Hybrid Visualization of the Medical Images Data Sets

Safa A. Najim
Computer Science Department
College of Science
Basrah University

Widad Abdulsamad Mansour
Biology Department
College of Science
Basrah University

ABSTRACT

This paper presents new method to visualize medical images data sets by using the properties of continuity and trustworthy dimensional reduction methods. Continuity and trustworthy dimensional reduction methods are well-known promising nonlinear methods are used to visualize different data sets, as medical images. However, their visualizations face the problem of *false colors* which lead the specialist to make wrong analysis of patient status. To overcomes these errors, we will combine these two methods in one to generate hybrid method has continuity and trustworthiness properties.
The proposed method produces best visualization by perfect preserving the corresponding color distances between visualization and original data sets in high-dimensional space. The application of hybrid method shows it is interested for visualizing medical images data sets. It has been compared with the continuity methods (as Isomap) and the trustworthy method (as curvilinear distance analysis (CDA)). The results proves the efficiency of of the proposed method in visualizing medical images data sets, where the *false colors* in the visualization are overcome as well as possible. The experiments shows the hybrid visualization has more chances to discover the true colors of the medical images data sets.

General Terms

Visualization, Medical images

Keywords

Visualization, Medical images, Dimensionality reduction, Isomap, CDA

1. INTRODUCTION

Medical imaging techniques are used to discover and understand the diseases development in human body. Because the image database has been rising during the last year, the useful information is required to help the specialist to evaluate the patient status [2]. For example, the most often method uses by a doctor is to look into images slices (slice by slice) to understand the disease development, and sometimes, it is difficult to have decision about that. In addition, if the disease is discover, it is difficult to know the disease severity or progression (early, medium or advance) [15]. Thus, visualizing medical images is a good way to generate accurate image summarizes the main information of all slices and give much better results. Color visualization by dimension reduction method is used recently to reduce a series of medical image slices into a right image [16] [9] [1]. Interpreting and analyzing the color image are easier than original slices [11]. This method tries to explore the right information to user by recognizing the similarities and differences in visual image. The user requires this summarized colour image, which represents the overview of all images, in order to start his process.

Each image in the data sets does not necessary have full knowledge, and sometimes, overall images do not help to explain the patient status. Feature dedication by dimension reduction methods has been used recently because it is successfulness in information visualization [17] [3] [5] [4]. Dimension reduction methods are used to reduce the dimensionality of original data sets into target dimension, where the hidden information in original medical images is revealed in the visual space. In visualization, the dimension reduction method projects the data sets into two-dimensional space (to generate gray image) or three-dimensional space (to generate color image). However, the visualization by dimensionality reduction is hard problem, and its efficiency depends on the used method.

There are numerous dimensionality reduction reduction methods have been proposed, such as well-known methods of multidimensional scaling (MDS) [12] and principal component analysis (PCA) [7], which they attempt to find projected space that preserve the corresponding distances in original and projected space [14]. Unfortunately, MDS and PCA are linear method, and their ability in visualization are limited. Thus, nonlinear dimensionality reduction methods, as Isomap [8], would be good choice to find satisfactory solution. Recently, Isomap has been widely used in visualization because it attempts to preserve the color distances as well to obtain highly efficient visualization [3] [10]. The efficency of Isomap in dealing with nonlinear structures is coming from using geodesic distance rather than euclidean distance.

In general, there are two classes of dimensionality reduction methods: continuity and trustworthy methods. Continuity methods, as Isomap, attempt to inherit the relations among global neighborhood points to local corresponding points in the projected space [8]. Although this procedure has advantages in strengthening the coherence of local corresponding points, *false neighborhoods* might occur in low-dimensional space [13]. In *false neighborhoods*, the farther away data points in the original space can be projected close by in the projected space; they cause projected space to be flattened. Trustworthy methods, as CDA [6], try to overcome the *false neighborhoods* problems by focusing on the projected space

distance to preserve the original distance. Although Isomap and CDA are used geodesic distance, the efficiency of CDA is better than Isomap because it is used perfect constraint to reduce the *false colours* possibility in visualization [6]. However, the *continuity errors* occurs in low-dimensional space. In *continuity errors*, the far away points in the projected space are neighbor points in the original space. *False colours* in the visualization of medical images are very dangerous problems than *continuity errors* because they lead the specialist to make wrong analysis of patient status.

