# A Multichannel Temporally Adaptive System for Continuous Cloud Classification From Satellite Imagery

Kishor Saitwal, *Student Member, IEEE*, Mahmood R. Azimi-Sadjadi, and Donald Reinke

*Abstract*—A two-channel temporal updating system is presented, which accounts for feature changes in the visible and infrared satellite images. The system uses two probabilistic neural network classifiers and a context-based predictor to perform continuous cloud classification during the day and night. Test results for 27 h of continuous classification and updating are presented on a sequence of Geostationary Operational Environmental Satellite 8 images. Further test results of the system on two new sets of data with 1–2 weeks time difference are also presented that show the potential of this system as an operational continuous cloud classification system.

*Index Terms*—Cloud classification, multispectral satellite imaging, probabilistic neural networks.

## I. INTRODUCTION

THE GEOSTATIONARY Operational Environmental Satellite 8 (GOES 8) provides frequent, high-resolution, visible, and infrared (IR) images over large areas [1]. It plays an important role in weather analysis and forecasting [2], both of which require accurate cloud detection and classification. However, due to the large volume of data received every day, an automated cloud classification method is needed in order to successfully classify sequences of images that are routinely available at a frequency of 1 h (locally, most geostationary systems will transmit data at 15–30-min frequencies; however, the global data that are used in this study are available at a common frequency of 1 h).

In processing sequences of GOES 8 satellite images, the issue of temporal feature changes is an important consideration. Owing to sun angle changes, the same cloud may look different in the visible image at different times of the day. On the other hand, because of diurnal changes of the temperature of the land and water areas, the features of cloud/no-clouds change in the IR channel. Although these feature variations may not be very prominent in the short-term, their effects over a longer period of time cannot be ignored. Thus, the performance of a fixed

classifier can drastically deteriorate as a consequence of these temporal variations in the feature space. This problem was first addressed in [3] where a temporal updating mechanism using probabilistic neural networks (PNNs) [4], [5] was proposed. Gaussian mixture models and maximum likelihood [6] were adopted to compute the parameters of the classifiers iteratively as new images are received. The system uses the visible and IR (Channel 4) data from GOES 8 satellite imagery to classify the images into ten different cloud/no-cloud classes. Promising results were reported on a sequence of GOES 8 images. However, the major shortcoming of this system is its inability to provide continuous cloud classification when the visible channel is not available during the nighttime. This problem also disrupts the temporal updating process for the next day due to discontinuity of the data.

To the best of our knowledge, there is no temporally adaptable continuous cloud classification system that can operate 24-h/day. This letter presents a new version of the temporal updating system in [3] that circumvents the problems of the previous version and provides a basis for an operational cloud classification system. The modified system offers a new way of combining the information in the visible and IR channels. The cloud images are classified into five different classes, namely 1) Land, 2) Water, 3) Low-level, 4) Middle-level, and 5) High-level clouds. During the daytime, two separate PNN classifiers are used for the two channels and their results are fused together to give the final classification results. During the nighttime, however, the PNN classifier for the IR channel alone can provide cloud classification for Middle-level and High-level clouds based upon the temperature features. The system can be applied to any other geostationary satellite imaging instrument (where data are available at a 1-h interval or better) and can easily be augmented to include other channels to aid in cloud classification. The classification accuracies of the system during the daytime are computed and compared with those of the previous system in certain areas labeled by expert meteorologists. However, during the nighttime, since Low-level clouds are very difficult to label, the system performance is mainly evaluated for the other four classes. The robustness of the overall system to variation in the training datasets has also been studied.

The organization of this letter is as follows. Section II gives the description of the proposed multichannel system. Test results of this new system, with and without a block-partitioning scheme are presented and discussed in Section III. Finally, Section IV provides concluding remarks on this work.
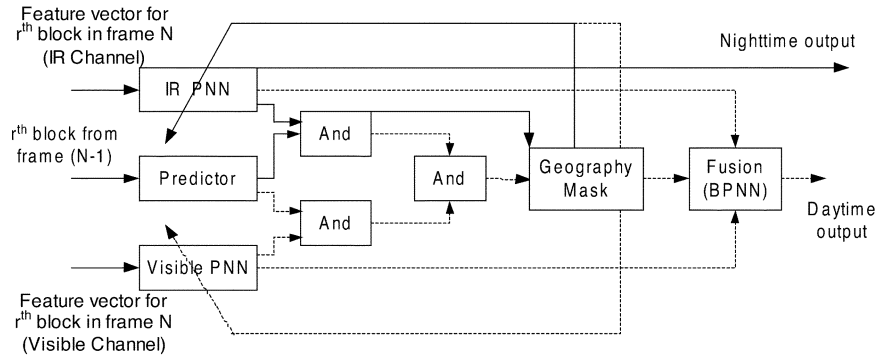
Fig. 1. Schematic diagram of proposed multichannel system. Solid lines indicate the infrared-only path, while dashed lines indicate the additional path that is used when visible data are available.

