Dear Prof. Warner

I would like to submit a revision of the manuscript entitled “Feature Clustering and Ranking for Selecting Stable Features from High Dimensional Remotely Sensed Data” for possible publication in the *International Journal of Remote Sensing*. While I realise that the original manuscript was officially rejected, I believe we have addressed the referees' concerns, as I describe below.

In response to concerns about the novelty of the manuscript, I would like to highlight the following (these aspects are now also more clearly stipulated in the revised manuscript):

1. The problem of instability that occurs when selecting features from data containing redundancy has not been explicitly addressed or quantified in a remote sensing context. Our method aims to remove feature redundancy, and therefore the resultant instability, through the clustering step. We demonstrate the success of this on remote sensing data, using a quantitative measure of stability.
2. While related methods exist, the feature clustering and ranking (FCR) method (which combines hierarchical clustering for redundancy isolation, and naive Bayes or mutual information criteria for identifying relevance) has not been previously published. The combination of techniques used in FCR has some specific advantages:
   1. Hierarchical clustering is deterministic, fast and does not require prior knowledge of the number of clusters.
   2. The relevance criteria do not make assumptions of linear dependence between features and class labels, as is the case with structured sparsity approaches.
3. Our method is distinct in that it allows factors such as computation time and measurement cost to be considered, in addition to relevance, when selecting features from correlated clusters.

I believe we have addressed the referees' comments around recently proposed feature selection approaches, structured sparsity regularisation, and the Chen et al (2017) paper, "Supervised Multiview Feature Selection Exploring Homogeneity and Heterogeneity With 𝓁1,2-Norm and Automatic View Generation". Specifically, we have differentiated our method from the general approach of structured sparsity regularisation, and the multiview feature selection (MultiviewFS) technique of Chen et el (2017). In summary:

1. Structured sparsity regularisation approaches minimise an objective function that assumes a linear dependence between features and class labels. FCR uses a ranking heuristic to select features that does not make assumptions of linearity.
2. MultiviewFS performs redundant feature clustering using a Euclidean distance measure, while FCR uses correlation. The Euclidean distance describes feature similarity or proximity, while correlation is a broader description of redundancy that captures general linear dependence.

Other points raised by the referees have also been considered in the revised manuscript.

I hope that you will reconsider this submission.

Sincerely,

Dugal Harris