M. Urban, T. Pajdla. Robust Wide Baseline Stereo from Maximally Stable Extremal Regions. Image and Vision Computing. Vol.22, issue 10. pages 761-767. 2002
- [Mez05] Mezghani, A., O'Shaughnessy, D.: Speaker verification using a new representation based on a combination of MFCC and formants. Canadian Conference on Electrical and Computer Engineering. pp. 1461--1464. (2005).
- [Moj07] Mojtaba Ahmadiéh Khanesar. A novel Binary Particle Swarm Optimization. 18th Mediterranean Conference on Control & Automation. 2007. T33-001.
- [New08] New, T.L., Li, H.: On fusion of timbre-motivated features for singing voice detection and singer identification. IEEE International Conference on Acoustics, Speech and Signal Processing.. pp. 2225--2228 (2008).
- [Ras08] Rashid, R.A., Mahalin, N.H., Sarijari, M.A., Abdul Aziz, A.A.; Security system using biometric technology: Design and implementation of Voice Recognition System (VRS). International Conference on Computer and Communication Engineering. pp. 898--902 (2008).
- [Shi99] Shi Y., Eberhart R. An empirical study of particle swarm optimization. IEEE Congress Evolutionary Computation. pp.1945-1949. Washington DC, 1999.
- [Smi97] Smith S., Brady J. "Susan – a new approach to low level image processing". International Journal of Computer Vision, Vol 23, nro.1, pp.45-78, 1997.
- [Ste97] Stephen M. Smith & Michael Brady. SUSAN (Smallest Univalve Segment Assimilating Nucleus) - A New Approach to Low Level Image Processing. International Journal of Computer Vision, vol.23, no. 1, pp. 45-78, 1997.
- [Zha08] Qi Zhang; Yurong Chen; Yimin Zhang; Yinlong Xu. SIFT implementation and optimization for multi-core systems. IEEE International Symposium on Parallel and Distributed Processing. 2008. Pages 1-8.