## II. MULTICHANNEL TEMPORAL UPDATING SYSTEM

The new multichannel temporal updating scheme, is shown in Fig. 1. As can be seen, the system is composed of several key components namely; two temporally adaptive PNN classifiers, a contextual-based predictor, comparators, a geographical mask and a final fusion system. These components and their functions are briefly described in the following subsections.

### A. Temporally Adaptive PNNs

PNN was introduced [5] to implement the Parzen nonparametric probability density function (PDF) estimation method. To overcome the structural and computational limitations (during the testing phase) of this network, Streit and Luginbuhl [6] introduced a modified version of the PNN that substantially reduces the number of neurons in the recognition layer by using Gaussian mixture models and the Expectation Maximization algorithm to estimate the network parameters. This type of PNN [3] is adopted in this letter.

Assume that the previous images up to frame $n - 1$ are correctly classified and the weights of the PNNs are updated to frame $n-1$. Now the new frame $n$, consisting of visible (Channel 1) and IR (Channel 4) images, arrives. These images are partitioned into nonoverlapping blocks (typically of size $8 \times 8$ corresponding to a region of size $32 \times 32$ km) and the spectral and textural features of each block are extracted [4] using singular value decomposition (SVD). The Fisher criterion is then used along with the forward floating sequential selection [7] to select three features with high discriminatory power from each channel. These features are then applied to the PNN classifier. Owing to the fact that the features of the clouds and background (land/water) do not change abruptly, there is rich temporal class contextual information between adjacent frames, which can be utilized to design a predictor (see Section II-B) that extrapolates its decision and provides classification results for the blocks of the current frame. The predictor's classification results can be used as "pseudotruth" for updating the PNNs.

The temporal updating of PNN is an on-line process that estimates the parameters of the PNN for every new frame so that it can more accurately represent the distribution of the temporally changed feature space. This process should satisfy two key requirements: a) the stability of those previously established categories and b) the flexibility to accommodate temporal changes of the data and new classes. The training samples for each PNN

at frame $n$ consist of two subsets, viz., $X_1$ and $X_2$, where $X_1$ includes all of those samples for which the class label is assumed to be known, while all the samples of unknown types form $X_2$. These subsets are determined by comparing the initial (nonupdated) results of the PNNs with those of the context-based predictor. That is, the classifiers perform preliminary classification of the data using the parameters that were updated based upon the data of the previous frame (1 h prior). The classification result of the last frame is also used as input to the predictor, which makes a prediction of the class of the current frame by using the contextual information in two consecutive frames. The feature vectors of the visible or IR channels corresponding to a particular block location are put in subset $X_1$ when the classification labels of the corresponding nonupdated PNN and the predictor agree, otherwise the sample is included in $X_2$. Since the subset $X_1$ contains labeled data with a relatively high level of confidence (pseudotruth), supervised learning is used to fine-tune the parameters of those Gaussian components that correspond to the selected class. This ensures the stability of the established classes. On the other hand, for subset $X_2$, since the class labels are unknown, an unsupervised learning is used to account for feature changes and provide flexibility needed in these situations. In the latter case, the parameters of all the Gaussian mixtures used to model the class distributions are slightly updated. Both learning mechanisms are implemented using an Expectation–Maximization (EM)-based updating strategy [3]. After the updating is completed, the block is reclassified by the updated PNN to give the final classification results. This process is repeated whenever a new frame arrives.

As can be seen in Fig. 1, the two PNNs share the same context-based predictor, the output of which is separately compared with those of the two PNNs, and the comparison results are used to update the parameters of the two PNNs separately. The PNN for the visible channel is trained for three different classes, viz., Land, Water, and Clouds, as this channel does not provide height information essential for differentiating all the five classes. The PNN for the IR channel, on the other hand, is trained to distinguish all the five different classes.

During the daytime, both the PNNs provide their separate classification results. Thus, to get the final classification results for each image block, one needs to fuse these results together. A multilayer back-propagation neural network (BPNN) was used to perform this decision-level fusion based upon the results of the two PNNs. The PNN outputs for the training data are used

for training of this fusion system, which combines the PNNs' outputs to classify the image blocks into five different classes. The fusion system output is used as a memory for the predictor, which predicts the class label for each block in the next frame on the basis of the results of the current frame. During the nighttime, the visible PNN and fusion system indicated by the dotted lines in Fig. 1 are shut-off and the IR PNN takes over to give the five-class classification results.

### B. Context-Based Predictor

The prediction block in Fig. 1 is designed to provide an initial guess of the class of the current data based on the previous classification results. In a sequence of satellite imagery, certain positions in the image tend to belong to the same class in the adjacent frame (temporal class dependency context). Additionally, most of the classes are likely to cover a relatively large area in one frame instead of appearing in isolated blocks (spatial class dependency context). Thus, there is a rich class dependency context in GOES imagery data, as the time interval between frames is relatively short (i.e., 1 h).

