

Sistemas de Visión Automática y Reconocimiento de Patrones Interfaces Avanzadas, Realidad Virtual y Aumentada

Javier Giacomantone, María José Abásolo, Oscar Bria, Federico Cristina, Sebastian Dapoto, Verónica Artola, Luciano Lorenti, Lucía Violini, Cristina Manresa-Yee, Ramon Mas Sansó, Marcelo Naiouf

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

{jog, mjabasolo, obria, fcristina, sdapoto, vartola, llorenti, lviolini, mnaiouf}@lidi.info.unlp.edu.ar

CONTEXTO

Esta línea de investigación y desarrollo (I/D) forma parte del Subproyecto “Tratamiento de imágenes digitales y video. Visión 3D”, dentro del Proyecto “Procesamiento Paralelo y Distribuido y Aplicaciones en Sistemas Inteligentes y Tratamiento de Imágenes y Video” del Instituto de Investigación en Informática LIDI, acreditado por la UNLP en el marco del Programa de Incentivos. Asimismo se integra al proyecto “A1/037910/11 Formación de Recursos Humanos e Investigación en el Área de Visión por Computador e Informática Gráfica (FRIVIG)”, financiado por la Agencia Española de Cooperación Internacional (AECID) dentro del Programa de Cooperación Interuniversitaria e Investigación Científica.

RESUMEN

Este trabajo describe una línea de investigación y desarrollo (I/D) y los resultados esperados de la misma. El objetivo principal es estudiar, desarrollar y evaluar métodos y algoritmos en sistemas de visión automática y reconocimiento automático de patrones. Los principales temas abordados son análisis de imágenes, clasificación supervisada y no supervisada, interfaces basadas en visión, realidad virtual y aumentada.

Palabras Clave: Visión automática, Reconocimiento de Patrones, Realidad Virtual y Aumentada.

1. INTRODUCCION

La línea de investigación y desarrollo descrita puede dividirse en un conjunto de temas de investigación. Estos pueden agruparse para facilitar su descripción y estudio en tres grupos temáticos. El objetivo primario es la investigación y desarrollo de modelos, métodos y algoritmos en cada conjunto de temas, cada uno caracterizado por abordar una etapa importante en un sistema de visión automática. El segundo objetivo que define esta línea es el análisis integrador del problema estudiado y un abordaje multidisciplinario.

1.1 Análisis de Imágenes

Se estudian métodos de segmentación, generación, selección y extracción de características en imágenes de intensidad y de rango. Para obtener los dos tipos de imágenes que nos permitan expresar las propiedades geométricas y dinámicas de objetos 3D se estudian modelos de visión estereoscópica e imágenes generadas a partir de cámaras de rango. Recientemente se ha adquirido equipamiento específico que permite obtener directamente imágenes de rango como la cámara MESA SR-4000. La cámara MESA utiliza el principio de tiempo de vuelo (TOF) para generar un mapa de distancias. Las cámaras basadas en tiempo de vuelo permiten resolver problemas de reconocimiento de patrones y requieren tanto la adaptación de métodos conocidos como la generación de nuevos mé-

todos para el tratamiento y clasificación de los datos obtenidos. Son de nuestro particular interés los métodos que permitan proponer alternativas para las interfaces con una computadora como la detección y clasificación de gestos o la detección en tiempo real [1][2].

1.2 Reconocimiento de Patrones

Reconocimiento automático de patrones es un área de investigación interdisciplinaria tanto en la investigación básica de métodos fundamentales [3] [4] [5], como en sus aplicaciones [6]. El objetivo principal de un sistema de reconocimiento automático de patrones es descubrir la naturaleza subyacente de un fenómeno u objeto, describiendo y seleccionando las características fundamentales que permitan clasificarlo en una categoría determinada. El tipo de objetos o fenómenos considerados en esta línea de trabajo pueden ser descriptos por un conjunto de características numéricas que definen patrones en un espacio n -dimensional. Los sistemas automáticos de reconocimiento de patrones permiten abordar problemas en informática, en ingeniería y en otras disciplinas científicas; por esto el diseño requiere de criterios de análisis conjuntos para validar los resultados [7] [8] y forman una etapa importante en un sistema de visión por computador. En particular estudiamos dos aspectos fundamentales, representación de objetos, generación de características, y la clasificación estadística de los mismos. Estas tareas se realizan a partir de la información obtenida de imágenes digitales adecuadamente procesadas. Los métodos de extracción de características intentan optimizar la clasificación generando características invariantes a distintas transformaciones [9], son particularmente importantes en el desarrollo de interfaces hombre-computadora a partir de imágenes que deben ser procesadas en tiempo real.

