

ANÁLISIS VISUAL DE DATOS MULTIDIMENSIONALES PROVENIENTES DE LAS CIENCIAS GEOLÓGICAS

M. Luján Ganuza¹, Antonella S. Antonini¹, Florencia Gargiulo², Gabriela Ferracutti², Ernesto A. Bjerg², Silvia M. Castro¹, Krešimir Matković³, Eduard Gröller⁴

⁽¹⁾ Laboratorio de I+D en Visualización y Computación Gráfica
Dpto. de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur (DCIC-UNS)
Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET)
Laboratorio de I+D en Visualización y Computación Gráfica,
(UNS-CIC Prov. de Buenos Aires)

mlg@cs.uns.edu.ar; antoo.antonini@gmail.com; smc@cs.uns.edu.ar

⁽²⁾ INGEOSUR, Dpto. de Geología, Universidad Nacional del Sur
florenciagargiulo@gmail.com; gferrac@uns.edu.ar; ebjerg@ingeosur-conicet.gob.ar

⁽³⁾ VRVis Zentrum für Virtual Reality und Visualisierung Forschungs-GmbH,
Donau-City-Strasse 11, 1220 Viena, Austria.

Matkovic@vrvis.at

⁽⁴⁾ Technische Universität Wien, Institut für Computergraphik und Algorithmen,
Favoritenstrasse 9-11 / E186, A-1040 Viena, Austria
groeller@cg.tuwien.ac.at

RESUMEN

Muchos de los problemas a los que debemos enfrentarnos en diversas áreas del conocimiento implican el tratamiento de grandes volúmenes de datos que requieren ser analizados desde múltiples

dimensiones. Al analizar datos multidimensionales el problema radica en considerar una gran cantidad de variables y sus relaciones simultáneamente. A medida que la cantidad de variables aumenta, la habilidad del usuario de entender la interacción o correlación entre éstas se ve severamente limitada si se emplean técnicas de visualización estándar, dado que éstas dimensiones no solo pueden referirse a escalas diferentes y representar distintos tipos de datos, sino que además, la calidad de estos datos puede no ser la esperada.

En este contexto, el objetivo de la presente línea de investigación es analizar la adecuación y/o la generación de técnicas de

visualización de datos multidimensionales, para responder a las necesidades de visualización, volumen de datos, dimensionalidad, complejidad y capacidad exploratoria presentes en los datos geológicos multidimensionales.

Palabras Clave: Datos Multidimensionales, Análisis Visual de Datos, Visualización de Datos Geológicos, Rendering de Volúmenes

CONTEXTO

Este trabajo se lleva a cabo con investigadores del Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab, UNS-CIC Prov. de Buenos Aires) del Departamento de Ciencias e Ingeniería de la Computación (DCIC-UNS), el Departamento de Geología de la UNS, el INGEOSUR CCT-CONICET, el VRVis Research Center y la Universidad Técnica de Viena (TU-Wien), Austria.

Los trabajos realizados bajo esta línea involucran a docentes investigadores, becarios doctorales y alumnos de grado.

1. INTRODUCCIÓN

La visualización es una herramienta muy valiosa que nos permite detectar, inferir y sacar conclusiones sobre las posibles relaciones existentes entre los datos. Esta herramienta es muy utilizada en diversas áreas de las ciencias, la ingeniería y la medicina, entre otras, en las cuales se generan volúmenes de datos cada vez más grandes y difíciles de analizar y comprender sin un soporte visual adecuado [1, 2, 3, 4, 5]. Es en estos casos en los que el aporte de la visualización a la exploración y entendimiento de grandes conjuntos de datos resulta altamente significativo.

Al aumentar la dimensionalidad de los datos, lograr una visualización eficiente de los mismos se convierte en una tarea cada vez más compleja, ya que los espacios multidimensionales son difíciles de entender, por su propia naturaleza, debido a la gran cantidad de dimensiones que se deben representar en un espacio muy limitado. El problema radica en la necesidad de considerar una gran cantidad de variables y sus relaciones simultáneamente. Muchas de las técnicas tradicionales sólo producen una visión parcial de los datos, ya sea por su dimensionalidad y/o su cantidad. Esto dificulta la detección, si fuese necesario, de las posibles relaciones que existen entre todas las variables que intervienen en un proceso, generando vacíos de conocimiento por la imposibilidad de tratar el fenómeno en toda su dimensionalidad.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En esta línea de investigación en particular, nos enfocaremos en un tipo específico de

datos multidimensionales, los datos geológicos.

