Review on Manuscript JDSSV-96-1-4-20201211

1 Summary of the contribution:

The authors present the grand tour as an enhancement to non-linear dimensionality reduction methods for cluster orientation tasks. Four case studies are presented to illustrate the practical usage of the method. The grand tour has been investigated in data visualisation for quite some time and has shown its effectiveness as a multiple purpose tool. The current manuscript adds an innovative addition to the research in this field and convincingly shows the capacity of this method in combination to non-linear dimensionality reduction techniques.

2 Major remarks

The manuscript is well written and presents the approach in a balanced way. Despite this, the goal and primary focus of the paper gets not always clear as it it switches focus between presenting an enhancement of NLDR methods, presenting a better alternative to NLDR methods for certain tasks (clustering) and describing the general functionality of interactions with grand tour representations. A revision should clarify the goals of the manuscript.

Secondly, the relevance of having a cluster orientation task at hand seems relevant for some arguments but not for others. Here also, the authors should clearly provide guidance to the reader on which data and research contexts are essential for certain conclusions.

Thirdly, the term high-dimensional data has gained fairly specific meaning in some contexts. Please clarify whether you use the term in its rather broad sense or whether you want to allude particularly to small-N-large-p situations and their resulting conditions.

Fourthly, the manuscript has been enhanced by links to video sequences of tour animations for the case studies. While these videos add a lot to the static description in the papers, additional explanations to the videos (voice-over or captions) would further ease the understanding of the animations and hence improve further the manuscript.

3 Minor remarks

p2. 2nd paragraph: clusters ...applying known labels. Do you refer to classification tasks here or do you want to deliberately distinguish this 'clustering' approach from classification?

p.3, section 2: DR is used as an abbreviation here. While it obviously refers to 'dimensionality reduction' the abbreviation should preferably be introduced.

p.$3_1$4: that the DR procedure to **produce** Y ...

p.$4^2$2: several subtleties to the to use .... Delete the second **to**!

p.8 bottom: Does the last sentence mean that the collected output is a set of bases?

p.$11^1$2: ...We see in this case that **it** is ..

p.$15^1$9. ...For clustering workflows the primary task is **to** verify

p.$15^2$1: ...Here a 'faithful' a embedding ... Delete the **a** before embedding.

p.15, Section 5.4, 2nd paragraph: What is **QC**?