Interpreting Distortions in Dimensionality Reduction by Superimposing Neighbourhood Graphs

Benoît Colange, Laurent Vuillon, Sylvain Lespinats, Denys Dutykh

Interpreting Distortions in Dimensionality Reduction by Superimposing Neighbourhood Graphs

Benot Colange[1,2*] Laurent Vuillon[2] Sylvain Lespinats[1] Denys Dutykh[2]

[1]Univ. Grenoble Alpes, INES, F-73375, Le Bourget du Lac, France
[2]Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LAMA, 73000 Chambry, France

ABSTRACT

To perform visual data exploration, many dimensionality reduction methods have been developed. These tools allow data analysts to represent multidimensional data in a 2D or 3D space, while preserving as much relevant information as possible. Yet, they cannot preserve all structures simultaneously and they induce some unavoidable distortions. Hence, many criteria have been introduced to evaluate a map's overall quality, mostly based on the preservation of neighbourhoods. Such global indicators are currently used to compare several maps, which helps to choose the most appropriate mapping method and its hyperparameters. However, those aggregated indicators tend to hide the local repartition of distortions. Thereby, they need to be supplemented by local evaluation to ensure correct interpretation of maps.

In this paper, we describe a new method, called MING, for "Map Interpretation using Neighbourhood Graphs". It offers a graphical interpretation of pairs of map quality indicators, as well as local evaluation of the distortions. This is done by displaying on the map the nearest neighbours graphs computed in the data space and in the embedding. Shared and unshared edges exhibit reliable and unreliable neighbourhood information conveyed by the mapping. By this mean, analysts may determine whether proximity (or remoteness) of points on the map faithfully represents similarity (or dissimilarity) of original data, within the meaning of a chosen map quality criteria. We apply this approach to two pairs of widespread indicators: precision/recall and trustworthiness/continuity, chosen for their wide use in the community, which will allow an easy handling by users.

Index Terms: Evaluation—Qualitative Evaluation; Non-Spatial Data and Techniques—Dimensionality Reduction; Visual Analysis and Knowledge Discovery—Visual Knowledge Discovery

1 INTRODUCTION

Dimensionality reduction aims at representing N points ξ_i, lying in a high dimensional metric data space, by a set of corresponding points $x_i = \Phi(\xi_i)$ in an embedding space of lower dimensionality (with Φ a mapping). To obtain this mapping, many techniques have been developed, including classical MDS [19], Isomap [18], DD-HDS [11] and tSNE [20]. In order to perform visual exploration of data, the embedding dimensionality is often chosen to be 2 or 3. This task mainly relies on the hypothesis that two similar data points must be embedded close from each other, whereas dissimilar points must be represented far apart. However, the mapping process often introduces distortions, which invalidate this premise. As a consequence, the proximity or remoteness of points on a map alone is not sufficient to infer information concerning structures (*e.g.* underlying manifold, clusters or outliers [8,13]) in the original data. The considered visualization must be enhanced with local evaluation, locating precisely the distortions measured through global indicators.

*benoit.colange@cea.fr

In the present article, we introduce a new local evaluation method called MING, for "Map Interpretation using Neighbourhood Graphs". This approach adapts global rank-based indicators (*e.g.* precision and recall [22] or trustworthiness and continuity [21]), designed for assessing overall performances of dimensionality reduction, to the needs of local evaluation. It displays on the map neighbourhood graphs, which link each point to its κ nearest neighbours, in order to exhibit distortions. On these graphs (considered successively in the embedding and data space), edges corresponding to reliable or unreliable neighbours are then distinguished using different colours, colour range being chosen so as to offer a local representation of the selected global indicators.

