

# CFD Data Differencing: Extraction and Representation for Uncertainty Visualization

**Zhanping Liu**  
Old Dominion University  
Norfolk, VA  
z1Liu@odu.edu

**David Kao**  
NASA Ames Research Center  
Moffett Field, CA  
david.l.kao@nasa.gov

## ABSTRACT

As data grows at exponential rates, the last two decades have seen widespread applications of scientific visualization to intuitive analysis of unknown complex patterns that may not always be properly recognized or effectively conveyed by non-visual computing methodologies such as data mining and machine learning (artificial intelligence). While scientific visualization provides deep insight into the data originating from multi-physics modeling and high-fidelity simulation of Computational Fluid Dynamics (CFD), one challenge that has garnered tremendous attention is uncertainty visualization. This paper presents a suite of techniques for *extracting, representing, and visualizing uncertainty of CFD data* (possibly tainted with noise) by addressing a *field of difference* that is derived and processed for any two cases of an ensemble of scalar data acquired on a 2D grid (or any slice cut through a volume of data), without assuming any statistical or simulation model, without any input from domain experts, and without any user interaction. By integrating multiple color maps that are visually discontinuous and significantly contrasted in hue, Hybrid color Map (HyMap) offers both qualitative visualization and relatively quantitative visualization to attain a focus+context representation. Incorporated with HyMap are Relief Image Synthesis (RIS) and Differencing by Line Integral Convolution (Diff-LIC). RIS takes a field of difference as a height surface and local gradients as normal vectors, employing an illumination model to produce an image of “terrain”. Diff-LIC extends LIC, a texture-based approach intended for flow data, to uncertainty visualization by creating an artificial unusual 2D vector field. These proposed techniques are part of our primitive representations on which composite algorithms may be further developed to visualize the uncertainty exhibited by an ensemble of results that a numerical simulation runs for a large number of times to yield. HyMap, RIS, and Diff-LIC are among the constituents of our CFD data Uncertainty Visualization (CUVis) system. Preliminary tests demonstrate that they are well suited for uncertainty visualization.

## ABOUT THE AUTHORS

**Dr. Zhanping Liu** ([www.zhanpingliu.org](http://www.zhanpingliu.org)) is an assistant professor with the department of Computational Modeling and Simulation Engineering at Old Dominion University. He was a 2014-summer USRA-sponsored faculty visitor at NASA Ames Research Center and a 2013-summer FRPP-sponsored faculty visitor at Argonne National Lab during his employment as an assistant professor with the department of Computer Science at Kentucky State University. Much earlier, Dr. Liu was a research staff member at the University of Pennsylvania, Kitware, Inc., and Mississippi State University after a post-doctoral position at the University of Iowa. He received the PhD degree in Computer Science from Peking University (2000) and the BS degree in Mathematics from Nankai University (1992). Dr. Liu’s research interests remain in scientific visualization, particularly vector (flow) visualization and parallel visualization.

**Dr. David Kao** is a researcher in the Advanced Computing Branch of the NASA Advanced Supercomputing (NAS) Division at NASA Ames Research Center. He has developed numerous collaborations (both internal and external to NASA) and created applications for scientific visualization in several disciplines. He has led several innovative software projects at NASA. Dr. Kao’s research focuses on scientific visualization, numerical flow visualization, and computer graphics and is a subtopic manager for NASA’s Small Business Innovation Research. He has won several NASA Group Achievement Awards, NASA Ames Honor Awards, and NASA Space Act Award for the software code Unsteady Flow Analysis Toolkit. He has served as an associate editor for the IEEE Transactions on Visualization and Computer Graphics and is a co-chair of the IS&T Visualization and Data Analysis Conference. Dr. Kao received the PhD degree in Computer Science from the Arizona State University.

# CFD Data Differencing: Extraction and Representation for Uncertainty Visualization

Zhanping Liu  
Old Dominion University  
Norfolk, VA  
z1Liu@odu.edu

David Kao  
NASA Ames Research Center  
Moffett Field, CA  
david.l.kao@nasa.gov

## INTRODUCTION

*Scientific Visualization* (SciVis) is an important application of computer graphics to intuitive, interactive, and insightful analysis of large-scale data arising from nearly all scientific and engineering disciplines. SciVis exploits data processing, pattern recognition, geometric mapping, image synthesis, graphical rendering, high-performance (parallel) computing, and domain knowledge to enable visual perception and explorative inspection of a vast amount of data, with one or multiple physical variables defined on a 2D/3D mesh (grid) of discrete sample points, in a way to help with cognitive reconstruction, in-depth understanding, and effective interpretation of significant structures as well as intricate features behind the scene towards scientific discovery. It has been a constituent stage in the *modeling-simulation-visualization-analysis* pipeline, to which researchers and engineers turn to tackle daunting problems. While visual data analysis plays a crucial role in Computational Fluid Dynamics (CFD), among many other areas (e.g., oceanography, atmosphere, and electromagnetics, to name only a few), one line of research that has attracted increasing attention in the past decade is *uncertainty visualization* (Brodie et al., 2012; Bonneau et al., 2014; Hullman et al., 2019), involving *uncertainty of data* (originating from data generation, acquisition, collection, and maintenance) and *uncertainty of visualization* (stemming from data processing, information derivation, feature extraction, content mapping, and graphical rendering).