The spatial-temporal class contextual information can be modeled by a Markov chain [3]. In order to simplify the modeling, a first-order Markov process is considered where the class of the current frame is solely determined based upon those in the previous one. Moreover, for a block $(i, j)$ in the current frame $n$, its spatial temporal neighborhood in frame $n - 1$ is assumed to contain nine blocks $(i - n, j - m), m, n \in [-1, 1]$. Two underlying Markov chains are assumed in [3]. The first one describes the spatial movement such as clouds drifting, while the second one describes the possible class changes. The class transition Markov chain is necessary; otherwise, the current block will always be one of the types that appeared in its spatial temporal neighborhood. This is not the case in real situations, as certain clouds may be generated, terminated, or evolved to other classes. Using this model, the class conditional probabilities are computed and then compared with those generated using the PNNs in the comparison block.

### C. Geographical Mask

In addition to the multichannel nature of this new system, there is yet another major difference in the updating of the PNNs in comparison with that of the previous system [3]. Geographical and topographical information (Land/Water mask) is embedded into the system to improve the classification accuracy and confidence of the updating process. This process is accomplished by exploiting the geography map when the output of the final AND operation is either Land or Water class. In this case, the geography map result is given the precedence over the AND results of the two channels. For instance, if the final AND operation labels a particular block as Land whereas the label according to the geography map should be Water, then this block is put into $X_1$ subset as a Water sample for the supervised updating of both PNNs. Since such blocks are labeled with 100% confidence, their inclusion in the supervised learning clearly improves the accuracy of the temporal updating process. This process is justified because Land and Water are more likely to be misclassified to each other owing to small temperature and reflectivity differences.
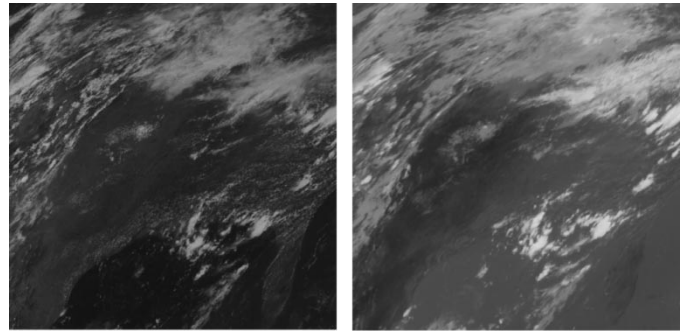


Fig. 2. Image pair at 18 UTC on July 23, 1998.

### III. TEST RESULTS

#### A. Daytime Operation

The performance of the proposed multichannel system was examined and compared with that of the original system [3] on a sequence of GOES 8 satellite images. One typical image pair, obtained at 18 UTC, July 23, 1998 is shown in Fig. 2. These images, of size $512 \times 512$ pixels (spatial resolution of 5 km/pixel), cover the Midwest and most of the Eastern part of the United States, extending from the Rocky Mountains to the Atlantic coast. The images cover mountains, plains, lakes, and coastal areas where clouds have some specific features that are tied to topography. Florida is located in the lower right, with the Gulf of Mexico in the lower center part of the image. These sequences are of particular interest because of the presence of a variety of cloud types. Certain high confidence cloud/no-cloud regions were identified and labeled for the initial training, validation, and testing of the systems based on visual inspection and other related information used by the expert meteorologists. Two experts labeled the daytime images and only those areas where the labeling results agreed were used for initial training, validation, and classification accuracy determination on the testing set.

The selected SVD feature vectors $(3 \times 1)$ for each $8 \times 8$ block of each channel were applied to the system. The goal was to classify each block into five classes as mentioned before. The labeled images for the visible and IR channels at 15 UTC from July 23 and 24, 1998 were used to get the training and validation dataset for the PNNs. Half of these labeled blocks were randomly chosen as the "training data" set, while the remaining blocks were used as the "validation data" set. The validation dataset is used to determine the optimally trained PNNs. That is, the training and validation processes were repeated for ten trials, and the classifiers that produced the best performance were selected in the multichannel system to classify 27 h of continuous data. A two-layer fusion BPNN with the structure 8-30-5 (i.e., eight inputs, 30 hidden-layer neurons, and five output neurons) was trained using the normalized PNNs outputs for the training set. The normalization was done in order to produce the *a posteriori* class condition probabilities for each class. The number of epochs used for the training of this system was 500, while the error goal was kept at $10^{-2}$. For given training and validation datasets, this BPNN structure was also trained and validated for ten different trials. The BPNN with best results on the validation dataset was used.

