

# A Data-Driven Approach for Mapping Multivariate Data to Color

Shengui Cheng, Wei Xu, Wen Zhong, and Klaus Mueller

Visual Analytics and Imaging (VAI) Lab, Computer Science Department, Stony Brook University

Computational Science Initiative, Brookhaven National Laboratory

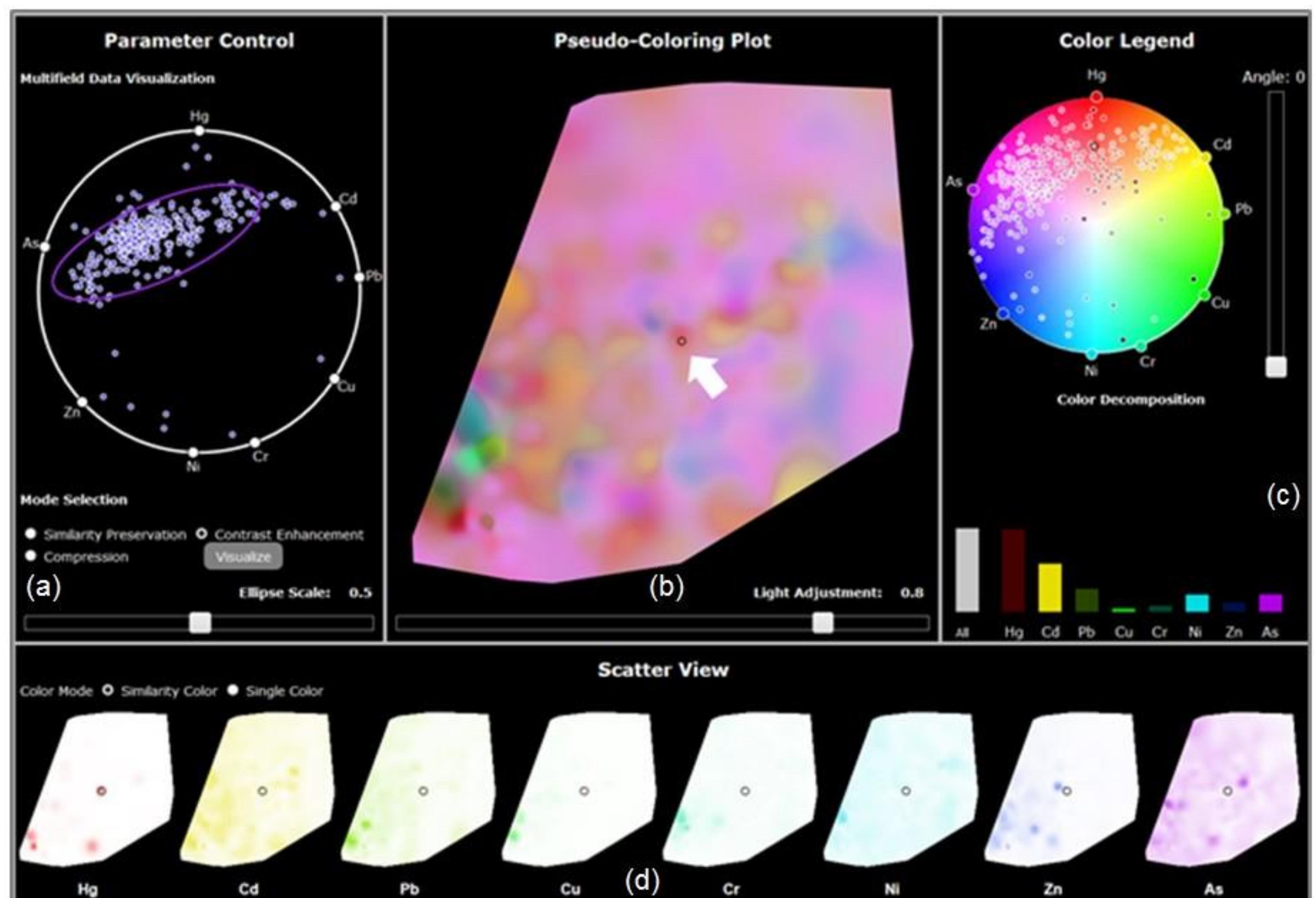
## Introduction

A wide variety of color schemes have been devised for mapping scalar data to color. Some use the data values to index a linear or bilinear color scale. Others assign colors to different materials which blend together in locations where these materials overlap. However, none of these methods scale to higher dimensional (multivariate) data. Likewise, schemes that take more artistic approaches through color mixing and style variations also face limits in the number of variables they can encode. We have developed a framework and system that fills this void and allows users to produce a single pseudo-colored plot from arbitrary multivariate data. It is a data driven method, which first gauges the (dis)similarity of the data attributes and then arranges them accordingly along the periphery of a convex 2D color space – we use the HSV color space. The data samples are then plotted into this space, and their color is obtained via generalized barycentric coordinate (GBC) interpolation.

