香港天文台
HONG KONG OBSERVATORY

Reprint 442

Pattern Recognition of Radar Echoes for Short-range Rainfall Forecast

E.S.T. Lai, P.W. Li,

C.M. Chan [1], M.C. Chu [1] & W.H. Wong [1]

Proceedings of International Conference on Pattern Recognition,

Barcelona, 3-8 September 2000

[1] *Department of Physics, The Chinese University of Hong Kong,*
Shatin, Hong Kong

Pattern Recognition of Radar Echoes for Short-range Rainfall Forecast

Edwin S.T. Lai, P.W. Li
Hong Kong Observatory, 134A Nathan Road, Kowloon, Hong Kong
stlai@hko.gcn.gov.hk
C.M. Chan, M.C. Chu, W.H. Wong
Department of Physics, The Chinese University of Hong Kong, Shatin, Hong Kong

Abstract

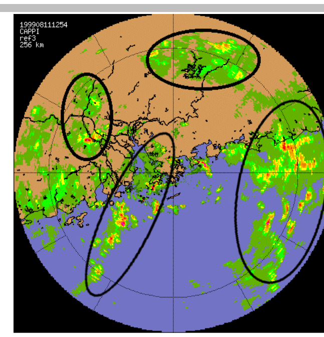
A four-layer feed-forward back-propagation Artificial Neural Network (ANN) is applied to weather radar echo maps of reflectivity data for the prediction of heavy rainfall events in the short-range of 1 to 2 hours. Inputs for the ANN are the cross correlations of statistical measures of a sequence of radar images. The ANN is trained to capture increasingly organized echo patterns that often are preludes to localized heavy rain. Results show that the ANN is able to achieve a success rate of 89% against a false alarm rate of 33%.

In parallel, a separate module utilizing Hough transform is developed to depict the lining up of echoes on the reflectivity maps. The module provides an objective analysis tool for forecasters to test the hypothesis that crossing or merging of echo lines, the so-called "X" patterns, would lead to enhanced convection at preferred locations.

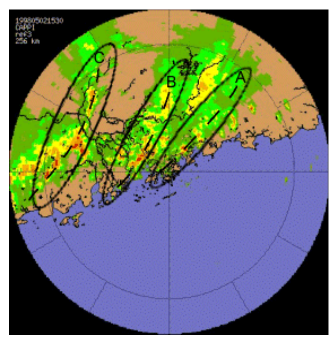
Working in tandem, the ANN helps to isolate specific sectors on the radar maps where organization is taking place so that the Hough Transform Module (HTM) can be meaningfully applied in the appropriate target areas. In turn, parameters derived from the HTM, along with the standard statistical measures, can be fed back into the ANN for further training and system enhancement in the identification of "X" patterns.

1. Introduction

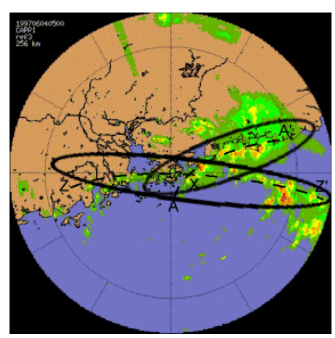
At the Hong Kong Observatory, automated radar image analysis is a crucial component in the development of a short range forecasting system for the prediction of rainstorms (i.e. SWIRLS - Short-range Warning of Intense Rainstorms in Localized Systems). By far, weather radar is still the most practical and reliable observational platform for monitoring rain clouds that have spatial resolution of several kilometres and a life span in the order of minutes. Data from the radar scan are processed and displayed for forecasters' reference every six minutes.



(a)



(b)



(c)

Figure 1 - Schematics of rainstorm evolutionary process: (a) Stage I - scattered and disjointed echoes; (b) Stage II - organized echo bands; (c) Stage III - crossing or merging of echo lines.