In this paper, the continuity and trustworthy ideas are combined in hybrid method to overcome the errors and as much as possible to produce highly quality visualization of the medical images data sets. The organization of the paper is as follows. In section 2, the related works are explained , and the continuity and trustworthy preservation ideas are discussed. Section 3 shows the proposed methodology in visualizing medical images. Section 4 conducts some experiments on the results of the proposed method in quantitative and qualitative manners. Finally, the conclusions in section 5.

2. RELATED WORK

The visualization by dimension reduction methods can be done by using continuity or trustworthy preservations.

2.1 Continuity Preservation

In continuity preservation, the methods try to preserve and hold the global structure of the original high-dimensional data. While these methods focus on computing global distance, the geodesic distance is an optimum method to acquire the main information. There are several methods have continuity property, but Isomap is a best global nonlinear dimension reduction method [8]. Isomap depends on the distance matrix to find low-dimensional space. Each entry in distance matrix represents the geodesic distance $r_{ij} = ||x_i - x_j||$ between two high-dimensional pairs of data points x_i and x_j. Geodesic distance is better than Euclidean distance in recognizing between near and far away neighbors, and it has perfect approximation to their distances. Isomap is continuity method attempts to preserve the geodesic distance r_{ij} in original space with their corresponding linear distance $d_{ij} = ||y_i - y_j||$ in projected space. Thus, the Isomap method will minimize the cost function in Equation 1.

$$\phi(Y) = \sum_{i<j} (r_{ij} - d_{ij})^2 \tag{1}$$

Isomap can be used to visualize medical images data sets by reducing the dimensionality of projected space into two or three.

2.2 Trustworthy Preservation

Curvilinear distance analysis (CDA) is a good example of the trustworthy preservation, which tries to preserve the local structure of the low-dimensional data [6]. CDA attempts to preserve the pairwise distances in low-dimensional space with the their corresponding pairwise distance in original high-dimensional space. It uses geodesic distance, as Isomap, to keep the preservation between two distances. It is try to minimize the cost function in Equation 2.

$$\phi(Y) = \sum_{i<j} (r_{ij} - d_{ij})^2 F(d_{ij}, \lambda_t) \tag{2}$$

$$r_{ij} = ||x_i - x_j|| \tag{3}$$

$$d_{ij} = ||y_i - y_j|| \tag{4}$$

where r_{ij} is the geodesic distance between data points i and j in the original high-dimensional space, and d_{ij} is the euclidean distance between data points i and j in the low-dimensional space. F is a bounded decreasing function, and allows CDA to preserve the distances on different scales depending on the time dependent value of λ_t which gradually decreased from largest possible value to zero during projection process, as in Equation 5.

$$F(d_{ij}, \lambda_t) = \begin{cases} 1, & \text{if } d_{ij} \leq \lambda_t \\ 0, & \text{if } d_{ij} > \lambda_t \end{cases} \tag{5}$$

3. METHODOLOGY: HYBRID PRESERVATION

The relation between continuity and trustworthy preservations is a positive relationship. The increasing in the *false errors* leads to increase in the *continuity errors*, and vice versa. Thus, the focusing on the one type of preservation will not help to discard the errors in the projected space.

To preserve the topological structure of the original space in the low-dimensional space, the distance differences between two spaces should be minimizing as much as possible. The general definition of our hybrid cost function in Equation 6.

$$Hybrid_f = \sum_{i \neq j} \alpha * Continuity_{(ij)} +$$
$$(1 - \alpha) * Trustworthy_{(ij)} \tag{6}$$

where α is a variable has the value in the range $(0 < \alpha < 1)$.

In the continuity part, the Equation 1 is multiplied by weight function C to generate Equation 7.

$$Continuity_{(ij)} = (r_{ij} - d_{ij})^2 C(r_{ij}, \lambda_t) \tag{7}$$

$$r_{ij} = ||x_i - x_j|| \tag{8}$$

$$d_{ij} = ||y_i - y_j|| \tag{9}$$

$$C(r_{ij}, \lambda_t) = \begin{cases} 1, & \text{if } r_{ij} \leq \lambda_t \\ 0, & \text{if } r_{ij} > \lambda_t \end{cases} \tag{10}$$

The role of the weight function C is to update the points coordinates depending on their relation in the high-dimensional space. Including Equation 10 in Equation 7 is to overcome the continuity errors.