TABLE I
CONFUSION MATRIX-PREVIOUS SYSTEM

|  | L | W | LLC | MLC | HLC |
|---|---|---|---|---|---|
| L | 107 | 0 | 236 | 6 | 0 |
| W | 0 | 14 | 59 | 0 | 0 |
| LLC | 0 | 0 | 240 | 5 | 0 |
| MLC | 0 | 0 | 4 | 494 | 6 |
| HLC | 0 | 0 | 0 | 79 | 54 |

TABLE II
CONFUSION MATRIX-PROPOSED SYSTEM

|  | L | W | LLC | MLC | HLC |
|---|---|---|---|---|---|
| L | 327 | 1 | 19 | 2 | 0 |
| W | 0 | 68 | 5 | 0 | 0 |
| LLC | 0 | 0 | 239 | 6 | 0 |
| MLC | 0 | 0 | 9 | 405 | 90 |
| HLC | 0 | 0 | 1 | 16 | 116 |

The performance of the proposed system was then compared with that of the previous system [3] on the testing dataset which included the expert labeled areas in all the other images from 16–22 UTC. Table I presents the confusion matrix of the previous system for the image pair shown in Fig. 2 for July 23 at 18 UTC. The diagonal elements indicate the number of blocks that are correctly classified, while the off-diagonal elements represent the number of misclassified blocks for each class. Note that the numbers on each row correspond to the classification results of the system for a particular class. The overall correct classification rate is around 70%. It can be observed from this confusion matrix that there are a lot of misclassifications particularly for Land, Water, and High-level clouds. Table II, on the other hand, presents the confusion matrix of the proposed two-channel system on the same image pair. The overall correct classification rate is around 89%, which indicates substantial improvement in the performance. Clearly, the classification accuracies for several classes except Middle-level clouds are increased substantially. The misclassification is mainly caused by the IR PNN, since the IR channel textural features for Middle-level clouds are quite similar to those of High-level clouds. In the single-channel system, combining both the visible and IR channel features resulted in a better performance on the Middle-level clouds while misclassifying thin High-level clouds as Middle-level clouds. One reason for the improved overall performance might be due to the fact that decomposing the system into two separate channels removes the two-channel feature interdependencies in the previous scheme [3]. Additionally, fusion of the two separate decisions can correct for some of the classification errors in either of the channels.

The gray-coded classified images that identify different cloud and no-cloud areas for both systems on the same image pair are shown in Fig. 3. Visual inspection of Fig. 3(a) and (b) reveals that there is substantial improvement in the results of the two-channel system [Fig. 3(b)] as compared to those of the previous system [Fig. 3(a)]. Fig. 3(c) gives the corresponding expert labeled image with the color map. Note that only those areas in which the labels assigned by the two meteorologists agreed
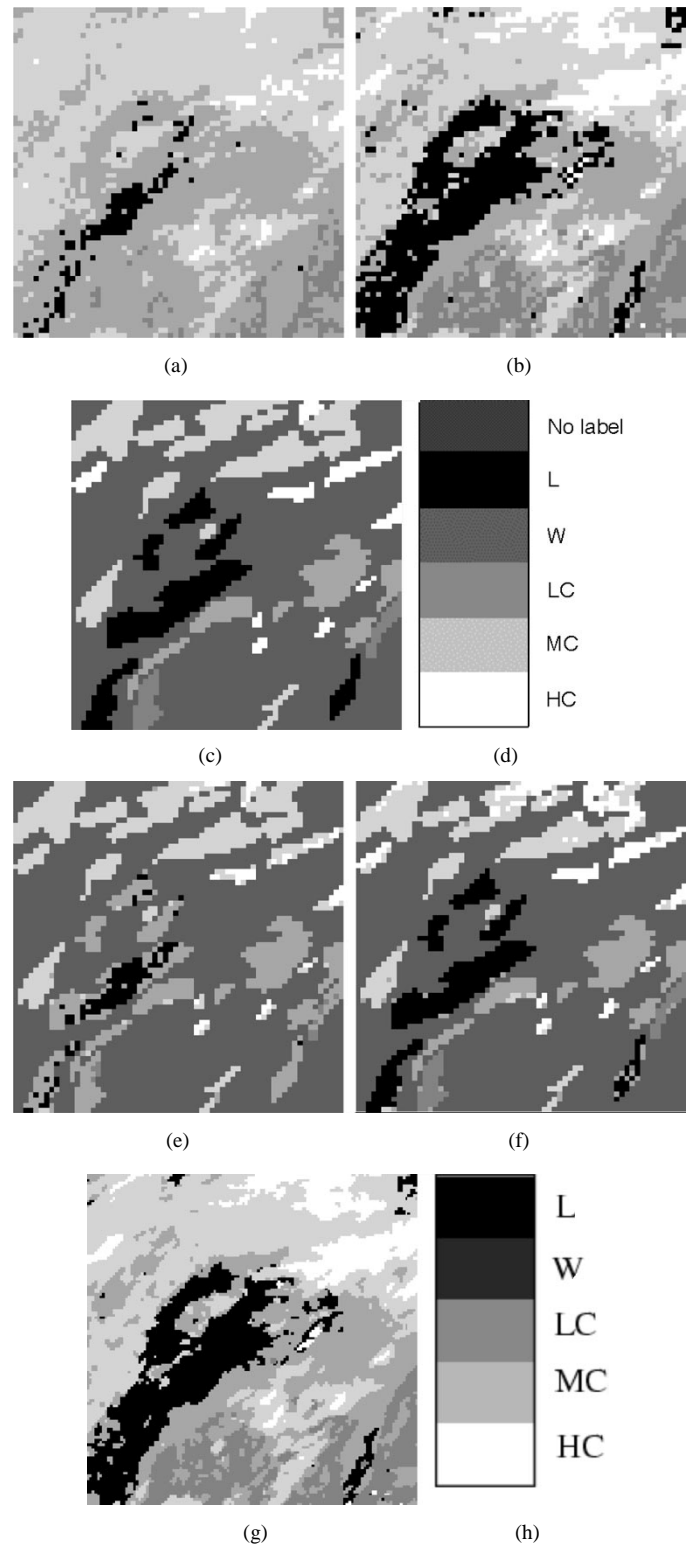