The ubiquitous occurrence of uncertainty hinders us from gaining an accurate comprehension of the data, more precisely the model(s) underneath and ultimately the phenomenon being investigated. In light of *data uncertainty*, CFD scientists usually choose to run a numerical simulation for multiple and even many times to produce *ensemble datasets* (or *cases*) in an attempt to improve their understanding of the model by analyzing the cases in an integrated manner. Of particular and emerging interest to CFD scientists is how to effectively and efficiently identify and convey the difference between an arbitrary pair of cases, in the form of two sets of scalar data, defined on a 2D Cartesian grid (e.g., each with a rectangular array of values on a uniformly-sampled cutting plane embedded in a volume of scalar data), *without any input of domain knowledge and without any user-view interaction*. By building on such a fundamental functionality, many ensemble cases may then be compared against one another, and furthermore, incorporated together and assimilated as a whole to facilitate us to reach an optimal approximation to the ground truth by uncertainty visualization.

Thus, our recent research is primarily focused on designing novel, atomic, general-purpose, and robust algorithms for extracting, processing, representing, and visualizing the difference between an arbitrary pair of cases of CFD scalar data mounted on a 2D Cartesian grid, *without assuming any statistical or simulation model, without collecting any input from domain experts, and without asking the user to operate on either the data or the view*. Possibly speckled with various sources of noise, the original data undergoes a series of processing steps through which a Field of Difference (FoD) is derived to encode both salient structures and fine features as the discrepancy between two cases of ensemble data. We present three methods for visualizing the FoD. A hybrid color map makes use of multiple visually disconnected color maps, with sufficient hue contrast in between, to reveal some levels of quantitative detail in addition to qualitative information. In combination with this special color map is relief image synthesis that we propose to extrude the plane of FoD to emulate a terrain-like 2.5-dimensional surface by adopting local gradients as normal vectors in a graphics illumination model. Also coupled with the hybrid color map is line integral convolution, a texture-based technique originally devised for flow visualization, which we adapt to show the pattern of signed differences by constructing a pseudo bi-variate orthogonally-oriented textile-like vector field from the FoD. Implemented as low-level “building blocks”, the aforementioned methods are integrated in a prototype system that we have developed and tested with sample CFD data. The present work, of tangible value to CFD

scientists, offers a necessary platform on which to design advanced composite algorithms for uncertainty visualization.

The remainder of this paper is organized as follows. Section 2 provides a concise survey of the literature on uncertainty visualization. Section 3 begins with an elucidation of the challenges that we need to take into account in order to address a real-world issue raised by CFD scientists. Then we present noise-suppressing feature-preserving generation of a scalar field of difference from any two cases of ensemble data, accompanied by three techniques that we propose for visualizing this output, i.e., hybrid color map, relief image synthesis, and an adapted version of line integral convolution. Results and discussions are given in section 4 to demonstrate the effectiveness of these methods. Section 5 concludes this paper with a brief summary and outlook on future work.

## RELATED WORK

Uncertainty visualization emerged in the late 1990s (Pang et al., 1997) and about one decade later it came to be one of the mainstream topics. There have been so many papers that we may not enumerate in breadth and depth, while three comprehensive literature reviews (Brodlie et al., 2012; Bonneau et al., 2014; Hullman et al., 2019) cover a plethora of algorithms and diverse applications. In reality, these algorithms do not pinpoint the difference between any two specific cases of data. Instead, they are dedicated to capturing and showing the overall distribution or behavior of many cases from a statistical perspective. They are good at drawing a high-level picture by digesting, fusing, and summarizing the entire ensemble data to reach a consensus. On the other hand, the associated “blending” operations tend to deemphasize, conceal, or even smear away the discrepancy that may exhibit between two cases of special interest to CFD scientists. In addition, these algorithms may not meet the stringent restrictions (as listed at the very beginning of paragraph 3, Section 1) that some CFD scientists impose. In some sense, this is a sign of gaps between visualization research and practical needs. In response to the restrictions and needs, the research that we present in this paper differs from nearly all previous work regarding the motivation, focus, methodology, and goal. Meanwhile, our work is still related but *loosely* to available research efforts, under the umbrella of uncertainty visualization. By sampling the literature, the review below carves out the context that our work is expected to touch in the next step.

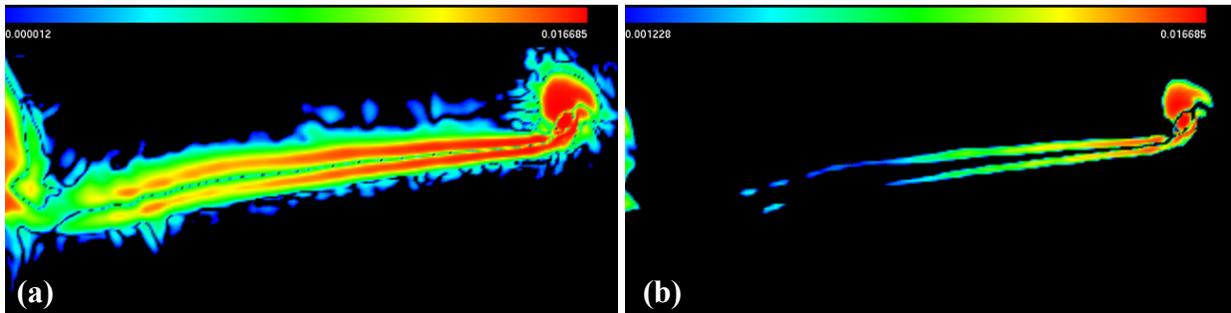
Whitaker et al. proposed to generate iso-contours for measuring the uncertainty of either a model or a simulation and then choose contour boxplots to reveal qualitative and quantitative information of an ensemble of iso-contour data (Whitaker et al., 2013). Guo et al. developed a prototype system, called ensemble Flow Line Advection and Analysis (eFLAA), for analyzing unsteady flow features across ensemble runs. A Lagrangian-based distance metric is employed to create a field of variation that encodes how ensemble runs deviate in particle traces (Guo et al., 2013). To deal with the variation of gradients in 2D scalar fields, Pfaffelmoser et al. adopted a color diffusion model for delineating the absolute and relative derivatives, a glyph-based approach for depicting the gradient orientation, and a contour tree for explorative visualization of uncertainty. Also investigated in this work are uncertainty parameters including gradients, confidence interval of the gradient magnitude, and confidence interval of the gradient orientation (Pfaffelmoser et al., 2013). Wu and Zhang developed a balanced planar hierarchical contour tree of uncertainty by which view occlusion and visual clutter are avoided during level-of-detail exploration (Wu & Zhang, 2013). Chen et al. applied projection (also known as dimensionality reduction) to multi-dimensional ensemble data for characterization, visualization, exploration, and analysis of the correlations between variables in the mean value and the overall distribution (Chen et al., 2015). To alleviate visual clutter and view occlusion, Liu et al. proposed to sample the ensemble data so as to select such a representative subset that is still able to preserve the spatial statistics of the original data very well (Liu et al., 2017). Sanyal et al. developed a tool for visualizing the uncertainty of ensemble data exported by a numerical weather model (Sanyal et al., 2010). In addition, they conducted a user study to evaluate four uncertainty visualization methods in comparison for 1D and 2D datasets (Sanyal et al., 2009).