The trustworthy part reduces the trustworthy errors by updating the coordinates of the neighbors points in the low-dimensional space, as in the Equation 11

$$Trustworthy_{(ij)} = (r_{ij} - d_{ij})^2 T(d_{ij}, \lambda_t) \tag{11}$$

$$r_{ij} = ||x_i - x_j|| \tag{12}$$

$$d_{ij} = ||y_i - y_j|| \tag{13}$$

$$T(d_{ij}, \lambda_t) = \begin{cases} 1, & \text{if } d_{ij} \leq \lambda_t \\ 0, & \text{if } d_{ij} > \lambda_t \end{cases} \qquad (14)$$

Thus, the proposed hybrid method will be defined in Equation 15. This definition will be applied on the medical data sets to reduce their dimensionality, where the points will be projected in the proper locations in the visualization of the original data sets.

$$Hybrid_f = \sum_{i \neq j} [\alpha * (r_{ij} - d_{ij})^2 \, C(r_{ij}, \lambda_t) +$$

$$(1 - \alpha) * (r_{ij} - d_{ij})^2 \, T(d_{ij}, \lambda_t)] \quad (15)$$

4. EXPERIMENTAL RESULTS

In this section, the performance of the proposed method in visualizing medical images is evaluated , and compare it with continuity and trustworthy dimension reduction methods. Two data sets are used: first one has dimension 146x190x6 and the second one has 175x160x6. It should be note, the comparisons between methods are done with assumption that they have the same geodesic distance matrix. The comparisons will base on visual quantity and quality.

4.1 Quantitative Comparison

Figure 1 shows the CIELAB color images of first data sets by using Isomap, CDA and Hybrid methods, and Figure 2 shows the visualization of the second data sets. We measured the efficiency of these visualizations by using residual variance. By computing the trustworthy histograms, as in Figures 3 and 4, we can see the Hybrid method achieves more accurate preserving color distance of visualization of both data sets. For first data sets, the trustworthy histograms show the Hybrid method has few number of false colors than Isomap and CDA, and worst visualization is generated by Isomap because the number of false colors is very high. This efficiency lead the Hybrid method to have higher efficiency measurement. In Figure 2, the same efficiency is also happened for second data sets in which Hybrid method is better in having less number of false colors so that it has better trustworthy histogram. Thus, the proposed method has higher trustworthy value in second data sets than Isomap and CDA.

4.2 Visual Comparison

In order to show the preserving color distances of the visualization with corresponding distances in the original medical images data sets, visual comparison among the visualizations of Isomap, CDA and Hybrid methods are done. The role of colors in color image is to recognize the image features. Thus, the distances among neighbors in the original space should be preserved as much as possible to have same color distances in the visualization. The overlapping among pixels causes the pixels in the visualization have false colors.