As intense convection tends to develop explosively, the forecasters have to keep a close watch on the radar monitors while at the same time assimilating weather information from other sources and attending to forecast formulation and warning procedures. If what transpires on the radar screen can be automatically and objectively analyzed, the forecasters will be in a better position to assess accurately and respond quickly as the rainstorm event unfolds. From forecasters' experience, isolated, scattered, unconnected and disjointed radar echoes (Stage I in Figure 1) that move along nicely with the prevailing flow will not pose a serious threat in terms of excessive rain. The likelihood of heavy rain will increase if convective activity persists or re-generates at preferred locations, or is lined up and advected along a corridor of sustained development. The detailed patterns may differ from case to case in a variety of ways; but in general the trend of increasing organization is unmistakable (Stage II of Figure 1). Occasionally, the forecasters may recognize the emergence of the "X" pattern signature in which lines of echoes cross or merge and convection becomes enhanced near the intersecting point (Stage III in Figure 1). The challenge for any automated pattern recognition algorithm is to reflect accurately the evolutionary processes from Stage I to Stage II, and possibly to Stage III as well, before degenerating back to Stage I.

2. The Artificial Neural Network

Artificial neural networks (ANNs) are known to be useful in pattern recognition. The application of ANN in operational meteorology has met with some degree of success in the classification of cloud systems on satellite images, for examples [1, 2, 3, 4]. Similar approaches have been attempted in Hong Kong [5] but experience in ANN application to radar images is relatively limited.

The ANN used in this study is a four-layer feed-forward back-propagation type. There are altogether four input neurons, two hidden layers and one output neuron. The output status of the ANN is either positive or negative in terms of echo organization. Following Weszka et al. [6], we use the lowest four moments of the radar reflectivity distribution to characterize a radar image. They are the mean m, standard deviation s, skewness s, and kurtosis k:

$$m \equiv \frac{1}{N}\sum_{i=1}^{N} z_i,$$

$$d \equiv \sqrt{\frac{1}{N}\sum_{i=1}^{N}(z_i - m)^2},$$

$$s \equiv \frac{1}{N}\sum_{i=1}^{N}\left(\frac{z_i - m}{d}\right)^3,$$

$$k \equiv \frac{1}{N}\sum_{i=1}^{N}\left(\frac{z_i - m}{d}\right)^4,$$

where z_i is the radar reflectivity at the i^{th} pixel, and N is the total number of pixels in the image. The size of the radar images used in this study is $N = 480 \times 480$ pixels.

The above moments characterize the spatial distribution of radar reflectivity on a given radar map at a given moment. The mean and standard deviation indicate echo intensity and the degree of concentration respectively. The skewness and kurtosis characterize the deviations from a Gaussian distribution, reflecting to some degree the organization of radar echoes. In Figure 2, the four moments in a typical rainstorm development scenario are plotted as functions of time. As the echoes become organized, the mean and standard deviation generally rise while the skewness and kurtosis drop. Armed with this knowledge, we compute the temporal correlations among the four moments. As input identifiers for the ANN, four cross-correlations out of the six possible combinations formed from the set $\{m, d, s, k\}$ are used, namely:

$$C_{md} \equiv \frac{1}{N_t \sigma_m \sigma_d}\sum_{j=1}^{N_t}(m_j - \bar{m})(d_j - \bar{d}),$$

$$C_{ds} \equiv \frac{1}{N_t \sigma_d \sigma_s}\sum_{j=1}^{N_t}(d_j - \bar{d})(s_j - \bar{s}),$$

$$C_{sk} \equiv \frac{1}{N_t \sigma_s \sigma_k}\sum_{j=1}^{N_t}(s_j - \bar{s})(k_j - \bar{k}),$$

$$C_{km} \equiv \frac{1}{N_t \sigma_k \sigma_m}\sum_{j=1}^{N_t}(k_j - \bar{k})(m_j - \bar{m}).$$

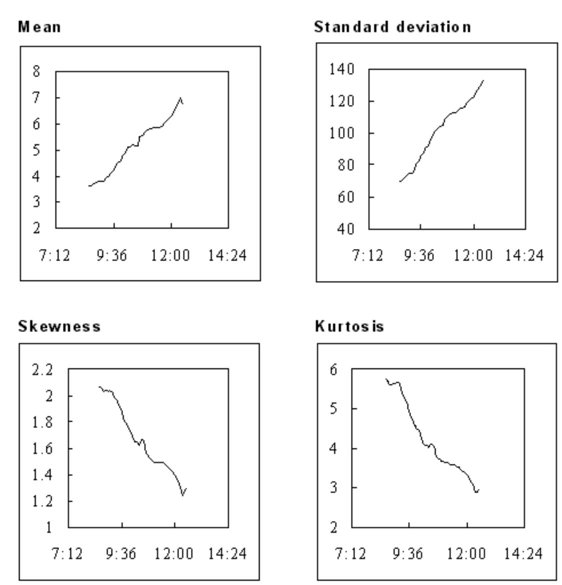


Figure 2 - For a developing rainstorm, both mean and standard deviation rise, but skewness and kurtosis drop.

