Visualizing Fuzzy Points in Parallel Coordinates

Michael R. Berthold, *Senior Member, IEEE,* and Lawrence O. Hall, *Fellow, IEEE*

Abstract—**Exploratory data analysis heavily relies on methods to visualize data and models in a user friendly and interpretable manner. We show how models consisting of a collection of fuzzy points can be visualized in parallel coordinates. In contrast to existing techniques that display only lines representing centroids or shaded areas showing the general variance of cluster centers, the proposed technique shows the spread of the fuzzy membership in each dimension in detail. This allows for a better interpretation of overlap in fuzzy rules.**

Index Terms—**Fuzzy rules, parallel coordinates, visualization.**

I. INTRODUCTION

VISUALIZATION has received increasing attention for the analysis of large data sets. Several recent articles have attempted summarizations of visualization techniques, a special issue on this topic [1] contains a number of introductory articles. In [2], an overview of visualization techniques can be found as well. Most methods discussed so far have mainly focused on presenting an overview of the data itself. However, in many instances it is important to not only present a view of the data (or a subset of it) but also show summarizations of data, such as cluster based models [3] or decision tables [4]. One problem of such visualizations is the dimension of the underlying feature space, which tends to be larger than three. For medium-dimensional features spaces (3–20), techniques such as principal component analysis, multidimensional scaling [5], or even just simple two-dimensional (2-D) scatter plots have been used frequently. However, such visualizations lose potentially important information. Parallel coordinates [6] have been introduced as an alternative way to visualize data in medium-dimensional feature spaces. In [7], an approach was discussed that shows hierarchical clusters in parallel coordinates through an interactive tool that enables one to zoom in and out of the underlying hierarchy. The shaded spread of lines in parallel coordinate space visualizes the extension of each cluster. We show how this technique can be extended to also visualize fuzzy points, where each such fuzzy point corresponds to an imprecise location in n-dimensional space. Alternatively, fuzzy points can also be viewed

M. R. Berthold is with the Tripos, Inc., Data Analysis Research Lab, San Francisco, CA 94080 USA (e-mail: berthold@tripos.com).

L. O. Hall is with the Department of Computer Science and Engineering, University of South Florida, Tampa, FL 33620 USA (e-mail: hall@csee.usf.edu).
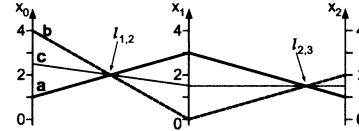
Fig. 1. Parallel coordinate depiction of three points on a line with $\mathbf{a} = (1.0, 3.0, 1.0)$, $\mathbf{b} = (4.0, 0.0, 2.0)$, and $\mathbf{c} = (2.5, 1.5, 1.5)$. The two intersection points at $l_{1,2} = (0.5, 2)$ and $l_{2,3} = (0.75, 1.5)$ uniquely describe the line going through all three points in the original Euclidean 3-D.
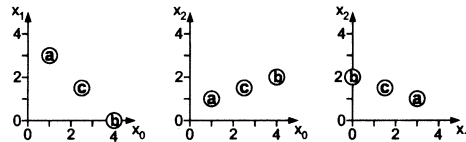

Fig. 2. Same points as in Fig. 1 shown in traditional 2-D projections.

as fuzzy clusters or as fuzzy rules in n dimensions. The shaded area in this case visualizes the degree of membership.

II. PARALLEL COORDINATES

An interesting approach to visualize high-dimensional data sets in two dimensions is *parallel coordinates* [6], [8]. Essentially, parallel coordinates transform multidimensional patterns into 2-D patterns. Visualization is facilitated by viewing the 2-D representation of the n-dimensional data points as lines crossing n parallel axes, each of which represents one dimension of the original feature space. This approach scales well with increasing n and has been incorporated into several data analysis tools such as Xmdv [9], XGobi [10], VisDB [11], and others [6].

The basis for parallel coordinates are a representation in which all n coordinate axes are lined up in parallel, next to each other. The distance between each adjacent axis is assumed to be equal to one. A point in n-dimensional space becomes a series of $n - 1$ connected lines in parallel coordinates which intersect each axis at the appropriate value for that dimension. A parallel coordinates example of three points in three dimensions, $\mathbf{a} = (1, 3, 1)$, $\mathbf{b} = (4, 0, 2)$, and $\mathbf{c} = (2.5, 1.5, 1.5)$, from a line is shown in Fig. 1. The same set of three points is shown in Fig. 2, this time using three traditional 2-D projections.