En esta línea de I/D se estudian problemas de clasificación supervisada y no supervisada. En clasificación supervisada donde el objetivo es optimizar la generalización del clasificador se estudian métodos basados en máquinas de soporte vectorial (SVM) [10] [11]. Las SVM implementan reglas de decisión complejas por medio de una función no lineal que permite mapear los puntos de entrenamiento a un espacio de mayor dimensión. En el nuevo espacio de características las clases son separadas por un hiperplano, siendo este el que maximiza la distancia entre este y los puntos de entrenamiento. Los problemas que permiten resolver las SVM son de distinta naturaleza como en sistemas biométricos [12][13], bioinformática [14], clasificación de imágenes de resonancia magnética funcional ruidosas [15][16] y correspondencia de puntos en imágenes estereoscópicas [17].

1.2 Interfaces avanzadas. Realidad Virtual y Aumentada

Realidad Virtual y Realidad Aumentada

La Realidad Virtual (RV) es un término que se aplica a un conjunto de experiencias sensoriales sintéticas, es decir generadas por computador, comunicadas a un operador o participante. La mayoría de las aplicaciones de realidad virtual son experiencias visuales donde el participante se ve inmerso e interactúa en un ambiente o escena virtual. Una de las principales aplicaciones de RV son los simuladores de entrenamiento.

La Realidad Aumentada (RA) agrega información sintética a la realidad. La diferencia principal entre RV y RA es que por una parte RV implica inmersión del participante en un mundo totalmente virtual y por otra parte la RA implica mantenerse en el mundo real con agregados virtuales.

A diferencia de las aplicaciones de RV las aplicaciones de RA generalmente necesitan la movilidad del usuario, incluso hacia ambientes externos- en inglés se denominan aplicaciones *outdoor*. En dichas aplicaciones de realidad aumentada puede ser necesaria conocer la posición global del participante utilizando dispositivos como GPS y brújulas digitales.

Interfaces basadas en Visión

Existen actualmente nuevos paradigmas de interacción que empiezan a estar presentes en todas las esferas de nuestra vida diaria: la computación ubicua, la realidad virtual y la realidad aumentada. Por ello los estudios de investigación para obtener nuevos sistemas de interacción basados en fuentes de percepción de información como el sonido, el tacto o la visión se han convertido en un campo en auge que pretende desarrollar interfaces más naturales, intuitivas, no invasivas y eficientes.

Hoy en día, las cámaras de bajo coste y con suficiente resolución se encuentran en la mayoría de las computadoras y en celulares. La información proveniente de la cámara se puede procesar mediante técnicas de visión por computador con fines de asistir la interacción persona-ordenador obteniendo las denominadas Interfaces basadas en visión (VBI). Turk and Kölsch [18] comentan las posibles funcionalidades que pueden ofrecer las VBIs: presencia y localización, identidad, expresión, gestos, foco de atención, postura del cuerpo y movimiento y actividad.

En el campo de la interacción persona-ordenador se utilizan pantallas táctiles (touchscreen), las cuales están superpuestas a un dispositivo de visualización, para la entrada táctil, el panel táctil. En algunos casos, las pantallas táctiles permiten la detección simultánea de varios puntos de contacto (multitáctiles). Esto ha hecho posible el desarrollo de metodologías de interacción que permiten al usuario una comunicación más natural e intuitiva con el ordenador. Los habituales gestos utilizados para seleccionar, arrastrar, rotar o escalar objetos visuales han evolucionado a versiones en las que el usuario utiliza ambas manos o varios dedos para realizar las mismas acciones, proveyendo de mecanismos que se han convertido en gestos estándares.