## CFD DATA UNCERTAINTY VISUALIZATION

In this section, we present our work on *CFD data Uncertainty Visualization (CUVis)*, involving extraction of a *Field of Difference (FoD)* from two cases of scalar data followed by visualization of the FoD with three methods.

## Extraction of Difference Fields

Given an arbitrary pair of cases of an ensemble data with respect to an unknown scalar variable (e.g., temperature, pressure, velocity magnitude, or vorticity magnitude) generated from an unknown CFD simulation, the two associated values at a grid point (one value for each case) tend not to be exactly the same and their (absolute) difference, as a derived value of our interest, is sensitive to the precision of a floating-point representation and especially to various unknown sources of noise. As a consequence, the values of difference (i.e., the values of the FoD) derived from the two source cases may exhibit a highly non-uniform distribution across a wide range — spanning multiple orders of magnitude (e.g.,  $10^{-6} \sim 10^{-1}$  and  $10^{-2} \sim 10^4$ ). Without any input from domain experts (e.g., any parameter of the CFD simulation), it is extremely difficult to select an appropriate *flooring threshold* (simply called *threshold* hereinafter), below which the value at a grid point of the FoD is replaced with zero. For example, a conservative threshold may cause negligible speckles or areas (of tiny difference) to tamper with the visualization (Figure 1.a), whereas a bold threshold may suppress or exterminate features (i.e., regions of considerable difference) in the view (Figure 1.b). It is worth mentioning that no any statistical model (e.g., normal distribution and exponential distribution, etc.) may be assumed to analyze either the source data or the FoD data for determining an appropriate threshold. The restriction on any user-view interaction (e.g., dragging a slider to adjust the threshold) further complicates this situation.



**Figure 1.** Two images generated via a traditional rainbow color map scheme by choosing (a) a conservative threshold, 0.000012, and (b) a bold threshold, 0.001228, respectively, to visualize a  $330 \times 146$  FoD of vorticity magnitude data, with deep blue for the minimum difference and full red for the maximum.

The use of a traditional color map offers a continuous depiction of the FoD. Such a visually appealing yet solely qualitative representation prevents the user from extracting in-depth information, e.g., the degree to which any two points or areas (of the FoD) can be distinguished or how much the two differ from each other. A relatively quantitative delineation through a special color map may help highlight significant features while keeping less important ones as the context. This is particularly necessary when the values of an FoD scatter in a wide range with multiple orders of magnitude.

Thus, two of the major tasks of CUVis consist in *feature-preserving data processing* for FoD generation and *relatively quantitative color mapping* for FoD visualization. Since a difference value indicates the deviation between two cases at the same point, the ultimate goal of CUVis is to remove noise and negligible speckles or areas with tiny differences, suppress unimportant regions with small differences, and accentuate features in the form of spots or areas with considerable differences.

CUVis begins by processing and filtering two cases of data on a per-point basis to discard unimportant points, i.e., assigning a value of zero to such points in both of the two cases. Specifically, any point where the two cases are both below a tiny threshold ( $EPSILON\_VALUE = 2 \times 10^{-6}$ ) is removed. This operation is meant to suppress the influence of noise. Also deleted is any point where the two cases are both below a tiny threshold ( $REGION\_OF\_FEATURE\_MIN\_VALUE = 1 \times 10^{-5}$ ) and at the same time the two local gradient magnitudes are both below a tiny threshold ( $POINT\_OF\_FEATURE\_MIN\_GRADIENT = 4 \times 10^{-6}$ ). This disposal is intended to ignore “flat” points with negligible source values. A flat point is considered irrelevant to such features as reflected by noticeable local changes. In addition, any point with the average (of the two cases), the absolute difference (of the two cases), and the ratio of the absolute difference to the average below three thresholds ( $MIN\_AVERAGE = 6 \times$

$10^{-6}$ ,  $MIN\_DIFFERENCE = 1 \times 10^{-5}$ ,  $MIN\_DIFF\_TO\_AVE\_RATIO = 0.01$ ), respectively and simultaneously, is discarded. The combination of these three conditions is aimed at skipping points with inadequate differences.