To prove the efficiency of Hybrid method results, three regions are selected p_1, p_2 and p_3, as in Figure 5, and compare them with their corresponding regions in the visualizations of the Isomap and CDA. Isomap and CDA did not preserve the distances between visualization and original space for the three selected regions, and the proposed method gets values more close to the original
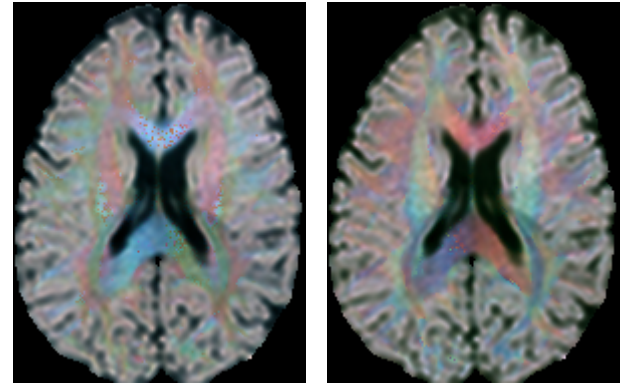
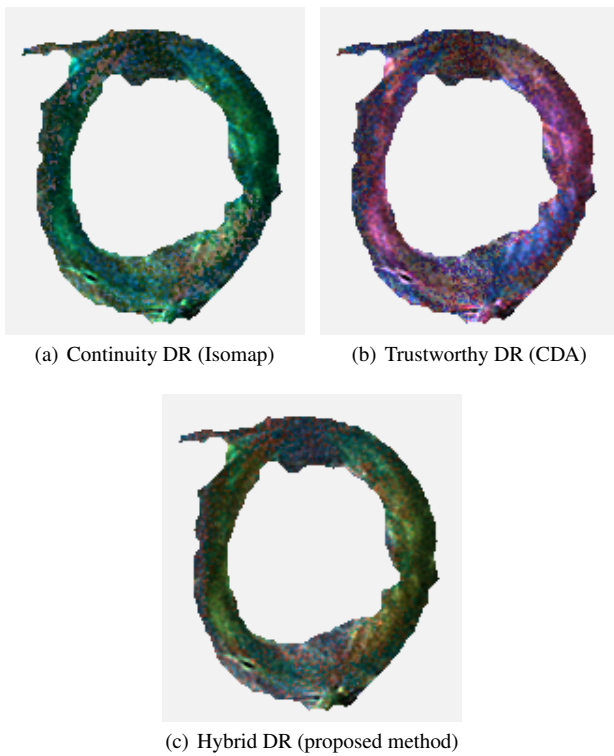(a) Continuity DR (Isomap)   (b) Trustworthy DR (CDA)
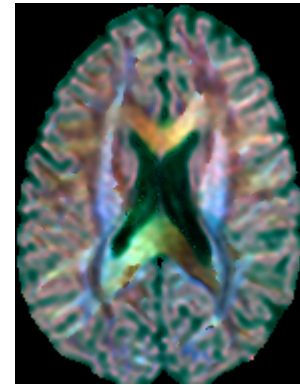
(c) Hybrid DR (proposed method)

Fig. 1: *CIELAB visualization of first medical images data sets by (a) Continuity dimension method (Isomap), (b) Trustworthy dimension method (CDA), (c) Hybrid dimension method (the proposed method).*

distances.

The regions p_1 and p_2 are neighbors in the original space, and they should be projected in the visualization as neighbors by having the similarity colors. The color distance between p_1 and p_2 in Isomap and CDA visualizations are 0.501 and 0.403, respectively, and these values are not preserved with their corresponding distance in the original space which is 0.254. Thus, p_1 and p_2 in Isomap and CDA visualizations gives wrong conception about what they represent. p_1 and p_2 in Hybrid visualization have the similarity colors where their color distance is 0.106, which is more close to their distances in the original space.

In the original space, the regions p_2 and p_3 are not neighbors, and they should have different colors in the visualizations. The proposed method has succeeded to prove that, but Isomap and CDA have drawbacks in this projection. The distance value between p_2 and p_3 in the original space is 0.576, and the colors distance between then in Isomap, CDA and Hybrid visualizations are 0.044, 0.163 and 0.534, respectively. The color distance in Hybrid method is better value to preserve their corresponding distance in the original space. The color distances are not preserved by Isomap and CDA methods. Thus, the proposed method achieves

(a) Continuity DR (Isomap)

(b) Trustworthy DR (CDA)

(c) Hybrid DR (proposed method)