Fig. 3. Classification results of the two systems for image pair in Fig. 2 (for 1-h interval imagery). (a) Original system results. (b) Two-channel system results. (c) Expert labeled image. (d) Gray map. (e) Masked (original system) results. (f) Masked (two-channel system) results. (g) Block partitioning results. (h) Gray map.

were selected. Fig. 3(e) and (f) gives the masked results of both systems against the expert labeled areas. As can be observed, the results of the new system much more closely match those of the expert labeled areas than those of the previous system.
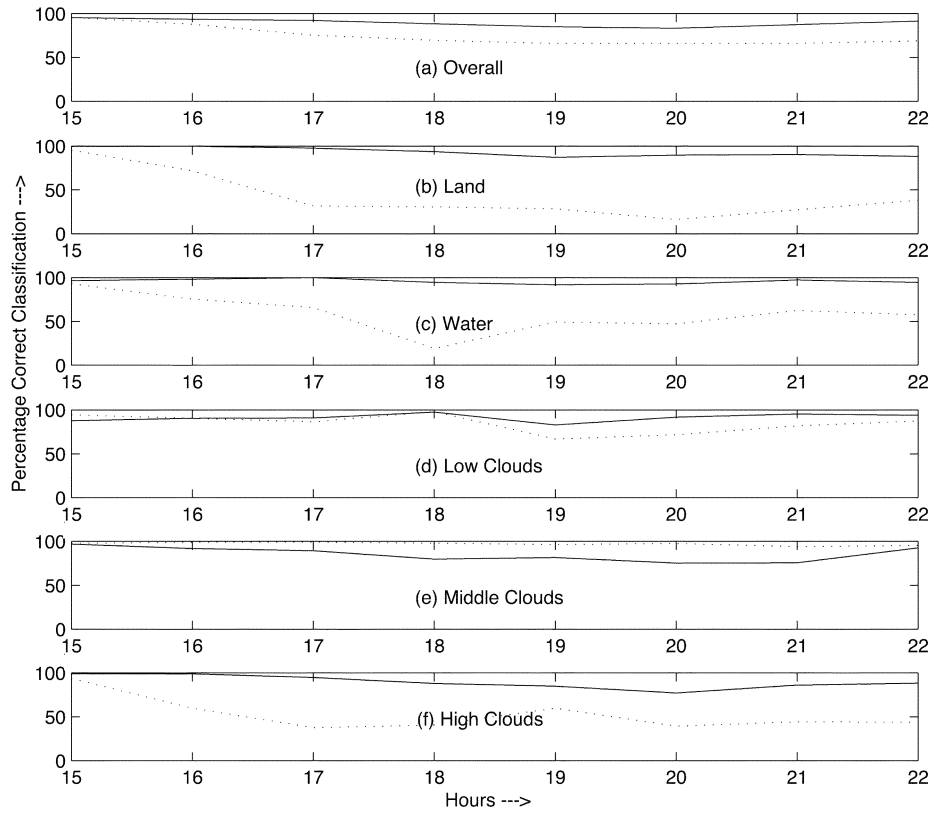
Fig. 4.   Two-channel system daytime continuous operation results.

The plot in Fig. 4(a) shows the percentage of correct classification accuracy $(P_{cc})$ for the overall system output versus time (in hours) for 8 h (15–22 UTC, July 23) of continuous temporal updating. The solid line corresponds to the multichannel system, while the dotted line is used for the previous system. Clearly, one can see that the temporal updating in the new system is more reliable and consistent than the previous system. The plots in Fig. 4(b)–(f) show the comparison between the two systems for the five classes separately. These plots, again, reveal that the proposed system has much better accuracy for all the classes except the Middle-level clouds due to the reason explained earlier. During daytime operation, the overall accuracy rate varies from 84% to 95%. The overall mean correct classification over a period of 8 h of continuous temporal updating is 90%.