1.1 Notations

Δ_{ij} is the distance in the data space between two points ξ_i and ξ_j, and ρ_{ij} is the rank of ξ_j in the neighbourhood of ξ_i, meaning that ξ_j is the ρ_{ij}^{th} nearest neighbour of ξ_i, with $\rho_{ii} = 0$ by convention. Conversely to distances, ranks are not necessarily symmetric (ρ_{ij} may be different from ρ_{ji}). Distances D_{ij} and ranks r_{ij} may be defined equivalently in the embedding space. In the scope of this article, the κ neighbourhood of the i^{th} point in a given space is defined as the set of its κ nearest neighbours. It is denoted as a set of those neighbours' indices: $\nu_i(\kappa)$ in the data space and $n_i(\kappa)$ in the embedding space.

1.2 Distortions as corruption of neighbourhood graphs

Herein, we define distortions as a defect of preservation of *neighbourhood relations*. At a scale κ, the relation (i, j) is said to exist in a given space if the point j belongs to the κ neighbourhood of point i in that space. Those relations may easily be interpreted as edges of neighbourhood graphs. For a specific mapping, the neighbourhood relation (i, j) may be qualified as reliable if it exists in both spaces, false if it exists only in the map, missed if it exists only in the data space, or non-existent if it exists in neither of those. Figure 1 sums up these various cases, κ neighbourhoods being represented as balls centred on i and whose radius is the distance to the κ^{th} neighbour.

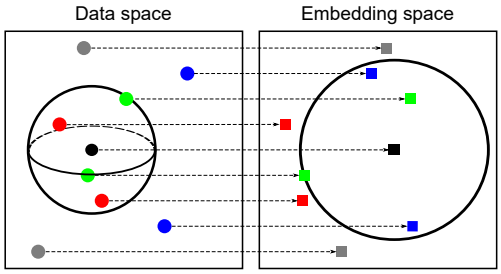

Figure 1: Preservation and distortions of 4-neighbourhood relations (i, j) for a point i (black), with reliable (green), false (blue), missed (green) and non-existent (grey) neighbours, and 4-neighbourhoods $\nu_i(4)$ and $n_i(4)$ materialized by the black sphere and circle.

2 Related work

Several indicators are currently used in order to estimate a map's quality. They are generally based on the comparison of:

- distances Δ_{ij} and D_{ij}, used in stress functions of many dimensionality reduction methods [4, 14, 19]
- ranks ρ_{ij} and r_{ij}, used in several rank-based quality criteria [9]. Due to their robustness to the phenomenon of norm concentration occurring in high dimensionality spaces, ranks tend to be more considered.

Global and local evaluation consider the distortions at three levels of aggregation.

1. Map-wise aggregation allows to produce global indicators for evaluating mappings (see Section 2.1). Those are computed by averaging point-wise values.
2. Point-wise aggregation of pairwise penalizations is often displayed by local evaluation methods, as mentioned in Section 2.2.
3. Pairwise penalizations are constructed for each neighbourhood relation (i, j), assessing its level of distortion. Though such penalizations are commonly established, previous works tend to either aggregate them, losing the pairwise information, or to display them for only one reference point at a time. In Section 3, we introduce MING to display the pairwise penalizations on neighbourhood graphs.

2.1 Global evaluation

In order to compare different mappings, quantitative measures have been developed, producing scalar indicators. Many of these go by pairs $(\mathcal{F}, \mathcal{M})$, with $\mathcal{F}$ penalizing false neighbours and $\mathcal{M}$ penalizing missed neighbours. We focus here on widely used rank based criteria: precision $\mathcal{P}$ and recall $\mathcal{R}$ [22], and trustworthiness $\mathcal{T}$ and continuity $\mathcal{C}$ [21]. However, the adopted framework, and thus our proposed method, may easily be extended to other rank based criteria, such as Mean Relative Rank Error (MRRE) [8], or, with minor changes, to distance based criteria (*e.g.* Curvilinear Component Analysis (CCA) [4] and Sammon's mapping [14] stress functions). At a given scale κ, the considered rank-based criteria are computed as detailed below for each pair of indicators $(\mathcal{F}, \mathcal{M})$.

