

Tratamiento de imágenes digitales y video. Visión y Reconstrucción 3D.

Claudia Russo, Federico Cristina, Sebastián H. Dapoto, Javier Vegas,
Verónica Artola, Marcelo Naiouf, Javier Giacomantone, Oscar Bria, María José Abásolo

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

{crusso, fcristina, sdapoto, jvegas, vartola, mnaiouf, jog, onb}@lidi.info.unlp.edu.ar, mjabasolo@uib.es

CONTEXTO

Esta línea de Investigación forma parte del Subproyecto “Tratamiento de imágenes digitales y video. Visión 3D”, dentro del Proyecto “Algoritmos Distribuidos y Paralelos. Aplicación a Sistemas Inteligentes y Tratamiento Masivo de Datos” del Instituto de Investigación en Informática LIDI, acreditado por la UNLP en el marco del Programa de Incentivos. Asimismo se integra a un proyecto financiado por la Agencia Española de Cooperación Internacional (AECI).

RESUMEN

La línea de investigación presentada se centra en el estudio y desarrollo de temas relacionados con software para tratamiento de imágenes y video, reconstrucción 3D y diseño y programación de interfaces gráficas en 3D, especialmente aplicados a problemas de tiempo real. Son estudiadas técnicas de análisis de tipologías edilicias en imágenes aéreas. Los temas estudiados abarcan aspectos clásicos referidos al tratamiento de imágenes tales como detección de bordes, segmentación de imágenes, reconocimiento de patrones, estimación de movimientos y flujo óptico, cálculos de trayectoria, velocidad y aceleración de objetos. En el área de visión 3D, el énfasis está puesto en el estudio, investigación y desarrollo de aplicaciones en campos como la arquitectura y la arqueología, donde resulta muy útil contar con métodos rápidos y flexibles para adquirir información 3D a partir de una escena real. El diseño de am-

bientes 3D es utilizado como interfaz gráfica de otras aplicaciones valiéndose del uso de Realidad Virtual.

Palabras Clave: Imagen, Video, Reconstrucción 3D, Procesamiento de Imágenes, Reconocimiento de Patrones, Visión, Interfaz gráfica, reconstrucción 3D.

1. INTRODUCCION

El procesamiento digital de señales (en particular el tratamiento de imágenes) adquiere un interés especial ya que constituye la base para muchas aplicaciones importantes [1]. Una señal es una variable de una o más dimensiones que toma valores de acuerdo a otra variable, como por ejemplo el tiempo o el espacio. Las señales pertenecientes al mundo real son continuas. Sin embargo, una computadora digital si bien puede trabajar con señales reales no puede hacerlo con señales continuas. Necesariamente las señales deben ser muestreadas y digitalizadas, y por lo tanto se convierten en una secuencia de números [2][3]. Una imagen es un caso particular de señal que puede ser digitalizada en forma de matriz.

El procesamiento y el análisis de imágenes digitales nace en el momento en que se dispone de recursos tecnológicos para captar y manipular, en forma de matrices de valores, gran cantidad de información espacial [2][3].

Muchas disciplinas entre las que se encuentran el reconocimiento de patrones en tiempo real, el tratamiento y transmisión de video en tiempo real y visión por computa-

dora, requieren un importante esfuerzo en la investigación de algoritmos paralelos aplicables en áreas tales como: robótica, industria manufacturera, ingeniería forestal y medicina. Esto motiva la investigación y el desarrollo en temas relacionados con software para tratamiento de señales, especialmente en problemas de tiempo real.

Cuando el objetivo se centra en extraer y clasificar objetos que aparecen en la imagen se pueden utilizar varias técnicas de análisis, entre ellas las de reconocimiento estadístico de patrones. En particular, en bioinformática y en imágenes médicas es necesario evaluar sistemas de clasificación binaria [22], para lo cual se utilizan técnicas que facilitan el diseño y la visualización comparativa de los resultados como las curvas ROC (Receiver Operating Characteristics) [23]. Estas técnicas representan en muchos casos una solución apropiada a problemas complejos del mundo real, como el tratamiento de imágenes funcionales obtenidas mediante resonancia magnética nuclear (fMRI), que permiten determinar la actividad del cerebro humano ante estímulos predefinidos en el paradigma del correspondiente estudio médico [23].