### B. Block Partitioning

The classification results of the proposed scheme in Fig. 3, however, reveals that there are a large number of isolated blocks that are misclassified. In particular, "cloud holes," which cover approximately 5% of the image, will not be correctly classified. To solve this problem, a block partitioning scheme similar to that in [8] was adopted here. This algorithm makes use of the classification results of $8 \times 8$ overlapping blocks to determine the label of the $4 \times 4$ subblocks. Each $4 \times 4$ subblock is classified four times (twice for boundary blocks). In this way, resolution is improved by a factor of four without the need to redesign the training and updating mechanisms. The probability vectors for each individual overlapping block are computed to determine the corresponding classification labels. This probability-based decision for the middle $4 \times 4$ subblock, $r(i,j)_{(4\times4)}$, can be expressed as

$$
c\left(r(i,j)_{(4\times4)}\right)
$$
$$
= \arg\max_{k} \sum_{m=0}^{1} \sum_{n=0}^{1} P\left(c_k \,|\, r(i-m, j-n)_{(8\times8)}\right) \quad (1)
$$

where $k = 1, 2, \ldots, 5$, i.e., one of the five classes and $c(r(i,j)_{(4\times4)})$ is the class label of block $r(i,j)_{(4\times4)}$. Note that subscripts $4 \times 4$ and $8 \times 8$ indicate the corresponding block sizes. The probabilities $P(c_k \,|\, r(i-m, j-n)_{(8\times8)})$ correspond to the *a posteriori* probabilities of the class membership for the participating overlapping $8 \times 8$ blocks. As a result, each $4 \times 4$ subblock is given the class label, which has the maximum total probability among the classes for the neighboring overlapping blocks.

Fig. 3(g) shows the classification results of the proposed multichannel system with the block partitioning algorithm for the image pair in Fig. 2. The overall correct classification rate for this result was found to be around 90%. This result clearly shows that almost all the isolated blocks are removed giving much smoother gray-coded areas and boundaries. Additionally, the resolution is substantially improved without considerable increase in the computational cost. This is accomplished because the updating is done based upon the classification results of $8 \times 8$ nonoverlapping blocks rather than the partitioned $4 \times 4$ subblocks. The computation time required for the block partitioning method was about 37.54 s, while the two-channel system alone took 18.3 s to generate the gray-coded image for one pair

of images. The time difference is mainly attributed to the increase in the number of blocks that had to be classified and the voting mechanism.

To show the potential of the proposed two-channel system for real operational cloud classification systems, we tested the temporal updating mechanism on two new sequences of images on July 5 (two weeks before) and July 31 (one week after). For the sake of simplicity, the temporal updating starts with the same previously initially trained system at 15 UTC and performs continuous classification and parameter adaptation for several consecutive hours. Fig. 5(a) and (b) and (e) and (f) show two pairs of visible and IR channel images at 18 UTC on July 5 and 31, respectively. Fig. 5(c) and (g) shows the corresponding gray-coded classification results of these images. Close evaluation of these results show the effectiveness of this two-channel system. It must be emphasized that much better results could have been obtained if the PNN classifiers were retrained at initial frame (15 UTC) every few days.

## C. Nighttime Operation

During the nighttime, the visible PNN and BPNN fusion system indicated by the dotted lines in Fig. 1 are shut-off and the IR PNN alone takes over to give the five-class classification results. When the IR PNN classifies a block as Land or Water, the geography mask is used to make the final decision. Due to unavailability of the visible channel data during the nighttime, Low-level clouds are not labeled as they are not distinguishable to the naked eye. Consequently, the experts labeled the areas into only four remaining classes for classification accuracy determination.

Since only the IR PNN makes the decision during the nighttime, it is important to study the robustness of the system to different initial training conditions. To introduce this variability in the datasets, half of the randomly picked (ten different times) samples from all the labeled blocks in the visible and IR images at 15 UTC from July 23 and 24, 1998 were used to form the training set, while the remaining half were used to form the validation set for selecting the best networks. Trained PNNs and BPNN fusion systems were then used to perform continuous daytime and nighttime updating over a period of 27 h from 15 UTC on July 23 to 17 UTC on July 24, and the classification accuracies were computed in the labeled areas (excluding the training or validation sets). The plots in Fig. 6 show the percent correct classification ($P_{cc}$) versus time (in hours) for individual classes and for ten different trials. Stars indicate the "mean" of the classifier outputs for a particular hour, while vertical lines indicate "standard deviations" from the mean values. Note that in these plots, Hour 1 corresponds to 15 UTC on July 23, 1998, while Hour 27 corresponds to 17 UTC on July 24, 1998. The overall correct classification rate for all the trials and the frames of data considered is around 91.42%. During the period between 23 UTC on July 23 and 12 UTC on July 24 the accuracies for Low-level clouds are not computed due to the lack of expert labeling. These plots show the effectiveness of this system for daytime as well as nighttime operations.

