Verification of a Multispectral Cloud Classifier

Jukka Iivarinen, Markus Peura, Ari Visa
Helsinki University of Technology
Laboratory of Computer and Information Science
Rakentajanaukio 2 C, FIN-02150 Espoo
Finland
email:Jukka.Iivarinen@hut.fi

Abstract

The multispectral cloud classification scheme is implemented to automate the interpretation of satellite images. The developed classifier is capable of classifying satellite images taken round the year, during day and night. The classifier is fully automatic, and it can be adapted to changing situations with new examples. The classification is done in two phases, the clouds are detected and they are classified. There is a need for numerical and objective methods in evaluating the classifier. The automatic evaluation scheme to check the quality of the classifier is based on the comparison between the SYNOP observations and the satellite observations. Comparisons with other published results show that the classifier is working relatively well. The improvement of wintertime classification is the main objective for future work.

1. Introduction

The high-resolution satellite images received from weather satellites are used for weather forecasting and medium term prediction. The NOAA-11 and the NOAA-12 satellites are available in the Nordic countries. The automatic interpretation of satellite images has been studied in Laboratory of Computer and Information Science in Helsinki University of Technology since 1991. The main interest has been the classification of clouds from satellite images.

There have been several studies concerning the cloud detection and the classification of cloud types from a satellite image since 1970's [1]. The bispectral method introduced by Reynolds and Haar in 1977 consisted of localizing the cloud cover by using a visible channel and then estimating the cloud top temperature and the height of clouds by using an infrared channel [16]. Statistical methods have been applied widely, for example in [3] [12] [11].

Saunders and Kriebel developed a method for detecting clear sky and cloudy radiances [17]. Five daytime and five night-time thresholds were used to detect a pixel in an image as cloud-free, partly cloudy, or cloud-filled. Karlsson and Liljas have used multispectral threshold algorithms to construct an operational classifier for cloud types, precipitation intensities, and snow extent [10]. The classification is made according to season, satellite, and sun elevation. An automatic cloud detection algorithm was developed by Derrien et al. [4]. The method is based on threshold tests applied to different combinations of channels. It has been on operational use at the Centre de Météorologie Spatiale (CMS) in France since 1990.

Texture measures have been introduced in cloud recognition by Kittler and Pairman in 1985 [14]. They used texture measures to detect clouds from the sea and then to discriminate between cloud types. Ebert developed a method for polar cloud classification [5] [6]. Ten texture and spectral features were combined with a maximum likelihood decision rule to identify regions of various surfaces and cloud types.

Some neural network based methods have also been applied to cloud classification since the end of the 1980's. The multilayer perceptron was used by Key et al. [13], Benediktsson et al. [2], and Lee et al. [15]. The use of neural network as a classifier was promoted by Welch et al. after the evaluation of several methods [22]. The Self-Organizing Map was used by Visa et al. [21][18] [20]. In this paper some part of the performance of the simplified version is verified. Results are reported here.

2. The Multispectral Classifier

The classifier developed in this study is a multispectral, vector-based classifier. It is further developed from the classifier based on the Self-Organizing Map (SOM) [21][18] [20].

The classification of a satellite image is done in two phases (Figure 1). In the first phase the clouds are separated from the surface (referred as *cloud screening*), and in the second phase the cloudy regions are further classified into ten different cloud types (referred as *cloud classification*).

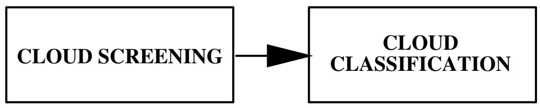


Figure 1: The classification is done in two phases. Only cloudy regions are considered in cloud classification phase.

The selected features are essential to the performance of a multispectral classifier. The features are found among the features suggested in literature. They are physically motivated. Feature selection was done with SOM [7]. The features used in cloud screening and in cloud classification are depicted in Table 1. Due to the absence of visible light at nights, there are different features for night classification. The selected features are spectal features and one texture feature.

At daytime features VIS1 (0.58-0.68 μm), VIS2 (0.725-1.1 μm), and INF1-INF2 must be normalized according to the sun zenith angle. The wavelength bands of channels INF1, INF2, and INF3 are 3.55-3.93 μm, 10.3-11.5 μm, and 11.5-12.5 μm, respectively.

In the cloud screening procedure a feature vector is extracted for each image pixel. The extracted feature vector $\mathbf{f}$ is compared with the codebook. The comparison is done in a best match way. The classification of cloudy regions is then accomplished. A new feature vector $\mathbf{g}$ is extracted and the classification is done as in the cloud screening procedure (Figure 2).