Pairwise penalizations $\mathcal{F}_{ij}(\kappa)$ and $\mathcal{M}_{ij}(\kappa)$ are defined for all penalized neighbourhood relations (i, j), accounting for their distortions. Those penalized relations are restricted for each point i to the set of κ neighbourhood relations existing at least in the embedding space $\{(i, j) \mid j \in n_i(\kappa)\}$ for false neighbours indicator $\mathcal{F}$, and to the set of neighbourhood relations existing at least in the data space $\{(i, j) \mid j \in \nu_i(\kappa)\}$ for missed neighbours indicator $\mathcal{M}$. Table 1 sums up the values of these pairwise penalizations for the two selected pairs of indicators. With these penalizations, we can distinguish reliable relations with null weights from distorted relations (false or missed) with strictly positive weights. To calculate global indicators, based on this pairwise information, point-wise (Equations 1 and 2) and map-wise (Equation 3) aggregations are then performed successively:

$$\mathcal{F}_i(\kappa) \triangleq \sum_{j \in n_i(\kappa)} \mathcal{F}_{ij}(\kappa), \tag{1}$$

$$\mathcal{M}_i(\kappa) \triangleq \sum_{j \in \nu_i(\kappa)} \mathcal{M}_{ij}(\kappa), \tag{2}$$

$$\mathcal{I}(\kappa) \triangleq 1 - \frac{1}{N} \sum_{i=1}^{N} \frac{\mathcal{I}_i(\kappa)}{C_{\mathcal{FM}}(\kappa)} \text{ for } \mathcal{I} \in \{\mathcal{F}, \mathcal{M}\}, \tag{3}$$

where $C_{\mathcal{FM}}(\kappa)$ is the upper bound of the point-wise values, so that criteria go from 0 for worst mapping (reversed ranks) to 1 for an ideal mapping.

Table 1: Pairwise penalizations and sets of considered neighbours indices for precision/recall and trustworthiness/continuity.

$(\mathcal{F}, \mathcal{M})$	$\mathcal{F}_{ij}(\kappa)$ for $j \in n_i(\kappa)$	$\mathcal{M}_{ij}(\kappa)$ for $j \in \nu_i(\kappa)$
$(\mathcal{P}, \mathcal{R})$	$\begin{cases} 1 \text{ if } j \notin \nu_i(\kappa) \\ 0 \text{ if } j \in \nu_i(\kappa) \end{cases}$	$\begin{cases} 1 \text{ if } j \notin n_i(\kappa) \\ 0 \text{ if } j \in n_i(\kappa) \end{cases}$
$(\mathcal{T}, \mathcal{C})$	$\begin{cases} \rho_{ij} - \kappa \text{ if } j \notin \nu_i(\kappa) \\ 0 \text{ if } j \in \nu_i(\kappa) \end{cases}$	$\begin{cases} r_{ij} - \kappa \text{ if } j \notin n_i(\kappa) \\ 0 \text{ if } j \in n_i(\kappa) \end{cases}$

2.2 Local Evaluation

Many existing evaluation methods tend to display point-wise estimations of local distortions [12, 16, 17]. Among them, CheckViz [10] provides simultaneously the values of $\mathcal{F}_i$ and $\mathcal{M}_i$ with a bi-dimensional colour scheme interpolated onto the background, distinguishing areas with false/missed neighbours, or both. Yet CheckViz fails to show which pairwise relations suffer from these distortions.

Other approaches focus on a reference point to present its true neighbourhood relations, by colouring the background (*e.g.* Proxilens [2, 6]) or correcting distances [17]. These unveil reliable relations and distortions but for only one point i at a time. Martins et al. [12] also proposed to focus on subsets of reference points for representing missed neighbours.

3 Map Interpretation using Neighbourhood Graphs (MING)

In the present section, we describe a local evaluation method that:

- is consistent with state of the art global rank based indicators,
- shows the reliability of κ neighbourhood relations between pairs of points (without any aggregation),
- can be directly applied to data visualization in a 3D embedding space, while partially alleviating depth ambiguity and apparent distances distortions that occur in 3D scatter plots [15].