Puede decirse que el procesamiento de imágenes como tal comienza en los años 1950-1960 debido a la combinación de dos hechos: por una parte la aparición de las computadoras digitales y por otra los *Programas Espaciales*. El objetivo inicial era muy concreto: mejorar la calidad visual de dichas imágenes. En este contexto se ha avanzado en el estudio de técnicas de eliminación de ruido, “blurring”, entre otras [5]. De esta manera, aparecen los métodos de restauración de imágenes que utilizan filtros inversos. Con las técnicas de eliminación de ruido y restauración se desarrollan las denominadas técnicas de mejora de imágenes, aunque tanto la eliminación de ruido como la restauración pueden entenderse como técnicas de mejora.

Asimismo, ha habido un creciente interés en el estudio del problema de la restauración de súper-resolución de secuencias de video. Mientras que en el problema tradicional de resolución se dispone de una única imagen de entrada, en la tarea de la súper-resolución se trabaja a partir de múltiples imágenes submuestreadas y degradadas, y de esta manera se puede tomar ventaja de datos espacio-temporales adicionales disponibles en la secuencia de video [2][3]. Un problema común e importante que surge en las comunicaciones visuales es la necesidad de crear una secuencia de imágenes de video con resolución mejorada a partir de un stream de video fuente de menor resolución [5][6][7][25].

Un tema más, que se suele considerar parte del procesamiento, es la compresión de imágenes. Puede citarse que el comienzo de la compresión de imágenes se sitúa en 1950 cuando se solicita la patente del código predictivo. Sin embargo, surge la pregunta de por qué se puede considerar la compresión parte del procesamiento de imágenes. La razón es simple: por una parte la compresión era, al menos inicialmente, una forma de procesamiento orientada a que las imágenes ocupasen menos espacio para transmisión y almacenamiento. Han surgido otras aplicaciones como la eliminación de artificios en imágenes y video comprimidos, un campo en continua expansión [8]. Uno de los aspectos más sorprendentes cuando se estudia la percepción humana es la capacidad del observador para determinar la estructura 3-D (Visión 3D) de los objetos a partir de patrones bidimensionales de luz [9][10].

Hoy en día existe un interés por comprender la capacidad de los algoritmos para poder derivar mediciones tridimensionales de alta precisión. Un tema actual es el de Reconstrucción 3D a partir de imágenes, generalmente capturadas por adquisidores ubicados en distintas posiciones. Una aplicación de esta línea de investigación es la

reconstrucción de piezas como por ejemplo antropológicas, dentarias, etc. [11].

Los dispositivos escáner 3D permiten obtener modelos 3D de objetos, esculturas, edificios, yacimientos arqueológicos, zonas urbanas y paisajes naturales. Utilizan medios ópticos (como por ejemplo láser) proyectados en el objeto de interés. Actualmente existen en el mercado diversos equipos de escáner 3D pero el costo de los mismos es muy elevado y su portabilidad es en muchos casos reducida. La estereovisión o visión estereoscópica es un método de reconstrucción 3D a partir de fotografías que no requiere de ningún equipo especial sino que puede funcionar con cámaras de fotos estándar. Cuando se intenta evaluar una aproximación computacional para la percepción artificial de formas 3D es necesario tener en cuenta dos hechos. Por una parte, que existen numerosos atributos de la estructura 3D que potencialmente podrían estar representados en el sistema visual (curvatura, distancia relativa, orientación local, entre otras) cuyas dificultades computacionales no son las mismas, y por otra, que para la evaluación de las diferentes aproximaciones computacionales es necesario examinar la validez de las hipótesis subyacentes. Puesto que existen numerosas escenas que producen la misma imagen, todos los análisis computacionales de la percepción 3D deben restringirse a un conjunto de posibles interpretaciones suponiendo una serie de restricciones más o menos reales [12].

Ciertamente los modelos de los objetos contienen más información de la que normalmente los sensores pueden extraer de una imagen. Por ello se suele trabajar con un modelo parcial que se puede extraer tanto del modelo del objeto como de la imagen observada. En función de ese modelo parcial es posible realizar el reconocimiento [13][14].