The dual of an n-dimensional line in Cartesian coordinates is a set of $n - 1$ points in parallel coordinates [6], [12], for the example in Fig. 1 these are indicated by $l_{1,2} = (0.5, 2)$ and $l_{2,3} = (0.75, 1.5)$, which uniquely describe a line in three dimensions.

The n-dimensional line in Cartesian coordinates can be represented by $n - 1$ linearly independent equations each of which results from equating a different pair of the following fractions [13]:

$$\frac{x_1 - a_1}{u_1} = \frac{x_2 - a_2}{u_2} = \cdots = \frac{x_n - a_n}{u_n}. \tag{1}$$
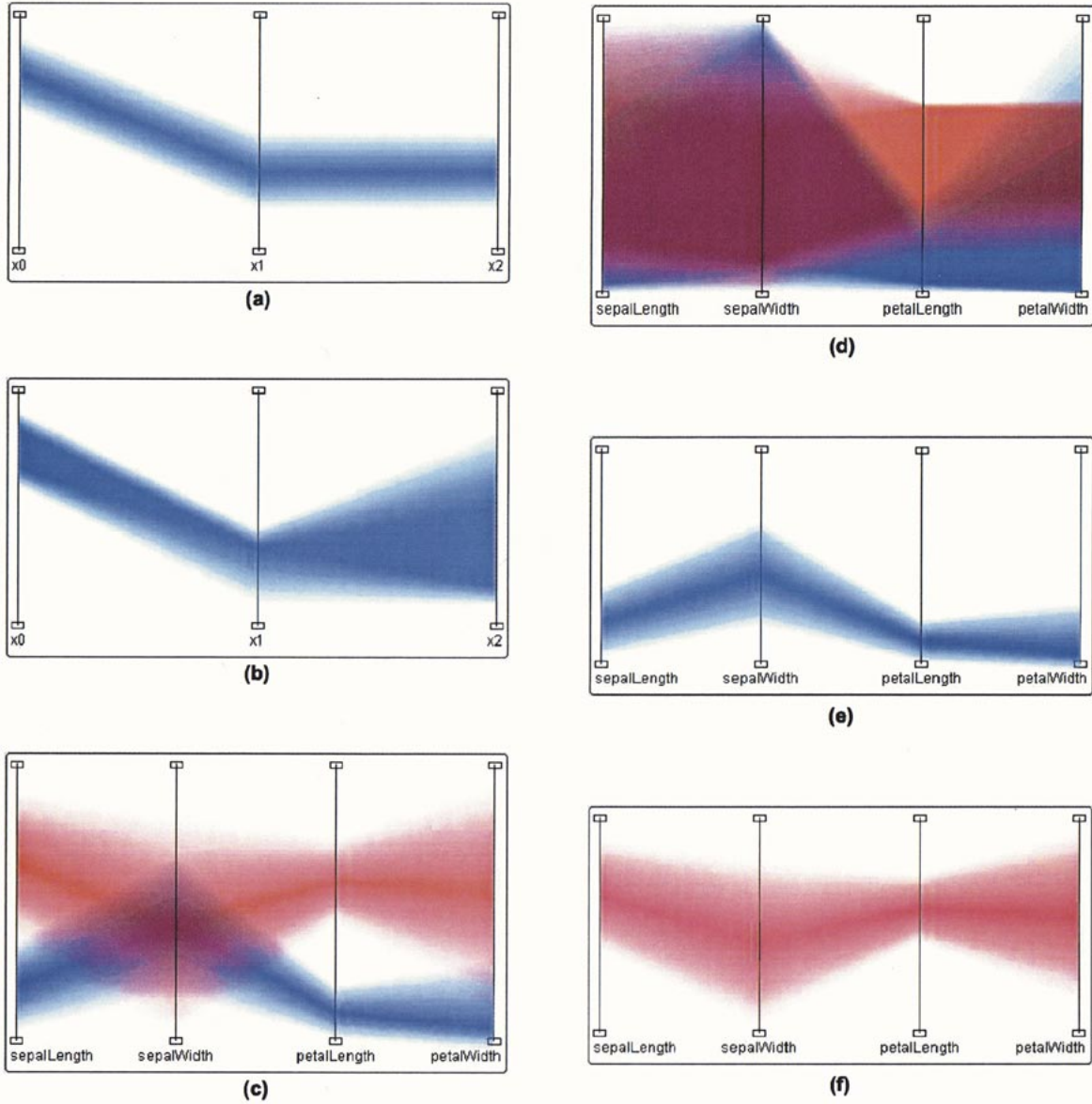
Fig. 3. (a) Fuzzy cluster displayed in parallel coordinates. The solid line indicates the centroid and the fading region shows the decline in membership value, which in this case comes from spherical membership functions. The resulting display is similar to the one used in [9] for visualization of hierarchical clusters. (b) Fuzzy rule with trapezoidal membership function displayed in parallel coordinates. The solid area indicates the core region ($\mu_{x_i} \equiv 1$ where μ_{x_i} is the membership function for dimension x_i). The fading region shows the (in this case) linear decline in membership value. (c) Two clusters for two different classes of the Iris data (blue: Iris-Setosa, red: Iris-Versicolor). Note how the cluster visualization presented in [9] only shows the cluster center and the spread as a shaded region with equal decline. (d) Two fuzzy rules which cover the most examples for classes Iris-Setosa (blue) and Iris-Versicolor (red) of the Iris data. Note how the fuzzy points indicate regions where patterns of each class were encountered (solid region) as well as the individual decline in membership value along each dimension (shaded regions). (e) and (f) Two clusters from (c) for two different classes of the Iris data shown separately. Note how the clusters suggest that these two classes can be separated along three attributes.