	Cloud Screening		Cloud Classification	
Feature	*Day*	*Night*	*Day*	*Night*
VIS1			x	
VIS2	x			
VIS1-VIS2	x			
INF2	x	x	x	x
INF1-INF2	x	x	x	x
INF2-INF3			x	x
(INF1-INF2)/VIS1			x	
var(INF2)	x	x	x	x

Table 1: Features used by the classifier.

The pooled form KNN algorithm ($K = 3$) [19] with the Hamming distance ($p = 10$) is used in cloud screening and cloud classification. The Hamming distance between the codebook vector $\mathbf{m_i}$ and the feature vector $\mathbf{f}$ is defined as

$$d_{Hamming}(\mathbf{m_i}, \mathbf{f}) = \sum_j d_j, \quad where \quad d_j = \begin{cases} 0 & \text{if } \left| \frac{m_i(j) - f(j)}{m_i(j)} \right| < \frac{p}{100} \\ 1 & \text{otherwise} \end{cases} \tag{1}$$

where p is a integer value that defines a small interval that fulfills the equation.

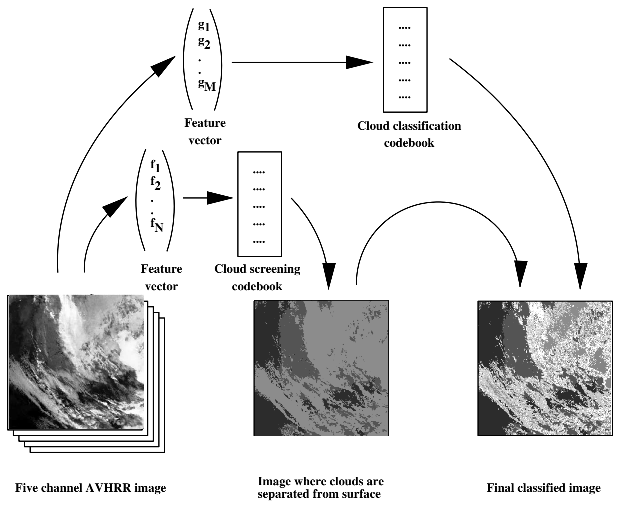


Figure 2: The classification procedure.

The cloud screening codebooks (a night and a day codebook) and the cloud classification codebooks (a night and a day codebook) are formed by a number of preclassified feature vectors. For each season there are these four codebooks, making a total of 16 codebooks. These codebooks can also be created by SOM but during this verification the preclassified samples are preferred to SOM generated codebooks.

2.1. Classes

The classes used in cloud screening are open sea, land, and cloud. In wintertime snow and ice are also classified. The cloudy areas are further classified to ten cloud types which are cirrus over land/sea, cirrus over low clouds, cirrus over middle clouds, cirrostratus, altostratus/altocumulus, stratus/stratocumulus, fog/stratus, cumulus, cumulonimbus, and nimbostratus.

3. Evaluation of the Classifier

One of the main problems in developing a cloud classifier is the lack of suitable evaluation material. Without systematic evaluation the performance of the classifier is not objectively determined. Experienced meteorologists may consider the classified images and tell their opinion about the classification. This is a very tedious and time-consuming procedure.

The availability of the systematic ground observations led to another evaluation scheme. The weather stations around the world collect weather information and store it as SYNOP records. The SYNOP standard is same throughout the world and it consists of a set of codes describing recent and current weather. However, the SYNOP observations differ totally from the satellite observations. This means that a perfect match between the observations is rare.

Five ground stations were selected to be used in evaluation. They were located in southern Finland, namely at the airports in Helsinki, Tampere, Maarianhamina, Kuopio, and Vaasa. These were selected in order to get the most reliable observation data. The observations in these stations are made every three hours starting at midnight. The classification is based on the five channel AVHRR data with resolution of 1.1 km at the satellite subpoint. The AVHRR data consist of visible, near-infrared, and infrared channels. The brightness temperature resolution is 0.5 K.

3.1. Comparisons

The SYNOP observations provide a covering collection of different information concerning for example the types of clouds. However, the comparison between the SYNOP observations and the satellite observations is not straightforward. The visible and infrared radiation measured by the AVHRR cannot penetrate most clouds. Because the satellite view of the earth is a top-down view, the cloud layers and surfaces under the top cloud layer cannot be observed. The SYNOP observations suffer from the quite similar problem, higher cloud layers cannot be seen through the low-level clouds. The SYNOP observations and the satellite observations are thus limited in the amount of information they contain about the vertical distribution of clouds. Another problem arises from the resolution difference between the SYNOP observations and the satellite observations. The resolution of a satellite image is too coarse to detect small details. Small clouds (the diameter less than 2 km) and small cracks in a cloud cover cannot be detected.