Here, j labels the time-ordered radar maps, running from 1 to N_t, the number of consecutive radar images used (with radar images coming in every six minutes, we take $N_t = 5$ to look for half-hourly trend); $\overline{m}$, $\overline{d}$, $\overline{s}$, $\overline{k}$ are respectively the means of m, d, s, k over N_t images; and σ_i is the standard deviation of the moment $i = m, d, s, k,$ e.g.

$$\sigma_m = \sqrt{\frac{1}{N_t}\sum_{j=1}^{N_t}(m_j - \overline{m})^2} \cdot$$

Note that the cross-correlations $\{C_{md}, C_{ds}, C_{sk}, C_{km}\}$ as defined above take on values ranging from -1 to 1.

In practice, a z_i threshold of 30dBZ is imposed to isolate and highlight the significant pattern. Radar pixels with z_i < 30dBZ are discarded. To gain more spatial information, the radar image is divided into 25 (5 x 5) equal-sized boxes. Input identifiers computed for each box will be fed to the ANN except for boxes with m not rising monotonically throughout the 5-image radar sequence.

The output neuron produces a number A between 0 and 1. In the present study, we take $A \geq 0.75$ as positive and $A \leq 0.25$ as negative. In the training set, there are altogether 20 rain events from the rainy seasons of 1997 and 1998, in which 12 show increasing organization and lead to heavy rain, and the remaining 8 weakening and bringing only light to moderate rain. For each event, a sequence of five consecutive radar images is extracted for ANN training. Starting from a set of random network parameters, the ANN is updated repeatedly according to the error back-propagation algorithm until the total error for the 20 cases drop below 0.2. All cases in the training set are correctly classified as either positive or negative by the ANN. The ANN is tested using six 5-image radar sequences taken from the same rainy seasons but outside the training set.

The test results are tabulated in Table 1. On the first column are the dates and times of the last radar image in each sequence indicated as "yymmddhhmm" (year/month/day/hour/minute). ANN outputs for all 25 boxes are displayed in the second column. Boxes failing to meet the pre-condition and hence receiving no ANN output are marked with an "X". "P" means positive, i.e. $A \geq 0.75$; "N" means negative, i.e. $A \leq 0.25$. On the third column, verifications by human subjective judgment are represented by similar notations of "P" and "N" for positive and negative classifications respectively. The notations in the fourth column have the following meanings:

PP = positive in both ANN and human assessment;
PN = positive in ANN but negative by human;
NP = negative in ANN but positive by human;
NN = negative in both ANN and human assessment.

Table 1 - Test results of the ANN on six rainstorm cases.

yymmddhhmm	ANN outputs	Human judgment	PP	PN	NP	NN
9704301330	X X P X X	X X N X X		1		
	N N X X X	N N X X X				2
	X X P X X	X X P X X	1			
	X P X X X	X P X X X	1			
	X X X X X	X X X X X				
9704301406	X X X X X	X X X X X				
	X P P X X	X N P X X	1	1		
	X X X X X	X X X X X				
	X P X X X	X P X X X	1			
	X X X X X	X X X X X				
9706140306	X X X N X	X X X N X				1
	X X X X X	X X X X X				
	X X N N X	X X N N X				2
	P N N N P	N N N N N		2		3
	X X X X X	X X X X X				
9706140230	X X X N X	X X X N X				1
	X X X X X	X X X X X				
	X N X P X	X N X P X	1			1
	X P N N X	X P N N X	1			2
	X X X X X	X X X X X				
9805021506	X X X X X	X X X X X				
	N X X N X	N X X N X				2
	N X P X X	N X P X X	1			1
	X X X X X	X X X X X				
	X X X X X	X X X X X				
9805021542	X N X N X	X N X N X				2
	X P X N X	X P X N X	1			1
	N X N X X	N X N X X				2
	X X X X X	X X X X X				
	X X X X X	X X X X X				
		Total:	8	4	0	20

From Table 1, skill scores in terms of NAP (No Alarm Probability), FAR (False Alarm Rate), CSI (Critical Success Index) and HSI (Heidke Skill Index) are computed:

NAP = NP / (PP+NP) = 0
FAR = PN / (PP+PN) = 0.33
CSI = PP / (PP+PN+NP) = 0.67
HSI = [(PP+NN)-R] / [(PP+PN+NP+NN)-R] = 0.71;

where R, random forecast, is given by

R = [(PP+PN)x(PP+NP)+(NP+NN)x(PN+NN)] / S; and
S = PP+PN+NP+NN.