Now, it may be assumed that the $n-1$ linearly independent equations are obtained from pairing the $n-1$ adjacent fractions, with no loss in generality. This yields

$$x_{i+1} = m_i x_i + b_i, \qquad i = 1, 2, \ldots, n-1 \qquad (2)$$

where $m_i = u_{i+1}/u_i$ represents the slope and $b_i = (a_{i+1} - m_i a_i)$ the intercept of the x_{i+1}-axis of the projected line on the x_i/x_{i+1}-plane. The dual point of the n-dimensional line in parallel coordinates, therefore, corresponds to the set of $n-1$ indexed points

$$\left(\frac{i}{1-m_i}, \frac{b_i}{1-m_i} \right), \quad \text{for} \quad i = 1, 2, \ldots, n-1. \qquad (3)$$

There are other nice results about the parallel coordinate representation [12], [14], [15] that are not germane to this paper.

III. VISUALIZING CLUSTERS

Instead of visualizing each data point it is often preferable to only show a summary of the data set. Visualizing clusters in parallel coordinates has been done by the tool described in [9]. In their scheme, the center of the cluster (which is a crisp point) is represented by a line in the parallel coordinates and the spread of the cluster is shown using a shaded area surrounding this line.

The degree of shading declines linearly toward the standard deviation of the cluster spread and does not carry additional information.

However, in the case of fuzzy clusters the shaded region can also be used to visualize the corresponding degree of membership for the entire cluster. Fig. 3(a) shows an example for such a display. The resulting display is similar to the cluster display presented in [9] because in this example the fuzzy membership function is singular and declines equivalently along all dimensions.

IV. VISUALIZING FUZZY POINTS

A fuzzy point represents an imprecise location in n-dimensional space. A set of one or more crisp points belong (partially) to a given fuzzy point. More broadly, fuzzy rules may be viewed as fuzzy points. The antecedent of a fuzzy rule applies to a set of nonfuzzy points in n-dimensions. In each dimension, the fuzzy rule is likely to apply to a broader region than a single fuzzified n-dimensional vector. Also, fuzzy rules can be created from fuzzy clusters [16]–[18]. Hence, a fuzzy cluster can also be viewed as a fuzzy point in n dimensions where the fuzzy range for each dimension is determined by the spread of the cluster in each dimension.

However, fuzzy rules do not need to be singular, that is membership functions do not have to be convex or show only one point of maximum degree of membership. Powerful algorithms to construct general fuzzy rules from data exist [19]–[24]. The resulting rules can have a plateau, i.e., an interval as the core region, where the degree of membership is maximal and the membership function does not need to decline linearly or even equally on all sides. The information about the shape of the membership function would be lost in the existing cluster visualization.

