Reviewer's report

Title: Computer-assisted Lip Diagnosis on Traditional Chinese Medicine Using Multi-class Support Vector Machines

Version: 2 Date: 6 March 2012

Reviewer: Yiguang Liu

Reviewer's report:

# Reviewer's report

- Major Compulsory Revisions

The authors developed a quantitative method for lip diagnosis of traditional Chinese medicine. Although the introduced work is interesting, the poor use of English language hinders the reader from comprehending easily the presented work, and there are several problems need to be considered and revised.

1. The paper needs serious editing in terms of English and scientific writing, and the specialized words in the paper should be used more carefully and appropriately. For example, more widely used descriptions of the four diagnostic methods in traditional Chinese method should be “inspection, olfaction & auscultation, interrogation and palpation”. Avoid using “simply speaking” and writing in the first person. Any abbreviations introduced in the paper for the first time are needed to give their full version, for example, “CHD” in the 4th paragraph of the introduction section.

2. There are many types of classification methods such as (A novel and quick SVM-based multi-class classifier, Pattern Recognition, Vol. 39, No. 11, Nov. 2006, 2258-2264; k-NS: a Classifier by the Distance to the Nearest Subspace, IEEE Transactions on Neural Networks, Vol.22, No.8, 2011, 1256-1268; Classification Using Distances from Samples to Linear Manifolds, Pattern Analysis and Applications ). Why the authors choose Support Vector Machine, and why not use other more common classifiers as the comparative methods, such as Naïve Bayesian, Bayesian networks, neural networks, etc. The authors should introduce the methods and compare them.

3. There are lots of feature selection methods, why the authors use recursive feature elimination? 9 selected features used in the experiment are chosen based on SVM, according to the reference [16] in the paper, the RFE depending on the classifier, therefore, I have a question that are the 9 selected features used by WSVM, kNN, and FLD same to the selected features of SVM?

4. The experimental dataset may be too small. As an extension work could the author give an extensive experimental results based on a larger dataset in this paper? The accuracies for Pale lip prediction are low and the Pale lip only has 12 instances, do they have any relation?

5. The medical meanings of the features used should be introduced in the paper
simply, and the affections of each type of features for lip image classification should be shown in the experiments. It is more helpful and valuable to the other researchers who have the related research interests.

6. The importance of feature selection to classification is a common view. In the experiment, the authors should analyze how the selected features effect on the accuracy.

7. Why the performance of WSVM is lower than SVM’s? According to the description in section 2.5 “the data is highly unbalanced and the sample size is rather small”, in the experiments, the performance of WSVM should be better than SVM’s. The authors should analyze the reasons.

8. The authors need conclude their work more carefully. What an explicit representation of the knowledge is? Does Lip diagnosis in TCM only indicate the four color lip classification? Could the experimental results sufficiently support the conclusion that the proposed approach is an effective technique for solving problems with multi-class in clinical practice of TCM?

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Needs some language corrections before being published

**Statistical review:** No, the manuscript does not need to be seen by a statistician.