For those points surviving the aforementioned filtering process, CUVis further classifies them based on the statistical distribution of the difference values and then divides them into several bands, of which each covers an order of magnitude. To retain up to 4 most significant bands, the least band(s) may be discarded if necessary. This selection of 4 most significant bands ( $10^n \sim 10^{n+4}$ ,  $-6 \leq n \leq 8$ : with  $10^2$  and  $10^4$  as the minimum possible ratio and the maximum possible ratio, respectively, for any two values within the four consecutive bands) is based on the observation that an entire range spanning no more than 4 orders of magnitude allows important features to be highlighted while retaining sufficient background or preventing loss of necessary information so as to provide a focus+context representation of an FoD.

One characteristic of this classification process is the adaptive generation of a very narrow “artificial gap” between any two consecutive bands, though all of the original difference values remain unchanged, e.g., in a possibly continuous distribution across the boundary of the two bands. The gap refers to a very small window that exactly touches the boundary of the higher band but still entirely resides within the lower band. Any value falling within the gap is logically “upgraded” from the lower band to the higher band. In practice, the size of the gap is determined on the fly by exploiting the statistical distribution of the values in the lower band. The purpose of such an artificial gap is to minimize the “sudden leap” effect that is incurred by discriminating two values in a very tight neighborhood, of which one is assigned to the lower band and the other is though assigned to the higher band. Suppose  $[10^n - \lambda, 10^n]$  is an artificial gap placed between two consecutive bands  $[10^{n-1}, 10^n]$  and  $[10^n, 10^{n+1}]$ . As the size of the gap,  $\lambda$  (possibly zero) should be identified such that a considerable number of values may fall within the very narrow gap  $[10^n - \lambda, 10^n]$ , while *only a small number of values or even none* exists within  $[10^n - \lambda - \gamma, 10^n - \lambda]$ , where  $\gamma$  is expected to be as large as possible. In other words, the main task is to find an optimal value of  $\lambda$  such that  $10^n - \lambda$  is able to divide the FoD values of the lower band into two groups in a “clear-cut” manner near the boundary of two bands. In this way, sudden leaps can be tremendously alleviated and even entirely concealed.

### Hybrid Color Map

We propose *Hybrid color Map* (HyMap) to map difference values to RGB (red, green, blue) colors based on the result of point filtering and classification. The data bands and the narrow artificial gaps between them allow HyMap to integrate several color maps that are visually disconnected and significantly contrasted in hue (Figure 2). While each color map offers a continuous representation of the associated band of difference values with a smooth transition of hue at a low level for qualitative visualization, the combination of multiple color maps creates a “discretized” delineation of the entire range with outstanding contours of hue at a high level for quantitative visualization. Thus, significant features within a higher band may be easily recognized by means of a color map considerably different from the maps used by other unimportant (lower) bands. Whenever necessary, data values of each band may be investigated and compared in detail by identifying those areas to which the associated color map is applied in the image. HyMap achieves both qualitative visualization and relatively quantitative visualization in a single view. It enables a focus+context analysis of an unknown FoD of which the range may be very wide.

One caveat of HyMap is that a value mapped to the right end of a color map (e.g., pure cyan in Figure 2) and another value mapped to the left end of the *next* color map (e.g., full red in Figure 2) might not be so “far away” from each other as shown by the distinction between the two respective hues. This illusion comes with the capability of HyMap in quantitative visualization of data. Fortunately, this side effect is reduced to the minimum in CUVis because the size of the artificial gap between the two associated bands is set adaptively by addressing the statistical distribution.

As a novel color map scheme, HyMap itself serves as a visualization technique for generating an image to show a field of difference between two cases of scalar data. Combined with HyMap are two other visualization methods, namely relief image synthesis and differencing by line integral convolution, which are described below.

### Relief Image Synthesis

By treating *absolute* differences (i.e., positive values) as height values and local gradients (with respect to the height) as on-surface normal vectors, *Relief Image Synthesis* (RIS) adopts a simplified illumination model to

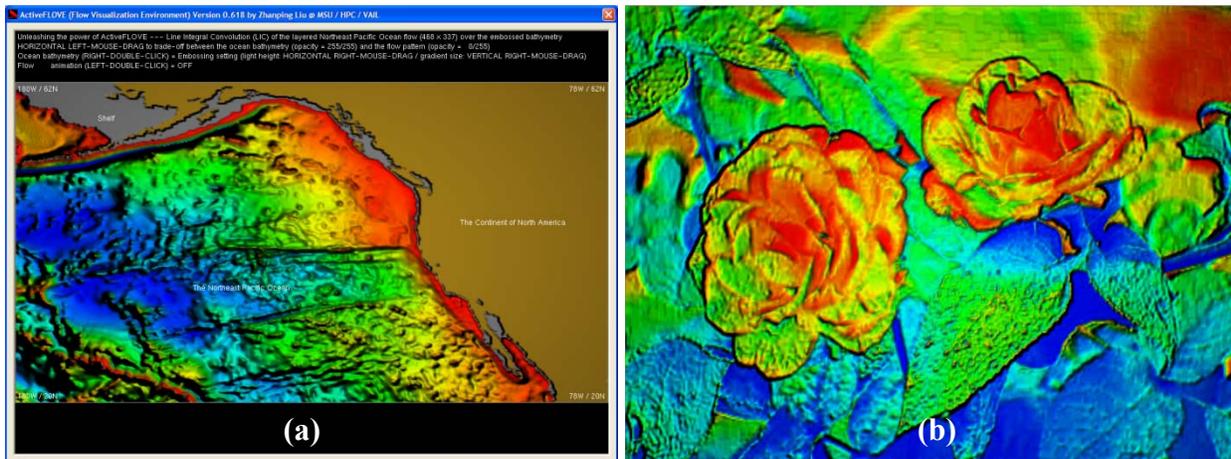
produce a terrain-like image (*without explicit construction of surface patches such as triangles or quadrilaterals*). The Phong model, well known in computer graphics, is adapted in CUVis to evaluate the local illumination for calculating the grey-scale *intensity* (brightness) value (0 ~ 255) for each data point of an FoD, involving ambient, diffuse, and specular elements, of which the first may just take a constant. Theoretically, the intensity at a point in 3D space is a function of the relationships between the viewer (eye), the light source (with only one light source in CUVis), and the 3D normal vector of the local tiny surface patch. If the view and light directions are both fixed, as is the case with our change to the Phong illumination model, the only remaining factor is the normal vector of the curved surface at the 3D point.



