

No Fotorealismo y Minería de Imágenes

Fernandez Jacqueline Guerrero Roberto

Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (LIDIC)

Departamento de Informática

Universidad Nacional de San Luis

Ejército de los Andes 950

D5700HHW - San Luis

Tel: +54 (2652) 420823

Fax: +54 (2652) 430224

e-mail: {jmfer, rag}@unsl.edu.ar

Resumen

Durante los últimos años, la proliferación de los medios digitales ha creado la necesidad del desarrollo de herramientas para la eficiente representación, acceso y recuperación de información visual. La minería de imágenes se ha convertido en una importante rama de investigación a causa del potencial que posee en descubrir patrones característicos a partir de un conjunto extenso de imágenes. El reconocimiento de patrones implica la identificación de relaciones invariantes en una colección de imágenes a modo de poder obtener información para una posterior clasificación. Por otro lado, cuando una persona observa imágenes surge una asociación natural como consecuencia de la información visual que estas brindan. El área de No fotorealismo estudia el desarrollo de técnicas y metodológicas que faciliten la transmisión de información mediante imágenes. Es de interés el estudio de las habilidades y posibles relaciones existentes entre estas dos áreas a modo de poder bosquejar técnicas o metodologías que permitan la colaboración entre ellas.

1- Introducción

En los últimos años, la industria de la computación se ha desarrollado rápidamente en todos sus campos: tecnológico y, de procesamiento, almacenamiento y acceso a la información. El avance tecnológico ha habilitado a la digitalización de datos; en la actualidad es extremadamente fácil la obtención y almacenamiento de grandes cantidades de datos, particularmente imágenes digitales.

La manipulación y administración de repositorios de imágenes es en sí misma un tema de importancia a ser tratado. Los sistemas de bases de datos de imágenes pretenden establecer un soporte para el almacenamiento y recuperación de las mismas. Las imágenes son recuperadas a partir de una imagen ejemplo entregada al sistema, donde las características típicas de las imágenes almacenadas son contrastadas con las de la imagen entregada en la consulta. Luego de la evaluación, las imágenes son ordenadas acorde con la similitud con la imagen de referencia.

La recolección de imágenes y su consecuente generación de repositorios se realiza en un amplio rango de ámbitos: comerciales, institucionales, militares. Su uso generalizado se debe a que las mismas son una fuente potencial de información a ser posteriormente analizada y procesada. Ello implica la identificación de las características presentes en una imagen y su consecuente evaluación. Una dificultad en esta tarea es disponer del conocimiento de todo el dominio de imágenes existentes para poder luego inferir información.

En consecuencia, existe actualmente una creciente demanda por sistemas que puedan en forma automática analizar imágenes y extraer información semánticamente relevante, al mismo tiempo que modelen relaciones y categorías en una colección de imágenes.

El análisis de la información entregada por un medio consiste en el análisis del contenido semántico del medio, o lo que es lo mismo, su división en objetos semánticos que lo conforman.

Si el medio es una imagen digital, el valor individual de los pixels de la misma no constituye un ente semántico en sí mismo, sino mas bien la combinación espacial de todos ellos . La combinación de los pixels habilita a un tratamiento global de la imagen permitiendo la identificación de características propias de color, forma o textura.

Por otro lado, la manera en que la información es presentada (es plasmada en el medio utilizado) facilita la identificación de características y así la extracción de información relevante. La misma información puede ser representada en un mismo medio mediante diferentes propiedades visuales dependiendo del objetivo final de la misma. Dichas propiedades visuales se verán reflejadas en restricciones en la generación de la imagen y consiguientemente en los valores de los pixels y sus interrelaciones.

Luego, lo anteriormente expresado podría plantearse como un problema general donde se deben tomar decisiones acorde con las características globales de una imagen (color, forma, textura) basadas en aspectos locales de la misma (pixels y sus interrelaciones).

Las áreas de Minería de Imágenes y Generación de Imágenes No Fotorealísticas se encuentran abocadas a la solución del mismo problema desde ópticas diferentes. En el caso de la primera, la intención es obtener información global a partir de información local, y en la segunda el poder brindar información global mediante la manipulación de la información local.