Extending the aforementioned visualization leads to a display where the color shading is used to display the degree of membership. Fig. 3(b) shows an example for such a display.

V. EXAMPLES

A. Iris Data

We have incorporated the parallel coordinates visualization technique into a tool which generates fuzzy rules from data [20]. Using the well-known Iris data set [25], we demonstrate how an extension such as presented above can offer additional, valuable insights into the generated models. The Iris data consists of 150 four-dimensional patterns describing three classes of Iris plants: Iris-Setosa, Iris-Versicolor, and Iris-Virginica. The four dimensions consist of measures for the petal and sepal, length and width. Most classification algorithms generate less than ten rules or clusters to sufficiently describe this data. In Fig. 3(c), two overlapping clusters obtained from a fuzzy c-means [26] partition into three clusters are shown. Note how only the cluster center is indicated and the spread of the clusters only gives an estimation of the spread of the underlying data.

We can also generate a set of fuzzy rules using the algorithm described in [20]. The two rules which cover the most examples of the classes Iris-Setosa and Iris-Versicolor are shown in Fig. 3(d). Note how the rule visualization allows the user to

judge not only the spread of the underlying patterns (solid area, i.e., the core of the rules) but also indicates areas where no conflicting information was encountered (shaded area, where degree of membership is <1).

The difference becomes more obvious when the two clusters or fuzzy rules are displayed separately [Figs. 3(e) and (f) and 4(a) and (b)]. The cluster display in Fig. 3(e) and (f) suggests three features for a possible separation between the two classes: sepal-length, petal-length, and potentially also petal-width. The display of the fuzzy rules in Fig. 4(a) and (b), however, clearly indicates that sepal-length is not an appropriate choice to separate the two classes, due to the heavy overlap of the two rules along this feature. This is due to the symmetric nature of the cluster display. It only shows the centroid of each cluster and a shaded area surrounding it that symbolizes the cluster's spread. The fuzzy rule display, in contrast, summarizes the encountered training instances along each dimension.

A different way to visualize fuzzy rules, as shown in Fig. 4(c), also demonstrates which features are useful. Here, all six rules generated for the Iris data are shown in an interlaced manner. Only the core-regions are displayed, that is, regions for which the training algorithm actually encountered training examples. The well-known fact that the Iris data can be separated best along the last two features is obvious from this display as well.

B. Ocean Satellite Images

In this example, we show how parallel coordinates can be used to separate out the fuzzy points of a specific class. In this case, the fuzzy points are fuzzy rules in five dimensions. We are searching for a rule that can separate Red Tide clusters from the rest, which represent other types of water.

The satellite data comes from a satellite used primarily to examine the ocean. The images are from the Coastal Zone Color Scanner (CZCS) and are of the West Florida shelf [27], [28]. The CZCS was a scanning radiometer aboard the Nimbus-7 satellite which viewed the ocean in six co-registered spectral bands 443, 520, 550, 670, 750 nm, and a thermal IR band. It operated from 1979 to 1986.

The features used were the 443-, 520-, 550-, and 670-nm bands and the pigment concentration value was derived from the lowest three bands. Atmospheric correction was applied to each image [29] before the features were extracted. A fast fuzzy clustering algorithm, mrFCM [30], was applied to obtain 12 clusters per image. There were five regions of interest in each image. These consist of red tide, green river, other phytoplankton blooms, case I (deep) water and case II (shallow) water; 25 images were ground-truthed by oceanographers [31] and 18 of these were used for training. The eighteen training images were clustered into 12 classes. Each class or cluster was labeled from the ground truth image as its majority class.

Similar to the experiments described in [32], the labeled cluster centers from the training images were then given to the same fuzzy rule generating tool used with the Iris data. It generated a set of fuzzy rules which are shown in Fig. 4(d) using the interlaced polygons display of the core or regions where membership is one.

