VISUALIZACIÓN DE DATOS MULTIDIMENSIONALES PROCEDENTES DE LAS GEOCIENCIAS

Antonella S. Antonini[1], M. Luján Ganuza[1], Florencia Gargiulo[2], Gabriela Ferracutti[2], Ernesto A. Bjerg[2], Silvia M. Castro[1], Krešimir Matković[3], Eduard Gröller[4]

(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)
antonella.antonini @cs.uns.edu.ar; mlg@cs.uns.edu.ar; 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) Institute of Visual Computing & Human-Centered Technology
Favoritenstrasse 9-11 / E186, A-1040 Viena, Austria
groeller@cg.tuwien.ac.at

RESUMEN

En diversas disciplinas, incluidas las Ciencias Geológicas, a menudo se generan grandes volúmenes de datos multidimensionales.

Debido a la dimensionalidad y/o volumen de los datos, muchas técnicas tradicionales solo pueden producir una vista parcial de los mismos [13, 14, 15, 16].

Al aumentar considerablemente el número de variables e intentar representar dichos datos en un espacio muy limitado, la tarea del usuario de visualizar; interactuar y analizar posibles relaciones, patrones y tendencias, se vuelve cada vez más compleja.

Por ende, es valioso disponer de un conjunto de metáforas visuales y técnicas de visualización para analizar datos con estas características.

El objetivo general es promover el desarrollo de tecnologías y soluciones en torno al análisis visual de datos multidimensionales. En particular, nos enfocaremos en la visualización de datos geológicos, con el fin de generar nuevos modelos, técnicas y herramientas para potenciar el análisis y la interpretación de los mismos.

Nos centraremos en la aplicación de estas técnicas al campo de la Geología, pero tenga en cuenta que estas técnicas se pueden utilizar en una variedad de disciplinas.

Palabras Clave: Visualización de Datos Multidimensionales; Análisis Visual; Visualización de Datos Geológicos

CONTEXTO

Esta línea de trabajo se lleva a cabo en el Departamento de Ciencias e Ingeniería de la Computación (DCIC-UNS). Trabajamos en estrecha colaboración con investigadores de reconocidos institutos de investigación nacionales e internacionales, incluidos el Institute of Visual Computing & Human-Centered Technology (TuWien,

Austria), VRVis (Center for Virtual Reality and Visualization, Austria), INGEOSUR-CONICET (Instituto Geológico del Sur) y el Departamento de Geología de la Universidad Nacional del Sur. Además, contamos con investigadores del Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab).

1. INTRODUCCIÓN

La visualización de datos está siendo cada vez más utilizada en diversas áreas ante la necesidad de mostrar adecuadamente conjuntos de datos y la dificultad de extraer información relevante de ellos.

Dentro de la visualización de datos, existen múltiples técnicas que permiten un mejor entendimiento de la información. Sin embargo, la representación de datos multivariables y multidimensionales requiere una visualización más compleja. Por ello, no es posible su representación de manera intuitiva utilizando una única gráfica. Estos datos se componen de multitud de variables que son importantes en conjunto y que necesitan una representación que ayude a analizarlos e inferir conclusiones sobre las posibles relaciones existentes entre ellos.

Nuestro objetivo y principal dificultad, es el estudio de técnicas que exploten al máximo las capacidades de percepción humana y permitan representar estos datos de una manera sencilla, evitando visualizaciones complejas y confusas.

En este contexto, exploramos nuevas técnicas multidimensionales y diversas visualizaciones correlacionadas que permiten al usuario el análisis visual de datos multidimensionales.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En particular trabajamos con datos provenientes de las Ciencias Geológicas, pero sin perder de vista el objetivo general de desarrollar técnicas de análisis de datos que sean útiles en diversas disciplinas.

Las diferentes líneas de investigación y desarrollo se orientan a la visualización de datos geoquímicos.

Un gran desafío para los geólogos consiste en caracterizar una región geológica en términos de su ambiente geotectónico. Para alcanzar este objetivo, se debe procesar y analizar una gran cantidad de información obtenida de rocas y minerales [10, 11, 12]. En particular, analizan los minerales que integran el grupo de los espinelos; un importante indicador petrogenético, que brinda información vital en lo referido al ambiente tectónico de las rocas presentes en un área, y proveen información valiosa acerca de los procesos que sufrieron las rocas que contienen a esos minerales.

Si bien existen técnicas de visualización que permiten analizar grandes conjuntos de datos multidimensionales provenientes de distintos contextos, en el análisis de datos geoquímicos sólo se trabaja con gráficos especiales que considera únicamente un conjunto reducido de estos atributos a la vez.

En este contexto, es importante diseñar y desarrollar métodos apropiados de visualización de datos multidimensionales que permitan a los geólogos considerar todos los datos obtenidos en el análisis.

Cabe destacar, que al trabajar juntamente con investigadores del Instituto Geológico del Sur (INGEOSUR-CONICET) y del Departamento de Geología de la Universidad Nacional del Sur, es posible validar los resultados obtenidos y verificar la utilidad de las nuevas técnicas propuestas.

3. RESULTADOS OBTENIDOS Y ESPERADOS

Desde hace tiempo trabajamos en diversas técnicas y herramientas para la visualización

de datos procedentes de las Geociencias [1, 2, 3, 4, 5].

En los últimos años, se implementó una herramienta web gratuita para la visualización de datos geológicos que integra la mayoría de los diagramas que utilizan los geólogos habitualmente para analizar las características químicas de los minerales e incorpora visualizaciones interactivas en 3D [6]. También proporciona vistas coordinadas e interacciones adecuadas para que los usuarios interactúen con sus conjuntos de datos.