A continuación se realiza una pequeña reseña de cada una de las áreas a modo de clarificar la modalidad de afrontar el problema por parte de cada una de ellas.

1.1 - Minería de Imágenes

La Minería de Datos es un proceso relacionado con descubrir patrones, asociaciones, anomalías y estructuras en volúmenes de datos estadísticamente significantes. El proceso consiste de dos pasos principales: el pre-procesamiento de los datos, durante el cual las características o atributos de alto nivel mas relevantes son extraídas de los datos de bajo nivel, y el reconocimiento de patrones, en el cual un patrón en los datos es reconocido usando estas características.

Luego de realizadas las dos etapas, los patrones identificados deben ser mostrados al usuario para la validación de los mismos. De la validación puede ocurrir la necesidad de refinar aún más los datos, existiendo la posibilidad de utilizarlos como realimentación para un nuevo proceso de minería. Consecuentemente, la minería de datos es básicamente un proceso iterativo e interactivo.[1,2,3,7,11]

La minería de datos es un área multidisciplinaria, tomando ideas de otras áreas tales como la inteligencia artificial y el aprendizaje de máquinas, la estadística, la computación de alta performance, el procesamiento de señales, el procesamiento de imágenes, la optimización matemática y el reconocimiento de patrones, entre otros [4,5,6,8,9,10]. A su vez, los campos de aplicación son también variados, entre ellos el censado remoto, la bio-informática, las imágenes por computadora, la astronomía, la minería web y la minería de textos.

Si bien en el proceso de minería gran parte de la atención se destina a la etapa de reconocimiento de patrones, la etapa de pre-procesamiento tiene mayor influencia sobre el éxito del proceso total. Sin embargo, y desafortunadamente, esta etapa depende del dominio de los datos y del problema en particular. Este es el caso de la extracción de información a partir de una imagen.

En la actualidad los desarrollos mas innovadores se realizan en el área de la tecnología satelital [1,2,3], donde el alto número de imágenes capturadas por sensado remoto hace imprescindible disponer de una herramienta automática para la extracción de contenido, clasificación y recuperación basada en el contenido. Otra área se relaciona con el análisis de imágenes médicas complejas y la extracción de información significativa [4,5,8]. También se puede mencionar el análisis de información multimedial de alto nivel sobre bases de datos multimediales[9,10], entre otros.

Una imagen digital es un arreglo rectangular de pixels donde cada una de sus celdas posee un valor numérico representando características físicas que la definen unívocamente. La extracción de información a partir de la misma consistirá en poder identificar los objetos que constituyen la misma.

En el mundo real, los bordes constituyen el contorno de un objeto. Un borde es el límite entre un objeto y su contexto, e indica el límite entre objetos que se superponen. Esto significa que si los bordes en una imagen pueden ser identificados certeramente y las regiones que ellos determinan clasificadas, todos los objetos pueden ser localizados y sus propiedades básicas tales como el área, perímetro y forma pueden ser medidos.

La piedra fundamental de toda aplicación destinada al manejo de información de contenido son los algoritmos de segmentación. Existen en la bibliografía numerosos métodos de segmentación basados en diferentes técnicas de procesamiento de imágenes [4,7,8]. El éxito de los métodos depende entre otros aspectos, del tipo y tamaño de imágenes analizadas. Las imágenes de gran tamaño implican problemas de eficiencia en cuanto al tiempo de proceso y calidad de los resultados. De la misma manera, las imágenes con alto nivel de detalle involucran una alta probabilidad de error en la identificación de los objetos semánticos que la conforman.

1.2 - No Fotorealismo

Tradicionalmente, el término Computación Gráfica ha sido asociado al término imagen; éste último como la réplica o imitación visual de un objeto o una escena. Es por ello que históricamente los desarrollos en el área se han centrado en la descripción geométrica tridimensional de los objetos, sus propiedades materiales y la simulación de las fuentes de luz, para luego combinarlos y plasmar así, una proyección bidimensional de la escena. Como resultado, el proceso de generación de imágenes se redujo a un proceso automático de simulación física [12,13].

Sin embargo, la mayoría de las imágenes no tienen como único propósito el representar solamente propiedades visuales de una escena; no son siempre una simple proyección unidireccional de una escena 3D a una descripción bidimensional. Estas pueden tener como objetivo el transmitir un mensaje educativo, estético o emocional, entre otros.