**Figure 2.** An example of HyMap comprising four disjoint color maps which are vastly contrasted in hue, with each for one band of data. As contours are well discerned in a HyMap image, scalar values can be quantitatively differentiated across orders of magnitude and qualitatively compared within each magnitude.

The terrain shape erected over the 2D rectangular domain of an FoD can be described by a discretized version of surface parameterization:  $x = i, y = j, z = H(i, j)$ , where  $i$  and  $j$  are indices of the 2D Cartesian grid of the FoD. Then, the  $x$ - and  $y$ -components of the 3D normal vector at the on-surface vertex right over *each* grid point can be approximated by the partial derivatives of  $z$  with respect to  $i$  and  $j$ , respectively. In CUVis, the  $z$ -component is controlled separately, which offers great flexibility of either increasing or decreasing the contrast in height. In fact, an experimentally optimal value is employed in CUVis for the  $z$ -component. Once the intensity value is determined for a data point of the FoD, the initial RGB color assigned by the HyMap is then modulated to yield a new RGB color. With shading effects “added” by a rectangular array of such new colors, the FoD plane is then visually “extruded” to emulate a terrain-like surface.

As a *pixel-based image synthesis* method (without resorting to any graphical element or OpenGL rendering), RIS applies to an arbitrary 2D field of scalar data defined on a Cartesian grid. Likewise, RIS also works with any image, of which the brightness of each pixel, obtained via a weighted sum of the RGB channels (e.g.,  $0.30 \times R + 0.59 \times G + 0.11 \times B$ ), can be reinterpreted as a pseudo height. Figure 3.a shows an RIS image generated from the bathymetry of the Northeast Pacific ocean (note that depth is negated height). Figure 3.b shows an RIS image created from an image of flowers.



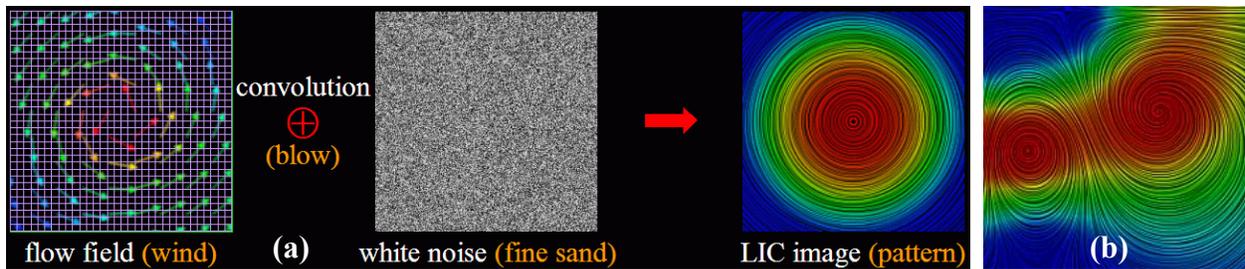
**Figure 3.** RIS for synthesizing a 3D surface from either (a) a 2D scalar data field such as the bathymetry of the Northeast Pacific ocean or (b) a 2D image of flowers.

### Differencing by Line Integral Convolution

Line Integral Convolution (LIC, Cabral & Leedom, 1993; Liu et al., 2012; Liu & Moorhead, 2016) is a texture-based technique for dense visualization of a 2D vector data (flow) field, of which each point carries two

intrinsically-coupled component values to encode the local direction. LIC applies a low-pass filter to white noise to convolve an array of texture elements along a 1.5-dimensional curve, i.e., along a streamline (the trajectory of a particle released into and driven by the flow) spawned from each pixel being synthesized. The basic idea is to expose the close correlation between the samples of each streamline, in the form of the associated consecutive pixels possessing similar intensity values, through a smearing operation on a dense aggregation of massless particles. This image synthesis procedure emulates what happens when a rectangular area of fine sands is blown by a gust of strong wind (Figure 4.a). In more detail, from the center of each pixel (the *target*) of the output LIC image, a seed is released to integrate a streamline in both negative and positive directions, but with the same length wherever possible. Then, the correlated pixels (as the contributors) are located for the target pixel (as the receiver). Next, white noise is indexed to get the texture values for these contributing pixels. Finally, convolution is performed on these values to calculate a normalized weighted sum as the value of the target pixel. This pipeline is repeated for each and every pixel to produce the entire LIC image. The strengths of LIC consist in a high-resolution, dense, and realistic representation of the flow (Figure 4.b).

*Differencing by LIC* (Diff-LIC) is a technique that we design by adapting LIC to show the signed difference between two cases (e.g.,  $A$  and  $B$ ) at each data point of an FoD despite it being a *scalar* field. The *signed* difference at a data point denotes whether the value in  $A$  is Greater Than (“>”, the GT case for a positive difference), Equal To (“=”, the ET case for a zero difference), or Less Than (“<”, the LT case for a negative difference) than in  $B$ . The vectors of a typical 2D flow field point to various directions, whereas in CUVis we create a unique 2D vector field based on the signed difference values (*scalar* values) of an FoD. Specifically, a signed difference value is mapped to vector  $(|diffVal|, 0.0)$  for the GT case, to vector  $(0.0, 0.0)$  for the ET case, and to vector  $(0.0, |diffVal|)$  for the LT case, where  $|diffVal|$  refers to the absolute difference which is mapped to the velocity magnitude of the artificial vector.