Table III gives the confusion matrix of the IR classifier alone system at 10 UTC (2:00 A.M.) on July 24. These results indicate
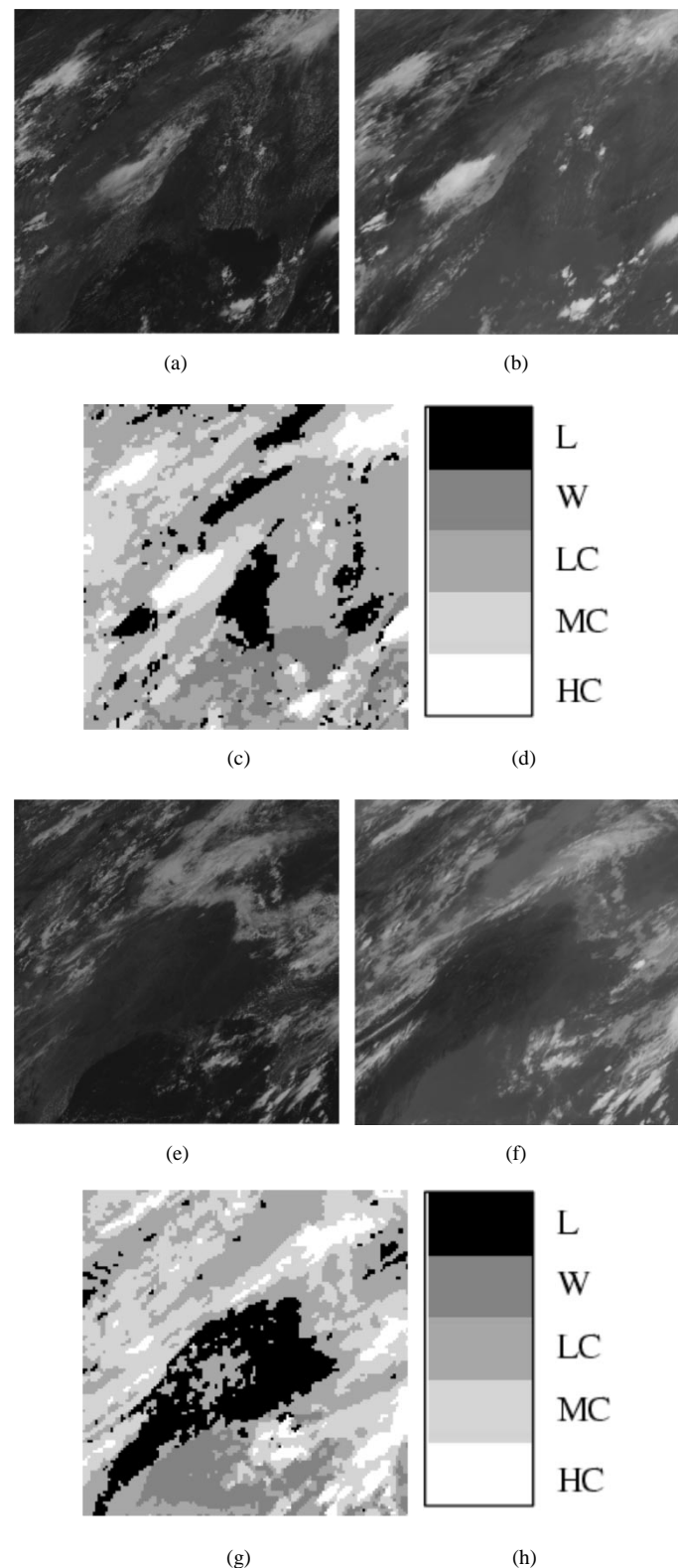


Fig. 5. Classification results of the two-channel system. (a) Visible channel image (July 5). (b) IR channel image (July 5). (c) Two-channel system results. (d) Gray map. (e) Visible channel image (July 31). (f) IR channel image (July 31). (g) Two-channel system results. (h) Gray map.