In this approach, we represent the penalized neighbourhood relations as edges of two directed graphs, each associated to one of the indicators. Using the interpretation of data visualization as an information retrieval task [22], the graph associated to the false neighbours indicator $\mathcal{F}$ is called the retrieval graph, defined as follows:

$$G^{\text{retr}}_{\mathcal{F}}(\kappa) \triangleq \left(V, E^{\text{retr}}(\kappa), W^{\text{retr}}_{\mathcal{F}}(\kappa)\right). \tag{4}$$

In Equation 4, V are the vertices of the graph (corresponding to the points), $E^{\text{retr}}(\kappa) \triangleq \left\{(i, j) \in V^2 \mid j \in n_i(\kappa)\right\}$ the edges corresponding to penalized neighbourhood relations for $\mathcal{F}$, *i.e.* relations existing at least in the embedding space (reliable and false), and $W^{\text{retr}}_{\mathcal{F}}(\kappa) \triangleq \{\mathcal{F}_{ij}(\kappa) \mid (i, j) \in E^{\text{retr}}(\kappa)\}$ the weights associated with these edges and corresponding to the pairwise penalizations defined for the indicator.

In the same way, the graph associated to the missed neighbours indicator $\mathcal{M}$ is called the relevance graph and may be defined by:

$$G^{\text{rel}}_{\mathcal{M}}(\kappa) \triangleq \left(V, E^{\text{rel}}(\kappa), W^{\text{rel}}_{\mathcal{M}}(\kappa)\right), \tag{5}$$

with $E^{\text{rel}}(\kappa) \triangleq \left\{(i, j) \in V^2 \mid j \in \nu_i(\kappa)\right\}$ the penalized edges for $\mathcal{M}$, *i.e.* relations existing at least in the data space (reliable and missed), and $W^{\text{rel}}_{\mathcal{M},ij}(\kappa) \triangleq \mathcal{M}_{ij}(\kappa)$ their associated weights.

The graphs are then visualized by successively superimposing them on the map (namely, positioning vertices $i \in V$ using coordinates of points in the embedding). Distortions introduced by the mapping are shown by colouring the edges depending on their weights. In order to stress the differences between the two graphs, the two kinds of penalizations are colour-coded differently, using *GnBu* for the retrieval graph and *OrRd* for the relevance graph (those

two colour maps from ColorBrewer [1] being colour-blind safe). Thus, reliable edges are displayed as white, while false neighbours are shown in shades of green and blue, and missed neighbours in shades of orange and red.

Neighbourhood relations being asymmetrical, directed edges (i, j) and (j, i) are different. Thus the line connecting points i and j on the map is split into two segments coloured independently. The colour closer from a point i corresponds to the weight of the edge (i, j) (in the direction from i to j), and equivalently, the color closer from point j shows the weight in the other direction (i, j). For relations existing only in one direction, as is common for outliers, the second part of the segment, associated with the non-existent direction is materialized by a dotted line. The best illustration of this may be found on Figure 4.

4 Experiments

4.1 Considered datasets

To illustrate our method, we consider two datasets. The first consists in 200 randomly selected samples from the handwritten digits dataset [5] (8×8 grey level images, seen as data points in a space of 64 dimensions). These are embedded in a 2D space using DD-HDS [11] (as shown in Figures 2, 3, 4 and 6). This dataset has a high intrinsic dimensionality, so that the mapping process has to introduce many distortions. The second is the coil-20 dataset, which consists in 128×128 images of 20 different objects continuously rotated around an axis. Consequently, its intrinsic dimensionality is intuitively very low, since the generation of images for one object has only one degree of freedom (which is the rotation angle). Thus, its embedding, performed using t-SNE, may be relatively good.

One key advantage of choosing these images datasets, is that we may use the actual images as markers for points on the map, allowing the reader to assess visually whether two items are similar or not in the data space (within the meaning of Euclidean distances). Note that for a better rendering of these images, the coil-20 data set has been sampled by removing one in two elements (artificially leading to a bigger angular step). Finally, the use of circular backgrounds, in the presented maps, intends to remind viewers that axes obtained by MultiDimensional Scaling have no directly accessible meaning [3]. Indeed, such a map may be translated, rotated or symmetrized without any effect on the information that it conveys.