Analysis of the Observation Area The observation area around the ground stations was approximated by a circle with a constant radius. Several reasons affected the choice of the radius. The observer's field of view is a funnel-shaped view. In low cloud conditions the field of view is smaller than in high cloud or clear conditions. Furthermore, the field of view can be limited by landscape, and the treatment of field of view may vary depending on the observer. The effect of winds must also be considered. It is proportional to the time difference between the ground observations and the satellite observations.

Maybe a true radius of a field of view would be around 10 to 15 kilometers. A study applied to 29 summertime images showed that a radius of 10 or 15 km was too small. Based on the study the radius of 25 km was decided to be used in continental stations, whereas the radius of 30 km was set for Vaasa and Maarianhamina. No adjustment was made according to different landscape conditions in different stations.

3.2. Cloud Detection

The overall cloudiness (SYNOP code N) is reported in octas so that clear sky corresponds to code 0 and cloud-covered sky to code 8. Code 9 is used only in situations where the cloudiness cannot be determined. The cloud screening procedure was made inside the observation circle. The amount of detected cloudiness was then mapped to the octal scale so that the value would be comparable with the SYNOP observation (Table 2). The mapping is slightly nonlinear.

Amount of cloudy pixels (upper limit) [%]	0	18.5	31.0	43.5	56.0	68.5	81.0	99.5	100
Code	0	1	2	3	4	5	6	7	8

Table 2: Mapping of the satellite-detected cloudiness to octal scale.

In addition to the problems discussed earlier there exist a couple more problems concerning the comparisons. The time of a satellite pass differs from the time of a ground observation. Only the satellite passes within 30 minutes of the ground observations are taken into account. It is difficult to determine the amount of cloudiness from the ground, i.e., it is hard to say how many octas on the sky is covered by clouds. Furthermore, the ground observations may differ depending on the observer.

The comparisons were made for all seasons from autumn 1991 to autumn 1993. Both the day and night situations are included in comparison. The results are depicted in Table 3. The total classification rate can be considered sufficiently good. The classification results with zero tolerance ($\Delta = 0$) cannot be considered as a comparable measure of the quality of the classifier. There are too many weak links in the decision scheme starting from a ground observer and ending to cloud movements. The possibility that the ground observation and the satellite observation will match is a small one. The more reliable results are obtained when using tolerance values of one and two octas between the SYNOP observation and the satellite observation. The results are only 4-8% better with the tolerance value $\Delta = 2$ than with $\Delta = 1$. The wintertime results are remarkably poor. The improvement of wintertime classification is the main objective for future work.

3.3. Rain Detection

The SYNOP code ww for current weather was found to be useful in rain detection. Nimbostratus (Ns) and cumulonimbus (Cb) clouds have high correlation with

Season	*Number of comparisons*	*Classification results* $\Delta = 0$	$\Delta = 1$	$\Delta = 2$
Summer	153	37%	82%	89%
Autumn	295	39%	76%	87%
Winter	295	19%	53%	66%
Spring	395	38%	71%	79%
Total	1138	33%	71%	80%

Table 3: Classification results of overall cloudiness. Δ stands for the maximum tolerance in octas between the SYNOP observation and the satellite observation.

occurred rain. Cumulus clouds (*Cu*) and thin cirrus clouds over land or sea (*Ci1*) indicate little correlation with occurred rain.

A simple thresholding test was used in detecting rain in satellite observations. If the amount of nimbostratus clouds was more than 5% inside the observation circle (in classified satellite image), it was raining. In comparing 153 summer and 65 autumn observations, the following results were reported. 64% of the occurred rain was detected, 87% of the fair weather was detected, and 82% of all the observations was classified correctly. Rain was reported in 23% of the observations.

Fog is also reported with code ww, so the test was further expanded to detect fog as well. First the rain was detected with threshold $Ns > 5\%$, and second the fair weather was further divided to fog and sunny weather with threshold $Fog > 1\%$. This test was only made with summertime samples. The results are depicted in Table 4.