El arte y oficio de crear una gráfica procura optimizar el resultado final acorde con un objetivo bajo ciertas restricciones establecidas por el medio utilizado, el contexto social y el estilo artístico, entre otros. Como caso particular, un artista puede plasmar una escena la cual aparentemente se asemeje a una representación tridimensional de la realidad pero que no exista ningún escenario objetivo que siguiendo las técnicas de simulación tradicionales pueda ser proyectado en una representación bidimensional. Lo que el artista pretende es lograr que su representación sea “convinciente” antes que “realista” [15,16]. La pretensión del artista añade un aspecto no considerado en la computación gráfica tradicional: la compleja interrelación entre una representación gráfica y el observador.

Transponiendo este concepto a la computación gráfica, y extendiéndolo, se podría pensar que el proceso de generación de una imagen se compone de un conjunto de estrategias de interacción entre el modelo, las técnicas de plasmado (rendering) y el usuario. La incorporación del usuario-observador al proceso implica la consideración de un conjunto de diversos aspectos cognitivos, físicos y sociales antes no considerados y que adquirirán mayor o menor relevancia dependiendo del contexto y finalidad a otorgar a una imagen.

En consecuencia, el destino final de una imagen establece restricciones en la generación de la misma, ya sea en términos de claridad, representación de sus cualidades o su plasmado en 2D.

Finalmente, se podría redefinir el proceso de generación de una imagen como un problema de optimización que tiene como objetivo producir la mejor representación gráfica para un propósito en

particular. El área de No Fotorealismo intenta reunir en uno, los tres aspectos antes aludidos mediante el estudio e incorporación de diferentes métodos, técnicas, medios o combinaciones de ellos al proceso de generación de una imagen [14,16,17,18,19].

En la actualidad, entre los desarrollos mas novedosos se encuentra la formulación de técnicas no fotorealísticas para la detección, realce y comunicación de formas de objetos en imágenes reales. Todo ello con el objeto de facilitar y aumentar la comprensión de imágenes confusas o de bajo contraste provenientes de escenarios complejos tales como la medicina (imágenes de órganos) o la mecánica (imágenes de partes) [18,19,20,21,22,23,24,25,26,27].

2 - Lo que se pretende

El área de minería de imágenes se centra en el análisis de imágenes para la identificación de características que permitan la obtención de un conocimiento a priori del tipo de información que puede ser requerido de una imagen.

El área de No Fotorealismo se centra en la generación de imágenes que transmitan información adicional al simple plasmado de pixels.

Desde un punto de vista computacional, las imágenes y la información contenida en ellas son un nexo que habilita a la resolución de un variedad de problemas colaborativamente, permitiendo una realimentación de conocimiento y enriquecimiento hacia nuevas líneas de trabajo, como ejemplo se podrían mencionar:

- La discusión de métodos No Fotorealísticos que permitan la rápida extracción y/o derivado de características y en consecuencia una evaluación semántica eficiente de las mismas.
- La investigación de técnicas que permitan aumentar la “alfabetización visual” del área de minería de imágenes, ello con el objeto de incorporar a la misma una nueva capacidad, la de observador.

3- Bibliografía

- [1] Selim A., Krzysztof K., Carsten T., Giovanni M., “*Interactive training of advanced classifiers for mining remote sensing image archives*” (Industry/government track posters), Proceedings of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 773-782 , [2004].
- [2] Ruofei Z., Zhongfei Z., Sandeep K., “*A data mining approach to modeling relationships among categories in image collection*” (Research track posters), Proceedings of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 749-754 , [2004].
- [3] Da D., Jianhua Z., Martin P., “*Visualisation and comparison of image collections based on self-organised maps*”, Proceedings of the second workshop on Australasian information security, Data Mining and Web Intelligence, and Software Internationalisation - Volume 32, pp. 97–102, [2004].
- [4] Mitchell T., Hutchinson R., Just M., Niculescu R.S., Pereira F., Wang X., “*Classifying Instantaneous Cognitive States from fMRI Data*”, Proc. 2003 American Medical Informatics Association Annual Symposium, pp. 465-469, [2003].
- [5] Hsu W., Lee M., Goh K., “*Image mining in IRIS: integrated retinal information system*”, ACM SIGMOD Record , Proceedings of the 2000 ACM SIGMOD international conference on Management of data, Volume 29 Issue 2, pp 593-596 [2000].
- [6] Chinrungrueng Ch., Toonkum P., “*Real-Time Speckle Reduction and Coherence Enhancement of Ultrasound Images Based on Anisotropic Savitzky-Golay Filters*”, Proceedings of International Conference on Systems, Man and Cybernetics (IEEE SMC 2004), October 2004, The Hague, The Netherlands.