4.2 Choice of κ

In MING, as in CheckViz, the user remains free to set the parameter κ, depending on the number of neighbours they judge relevant. Several values may also be used successively during users interaction with the representation, in order to get information for different scales, and on the hierarchical organization of data. However, with too high values of κ, MING figures tends to be less readable due to visual clutter, especially for the relevance graph, which may be considerably entangled. A similar phenomenon occurs for severely distorted mappings, which provide very few reliable neighbourhood relations.

For the digits dataset, a setting of $\kappa = 10$ has been selected, so that the majority of edges remains intra-classes. For the coil-20 data set however, points form continuums, so that only a few closest neighbours are sufficient to illustrate the quality of the mapping. Thus, we used a value of $\kappa = 4$ for this dataset.

4.3 MING interpretation guidelines

For the sake of comparison, Figure 2 shows CheckViz on the background of the same map used for illustrating MING. Results are coherent with MING about neighbourhood distortions for each point, but it fails to inform on which pairwise relations are distorted.

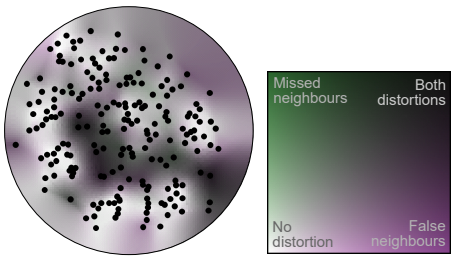

Figure 2: CheckViz with trustworthiness and continuity weights and $\kappa = 10$, for 200 handwritten digits data mapped using DD-HDS.

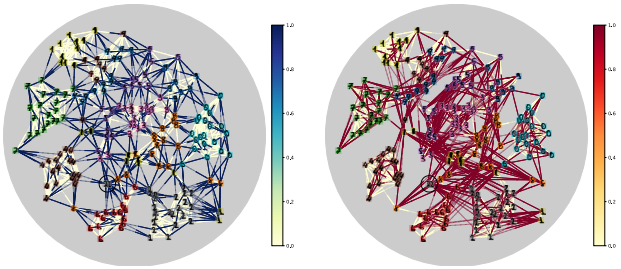

Figure 3: From left to right: MING retrieval graph $G^{\mathrm{retr}}_{\mathcal{P}}(10)$ and relevance graph $G^{\mathrm{rel}}_{\mathcal{R}}(10)$ with precision and recall weights, for 200 handwritten digits data mapped using DD-HDS.

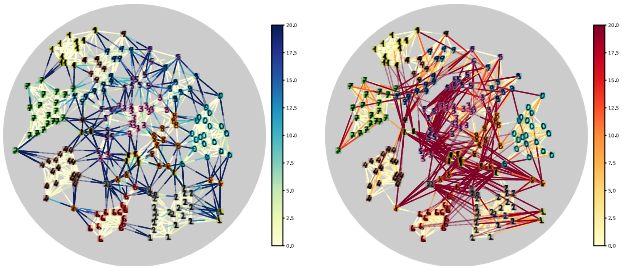

Figure 4: From left to right: MING retrieval graph $G^{\mathrm{retr}}_{\mathcal{T}}(10)$ and relevance graph $G^{\mathrm{rel}}_{\mathcal{C}}(10)$ with trustworthiness and continuity weights, for 200 handwritten digits data mapped using DD-HDS.