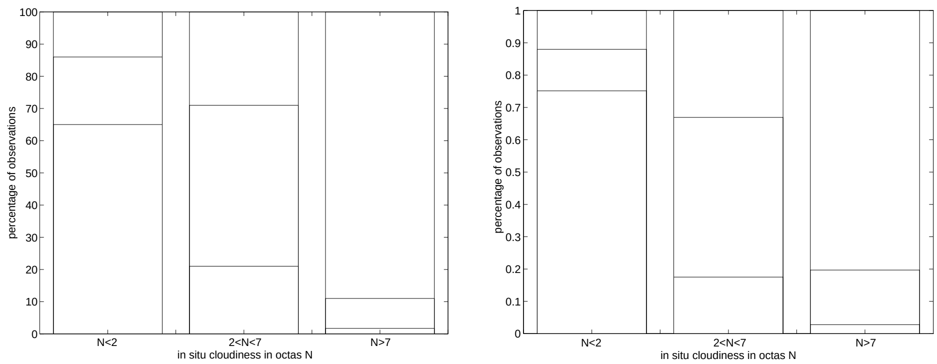
Current weather ww		*Classification results*
1–9; 13–39	Fair	77%
10–12; 40–49	Fog	50%
50–99	Rain	69%

Table 4: Classification results of current weather. Summer samples only.

4. Discussion

The overall cloudiness has also been compared by Karlsson [8] [9] and by Derrien et al. [4]. Karlsson compared approximately 2900 satellite images to the ground observations. The radius of the observation circle was 15 km, 20 km, and 30 km for low clouds, middle clouds, and high clouds, respectively. The total fractional cloud cover was determined by an equation the coefficients of which were determined by experiments. The used classifier was based on different threshold tests that varied with a sun elevation, a satellite, and a season. The total number of empirically determined thresholds was around 20.000. The total comparison result (including all four seasons, the day and night situations) with the tolerance of 2 octas was 81.4%. The results for summer conditions (15 May - 15 Sep) and for winter conditions (15 Nov - 15 Mar) were 83.6% and 78.1%, respectively. The obtained results are comparable with our results (total 80%, summer 89%, and winter 66%).

The cloud detection scheme introduced by Derrien et al. has been on operational use since February 1990. The scheme is based on thresholds and it is applied to the NOAA-11 AVHRR imagery. The thresholds have been empirically determined. They are applicable over Western Europe. The daytime satellite-derived cloudiness has been compared from April 1990 to February 1992 with ground observations. The observation area was approximated with a 20 km x 20 km square. The format of the announced results is different from ours. In Figure 3a are the results from the French experiment (for daytime images only) and in Figure 3b are our results formatted in the same way.



(a) (b)

Figure 3: (a) French results (for daytime images only) and (b) our results on comparison of overall cloudiness. Each bar is divided in three areas that are (from bottom to top) satellite-based cloudiness $< 20\%$, satellite-based cloudiness $> 20\% \wedge < 80\%$, satellite-based cloudiness $> 80\%$.

There is an evident need for numerical and objective methods in evaluating the classifier. The quality of the classifier cannot be determined without these methods. The subjective evaluation method used earlier gave us some information on limitations and problems met in the classifier, but long-term decisions, especially decisions concerning the quality of different classifier versions, were hard to make. The evaluation was made by looking ten or twenty classified images and making notes about the problems that were found. After some minor changes in classification the same procedure was repeated. There were three main problems in this method that suggested a better and automatic evaluation method: visual evaluation of the classified images was slow (maybe couple of images per hour), results were hard to compare, and different situations were impossible to cover with ten or twenty images. After the testing of different evaluation methods the comparison of overall cloudiness between ground observations and satellite observations was found to be a suitable method.

5. Conclusions

The multispectral cloud classifier is implemented to automate the processing of satellite images. The classifier is fully automatic, and it can be adapted to changing situations with new examples. The quality of the classifier has been verified with hundreds of images. In addition to the visual inspection, the automatic evaluation scheme is developed. The comparisons with other published results show that the classifier is working relatively well. The improvement of wintertime classification is the main objective for future work.

Acknowledgement

The authors wish to thank the Finnish Meteorological Institute for the interesting application problem and the Technology Development Centre of Finland for financial support (TEKES's grant 4005/94).

References

[1] R. K. Anderson and N. F. Veltishchev. The use of satellite pictures in weather analysis and forecasting. Technical Report WMO-No. 333, World Meteorological Organization, 1973.

[2] J. A. Benediktsson, P. H. Swain, and O. K. Ersoy. Neural Network Approaches Versus Statistical Methods in Classification of Multisource Remote Sensing Data. *IEEE Transactions on Geoscience and Remote Sensing*, 28(4):540–552, July 1990.