En 2022 se desarrolló una aplicación de libre acceso [7] que integra un conjunto de técnicas de visualización coordinadas en un único dashboard. La herramienta puede ser utilizada en diversas disciplinas y brinda al usuario la posibilidad de visualizar un mismo conjunto de datos con diferentes técnicas de visualización y luego reutilizar la configuración de vistas con otros conjuntos de datos. VISUEL es totalmente interactivo y admite interacciones tradicionales como filtrado, selección, brushing, y zoom, entre otros. También permite al usuario configurar la representación visual de sus datos, eligiendo el color y la forma de las representaciones.

Además, en el último año llevamos a cabo una revisión sistemática de las técnicas de visualización presentes en la literatura y definimos una taxonomía para organizar las contribuciones recopiladas [8]. El trabajo realizado propone una clasificación completa de las técnicas y las interacciones asociadas; y ayuda a identificar áreas que no están suficientemente exploradas y son candidatas a ser investigadas.

También, desarrollamos una librería llamada np-GLCVis [9] que admite las Non-Paired General Line Coordinates (npGLC). Las Coordenadas de Líneas Generales (del inglés General Lines Coordinates o GLC) son representaciones visuales reversibles que conservan información multidimensional para el descubrimiento del conocimiento. La librería presenta una colección de métodos de visualización e interacciones tradicionales, diseñada para experimentar con técnicas npGLC en el desarrollo de aplicaciones de visualización.

Actualmente, nos encontramos trabajando en el diseño e implementación de diversas herramientas de visualización que integran técnicas multidimensionales de análisis visual con modelos de Machine Learning. Nuestro objetivo es facilitar el trabajo de los geólogos en el proceso de categorización de una muestra de datos para determinar a qué ambiente tectónico corresponde.

4. FORMACIÓN DE RECURSOS HUMANOS

A continuación, se detallan los proyectos de investigación vinculados con las líneas de investigación presentadas y las becas adquiridas en centros de investigación científica.

Proyectos vigentes

• PGI SeCyT-UNS (24//N048) *"Análisis Visual de Datos"*. Directora: Silvia M. Castro.

• PGI SeCyT-UNS (24/ZN38) *"Tecnologías Inmersivas y Visualización Situada aplicadas a Geociencias"*. Directora: Dra. María Luján Ganuza.

• PIBAA - CONICET (2872021010 0824CO) "Análisis Visual de Datos Multidimensionales sin Pérdida de Información". Directora: Dra. M. Luján Ganuza.

• PICT (2017-1246) *"Análisis Visual de Datos en Geociencias"*. Directora: Silvia M. Castro.

Proyectos finalizados

• PDTS, Proyecto de Desarrollo Tecnológico y Social: *"Desarrollo de un dashboard web integrado para el análisis visual de datos geoquímicos para la identificación de recursos naturales no renovables"*. Directora: Dra. Silvia Castro.

Beca Posdoctoral finalizada

M. Luján Ganuza. Título del plan propuesto: *"Análisis Visual de Datos Multidimensionales en Espacios Ad-Hoc"*. Beca Postdoctoral 2018 CONICET. Comisión Nacional de Investigaciones Científicas y Técnicas de la República Argentina (CONICET).

Beca Doctoral vigente

Antonella S. Antonini. Título del plan propuesto: *"Análisis Visual de Datos Multidimensionales"*. Beca doctoral 2019 CONICET. Comisión Nacional de Investigaciones Científicas y Técnicas (CONICET) de la República Argentina. Adjudicada a partir de abril de 2019 y por un término de 5 años.

5. BIBLIOGRAFÍA

[1] 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. Mineral. Petrol. 109, 2 (2015), 153-160.

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

[3] 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 TVCG. 20, 12 (2014), 1913–1922.

[4] Ganuza, M. L., Ferracutti, G., Gargiulo, M. F., Castro, S. M., Bjerg, E. A., Gröller, E., Matkovic, K. Interactive visual categorization of spinel-group minerals. Proc. of the Spring Conf. on Comput. Graph. (2017).

[5] Ganuza, M. L., Gargiulo, M. F., Ferracutti, G., Castro, S. M., Bjerg, E. A., Gröller, E., 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), 197–198.

[6] Antonini, A. S., Ganuza, M. L., Ferracutti, G., Gargiulo, M. F., Matković, K., Gröller, E., Bjerg, E. A & Castro, S. M. (2021). Spinel web: an interactive web application for visualizing the chemical composition of spinel group minerals. Earth Science Informatics, 1-8.

[7] Antonini, A. S., Ganuza, M. L., & Castro, S. M. (2022). VISUEL - A Web Dynamic Dashboard for Data Visualization. Journal of Computer Science and Technology, 22(1), e03.

[8] Antonini, A.S., Luque, L., Ganuza, M.L. et al. Toward a taxonomy for 2D non-paired General Line Coordinates: a comprehensive survey. Int J Data Sci Anal (2022).

[9] Luque, L.E., Ganuza, M.L., Antonini, A.S., Castro, S.M. (2021). npGLC-Vis Library for Multidimensional Data Visualization. In: Naiouf, M., Rucci, E., Chichizola, F., De Giusti, L. (eds) Cloud Computing, Big Data & Emerging Topics. JCC-BD&ET 2021. Communications in Computer and Information Science, vol 1444. Springer, Cham.

[10] Barnes, S. J., and Roeder, P. L. The range of spinel compositions in terrestrial mafic and ultramafic rocks. J. Petrol. 42, 12 (2001), 2279–2302.

[11] Lindsley, D.H. Oxide minerals: petrologic and magnetic significance. Review in Mineralogy 25 (1991).

[12] Roeder, P. L. Chromite; from the fiery rain of chondrules to the kilauea iki lava lake. The Canadian Mineralogist 32, 4 (1994), 729–746.

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

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

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

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