**Figure 4.** (a) The basic idea of LIC, analogous to the wind blowing across a rectangular area of fine-grained sands. (b) An image produced by LIC for a 2D vector field with two vortices.

When visualized, such a pseudo 2D vector field exhibits a textile-like pattern, which is governed by the signed difference field and hence does not look like a normal flow. However, it reflects the fluctuation of the values in case  $A$  versus the counterpart in case  $B$ . In more detail, the GT and LT cases contribute to horizontal and vertical streaks, respectively, in the Diff-LIC image. The ET case (essentially for velocity-vanishing positions called critical points) does not lead to any pattern but instead simply leaves voids (in fact black dots) in the output image. Thus, what is visualized by Diff-LIC is an artificially-fabricated bi-variate orthogonally-oriented vector field.

## RESULTS AND DISCUSSIONS

We have developed a prototype system, the namesake of our CUVis work, to integrate the proposed algorithms, among a series of other auxiliary functionalities such as Plot3D data reader, image writers, classification map, histogram viewer, and in-view image controller. CUVis is a cross-platform uncertainty visualization toolkit, with low-level graphics rendering based on GLUT for Unix and Microsoft Windows but on FLTK for Linux Red Hat. The core of CUVis is made up of 3,500+ lines of new C/C++ code. In fact, CUVis takes advantage of some C++ classes of ActiveFLOVE (Liu & Moorhead, 2016), a data visualization package developed in previous work. These legacy C++ classes include LIC2D (for texture-based visualization of 2D vector data), DataMap (for nonlinear conversion of scalar data), ColorMap (for a wide variety of color maps), ColorWheel (for a color-wheel map), ImageEmboss (for pixel-based synthesis of height surface from 2D scalar data), BMPWriter (for image output on Microsoft Windows), and PPMWriter (for image output on Unix/Linux).

Since CUVis does not need, and should not be provided with, any user interaction except for a command-line specification of two data files (currently in Plot3D format, with each taken for one case) to be loaded, a concise caption is shown above the visualization space to offer some necessary information about data processing and feature extraction. Also displayed in the view is a HyMap color bar legend between the caption and the visualization area, with the minimum and maximum values of each band, for both qualitative exploration and quantitative examination of a field of difference.

In order to test CUVis on the proposed functionalities and the practical effectiveness in extracting and visualizing 2D difference fields, we selected four pairs of cases (labeled as A-B, C-D, E-F, and G-H in this section to simplify the description) of data, with velocity magnitude and vorticity magnitude as two scalar variables uniformly sampled on a  $330 \times 146$  cutting plane in a volume. These datasets were exported from a CFD simulation based on the European High Lift Programme (EUROLIFT) DLR (German Aerospace Center) F11 high-lift configuration (with brackets on and off, Rudnik & Huber, 2013).

Figure 5.a-c are results of visualizing the difference field between cases A-B in vorticity magnitude (with brackets on) by HyMap, HyMap plus RIS, and HyMap plus Diff-LIC, respectively, with four color bands detected and utilized. As indicated by the small green-to-yellow patch, the largest differences happen near the center of the rightmost “head”, followed by those at red-to-magenta, blue-to-cyan, and grey-to-white areas in decreasing order. With the four color bands, we notice that the differences rapidly attenuate from two stripes, near the middle of the skeleton of the curve body, not only toward the peripheral parts but also toward the “ridge line” of the body. On the other hand, the differences are pretty similar, or vary only slightly, along the skeleton. The salient staircase or cascading pattern means that the differences span four orders of magnitude, i.e., a huge range of intensity, though across a very narrow distance in the direction orthogonal to the skeleton. Please note that this kind of quantitative information, available by HyMap, cannot be gained via traditional color maps. Figure 5.b outperforms Figure 5.a by emulating a height surface to emphasize the differences at one area in comparison with another, either along or perpendicular to the skeleton. Figure 5.c exposes more detailed information than obtained in Figure 5.a by showing the sign of the difference at each data point. Case A is less than case B in nearly all regions above the “ridge line”, as reflected by vertical streaks therein. However, case A is greater than case B in nearly all regions below the “ridge line”, as indicated by horizontal streaks. Streaks are ambiguous or uncertain along the “ridge line” regarding the orientation, implying the fluctuation of the data values there in case A versus those in case B.

Figure 5.d depicts the difference field between cases C-D in vorticity magnitude (with brackets on) by HyMap+RIS with three color bands. Figure 5.e delineates the difference field between cases E-F in velocity magnitude (with brackets off) by HyMap+RIS with two color bands. These two images demonstrate the ability of CUVis to handle difference fields with an adaptive number of color bands automatically determined from the distribution of differences in the domain. The differences in Figure 5.d (between cases C-D) change more rapidly, particularly in the rightmost “head” area, but take one order of magnitude fewer than those in Figures 5.a-c (between cases A-B). The differences in Figure 5.e (with respect to velocity magnitude) vary far more rapidly than those in Figures 5.a-d, as shown in the rightmost “head”. In particular, the largest differences scatter in several regions which though are not along the “ridge line”.

Figure 5.f is a snapshot of visualizing the difference field between cases G-H in vorticity magnitude (with brackets off) by HyMap+RIS with four color bands. This image demonstrates the capability of CUVis in accentuating significant features (along the skeleton of the curve body) as the focus and displaying intricate, relatively unimportant patterns (distributed in the rest of the field) as the context. This focus+context visualization allows CFD scientist to capture and inspect regions of interest (e.g., those with largest differences or those with the differences rapidly varying) while finding some elusive correlations that might exist between the focus and the context.