Los trabajos desarrollados por los geólogos, tanto de campo como de laboratorio, implican grandes desafíos en lo que respecta al procesamiento y análisis de los datos que se obtienen en cada uno de estos contextos. Una gran cantidad de estos datos son indudablemente datos multidimensionales.

Los geólogos analizan grandes volúmenes de datos provenientes de análisis geoquímicos. Estos análisis arrojan una gran cantidad de atributos para cada análisis efectuado y éstos tienen asociados dominios espaciales característicos como los prismas o los tetraedros. Debido a las distintas características de los datos involucrados, no hay una única metáfora visual que pueda satisfacer todas las necesidades analíticas. Por esta razón, resulta de suma utilidad contar con un conjunto de metáforas visuales y de técnicas de visualización asociadas que permitan analizar estos datos. Estas representaciones deben poder comunicar efectivamente las características del espacio de información así como incentivar el descubrimiento. En este contexto, y en base a los trabajos realizados y a la experiencia obtenida hasta el momento en lo que respecta a Visualización de Datos Geológicos, se propone avanzar en el conocimiento sobre el tema de análisis visual de datos multidimensionales constituidos principalmente, aunque no exclusivamente, por datos geoquímicos, isotópicos, petrológicos, mineralógicos y estructurales.

Las técnicas tradicionales de visualización, tales como scatter plots, histogramas, etc., sólo producen una visión parcial de la información. Actualmente existen técnicas de visualización utilizadas en distintos contextos, que permiten analizar grandes conjuntos de datos multidimensionales y que han probado ser de gran utilidad [6, 7, 8].

En este contexto, se plantea la necesidad de contribuir en el desarrollo de tecnologías y soluciones en torno al análisis visual de datos multidimensionales. Particularmente, nos centraremos en la visualización de datos geológicos, con el objetivo de desarrollar nuevos modelos, técnicas y herramientas que permitan a profesionales de las Geociencias explorar efectiva y eficientemente sus espacios de datos. Si bien nos enfocaremos en la aplicación de estas nuevas tecnologías en el campo de la Geología, estos métodos redundarán muy probablemente en soluciones extensibles a distintas disciplinas en las que también deba obtenerse información a partir de conjuntos de datos multidimensionales.

3. RESULTADOS OBTENIDOS Y ESPERADOS

Los autores de este artículo conforman un grupo interdisciplinario integrado por profesionales de Ciencias de la Computación y de Ciencias Geológicas, que ha trabajado en temas relacionados con la Visualización de Datos aplicada a Geociencias desde hace varios años [9, 10, 11]. Además, con la colaboración del Institute of Computer Graphics and Algorithms de la Universidad Técnica de Viena (TU-Wien) y del VRVis Research Center de Viena, se han desarrollado nuevas técnicas de visualización para datos geológicos aplicados a la mineralogía y a la prospección de recursos naturales [12, 13, 14].

Específicamente, se plantea el diseño y desarrollo de técnicas y herramientas para la visualización de datos geoquímicos que contribuyan efectivamente a lograr una mejor comprensión de la interacción entre los procesos geológicos y las composiciones de minerales y de rocas en la configuración geológica de una región en particular. Se propone utilizar todos los atributos disponibles de los datos y metadatos

generados a partir de éstos, para validar y mejorar técnicas utilizadas en el contexto de datos altamente dimensionales (por ejemplo, datos de química mineral y datos geoquímicos de rocas). Para alcanzar este objetivo, nos centraremos en dos líneas principales:

- Diseño y desarrollo de técnicas de visualización para datos multidimensionales en el contexto de datos geológicos. Analizar técnicas y métodos de análisis visual para incorporar más de 8 miembros finales (y hasta un total de 22), que son los que se consideran actualmente y también datos geoquímicos de elementos mayoritarios. Esto se corresponde con el diseño y desarrollo de nuevas técnicas de visualización para el análisis de minerales del grupo de los espinelos.
- Extensión de las técnicas avanzadas de clasificación de espinelos. Extender al espacio 3D las técnicas avanzadas de clasificación de los minerales del grupo de los espinelos desarrolladas en nuestro trabajo previo [13, 14] y definir las interacciones asociadas. Esto también contribuirá significativamente al desarrollo de tecnologías y soluciones en torno a la generación de visualizaciones interactivas en 3D.

4. FORMACIÓN DE RECURSOS HUMANOS

En lo concerniente a la formación de recursos humanos se detallan las tesis finalizadas y en desarrollo relacionadas con las líneas de investigación presentadas, así también como los becarios y los proyectos de investigación vinculados.