Utilizing this display and the sliders upon each axis, we then attempt to derive a rule which will classify all of the red tide
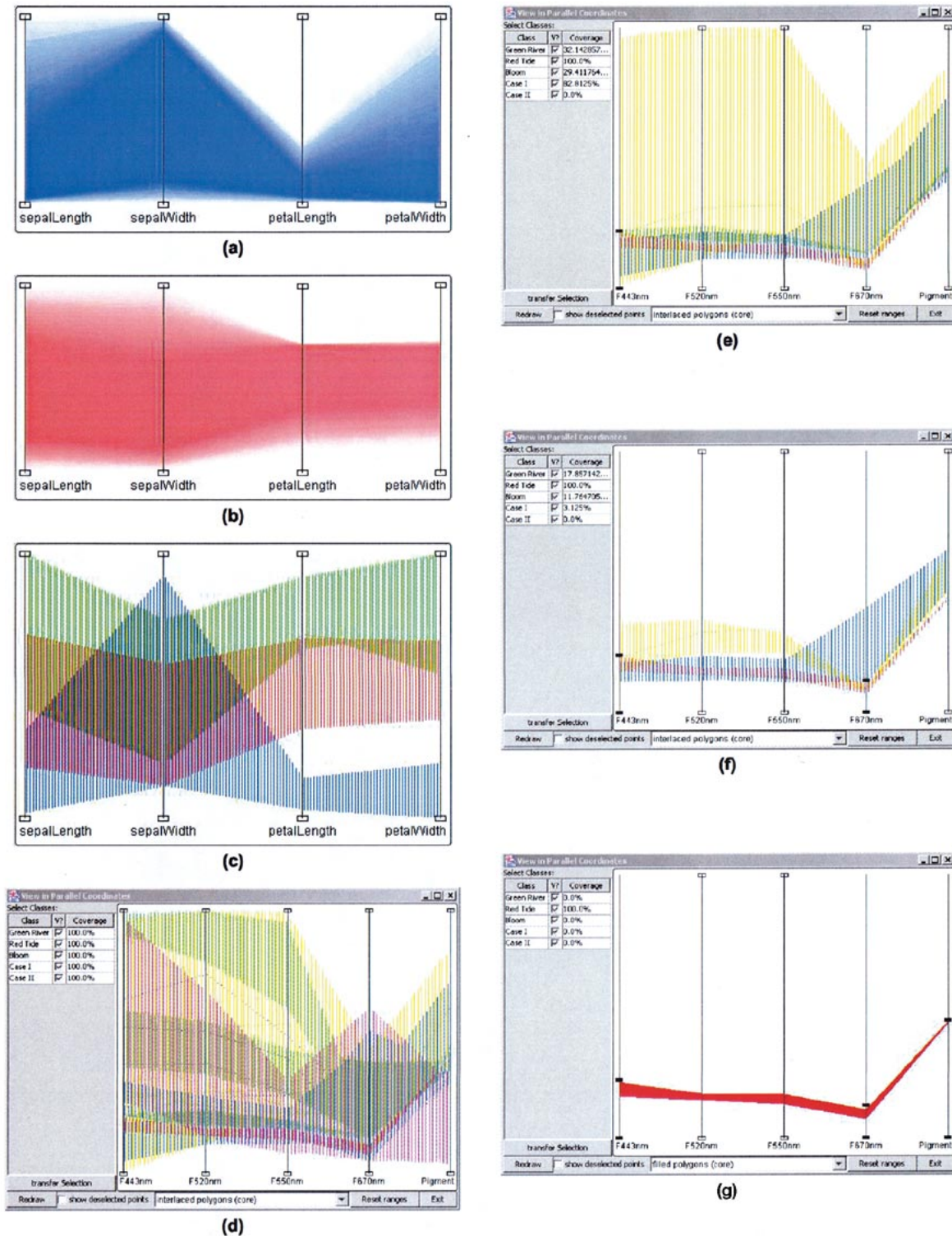
Fig. 4. (a) and (b) Two fuzzy rules from Fig. 3(d) for two different classes of the Iris data shown separately. Note how the fuzzy rules clearly indicate that the feature sepal-length should not be used to separate the two classes because of extensive overlap. (c) Interlaced display of all six fuzzy rules generated for the Iris data set (blue: Iris-Setosa, red: Iris-Versicolor, green: Iris-Virginica). Here, only the core region ($\mu \equiv 1$) is visualized. (d) Parallel coordinate view of the satellite image fuzzy rule cores. (e) and (f) Restricting the 443 nm range to $\leq$54.45 nm (e) and subsequentially the 670 nm band to $\leq$27.68 nm (f) isolates the red tide rules quite well. (g) Finally, restricting the pigment band ($\leq$112.71) will allow the extraction of a rule which identifies red tide clusters with no false positives.