4.3.1 Precision and recall weights

Figure 3 shows the retrieval and relevance graphs for precision and recall weights. On the retrieval graph (left panel), the visually discernible group of 2 at the bottom right of the map is confirmed by the numerous white edges connecting points together. Conversely, this group is linked to the adjacent 6, 8 and 1 by a majority of blue edges, which shows that they are clearly separated in the data space. Note that without the superimposed graphs, a visual clustering could easily lead to wrongly splitting this group of 2 along its central vertical gap. A similar reasoning allows to locate the isolated pair of 2 circled on Figure 3 and lost between 4, 3, 8 and 6. Moreover, on the relevance graph, these two 2 are linked to the rest of their class by many red edges, indicating that they have been torn apart by the mapping. Thus, an analyst could safely conclude with MING that all the 2 are part of the same cluster, and should have been mapped together. In the same way, we may identify the 1 (top left), 4 (bottom left) and $4 \cup 9$ (top) clusters. In the mapping, legitimate proximities among distinct classes, supported on Figure 3 by visual similarity of digits, are also presented by the retrieval graph (left panel), for example between the 1 and the $4 \cup 9$ clusters, or inside the latter. On another note, the relevance graph (right panel) shows points placed between several attracting clusters mapped at distant

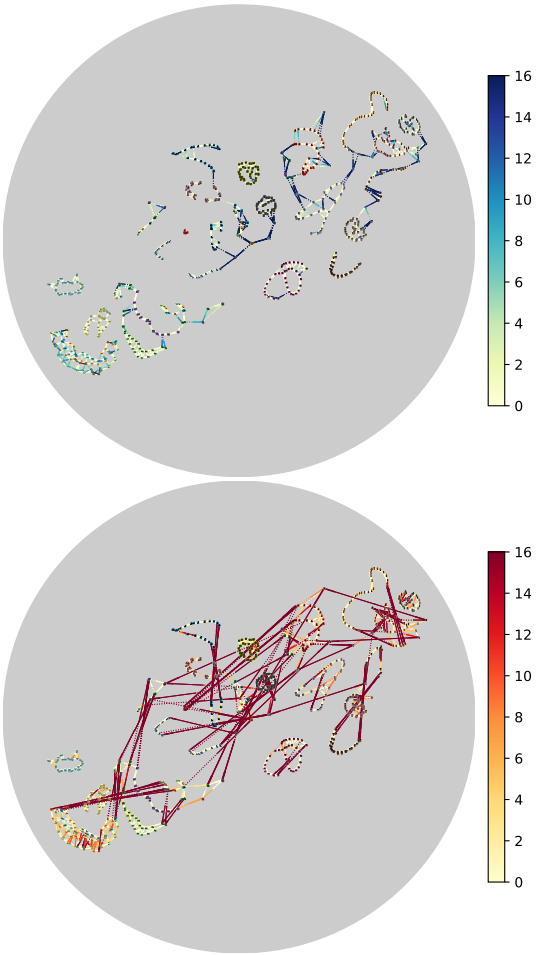

Figure 5: From top to bottom: MING graphs $G_{\mathcal{T}}^{\mathrm{retr}}(4)$ and $G_{\mathcal{C}}^{\mathrm{rel}}(4)$ with trustworthiness and continuity weights, for half of the coil-20 dataset mapped with t-SNE.

locations, such as the isolated 4 linked by numerous red edges to the $4 \cup 9$ and 4 clusters.

4.3.2 Trustworthiness and continuity weights

Since precision and recall weights are binary, slight rank distortions around the cut-off κ (*e.g.* tenth neighbour becoming the eleventh) are displayed the same way as severe distortions. Thereby, some undesired blue and red edges appear inside clusters on Figure 3. To allow a softer behaviour, we use trustworthiness and continuity weights (as presented in Table 1). In order to increase colour gradient, the colour map is saturated at 20 (namely, all penalizations above this limit are represented with the colour corresponding to maximum penalization), so that rank variations beyond the 30 neighbourhood (3κ) are ignored. We may note that the choice of colour saturation is critical in the perception of penalizations, a lower saturation inducing a higher perceived severity for an identical distortion. With this weighting, presented by Figure 4, edges inside clusters turn to lighter shades (green for formerly blue edges, and orange for red), as may be observed by zooming in on the cluster of 0. Hence, severely distorted neighbourhood relations tend to stand out, while light distortions are still noticeable through a thorough observation.