Fig. 2: *CIELAB visualization of second second images data sets by (a) Continuity dimension method (Isomap), (b) Trustworthy dimension method (CDA), (c) Hybrid dimension method (the proposed method).*
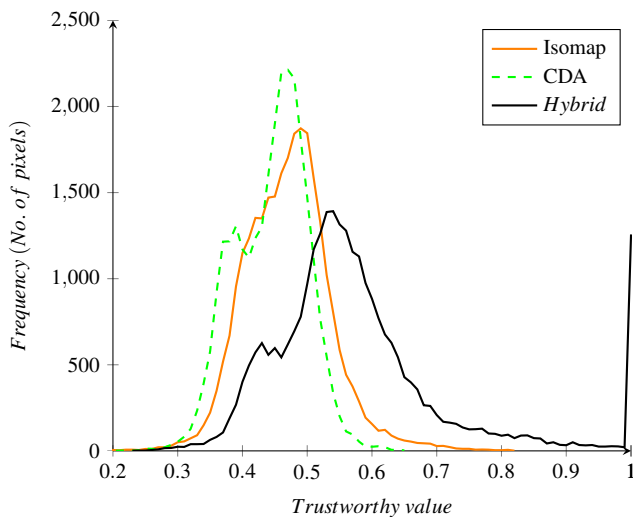
Fig. 3: *The efficiency measurement of Hybrid's visualization shows it has few number of false colors than the visualizations of the Isomap and CDA, so that Hybrid method gets better trustworthy histogram of the first medical images data sets.*
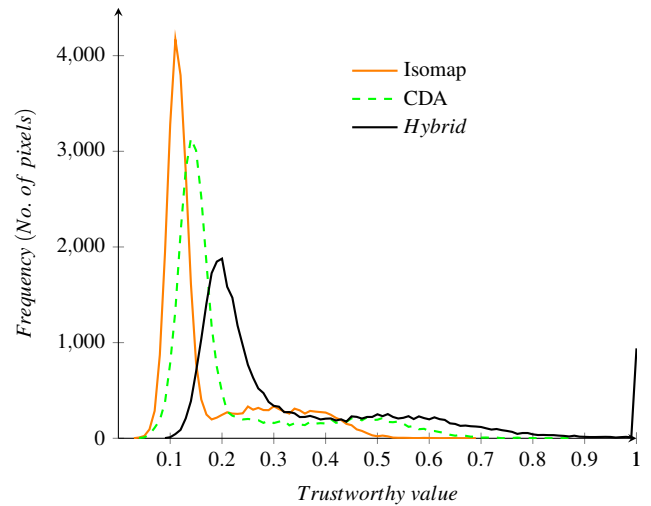
Fig. 4: *The efficiency measurement of the visualizations by using Isomap, CDA and Hybrid methods shows the visualization by Hybrid method gets lower number of the false colors than the visualizations by Isomap and CDA of the second data sets.*

more accurate trustworthy colors distances in it is visualization.

The visual comparison demonstrates that, although continuity and trustworthy dimension reduction methods might generate satisfactory visualization of medical images data sets, the Hybrid method visualization is more preferable for revealing perfect color display overcomes the *false neighborhood* errors.
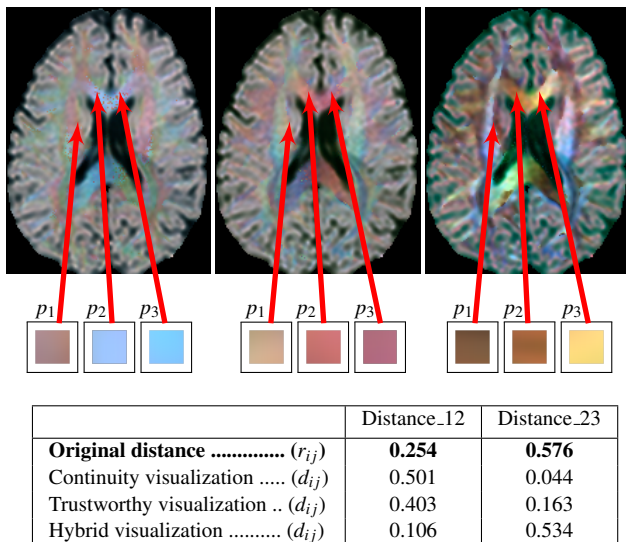
5. CONCLUSION

The Hybrid method is suggested in the paper, which has the continuity and trustworthy preservations, for visualizing two medical images data sets. The Hybrid method preserves the original relations among points in it's visualization, where the neighbors points have the similarity colors in the visualization.

The results showed the proposed method finds efficient visualization that represents most information in the original data sets. The efficient visualization is more better to understand by specialist to make good treatment of patient status. The measurement found the continuity and trustworthy visualizations lost more information by having huge false colors. The Hybrid visualization has overcame the *false colors* as well as possible, where they are less than those appear in the visualizations by Isomap and CDA. Hybrid visualization can discover more things which are invisible by continuity and trustworthy dimension reduction methods.

6. REFERENCES

[1] Samuel Gerber, Tolga Tasdizen, P. Thomas Fletcher, Sarang Joshi, Ross Whitaker, and the Alzheimers Disease Neuroimaging Initiative (ADNI). Manifold modeling for brain population analysis. *Medical Image Analysis*, 14:643–653, 2010.