clusters and no other clusters. The sliders are used to constrain the displayed rules to specific ranges on the corresponding attributes. That is, they are used to form a conjunctive crisp rule with each attribute lying within the displayed ranges. By suc-

cessively limiting the range of the rules in the 443–nm [see Fig. 4(e)], then 670-nm [see Fig. 4(f)], and finally pigment bands [see Fig. 4(g)], we derive a new crisp rule which models red tide. It correctly classifies all of the clusters in the training set that are
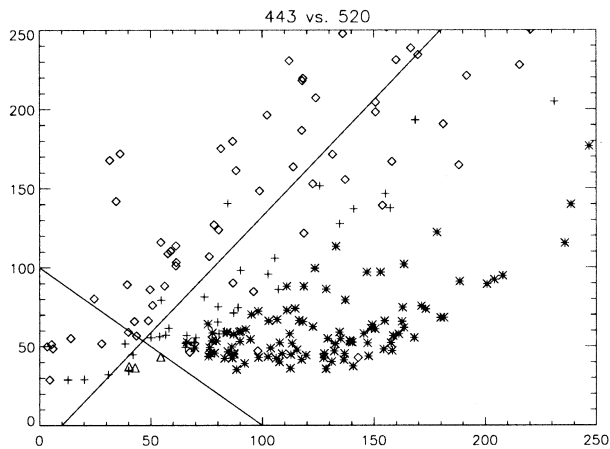
Fig. 5. 2-D projection with nonaxis parallel lines which split off most of the red tide cluster centroids. The red tide clusters are the triangular shaped objects on the graph, the +'s are other phytoplankton blooms (including green river), the stars are case I water and the diamond like objects are case II water.

red tide, while making no errors. This crisp rule was not previously known to us.

The generated rule actually better fits the training data than previous work. In [27] and [31], red tide (triangle shaped objects) were separated from most other fuzzy cluster centroids by using nonaxis parallel cuts in a two dimensional projection of the training images as shown in Fig. 5. The lines can be used to build a rule to separate most of the red tide from the rest of the clusters.

Our parallel coordinates display allows only axis parallel boundaries between classes and off axis boundaries were used in Fig. 5. However, the use of an extra attribute actually results in a model that better fits the training data. We recognize that mileage may vary on the test data, but the point here is to show the utility of parallel coordinates in deriving new information from fuzzy points.

VI. OUTLOOK

Techniques such as the one discussed here—which visualize models rather than the entire data set—make the exploration of very large data sets possible. Future work will also focus on how a parallel coordinate display can be used to adjust existing, and create new, fuzzy rules to model a given set of examples. As more rules are introduced or existing ones are altered the coverage can then be visually inspected. This would represent an extension along the lines of the multidimensional brushing technique presented in [7].

VII. CONCLUSION

We have demonstrated how an existing method to visualize cluster centers and their spread can be extended and used to visualize fuzzy rules, or more generally fuzzy points in more than 3-D spaces. The presented methodology uses parallel coordinates and provides visualization of the degree of memberships through different degrees of shading. The ability to inspect the degrees of membership in individual dimensions allows the user to investigate areas of overlap. We believe that such visualizations of fuzzy rules are a promising approach to exploratory

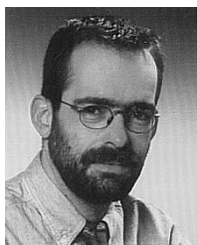data analysis and will also lead to interactive ways to adjust and create fuzzy rule sets.