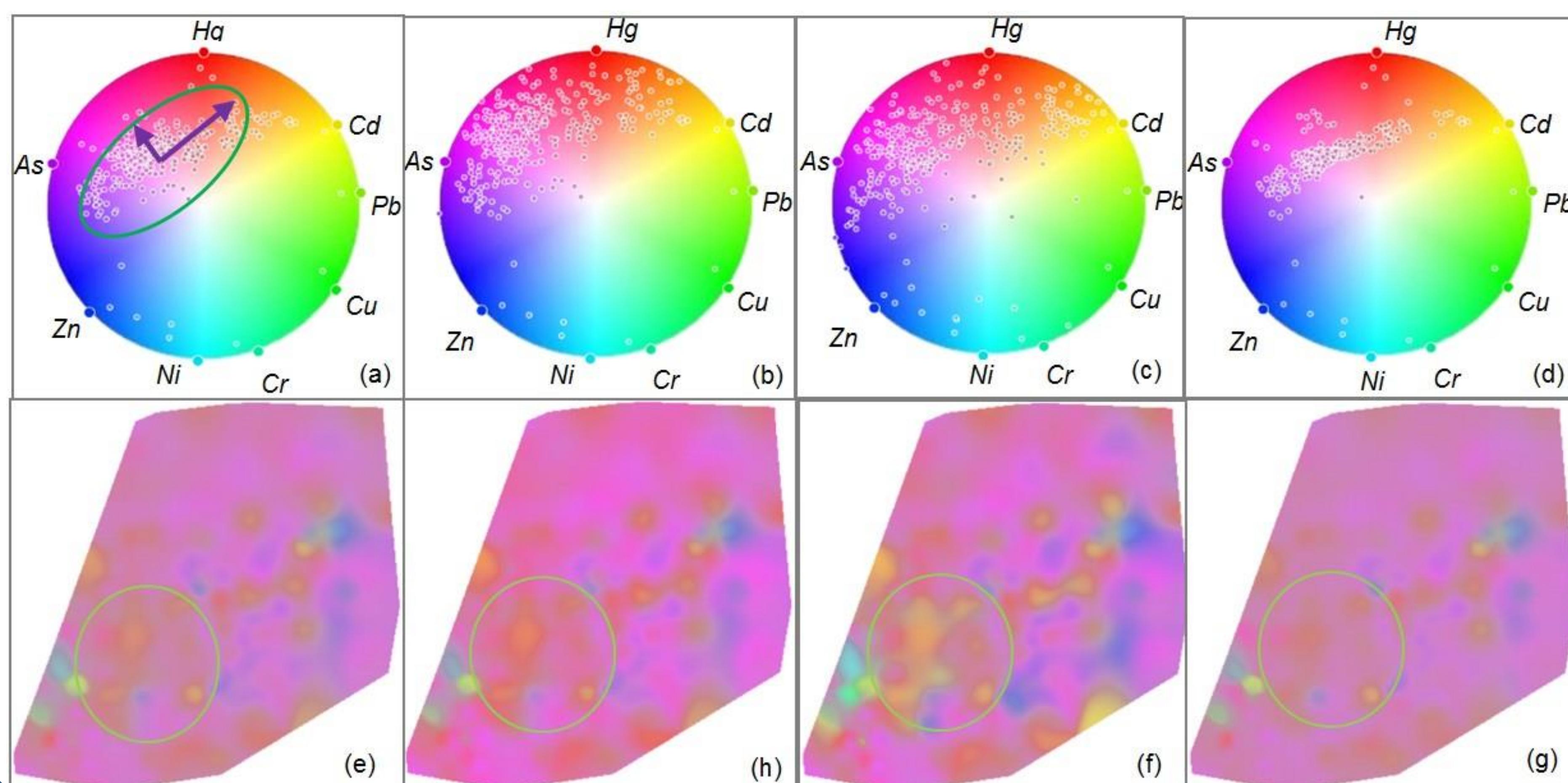
## Interface

Our interface consists of five parts. The parameter control panel (a) plots the data into a circular multivariate information display. The nodes on the periphery are the attributes arranged according to their similarity given the data. Data clusters exhibit similar attribute vectors. An elliptical region can be configured to stretch the color map and so enhance contrast for the data inside it. The domain plot (b) displays the pseudo-colored data. The color legend (c) is similar to control panel (a) but now shows the data points in the context of the HSV color map. The bar chart below shows the attribute values for a specific data point pointed to by mouse-over in the other plots. The scatter view (d) displays a set of heat maps, one per attribute and colored according to the primary color the attribute has in legend (b).

The white arrow in the pseudo-colored domain plot in the center points to a small red-colored region (outlined with a small black circle). This region is also shown in all other information displays, marked with a black circle as well – the color legend on the right and the individual domain plots for each attribute on the bottom (labeled scatter view)



## Contrast Enhancement



To improve the contrast among data features we allow users to manipulate the color map via various contrast enhancement schemes. One of these schemes warps dense data areas (as those shown inside the ellipse) to the full HS (Hue Saturation) space which increases the color contrast of these data, but compresses the contrast for the outliers outside the ellipse (which already have sufficient contrast). Conversely, the compression coloring scheme restricts the dense points to an even smaller area and emphasizes the outliers.

Panels (a)-(d) show the legend and domain map, respectively, for the default coloring. Panels (e)-(f) and (g)-(h) show our color/contrast preserving enhancement coloring, respectively. Finally panels (d)-(h) show the compression coloring.

Contact: {shecheng, wxu, wezzhong, mueller}@cs.stonybrook.edu

This research was partially supported by NSF grant IIS 1527200, the MSIP Korea, under the "IT Consilience Creative Program (ITCCP)" supervised by NIPA, and BNL LRD grant 16-041.