[2] Patric Hagmann, Lisa Jonasson, Philippe Maeder, Jean philippe Thiran, Van J. Wedeen, and Reto Meuli. Understanding diffusion mr imaging techniques: From scalar diffusion-

	Distance_12	Distance_23
Original distance .... (r_{ij})	**0.254**	**0.576**
Continuity visualization ..... (d_{ij})	0.501	0.044
Trustworthy visualization .. (d_{ij})	0.403	0.163
Hybrid visualization .... (d_{ij})	0.106	0.534

Fig. 5: *The regions p_1 and p_2 are neighbors in the original space, and the region p_3 is far away of them (see the above table). Isomap and CDA cannot preserve the distance between p_1 and p_2 in the visualization, where p_2 has false colors by getting the same colors of p_3, as in the second column in the above table. Hybrid method preserves the distances between points by preserving the neighborhood relation between p_1 and p_2, where they obtained the similarity colors. Their colors are different than p_3 because they are projected far away to p_3 in the visualization.*

weighted imaging to diffusion tensor imaging and beyond. *Radio Graphic*, 26:205–223, 2006.

[3] Ghassan Hamarneh, Senior Member, Chris McIntosh, and Mark S. Drew. Perception-based visualization of manifold-valued medical images using distance-preserving dimensionality reduction. *IEEE Transactions on Medical Imaging.*, 30:1314–1327, 2011.

[4] Jihun Hamma, Dong Hye Ye, Ragini Verma, and Christos Davatziko. Gram: A framework for geodesic registration on anatomical manifolds. *Medical Image Analysis*, 14:633–642, 2010.

[5] Zhanli Hu, Jing Zou, Jianbao Gui, Junyan Rong, Yanming Li, Dongxing Xi, and Hairong Zheng. Real-time visualization and interaction of threedimensional human ct images. *Computers*, 5:1335–1342, 2010.

[6] Michel Verleysen John Aldo Lee, Amaury Lendasse. Nonlinear projection with curvilinear distances: Isomap versus curvilinear distance analysis. *Neurocomputing*, 57:49–76, 2004.

[7] I.T. Jolliffe. *Principal Component Analysis*. Springer Verlag, New York, 2002.

[8] Vin de Silva Joshua B. Tenenbaum and John C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290:2319–2323, 2000.

[9] Aljabar P., Wolz R., Srinivasan L., Counsell S. J., Rutherford M. A., Edwards A. D., Hajnal J. V., and Rueckert D. A combined manifold learning analysis of shape and appearance to characterize neonatal brain development. *IEEE Transaction on Medical Imaging*, 30:2072–2086, 2011.

[10] Parmeshwar Khurd Ragini Verma and Christos Davatzikos. On analyzing diffusion tensor images by identifying manifold structure using isomaps. *IEEE Transactions on Medical Imaging.*, 26:772–778, 2007.

[11] Richard Souvenir and Robert Pless. Image distance functions for manifold learning. *Image and Vision Computing*, 25:365–373, 2007.

[12] M. Steyvers. Multidimensional scaling. *Encyclopedia of Cognitive Science*, 2002.

[13] Michal Aupetit Sylvain Lespinats. Checkviz: Sanity check and topological clues for linear and non-linear mappings. *Comput. Graph. Forum*, 30:113–125, 2011.

[14] Silva D.D. Langford J. Tenenbaum, J. A global geometric framework for nonlinear dimensionality reduction. *Science 290*, 5500:23192323, 2000.

[15] Adam ?wito?ski, Marcin Michalak, Henryk Josi?ski, and Konrad Wojciechowski. Detection of tumor tissue based on the multispectral imaging. *In Computer Vision and Graphics, Lecture Notes in Computer Science, Part 2, of Springer*, 6375:325–333, 2010.

[16] Chris M. Clark Yong Fan, Nematollah Batmanghelich and Christos Davatzikos. Spatial patterns of brain atrophy in mci patients, identified via high-dimensional pattern classification, predict subsequent cognitive decline. *NeuroImage*, 39:1731–1743, 2008.

[17] Xin Zhao and Arie E Kaufman. Multi-dimensional reduction and transfer function design using parallel coordinates. In *Volume Graphics*, 2010.