[3] M. Debois, G. Seze, and G. Szejwach. Automatic classification of clouds on meteosat imagery: application to high-level clouds. *Journal of Applied Meteorology*, 21:401–412, 1982.

[4] M. Derrien, B. Farki, L. Harang, H. LeGléau, A. Noyalet, D. Pochic, and A. Sairouni. Automatic Cloud Detection Applied to NOAA-11/AVHRR Imagery. *Remote Sensing of Environment*, 46(3):246–267, 1993.

[5] E. E. Ebert. A Pattern Recognition Technique for Distinguishing Surface and Cloud Types in the Polar Regions. *Journal of Climate and Applied Meteorology*, 26:1412–1427, 1987.

[6] E. E. Ebert. Analysis of Polar Clouds from Satellite Imagery Using Pattern Recognition and a Statistical Cloud Analysis Scheme. *Journal of Applied Meteorology*, 28:382–399, May 1989.

[7] J. Iivarinen, K. Valkealahti, A. Visa, and O. Simula. Feature Selection with Self-Organizing Feature Map. In *International Conference on Artificial Neural Networks*, Sorrento, Italy, May 26-29 1994.

[8] K.-G. Karlsson. Use of AVHRR-derived cloudiness and precipitation in operational weather forecasting and in model studies. In *Proceedings of the Central Symposium of the 'International Space Year' Conference*, pages 211–216, Munich, Germany, March 30 - April 4 1992.

[9] K.-G. Karlsson. Comparison of Operational AVHRR-Based Cloud Analyses with Surface Observations. In *Proceedings of 6th AVHRR Data Users' Meeting*, pages 223–229, Belgirate, Italy, June 29 - July 7 1993.

[10] K.-G. Karlsson and E. Liljas. The SMHI model for cloud and precipitation analysis from multispectral AVHRR data. Technical Report 10, Swedish Meteorological and Hydrological Institute, August 1990.

[11] J. Key. Cloud Cover Analysis with Arctic Advanced Very High Resolution Radiometer Data, 2. Classification with Spectral and Textural Measures. *Journal of Geophysical Research*, 95(D6):7661–7675, May 20 1990.

[12] J. Key and R. G. Barry. Cloud Cover Analysis with Arctic AVHRR Data, 1. Cloud Detection. *Journal of Geophysical Research*, 94(D15):18521–18535, December 20 1989.

[13] J. Key, J. A. Maslanik, and A. J. Schweiger. Classification of Merged AVHRR and SMMR Arctic Data with Neural Networks. *Photogrammetric Engineering & Remote Sensing*, 55(9):1331–1338, September 1989.

[14] J. Kittler and D. Pairman. Contextual Pattern Recognition Applied to Cloud Detection and Identification. *IEEE Transactions on Geoscience and Remote Sensing*, GE-23:855–863, November 1985.

[15] J. Lee, R. C. Weger, S. K. Sengupta, and R. M. Welch. A Neural Network Approach to Cloud Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 28(5):846–855, September 1990.

[16] D. W. Reynolds and T. H. V. Haar. A bi-spectral method for cloud parameter determination. *Monthly Weather Review*, 105:446–457, 1977.

[17] R. W. Saunders and K. T. Kriebel. An improved method for detecting clear sky and cloudy radiances from AVHRR data. *International Journal of Remote Sensing*, 9(1):123–150, 1988.

[18] O. Simula, A. Visa, and K. Valkealahti. Operational Cloud Classifier Based On the Topological Feature Map. In *Proceedings of the International Conference on Artificial Neural Networks*, pages 899–902, Amsterdam, September 13-16 1993.

[19] C. W. Therrien. *Decision, Estimation and Classification.* Wiley, New York, 1989.

[20] A. Visa, K. Valkealahti, J. Iivarinen, and O. Simula. Experiences from Operational Cloud Classifier Based on Self-Organising Map. In *SPIE Vol. 2243 Applications of Artificial Neural Networks V*, volume 2243, pages 484–495, Orlando, Florida, April 5-8 1994.

[21] A. Visa, K. Valkealahti, and O. Simula. Cloud Detection Based on Texture Segmentation by Neural Network Methods. In *Proceedings of IEEE International Joint Conference on Neural Networks*, volume 2, pages 1001–1006, Singapore, November 18-21 1991.

[22] R. M. Welch, S. K. Sengupta, A. K. Goroch, P. Rapindra, N. Rangaraj, and M. S. Navar. Polar Cloud and Surface Classification Using AVHRR Imagery: An Intercomparison of Methods. *Journal of Applied Meteorology*, 31:405–419, May 1992.