- [7] Jurisica I., Glasgow J., “*Extending case-based reasoning by discovering and using image features in IVF*”, Proceedings of the 2000 ACM symposium on Applied computing, pp. 52-59, [2000].
- [8] Megalooikonomou V., Davatzikos Ch., Herskovits E., “*Mining lesion-deficit associations in a brain image database*” Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 347-351, [1999].
- [9] Keiji Y., “*Managing images: Generic image classification using visual knowledge on the web*”, Proceedings of the eleventh ACM international conference on Multimedia, pp. 167-176, November 2003.
- [10] Kosala R., Blockeel H., “*Web mining research: a survey*”, ACM SIGKDD Explorations Newsletter, Volume 2 Issue 1, pp. 1-15, June 2000.
- [11] Wang Y., Makedon F., Ford J., Shen L., Goldin D., “*Image and video digital libraries: Generating fuzzy semantic metadata describing spatial relations from images using the R-histogram*”, Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries, pp. 202-211, [2004].
- [12] Foley J. D., Van Dam A., Feiner S. K., And Hughes J. H., *Computer Graphics Principles and Practice, 2nd Edition*, Addison-Wesley, [1990].
- [13] Watt A. And Policarpo F., *The Computer Image*, Addison-Wesley, [1998].
- [14] NPAR Discussion. First International Symposium on Non-Photorealistic Animation and Rendering, [2000].
- [15] J. Willats, *Art and Representation*, Princeton U. Pr., [1997].
- [16] Gooch and Gooch, *Non Photorealistic Rendering*, A. K. Peters Ltd., [2000].
- [17] P. Debevec, C. Taylor and J. Malik, “*Modeling and rendering architecture from photographs: A hybrid geometry- and image-based approach*”, in Proceedings of SIGGRAPH 96, [1996].
- [18] Saito T., Takahashi T., “*Comprehensible rendering of 3D shapes*”, in Proceedings of SIGGRAPH 90, [1990].
- [19] Shiraishi M., Yamaguchi Y., “*An algorithm for automatic painterly rendering based on local source image approximation*”, in proceedings of NPAR 2000, [2000].
- [20] P. Sander, X. Gu, S. Gortler, H. Hoppe and J. Snyder, “*Silhouette clipping*”, in Proceedings of SIGGRAPH 2000, [2000].
- [21] J. Elder and R. Godberg, “*Image editing in the contour domain*”, in IEEE Trans. On Pattern Analysis and Machine Intelligence, 23(3), [2001].
- [22] Gooch A., Gooch B., Shirley P., Cohen E., “*A non-photorealistic lighting model for automatic technical illustration*”, in Proceedings of SIGGRAPH 98, [1998].
- [23] Raskar R., Tan K., Feris R., Yu J., Turk M., “*Non-photorealistic Camera: Depth Edge Detection and Stylized Rendering using Multi-Flash Imaging*”, in Proceedings of SIGGRAPH 98, [2004].
- [24] Akers D., Losasso F., Klingner J., Agrawala M., Rick J., Hanrahan P., “*Conveying Shape and Features with Image-Based Relighting*” in IEEE Visualization, [2003].
- [25] De Carlo D., Finkelstein A., Rusinkiewicz S., Santella A., “*Suggestive contours for conveying shape*”, ACM Trans. Graph., 22(3), 848-855, [2003].
- [26] Raskar R., Ilie A., Yu J., “*Image Fusion for Context Enhancement and Video Surrealism*”, In Proceedings of NPAR, [2004].
- [27] Shroeder M., Moser S., “*Automatic color correction based on generic content-based image analysis*”, Ninth Color Imaging Conference, pp. 129-139, [2001].