Los displays de grandes dimensiones también se han hecho más y más necesarios por el auge del trabajo cooperativo en un mismo entorno. Han surgido nuevos términos para describir estos dispositivos: superficies (surfaces), mesas multitáctiles (multi-touch tables), *tabletops*. La utilización de cámaras y el posterior tratamiento de las imágenes mediante algoritmos de visión por computador se emplea en la construcción de mesas multitáctiles de bajo coste. En

un contexto educativo, las mesas multitáctiles permiten a los estudiantes interactuar con objetos digitales en tareas colaborativas. En el caso de la educación especial, la tecnología puede proporcionar a los usuarios oportunidades de aprender, compartir información y ganar independencia. Se dedican esfuerzos para adaptar y desarrollar aplicaciones a las superficies multitáctiles para usuarios con limitaciones cognitivas y dificultades sociales. En [19] se describe una aproximación para la construcción de una mesa multitáctil de bajo coste pensando sobre todo en la seguridad de uso y en la accesibilidad y confort para usuarios discapacitados.

2. LINEAS DE INVESTIGACION y DESARROLLO

- Estudio de métodos para el tratamiento de datos en cámaras basadas en tiempo de vuelo.
- Clasificación supervisada. Discriminadores lineales y no lineales.
- Selección y Extracción de características.
- Segmentación, representación y clasificación.
- Análisis de características de textura, color y formas.
- Máquinas de soporte vectorial. Kernels y algoritmos de optimización.
- Criterios de evaluación de desempeño en sistemas de clasificación automática.
- Simuladores de entrenamiento de Realidad Virtual
- Aplicaciones de Realidad Aumentada en ambientes externos
- Interfaces Basadas en Visión

3. RESULTADOS OBTENIDOS Y ESPERADOS

- Se desarrolló un sistema, hardware y software, que permite la reconstrucción de

modelos 3D basado en la combinación de cámaras estereoscópicas y la utilización de luz estructurada [20].

- Se estudiaron y propusieron métodos para detección en series temporales de fMRI [15]
- Se realizó la evaluación de rendimiento en sistemas de reconocimiento de patrones supervisados y de clasificación binaria [16].
- Se desarrolló una solución para detección de construcciones en imágenes aéreas de baja calidad [21].
- Actualmente se están realizando pruebas con equipamiento de alto coste recientemente adquirido como la cámara de tiempo de vuelo MESA SR4000. Se espera en un corto a mediano plazo poder realizar reconstrucciones y utilizarlas para evaluar los prototipos escáneres desarrollados.
- Construcción de una mesa multitáctil basada en visión por computador para su uso en educación especial [19].
- Pruebas de aplicaciones de realidad aumentada en dispositivos móviles como tablets PC
- Desarrollo de un simulador de barcos utilizando realidad virtual, en conjunto con el instituto Pladema de la UNICEN, con fines educativos y de difusión a la comunidad.

4. FORMACION DE RECURSOS HUMANOS

En esta línea de I/D la formación de recursos humanos es uno de los objetivos principales. Las Universidades que participan del proyecto AECID “FRIVIG: Formación de Recursos Humanos e Investigación en el área de Visión por Computadora e Informática”: Universidad de Islas Baleares (España), Universidad Nacional del Sur (Argentina) y la Universidad Nacional de La Plata han estructurado con sede en la UNLP una Especialización en Compu-

tación Gráfica, Imágenes y Visión por Computadora, con la participación de docentes-investigadores de las tres instituciones y de otras Universidades de Argentina y España con experiencia y formación en los temas propuestos. Dichos cursos han sido aprobados en la Facultad de Informática de la UNLP.

Recientemente dentro del marco de la XV Escuela Internacional de Informática de CACIC 2011 se dictó, en forma conjunta profesores argentinos y españoles, un curso de Interfaces Avanzadas, Realidad Virtual y Realidad Aumentada [22].

En el marco de esta línea de investigación hay dos investigadores realizando su doctorado y tres alumnos desarrollando tesinas en sistemas de visión por computadora y realidad aumentada.

5. AGRADECIMIENTOS

Esta línea de investigación es parcialmente financiada por el proyecto “A1/037910/11 Formación de Recursos Humanos e Investigación en el Área de Visión por Computador e Informática Gráfica (FRIVIG)” de la Agencia Española de Cooperación Interuniversitaria e Investigación Científica.