REFERENCES

[1] D. A. Keim, "Information visualization and visual data mining," *IEEE Trans. Visual. Comput. Graph.*, vol. 8, pp. 1–8, Jan.–Mar. 2002.
[2] D. Keim and M. Ward, "Visualization," in *Intelligent Data Analysis, An Introduction*, 2nd rev. ed, M. R. Berthold and D. J. Hand, Eds. New York: Springer-Verlag, 2002.
[3] J. Vesanto, "SOM-based data visualization methods," *Intell. Data Anal.*, vol. 3, no. 2, pp. 111–126, 1999.
[4] R. Kohavi and D. Sommerfield, "Targeting business users with decision table classifiers," in *4th ACM SIGKDD Conf. Knowledge Discovery Data Mining*, 1998, pp. 249–253.
[5] J. Meulman, *A Distance Approach to Nonlinear Multivariate Analysis*. Leiden, The Netherlands: DSWO Press, 1986.
[6] A. Inselberg, "Multidimensional detective," in *Proc. IEEE Symp. Information Visualization*, 1997, pp. 100–107.
[7] Y.-H. Fua, M. O. Ward, and E. A. Rundensteiner, "Structure-based brushes: A mechanism for navigating hierarchically organized data and information spaces," *IEEE Trans. Visual. Comput. Graph.*, vol. 6, pp. 150–159, Apr.–June 2000.
[8] A. Inselberg and T. Avidan, "The automated multidimensional detective," presented at the IEEE Symp. Information Visualization, San Francisco, CA, 1999.
[9] Y.-H. Fua, M. O. Ward, and E. A. Rundensteiner, "Hierarchical parallel coordinates for exploration of large datasets," in *Proc. IEEE Conf. Visualization*, 1999, pp. 43–50.
[10] D. F. Swayne, D. Cook, and A. Buja, "Xgobi: Interactive dynamic data visualization in the x window system," *J. Comput. Graph. Statist.*, vol. 7, no. 1, 1998.
[11] D. A. Keim and H.-P. Kriegel, "Visualization techniques for mining large databases," *IEEE Trans. Knowl. Data Eng.*, vol. 8, pp. 923–938, Dec. 1996.
[12] S.-Y. Chou, S.-W. Lin, and C.-S. Yeh, "Cluster identification with parallel coordinates," *Pattern Recogn. Lett.*, vol. 20, pp. 565–572, 1999.
[13] A. Inselberg and B. Dimsdale, "Multidimensional lines I: Representation," *SIAM J. Appl. Math.*, vol. 54, no. 2, pp. 559–577, 1994.
[14] ——, "Multidimensional lines II: Proximity and applications," *SIAM J. Appl. Math.*, vol. 54, no. 2, pp. 578–596, 1994.
[15] C. Gennings, K. S. Dawson, W. H. Carter, and R. H. Myers, "Interpreting plots of a multidimensional dose-response surface in a parallel coordinate system," *Biometrics*, vol. 46, pp. 719–735, 1990.
[16] S. Chiu, "Fuzzy model identification based on cluster estimation," *J. Intell. Fuzzy Syst.*, vol. 2, no. 3, 1994.
[17] M. Delgado, A. F. Gomez-Skarmeta, and F. Martin, "A fuzzy clustering-based rapid prototyping for fuzzy rule-based modeling," *IEEE Trans. Fuzzy Syst.*, vol. 5, pp. 223–233, Apr. 1997.
[18] M. Friedman and A. Kandel, *Introduction to Pattern Recognition : Statistical, Structural, Neural and Fuzzy Logic Approaches*. Singapore: World Scientific, 1998.
[19] M. R. Berthold and K.-P. Huber, "Constructing fuzzy graphs from examples," *Intell. Data Anal.*, vol. 3, no. 1, pp. 37–54, 1999.
[20] M. R. Berthold, "Mixed fuzzy rule formation," Int. J. Approx. Reason., no. 32, pp. 67–84, 2003, to be published.
[21] J.-S. R. Jang and C.-T. Sun, "Neuro-fuzzy modeling and control," *Proc. IEEE*, vol. 83, pp. 378–406, Mar. 1995.
[22] Z. Chi and H. Yan, "ID3-derived fuzzy rules and optimized defuzzification for handwritten numeral recognition," *IEEE Trans. Fuzzy Syst.*, vol. 4, pp. 24–31, Feb. 1996.
[23] O. Cordon, M. J. Del Jesus, F. Herrera, and M. Lozano, "MOGUL: A methodology to obtain genetic fuzzy rule-based systems under the iterative rule learning approach," *Int. J. Intell. Syst.*, vol. 14, no. 11, pp. 1123–1153, Nov. 1999.
[24] M. Figueiredo and F. Gomide, "Design of fuzzy systems using neuro-fuzzy networks," *IEEE Trans. Neural Networks*, vol. 10, pp. 815–827, July 1999.