## CONCLUSIONS AND FUTURE WORK

We have presented our work on extraction, representation, and visualization of the field of difference between an arbitrary pair of cases of ensemble scalar data defined on a 2D Cartesian grid, for intuitive analysis of the uncertainty issue that occurs with multiple runs of a CFD simulation. The main contributions include noise-suppressing feature-preserving generation of difference fields, hybrid color map with quantitative cueing besides qualitative encoding in support of focus+context visualization, relief image synthesis for extruding a terrain-like

surface to accentuate the pattern of variance, and differencing by line integral convolution to reveal signed changes in more detail. These methods that we propose for CFD data uncertainty visualization are integrated in a prototype namesake system CUVis. Preliminary tests show that CUVis can help CFD scientists detect where two runs of a simulation deviate as well as where most radical discrepancies are located and then pinpoint what factors of the simulation cause the distribution of variance. This visual analysis paradigm speeds up the process of figuring out and fine-tuning critical parameters of the underlying physical model.

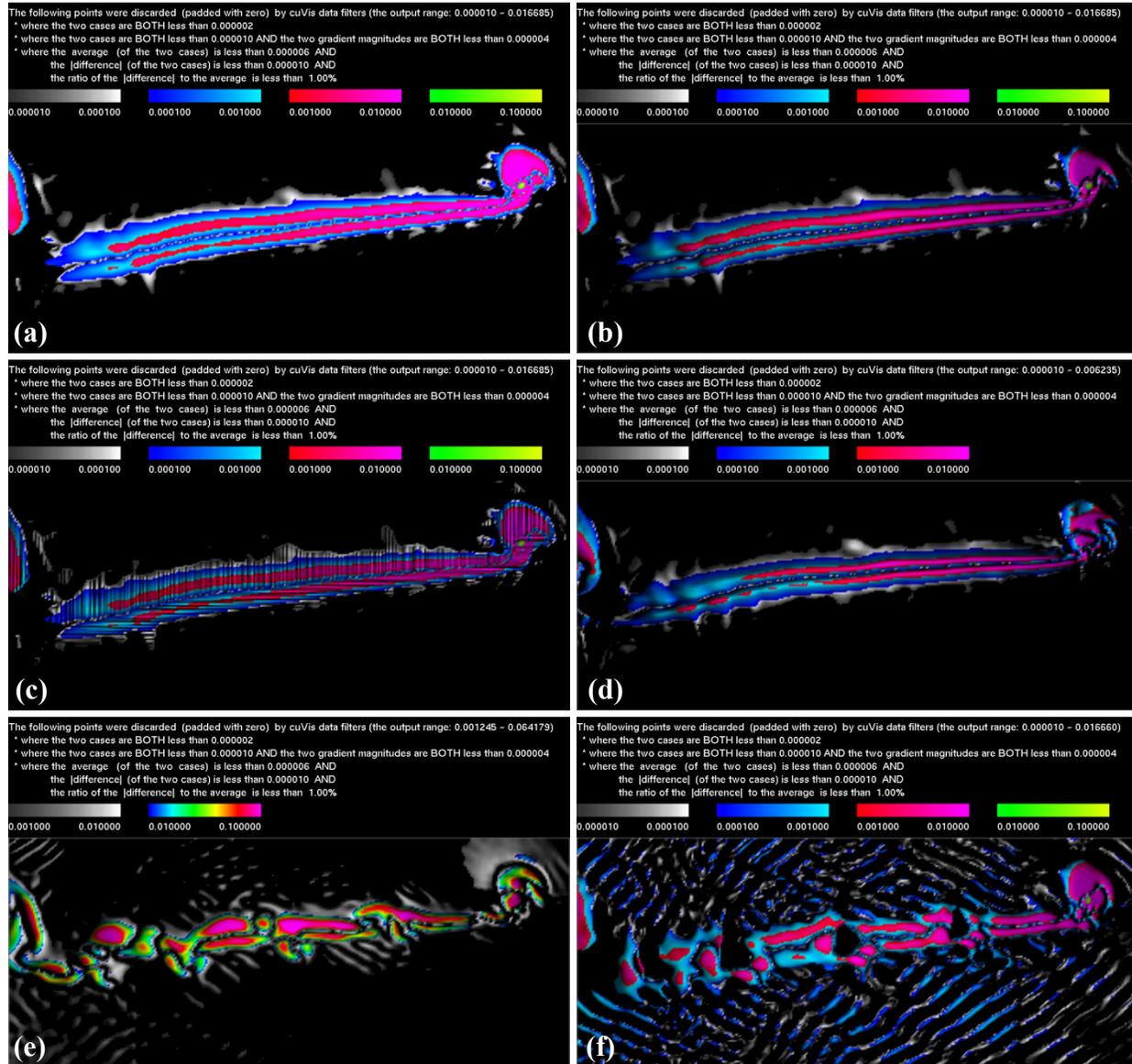


Figure 5. Use of CUVis for visualizing 4 difference fields (for cases A-B, C-D, E-F, and G-H, respectively) of the F11-CFD simulation data, with velocity magnitude and vorticity magnitude defined on a  $330 \times 146$  Cartesian grid. (a) Visualization of the field of difference between cases A-B in vorticity magnitude by HyMap with 4 color bands. (b) Visualization of the field of difference between cases A-B in vorticity magnitude by HyMap+RIS with 4 color bands. (c) Visualization of the field of difference between cases A-B in vorticity magnitude by HyMap plus Diff-LIC with 4 color bands. (d) Visualization of the field of difference between cases C-D in vorticity magnitude by HyMap+RIS with 3 color bands. (e) Visualization of the field of difference between cases E-F in velocity magnitude by HyMap+RIS with 2 color bands. (f) Visualization of the field of difference between cases G-H in vorticity magnitude by HyMap+RIS with 4 color bands.

