

## ***Interactive comment on “Block based cloud classification with statistical features and distribution of local texture features” by H.-Y. Cheng and C.-C. Yu***

**Anonymous Referee #1**

Received and published: 21 January 2015

In the present paper the cloud classification problem is considered, based on all-sky images. The proposed methodology follows the basic stages of a classical pattern classification system, with some extensions that are suitable for the specific problem at hand. Specifically, an image is first divided into blocks via the application of a rectangular grid on it. For the representation of each block, classical statistical spectral and textural features are combined with local texture features to form the corresponding feature vector. Then Principal Component Analysis (PCA) is performed and the most significant components are retained (selected in terms of an objective criterion). Then, on the reduced dimensionality data set, classification is performed using k-NN,

C4686

SVM and a variation of the Bayesian classifier with normal class distributions. After the completion of the training phase, an unknown (whole) image is classified as follows. After the segmentation and the classification of each block of the image, the obtained results from each block are combined, in order to identify whole regions in the image having the same type of clouds. Finally, classification of the whole image to a single cloud class is carried out according to a vote (majority) rule.

The paper exhibits an acceptable degree of novelty. It is well written, well structured and easy to follow. The references are adequate and the English are at an acceptable level (although there is still room for improvement).

However, there is an issue concerning the experiments. Specifically, concerning figure 10, it is stated in the first two lines of p. 16 “To compare the effect ... using different features and classifiers”. This sentence implies that the accuracy results shown in figure 10 are based on the classification of the training set. Nevertheless, this strategy leads to overestimated classification rates, since the same data were used for the classifier training. Unless techniques not explicitly mentioned (like “leave-one-out”) have been used... However, if the later is not the case, the corresponding experiments should be repeated. In any case, it would be better if a set of blocks of the testing images were obtained for this experiment.

Some additional comments follow:

Comparison with other works seems to be missing from the experimental verification of the method. However, some of the variants considered (e.g. only statistical features and k-NN) seem to have been proposed in previous works. Please mention explicitly some of these here. In addition, the comparison with more relevant works will make the results obtained by the proposed method more reliable.

Are the data set publicly available or it has been gathered by the authors? Some more information on this should be given.

C4687

What is the block size of the segmented images? Please mention.

In section 2.6 what is the value of the threshold for revising the classification of cloud initially classified as “stratus” to “cumulus”? Please mention.

Which classifier was involved in measuring the classification accuracy in fig. 9? Please mention.

It is implied that the images used are RGB. It would not harm if this were written somewhere in the text.

In page 5, lines 97 and 97: Replace the two “and” with “or”.

Line 161: Replace “circular” with “circularly”.

Line 184: Replace “the linear combination” with “linear combinations”.

Line 185: Replace “would” with “will”. This appears also in other places in the text. Please check.

Line 186: Replace “principal components ... larger variances” with “most of the data variability to the first (often few) principal components”.

Line 190: Replace “are the principal components” with “define the principal component directions”.

Line 208-209: The sentence contained in these two lines should move to line 207, before “For simplicity...”, taking care to make the appropriate syntactical corrections.

Line 218: The font size of the mathematical symbols should be larger. The same appears in other places in the text. Please check.

Line 232: Replace “Use” with “Using”.

Line 240: Insert before the full stop “as kernel functions”.

Figures 11 and 12 are erroneously referred to 1 and 2. Please correct.

C4688

Line 350: Replace “would” with “were shown to”.

---

Interactive comment on Atmos. Meas. Tech. Discuss., 7, 11771, 2014.

C4689