[25] R. A. Fisher, "The use of multiple measurements in taxonomic problems," in *Annual Eugenics, II*. New York: Wiley, 1950, vol. 7, pp. 179–188.

[26] N. R. Pal and J. C. Bezdek, "On cluster validity for the fuzzy c-means model," *IEEE Trans. Fuzzy Syst.*, vol. 3, pp. 370–379, June 1995.

[27] M. Zhang, L. O. Hall, and D. B. Goldgof, "Knowledge-based classification of czcs images and monitoring of red tides off the west florida shelf," in *Proc. 13th Int. Conf. Pattern Recognition*, vol. B, 1996, pp. 452–456.

[28] M. Zhang, L. O. Hall, D. B. Goldgof, and F. E. Muller-Karger, "Fuzzy analysis of satellite images to find phytoplankton blooms," in *Proc. IEEE Int. Conf. Systems, Man Cybernetics*, 1997, pp. 1430–1435.

[29] H. R. Gordon, D. K. Clark, J. L. Mueller, and W. A. Hovis, "Phytoplankton pigments derived from the nimbus-7 CZCS: Comparisons with surface measurements," *Sci.*, vol. 210, pp. 63–66, 1980.

[30] T. W. Cheng, D. B. Goldgof, and L. O. Hall, "Fast fuzzy clustering," *Fuzzy Sets Syst.*, vol. 93, pp. 49–56, 1998.

[31] M. Zhang, L. O. Hall, F. E. Muller-Karger, and D. B. Goldgof, "Knowledge guided classification of coastal zone color images off the west florida shelf," *Int. J. Pattern Recogn. Artif. Intell.*, vol. 14, no. 8, pp. 987–1007, 2000.

[32] W. Yao, L. O. Hall, D. B. Goldgof, and F. Muller-Karger, "Finding green river in seawifs satellite images," in *Proc. Int. Conf. Pattern Recognition*, vol. 2, 2000, pp. 307–310.

Lawrence O. Hall (S'85–M'85–SM'98–F'03) received the B.S. degree in applied mathematics from the Florida Institute of Technology, Melbourne, and the Ph.D. degree in computer science from the Florida State University, Tallahassee, in 1980 and 1986, respectively.

He is currently a Professor of Computer Science and Engineering at the University of South Florida, Tampa. His research interests lie in distributed machine learning, pattern recognition, and integrating AI into image processing. The exploitation of imprecision with the use of fuzzy logic in pattern recognition, AI, and learning is a research theme. He has authored over 140 publications in journals, conferences and books. He coedited the 2001 joint North American Fuzzy Information Processing Society (NAFIPS), IFSA Conference Proceedings.

Dr. Hall received the IEEE SMC Society Outstanding Contribution Award in 2000. He is a Past President of NAFIPS and the current Vice President for Membership of the IEEE Systems, Man, and Cybernetics Society. He is the Editor of the IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics, as well as an Associate Editor for the IEEE Transactions on Fuzzy Systems, the *International Journal of Intelligent Data Analysis*, and *The Handbook of Fuzzy Logic*.

Michael R. Berthold (S'96–M'97–SM'03) received the Diploma and Ph.D. degree in computer science, both from the University of Karlsruhe, Karlsruhe, Germany, in 1992 and 1997, respectively.

He is currently Director of Tripos, Inc.'s Data Analysis Research Lab, San Francisco, CA. From 1997 to 2000, he was a Research Fellow at the Berkeley Initiative in Soft Computing (BISC) and a Lecturer at the University of California, Berkeley. He was a Visiting Researcher at Carnegie Mellon University, Pittsburgh, PA, from 1991 to 1992 and at Sydney University, Sydney, Australia, in 1994. He also worked as a Research Engineer at Intel Corporation, Santa Clara, CA, in 1993. His research interests include neural networks, fuzzy logic, and explorative data analysis. His a coeditor (together with D. J. Hand) of the textbook *Intelligent Data Analysis: An Introduction* (New York: Springer-Verlag, 2003), which just appeared in a second, completely revised edition.

Dr. Berthold is an Associate Editor of the IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics and the *International Journal of Intelligent Data Analysis*. He is currently President of the North American Fuzzy Information Processing Society (NAFIPS).