**IJRS Review**

Comments to the Author  
1-      The novelty of the proposed method is limited. (This paper proposes the use of clustering and ranking of them).  
2-      The grammatical errors must be corrected. For example, in page 2, lines 45-48, “increase” should be changed to “is increased”.  
3-      Representation of just a simple feature extraction such as PCA in introduction is not sufficient. More recently proposed feature extraction methods should be referred in introduction.  
4-      It is suggested that introduce the recently proposed feature selection methods in introduction and also compare your method with them.  
  
  
Referee: 2  
  
Comments to the Author  
The authors propose a new feature selection method which can reduce redundancy when using feature ranking scheme. The idea is interesting.  
  
However, there are a major problem recorded as follows: There has been a published article with similar ideas, i.e.,  X. Chen, G. Zhou, Y. Chen, G. Shao and Y. Gu, "Supervised Multiview Feature Selection Exploring Homogeneity and Heterogeneity With $\ell\_{1,2}$ -Norm and Automatic View Generation," in IEEE Transactions on Geoscience and Remote Sensing, vol. 55, no. 4, pp. 2074-2088, April 2017. The published paper involves not only the redundancy but also the relevance with feature clustering beforehand. Please clarify the difference with it, and the innovation of this manuscript.  
  
In addition, there are two minor problems in this paper:  
1) The categories of feature selection methods may be incomplete in page 4. Recently, sparse regularization methods emerge. They are different from the traditional filter, wrapper and embedded models. Some papers have talked about the trend. In remote sensing filed, it is mentioned in a paper , X. Chen and Y. Gu, "Class-Specific Feature Selection With Local Geometric Structure and Discriminative Information Based on Sparse Similar Samples," in IEEE Geoscience and Remote Sensing Letters, vol. 12, no. 7, pp. 1392-1396, July 2015.  
In other fields, Nie, F., Huang, H., Cai, X. and Ding, C.H. have discussed it in Efficient and robust feature selection via joint ℓ2, 1-norms minimization. In Advances in neural information processing systems (pp. 1813-1821).  
  
2) In page 8, the authors mentioned that there are two categories in feature selection procedures. The third category is to select informative by sparse regularization, not ranking, after  use clustering. It incorporates the trade-off between redundancy and relevance, such as Chen, X., Liu, W., Su, F. and Zhou, G., 2017. Semisupervised Multiview Feature Selection for VHR Remote Sensing Images With Label Learning and Automatic View Generation. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.  
  
3) One of the key words is big data. Unfortunately, the data used in this paper seems not big. Please reconsider this key word.

**My Comments**

They want me to mention their methods and clarify the differences with it (innovation of my method).

They want me to talk about Sparse Regularisation compared to traditional filter, wrapper etc methods.

They want me to talk about Sparse Regularisation as an option following clustering as opposed to ranking.

None of the above necessarily require me to do an actual numerical comparison or say that there is something fundamentally wrong with my method.