6. BIBLIOGRAFIA

- [1] Penne et al. Robust real time 3-D respiratory motion detection using time of flight cameras. *International Journal of Computer Assisted Radiology and Surgery* v.3 pp.427-431. (2008).
- [2] Kollors, A., Penne, J., Hornegger, J., Barker A. Gesture Recognition with a Time-of-Flight camera. *International Journal of Intelligent Systems Technologies and Applications*, v. pp.334-343, (2008).
- [3] Fukunaga K. *Introduction to Statistical Pattern Recognition*. 2ed. Academic Press, 1990.
- [4] Batagelj V, Bock H, Ferligoj A. *Data Science and Classification*. Springer, 2006.
- [5] Devijer P. A., Kittler, J. *Pattern Recognition, A Statistical Approach*. Prentice Hall, 1982.

- [6] Devijver P. A., Kittler, J. Pattern Recognition: theory and applications. Springer, 1986.
- [7] Kim H.Y., Giacomantone J. O., Cho, Z. H. Robust Anisotropic Diffusion to Produce Enhanced Statistical Parametric Map, Computer Vision and Image Understanding, v.99, p.435-452 (2005).
- [8] Kim H.Y., Giacomantone J. O., A New Technique to Obtain Clear Statistical Parametric Map by Applying Anisotropic Diffusion to fMRI, IEEE, International Conference on Image Processing. Proceedings, Genova, Italy, v.3, p.724-727 (2005).
- [9] Lowe, D. G. Distinctive image features from scale invariant keypoints. International Journal of Computer Vision, 60 (2), 2004.
- [10] Cortes C., Vapnik V. Support vector networks. Machine Learning v.20, p.273-297 (1995).
- [11] Vapnik, V. The Nature of Statistical Learning Theory. Springer (1995).
- [12] Lei, Z., Yang, Y, Wu Z. Ensembles of Support Vector Machines for Text/Independent Speaker Recognition. IJCSNS v.6 n5 p.162-167 (2006).
- [13] Jain, A. K., Ross, A., Prabhakar, S. An Introduction to Biometric Recognition, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 14, No.1, January 2004.
- [14] Li, Y., Li, J. Predicting Subcellular Localization of Proteins using Support Vector Machines with N-Terminal Amino Composition, ADMA 2005, LNAI 3584, pp. 618-625.
- [15] Javier Giacomantone, Tatiana Tarutina. Diffuse Outlier Detection Technique for Functional Magnetic Resonance Imaging. Computer Science and Technology Series. XVI Argentine Congress of Computer Science Selected Papers. Edulp. Pp. 255-265 (2011).
- [16] Javier Giacomantone, Armando De Giusti, ROC performance evaluation of RADSPM technique, XIV Congreso Argentino de Ciencias de la Computación (CACIC), Chilecito (2008).
- [17] Pajares G., De la Cruz J. Stereovision matching through support vector machines. Pattern Recognition Letters 24, p. 2575-2583 (2003).
- [18] M. Turk, M. Kölsch (2005). Perceptual interfaces. En: G. Medioni and S.B Kang, editors, *Emerging Topics in Computer Vision*. Pearson Education
- [19] C. Manresa Yee, R. Mas Sansó, G. Moyà, M. J. Abásolo, J. Giacomantone. Interactive multi-sensory environment to control stereotypy behaviours. XVII Congreso Argentino de Ciencias de la Computación (CACIC), La Plata (2011).
- [20] Federico Cristina, Sebastián H. Dapoto, Javier Vegas, Verónica Artola, Claudia Russo, María José Abásolo, Armando De Giusti. 3D Scanner Development with Stereoscopic Cameras and Laser Illumination. IADIS Computer Graphics and Visualization CGV 2008. ISBN: 978-972-8924-63-8, p. 249-253.
- [21] Federico Cristina, Sebastián H. Dapoto, Verónica Artola, Claudia Russo, Graciela Viegas, Detección de edificios en imágenes urbanas, XV Congreso Argentino de Ciencias de la Computación (CACIC), San Salvador de Jujuy (2009).
- [22] C. Manresa Yee, M. J. Abásolo, R. Mas Sansó y M. Vénere. Realidad virtual y realidad aumentada. Interfaces Avanzadas. Edulp ISBN 978-950-34-0765-3 (2011).