It is worth emphasizing that all of the proposed methods are not limited to the scope of difference fields. Instead, they can directly apply to either 2D scalar data fields or 2D profiles of a scalar volume data (e.g., acquired from oceanographic-atmospheric simulation and medical CT / MRI / ultrasound scan) projected on arbitrary cutting planes. The present work paves the way for our further research and development in uncertainty visualization. Our immediate plan is to test CUVis with diverse CFD simulation data and then collect feedback from domain experts to decide how CUVis will be improved. Certainly, we will extend CUVis with the capability for processing, visualizing, and analyzing an ensemble of many (e.g., hundreds of) datasets in an integrated way.

## ACKNOWLEDGEMENTS

This work was supported by the NASA Advanced Supercomputing (NAS) Division at NASA Ames Research Center through the NASA-USRA (Universities Space Research Association) program. The first author would like to express sincere thanks to NASA and USRA for sponsoring the 10-week faculty visit in a summer. The authors thank Ralf Rudnik, Kerstin Huber, and Thomas Pulliam for the help with access to the F11 CFD simulation data. Thanks also go to anonymous reviewers for valuable comments and suggestions.

## REFERENCES

- Bonneau, G., Hege, H., Johnson, C., Oliveira, M., Potter, K., Rheingans, P., & Schultz, T. (2014) Overview and State of the Art of Uncertainty Visualization. *Scientific Visualization (Springer Publisher)*, 3–27.
- Brodlie, K., Osorio, R., & Lopes, A. (2012). A Review of Uncertainty in Data Visualization. *Expanding the Frontiers of Visual Analytics and Visualization (Springer Verlag Publisher)*, 81-109.
- Cabral, B., & Leedom, C. (1993). Imaging Vector Fields Using Line Integral Convolution. *Proceedings of ACM SigGraph '93*, 263-270.
- Chen, H., Zhang, S., Chen, W., Mei, H., Zhang, J., Mercer, A., Liang, R., & Qu, H. (2015). Uncertainty-Aware Multidimensional Ensemble Data Visualization and Exploration. *IEEE Transactions on Visualization and Computer Graphics*, 21(9), 1072-1086.
- Guo, H., Yuan, X., Huang, J., & Zhu, X. (2013). Coupled Ensemble Flow Line Advection and Analysis. *IEEE Transactions on Visualization and Computer Graphics*, 19(12), 2733-2742.
- Hullman, J., Qiao, X., Correll, M., Kale, A., & Kay, M. (2019). In Pursuit of Error: A Survey of Uncertainty Visualization Evaluation. *IEEE Transactions on Visualization and Computer Graphics*, 25(1), 903-913.
- Liu, L., Boone, A., Ruginski, I., Padilla, L., Hegarty, M., Creem-Regehr, S., Thompson, W., Yuksel, C., & House, D. (2017). Uncertainty Visualization by Representative Sampling from Prediction Ensembles. *IEEE Transactions on Visualization and Computer Graphics*, 23(9), 2165-2178.
- Liu, Z., Cai, S., Swan, E., Moorhead, R., Martin, J., & Jankun-Kelly, T. J. (2012). A 2D Flow Visualization User Study Using Explicit Flow Synthesis and Implicit Task Design. *IEEE Trans. Graphics & Vis.*, 18(5), 783-796.
- Liu, Z., & Moorhead, R. (2016). High-Performance Flow Visualization for Effective Data Analysis. *Journal of Flow Visualization and Image Processing*, 23(1-2), 41-57.
- Pang, A., Wittenbrink, C., & Lodha, S. (1997). Approaches to Uncertainty Visualization. *The Visual Computer*, 13(8), 370-390.
- Pfaffelmoser, T., Mihai, M., & Westermann, R. (2013). Visualizing the Variability of Gradients in Uncertain 2D Scalar Fields. *IEEE Transactions on Visualization and Computer Graphics*, 19(11), 1948-1961.
- Rudnik, R., & Huber, K. (2013). The HiLiftPW-2 Test Data Based on the European High Lift Programme (EUROLIFT) DLR F11 High-Lift Configuration. Available at the Web Site of the 2nd AIAA CFD High Lift Prediction Workshop (HiLIFTPW-2, <https://hiliftpw.larc.nasa.gov/index-workshop2.html>).
- Sanyal, J., Zhang, S., Bhattacharyya, G., Amburn, P., & Moorhead, J. (2009). A User Study to Compare Four Uncertainty Visualization Methods for 1D and 2D Datasets. *IEEE Transactions on Visualization and Computer Graphics*, 15(6), 1209-1218.

- Sanyal, J., Zhang, S., Dyer, J., Mercer, A., Amburn, P., & Moorhead, J. (2010). Noodles: A Tool for Visualization of Numerical Weather Model Ensemble Uncertainty. *IEEE Transactions on Visualization and Computer Graphics*, 16(6), 1421-1430.
- Whitaker, R., Mirzargar, M., & Kirby, R. (2013). Contour Boxplots: A Method for Characterizing Uncertainty in Feature Sets from Simulation Ensembles. *IEEE Transactions on Visualization and Computer Graphics*, 19(12), 2713-2722.
- Wu, K., & Zhang, S. (2013). A Contour-Tree Based Visualization for Exploring Data with Uncertainty. *International Journal for Uncertainty Quantification*, 3(3), 203-223.