Es de interés en el área de procesamiento de imágenes el tratamiento de video en

tiempo real para el análisis de información proveniente de la trayectoria de los objetos capturados [15][16][17][18][24].

El volumen de datos y el computo intensivo requerido motivan el estudio de la paralelización de las aplicaciones.

2. LINEAS DE INVESTIGACION y DESARROLLO

- Captura y procesamiento de señales.
- Análisis de señales en tiempo real.
- Tratamiento de imágenes. Técnicas para la adquisición, mejorado y segmentación.
- Análisis de formas. Parametrización.
- Extracción de características invariantes a rotación y escala.
- Técnicas de análisis de tipologías edilicias. Segmentación y clasificación.
- Visión estereoscópica.
- Reconstrucción de modelos 3D basado en la combinación de cámaras estereoscópicas y la utilización de luz estructurada.
- Técnicas de estimación de movimientos, cálculos de trayectoria, velocidad y aceleración de objetos.
- Interfaz gráfica 3D para el diseño de modelos de simulación. Realidad Virtual.
- Herramientas de Realidad Virtual. Plug-In y API para Java. Diseño y programación de Interfaces Gráficas en 3D.
- Compresión de imágenes y video. Algoritmos para tratamiento de imágenes y video comprimido en tiempo real.
- Migración de algoritmos a arquitecturas multiprocesador y soluciones paralelas.

3. RESULTADOS OBTENIDOS /ESPERADOS

- Desarrollar soluciones a problemas específicos de procesamiento de imágenes y video, en particular en tiempo real [26].
- Se han desarrollado prototipos específicos en el ámbito del entrenamiento deportivo, como el análisis de la trayectoria de un balón [24][26].
- Se encuentra en etapa de desarrollo la reconstrucción 3D del tiro capturado en un ambiente de simulación. Desarrollos realizados para la Empresa CN Sports.
- Se ha desarrollado un sistema de reconocimiento automático de texto Braille.
- Se desarrollo de 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 [27][19].
- Desarrollar una herramienta de Realidad Virtual. Plug-In y API para Java. Diseño y programación de Interfaces Gráficas en 3D. La herramienta permitirá diseñar ambientes 3D para ser utilizados como interfaz gráfica de otras aplicaciones [20][21][28].
- Implementación de un scanner 3D móvil, hardware y software, mediante el uso de técnicas de estimación de desplazamiento y herramientas de visualización 3D para dispositivos móviles.
- Desarrollar una solución que permita detectar construcciones en imágenes aéreas de baja calidad. Esta investigación se lleva a cabo junto a la Facultad de Arquitectura (FAU-UNLP).

4. FORMACION DE RECURSOS HUMANOS

En esta línea de I/D se espera concluir una tesis doctoral en 2009, se encuentra en desarrollo una tesina sobre una herramienta

gráfica 3D de realidad virtual y esta prevista la realización de tesinas y tesis de postgrado en procesamiento de imágenes. Se han concluido como resultado de este proyecto las siguientes tesinas: Reconocimiento de texto braille, Segmentación de Imágenes Médicas y Reconstrucción 3D de trayectorias en entrenamiento deportivo.

5. BIBLIOGRAFIA

[1] Murat Tekalp, Digital Video Processing, Prentice Hall, 1995.

[2] Rafael C. Gonzalez, Richard E. Woods, Digital Image Processing (2nd Edition) (Hardcover), Prentice Hall; 2nd edition 2002).

[3] John C. Russ, The Image Processing Handbook, Fourth Edition (Hardcover), CRC Press; 4th edition 2002

[4] Kenneth R. Castleman, Digital Image Processing (Hardcover), Prentice Hall; 1st edition, 1995.

[5] Molina, R., Katsaggelos, A. K., y Mateos, J., "Multichannel Image Restoration using Compound Gauss-Markov Random Fields ", ICASSP 2000, 2000.

[6] Molina, R., Nuñez, J., Cortijo, F. y Mateos, J., "Image Restoration in Astronomy. A Bayesian Approach", IEEE Signal Processing magazine, 2001.