The coil-20 dataset forms cyclic continuums in the data space, the image of an object changing just slightly from one angular step to another. However, for some classes such as the different models of cars, images are more similar between classes for a specific angle, than they are for the same class with very different rotations. Thus, the mapping cannot perfectly preserve those cycles. Figure 5 presents MING graphs for the coil-20 dataset (with colour saturation at a penalization of 16, namely for ranks beyond 5κ). On the retrieval graph (top panel), white edges highlight well-represented parts of continuums, but also reliable proximities between different classes. Blue edges show isolated images torn from their group and cycles not closing well (*e.g.* souvenir pigs) or placed too close from another different class (such as the cats). On the relevance graph (bottom panel), red edges show the right way of closing cycles (*e.g.* for pigs or cats), and link isolated images to their main group.

4.4 Edge bundling

In order to reduce the visual clutter occasioned by the method, edge bundling was experimented with a variant of the KDEEB algorithm [7], using the previously considered embedding of digits data for the sake of comparison. For bundling together edges whose weights individually convey information, we segregated them in four groups (for each graph) by their associated penalizations value, using the bins $\{0\}$, $]0;10]$, $]10;20]$ and $]20;+\infty[$. We then drew the corresponding bundles using the central colour of each group, resulting in the display shown Figure 6.

The obtained visualization is effectively less cluttered, with the structure of frontiers between clusters remaining visible on the retrieval graph. However, the relevance graph suffers from confusion between long red edges, sometimes losing the local information about their source and target. It is for example the case for the edges linking the aforementioned isolated pair of 2 to their main cluster, which are diverted when crossing orthogonal edges starting from the cluster of 6. Hence, density-based bundling may not be the best suited solution to visual clutter.

Another simple way of reducing the impact of visual clutter could be to interactively draw the edges sorted by increasing penalization. By tuning a penalization threshold, the analyst could first discover the reliable edges, and then add on top of them the more distorted edges, which tend to induce the most clutter in the relevance graph.

5 Conclusions

In this paper, we proposed a new method for interpreting mappings while accounting for distortions by superimposing two colour-coded graphs. These graphs allow to diagnose reliable and distortion-induced neighbourhood relations, while quantifying those distortions through a colour scale based on classical map evaluation criteria. Therefore, our pairwise display allows to interpret locally the map-wise results obtained for global indicators. Future work may extend MING to other pairs of indicators, such as MRRE or CCA and Sammon stresses, or search for ways of reducing visual clutter with the minimal loss of information, especially for missed neighbours edges.

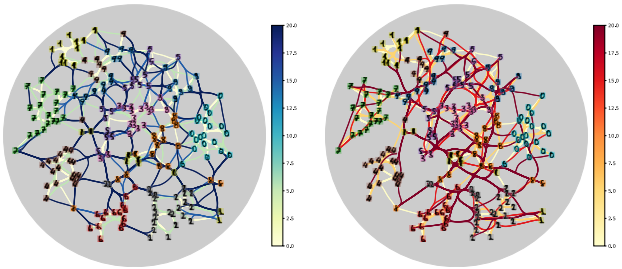

Figure 6: Edge bundling with a variant of KDEEB on the trustworthiness and continuity weighted retrieval and relevance graphs for the digits dataset.

REFERENCES

[1] ColorBrewer: Color Advice for Maps.

[2] M. Aupetit. Visualizing distortions and recovering topology in continuous projection techniques. *Neurocomputing*, 70(7-9):1304–1330, Mar. 2007. doi: 10.1016/j.neucom.2006.11.018

[3] F. Degret and S. Lespinats. Circular background decreases misunderstanding of multidimensional scaling results for naive readers. In *MATEC Web of Conferences*, vol. 189, p. 10002. EDP Sciences, 2018.

[4] P. Demartines and J. Hrault. Curvilinear component analysis: A self-organizing neural network for nonlinear mapping of data sets. *IEEE Transactions on neural networks*, 8(1):148–154, 1997.