Tesis Finalizada: “Interacciones en Visualización”, tesis de doctorado en Ciencias

de la Computación. Alumna: M. Luján Ganuza. Directora: Dra. Silvia Castro.

Proyecto Final Finalizado: “Visualización de Datos Geológicos en la Web”. Proyecto final de la Ingeniería en Sistemas de Computación. Alumna: Antonella S. Antonini. Directora Silvia Castro. Codirectora: M. Luján Ganuza.

Proyecto Final Finalizado: “Técnicas 3D para Visualización de Datos Geológicos en la Web”. Proyecto final de la Ingeniería en Sistemas de Computación. Alumno: Gonzalo Picorel. Directora Silvia Castro. Codirectora: M. Luján Ganuza.

Becaria: Antonella S. Antonini. Título del plan propuesto: “Visualización de Datos Geológicos en la Web”. Beca de entrenamiento 2018. Comisión de Investigaciones Científicas de la Provincia de Buenos Aires. Adjudicada a partir de abril de 2018 y por un término de 12 meses.

Proyecto: PGI 24/N037 “Análisis Visual de Grandes Conjuntos de Datos”. Directora: Silvia M. Castro.

Proyecto: PICT 2017-1246 “Análisis Visual de Datos en Geociencias”. Directora: Silvia M. Castro.

Becaria: M. Luján Ganuza. Título del plan propuesto: “Análisis Visual de Datos Multidimensionales en Espacios Ad-Hoc”. Beca Posdoctoral 2018 CONICET. Comisión Nacional de Investigaciones Científicas y Técnicas de la República Argentina. Adjudicada a partir de abril de 2019 y por un término de 24 meses.

5. BIBLIOGRAFÍA

[1] Cleveland, W. S. Visualizing Data. Hobart Press, 1993.

[2] Card, S. K., Mackinlay, J. D., and Shneiderman, B. Readings in information visualization: using vision to think. Morgan Kaufmann, 1999.

[3] Ware, C. Information Visualization: Perception for Design. Morgan Kaufmann, 2004.

[4] Spence, R. Information Visualization: Design for Interaction. 2007.

[5] Ward, M. O., Grinstein, G., and Keim, D. Interactive data visualization: foundations, techniques, and applications. CRC Press, 2010.

[6] J Pastizzo, M., Erbacher, R., and Feldman, L. Multidimensional data visualization. 158–62.

[7] Everitt, B., and Hothorn, T. Looking at multivariate data: Visualisation. In An Introduction to Applied Multivariate Analysis with R. Springer, 2011, pp. 25–60.

[8] Etemadpour, R., Linsen, L., Paiva, J. G., Crick, C., and Forbes, A. G. Choosing visualization techniques for multidimensional data projection tasks: A guideline with examples. In International Joint Conference on Computer Vision, Imaging and Computer Graphics (2015), Springer, pp. 166–186

[9] Ganuza, M. L., Castro, S. M., Martig, S. R., Ferracutti, G., and Bjerg, E. Mineral compositions visualization implementing the spinel prism. *Proceedings Congreso Argentino de Ciencias de la Computación. CACIC 2009* (2009), 576–585.

[10] Ganuza, M. L., Castro, S. M., Ferracutti, G., Bjerg, E. A., and Martig, S. Spinelviz: An interactive 3d application for visualizing spinel group minerals. *Computers & Geosciences* 48 (2012), 50–56.

[11] Ferracutti, G. R., Gargiulo, M. F., Ganuza, M. L., Bjerg, E. A., and Castro, S. M. Determination of the spinel group end-members based on electron microprobe analyses. *Mineralogy and Petrology* 109, 2 (2015), 153–160.

[12] Ganuza, M. L., Ferracutti, G., Gargiulo, M. F., Castro, S. M., Bjerg, E. A., Gröller, E., and Matkovic, K. The spinel explorer - interactive visual analysis of spinel group minerals. *IEEE Trans. Vis. Comput. Graph.* 20, 12 (2014), 1913–1922.

[13] Ganuza, M. L., Gargiulo, M. F., Ferracutti, G., Castro, S. M., Bjerg, E. A., Gröller, E., and Matkovic, K. Interactive semi-automatic categorization for spinel group minerals. In *2015 IEEE Conference on Visual Analytics Science and Technology, VAST 2015, Chicago, IL, USA, October 25-30, 2015* (2015), pp. 197–198.

[14] Ganuza, M. L., Ferracutti, G., Gargiulo, M. F., Castro, S. M., Bjerg, E. A., Gröller, E., and Matkovic, K. Interactive visual categorization of spinel-group minerals. *Proceedings of the Spring Conference on Computer Graphics* (2017)