that the system has very good performance during the nighttime. Some of the Land and Water blocks were misclassified as Low-level clouds, due to the fact that they might be covered with Low-level clouds.
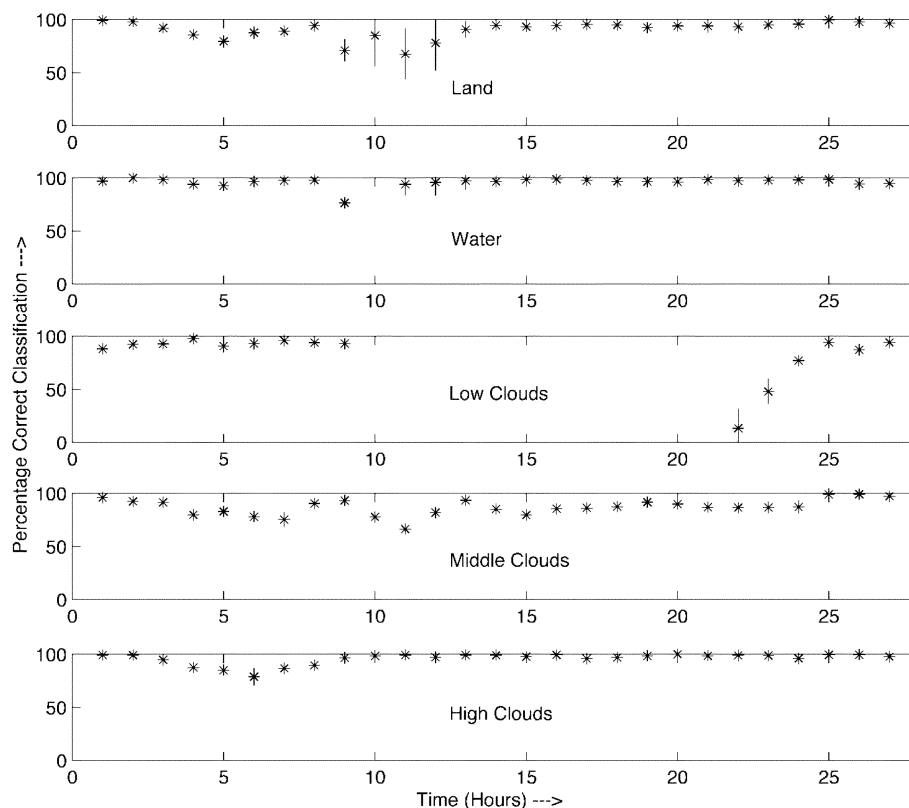
Fig. 6.    Robustness of classification accuracies over 27 h of temporal updating from 15 UTC on July 23 to 17 UTC on July 24.

TABLE  III
CONFUSION MATRIX AT 10 UTC (2:00 A.M.) ON JULY 24 (X: LABELED
DATA NOT AVAILABLE)

|      | L   | W   | LC  | MC  | HC  |
|------|-----|-----|-----|-----|-----|
| L    | 205 | 0   | 14  | 1   | 0   |
| W    | 0   | 184 | 7   | 0   | 0   |
| LC   | X   | X   | X   | X   | X   |
| MC   | 0   | 0   | 2   | 480 | 65  |
| HC   | 0   | 0   | 0   | 0   | 135 |

## IV. CONCLUSION

A multichannel temporal updating system is introduced in this letter, which makes continuous day/night processing feasible. Temporal updating takes place in each channel individually and the results are fused together using a BPNN system. During the daytime both channels participate in decision making, while during the night the decision is made solely based upon the IR channel. Geographical and topographical information is also exploited to aid the classification and updating of Land and Water classes. A block-partitioning scheme is employed to improve the performance of the proposed system and increase the resolution of the final classification result by a factor of four. This is accomplished without a drastic increase in the computational time. The daytime multichannel system gave much better results than those of the previous system, while the proposed nighttime system preserved the continuity of the classification (except for low-level clouds).

The preliminary results in this letter show the great promise of the proposed multichannel system, which provides a foundation toward building an operational cloud classification system with temporal adaptability. Surface observations may also be used not only to aid in the temporal updating but also for the purpose of evaluating the performance of the system. This system can be applied to various other satellite imaging problems that involve multichannel processing.

## REFERENCES

[1] J. J. Simpson and J. I. Gobat, "Improved cloud detection in GOES scenes over land," *Remote Sens. Environ.*, vol. 52, pp. 36–54, 1995.

[2] S. Q. Kidder and T. H. Vonder-Haar, *Satellite Meteorology: An Introduction.*   New York: Academic, 1995.

[3] B. Tian, M. R. Azimi-Sadjadi, T. H. Vonder Haar, and D. Reinke, "Temporal updating scheme for probabilistic neural network with application to satellite cloud classification," *IEEE Trans. Neural Networks*, vol. 11, pp. 903–920, July 2000.

[4] B. Tian, M. Shaikh, M. R. Azimi-Sadjadi, T. H. Vonder-Haar, and D. Reinke, "A study of cloud classification with neural networks using spectral and textural features," *IEEE Trans. Neural Networks*, vol. 10, pp. 138–151, Jan. 1999.

[5] D. F. Specht, "Probabilistic neural network," *Neural Networks*, vol. 3, pp. 109–118, 1990.

[6] R. L. Streit and T. E. Luginbuhl, "Maximum likelihood training of probabilistic neural networks," *IEEE Trans. Neural Networks*, vol. 8, pp. 114–132, Jan. 1997.

[7] A. Jain and D. Zongker, "Feature selection: Evaluation, application and small sample performance," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, pp. 153–158, Feb. 1997.

[8] P. M. Tag, R. L. Bankert, and L. R. Brody, "An AVHRR multiple cloud-type classification package," *J. Appl. Meterolol.*, vol. 39, pp. 125–134, Feb. 2000.