[5] D. Dua and E. Karra Taniskidou. UCI machine learning repository, 2017.

[6] N. Heulot, M. Aupetit, and J.-D. Fekete. Proxilens: Interactive exploration of high-dimensional data using projections. In *VAMP: EuroVis Workshop on Visual Analytics using Multidimensional Projections*. The Eurographics Association, 2013.

[7] C. Hurter, O. Ersoy, and A. Telea. Graph Bundling by Kernel Density Estimation. *Computer Graphics Forum*, 31(3pt1):865–874, 2012. doi: 10.1111/j.1467-8659.2012.03079.x

[8] J. A. Lee and M. Verleysen. *Nonlinear Dimensionality Reduction*. Information Science and Statistics. Springer-Verlag, New York, 2007.

[9] J. A. Lee and M. Verleysen. Quality assessment of dimensionality reduction: Rank-based criteria. *Neurocomputing*, 72(7):1431–1443, Mar. 2009. doi: 10.1016/j.neucom.2008.12.017

[10] S. Lespinats and M. Aupetit. CheckViz: Sanity Check and Topological Clues for Linear and Non-Linear Mappings. *Computer Graphics Forum*, 30(1):113–125, Mar. 2011. doi: 10.1111/j.1467-8659.2010.01835.x

[11] S. Lespinats, M. Verleysen, A. Giron, and B. Fertil. DD-HDS: A method for visualization and exploration of high-dimensional data. *IEEE Transactions on Neural Networks*, 18(5):1265–1279, 2007. doi: 10.1109/TNN.2007.891682

[12] R. M. Martins, D. B. Coimbra, R. Minghim, and A. Telea. Visual analysis of dimensionality reduction quality for parameterized projections. *Computers & Graphics*, 41:26–42, June 2014. doi: 10.1016/j.cag.2014.01.006

[13] D. Sacha, L. Zhang, M. Sedlmair, J. A. Lee, J. Peltonen, D. Weiskopf, S. C. North, and D. A. Keim. Visual interaction with dimensionality reduction: A structured literature analysis. *IEEE transactions on visualization and computer graphics*, 23(1):241–250, 2017.

[14] J. W. Sammon. A Nonlinear Mapping for Data Structure Analysis. *IEEE Transactions on Computers*, C-18(5):401–409, May 1969. doi: 10.1109/T-C.1969.222678

[15] M. Sedlmair, T. Munzner, and M. Tory. Empirical Guidance on Scatterplot and Dimension Reduction Technique Choices. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2634–2643, Dec. 2013. doi: 10.1109/TVCG.2013.153

[16] C. Seifert, V. Sabol, and W. Kienreich. Stress maps: analysing local phenomena in dimensionality reduction based visualisations. In *Proceedings of the 1st European Symposium on Visual Analytics Science and Technology (EuroVAST10), Bordeaux, France*, vol. 1, 2010.

[17] J. Stahnke, M. Drk, B. Mller, and A. Thom. Probing Projections: Interaction Techniques for Interpreting Arrangements and Errors of Dimensionality Reductions. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):629–638, Jan. 2016. doi: 10.1109/TVCG.2015.2467717

[18] J. B. Tenenbaum, V. De Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *science*, 290(5500):2319–2323, 2000.

[19] W. S. Torgerson. Multidimensional scaling: I. Theory and method. *Psychometrika*, 17(4):401–419, 1952.

[20] L. van der Maaten and G. Hinton. Visualizing data using t-SNE. *Journal of machine learning research*, 9(Nov):2579–2605, 2008.

[21] J. Venna and S. Kaski. Neighborhood preservation in nonlinear projection methods: An experimental study. In *International Conference on Artificial Neural Networks*, pp. 485–491. Springer, 2001.

[22] J. Venna, J. Peltonen, K. Nybo, H. Aidos, and S. Kaski. Information retrieval perspective to nonlinear dimensionality reduction for data visualization. *Journal of Machine Learning Research*, 11(Feb):451–490, 2010.