[7] Borko Furht, Joshua Greenberg, Raymond Westwater , "Motion Estimation Algorithms for Video Compression (The International Series in Engineering and Computer Science) (Hardcover)", Kluwer International, 1996.

[8] Peter Symes, "Digital Video Compression (with CD-ROM)(Paperback)", McGraw Hill, 2004.

[9] Yi Ma, Stefano Soatto, Jana Kosecka, S. Shankar Sastry , "An Invitation to 3-D Vision (Hardcover)", Springer, 2003.

[10] Trucco, Alessandro Verri, "Introductory Techniques for 3-D Computer Vision (Paperback)", Prentice Hall; 1998.

- [11] Emanuele Trucco, Alessandro Verri, *Introductory Techniques for 3D Computer Vision*, Prentice Hall – 1998.
- [12] Arturo de la Escalera, *Visión por Computador, Fundamentos y Métodos*, Prentice Hall, 2001.
- [13] Ramesh Jain, Rangachar Kasturi, Brian G. Schunck, *Machine Vision*, Mac Graw-Hill Internacional Edition, 1995.
- [14] Akihito Yamada, Yoshiaki Shirai and Jun Miura. *Tracking Players and a Ball in Video Image Sequence for Retrieving Scenes in Soccer Games*, Dept. of Computer-Controlled Mechanical Systems, Osaka University. 2-1, Yamadaoka, Suita, Osaka 565-0871, Japan, 2000.
- [15] Alvaro Castromán & Ernesto Copello. *Fútbol de Robots Uruguayo para Torneos*. Tesis de Grado de la Carrera de Ingeniería en Computación. Facultad de Ingeniería - Universidad de la República – 2004.
- [16] Jinchang Ren, James Orwell and Graeme A. Jones. *Digital Imaging, Estimating the Position of a Football from Multiple Image Sequences.*, Research Center, Kingston University. Penrhyn Road, Kingston upon Thames, Surrey, KT1 2EE, UK, 2003.
- [17] Jinchang Ren, James Orwell, Graeme Jones, Ming Xu. *Digital Imaging, Real-time 3D Soccer Ball Tracking from Multiple Cameras*, Research Centre, Kingston University. Surrey, KT1 2EE, U. K., 2004.
- [18] Xiao-Feng Tong, Han-Qing Lu, Qing-Shan Liu. *An Effective and Fast Soccer Ball Detection and Tracking Method*, National Lab. of Pattern Recognition, Inst. of Automation, Chinese Academy of Sciences – 2003.
- [19] Ron Kimmel, M. Bronstein, A. Bronstein. *"Numerical Geometry of Images : Theory, Algorithms, and Applications"*. Springer.
- [20] Daniel Selman. *"Java 3D Programming"*. Hanning.(A guide to key concepts and effective techniques).
- [21] Andrew Davison, "Killer Game Programming in Java". O'Reilly. (Java Gaming & Graphics Programming)
- [22] 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).
- [23] 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).
- [24] Federico Cristina, Sebastián H. Dapoto, Claudia Russo, Armando De Giusti, María José Abasolo. *Mobile path and spin 3D tracking and reconstruction*. *IV International Conference on Articulated Motion and Deformable Objects 2006*. AMDO 2006, Springer, Volumen: LNCS 4069. ISBN: 978-3-540-36031-5, ISSN:1611-3349, p.120-131.
- [25] Federico Cristina, Sebastián H. Dapoto, Claudia Russo, Oscar Bria. *High Resolution Images from Low Resolution Video Sequences*. *Journal of Computer Science & Technology*. Vol. 5, No 1, April 2005. ISBN: 1666-6038, p. 30-36.
- [26] Federico Cristina, Sebastián H. Dapoto, Claudia Russo. *A Lightweight Method for Computing Ball Spin in Real Time*. *Journal of Computer Science & Technology*. Vol. 7, No 1, April 2007. ISBN: 1666-6038, p. 34-38.
- [27] 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.
- [28] Federico Cristina, Sebastián H. Dapoto, Javier Vegas, Verónica Artola. *Herramienta de realidad virtual para diseño de interfaces gráficas 3D en Java*. *XIV Jornadas de Jóvenes Investigadores AUGM 2007*.