Good performance is characterized by low NAP and FAR but high CSI and HSI. For the test cases, NAP is found to be very good as no significant events have been missed. FAR is not bad but can be better. Both CSI and HSI are considered to be reasonable, with positive values in the latter meaning that ANN out-performs the low-skill random forecast.

3. The Hough Transform Module

By Stage III, the radar echoes are organized in such ways that they tend to congregate into rainbands. Convective patterns associated with many severe weather systems, such as tropical cyclones, warm or cold fronts, and squall lines, all have banding features that on the scale of the radar images can be approximated into straight lines. Hough transform has been recognized as one of the more efficient algorithms for identifying linear patterns on 2-dimensional images [7]. It has nice properties such as: (i) fast - computation time increases only linearly instead of quadratically with the number of pixels; (ii) high degree of tolerance - the line width, angular resolution and brokenness can be specified arbitrarily; (iii) flexible - the region of interest can be specified at any location inside the image.

In this study, the Hough transform module (HTM) is so designed to facilitate operational and research development. To highlight the prominent features on the radar images, echoes on the transformed space are multiplied by a weighting factor which is proportional to the respective pixel intensity value, i.e., z_i^{α}, and α is arbitrarily set to be 3. The region of interest is movable rather than fixed at the centre of the radar image. Once the centre and range are specified, the HTM will search for echo lines within the specified region of interest and ignores pixels outside. This serves to overcome the scaling problem and users can then zoom in on specific rain systems, be it local scale features of tens of kilometres in length or mesoscale features stretching to hundreds of kilometres. In fact, through a graphical user interface, users can adjust the parameters to optimize the line identification processes.

4. Concluding Remarks

Both ANN and HTM have shown initial promises in depicting the evolution of rainstorms, the former in terms of echo organization and the latter in the lining up of echoes. Training of the ANN will be a continuous process, and in time the system should become even more robust and resilient. HTM's application in operational forecasting will require further parameter-tuning, e.g. threshold setting, line attributes, etc., for optimization and customization purposes.

An effective way of combining ANN and HTM in rainstorm detection will be further explored. Figure 3 is an example of what can be done: through ANN, an area of positive echo organization is highlighted; and HTM is then applied to the highlighted area for identification of linear patterns. A more focussed approach is not simply a matter of saving computation time; it also draws forecasters' attention to the main area of activity and keeps track of the evolving pattern between successive images.

But ultimately, we hope to make better use of ANN and HTM for short-term prediction of enhanced convection in association with the notorious "X" pattern. Echo growth and decay for the purpose of quantitative precipitation forecast is still an area of active research in the SWIRLS development programme. By integrating ANN and HTM into a reliable objective tool in "X" pattern identification, rainfall enhancement with respect to the signature pattern can then be systematically analyzed and studied. If the popular hypotheses nominally linked with the so-called "X" pattern can be confirmed and suitably quantified, a positive impact on the rainfall forecast algorithm would logically follow.

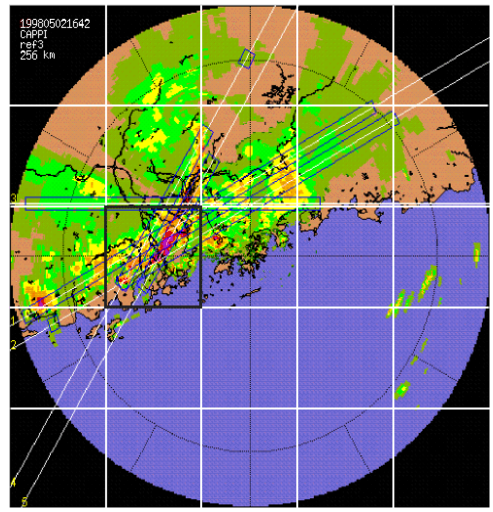


Figure 3 - Typical output from the combined ANN and HTM. Radar echoes appear to be merging within the highlighted section to the left of the central box.

5. References

[1] Kay, J., J.A. Maslanik and A.J. Schweiger, "Classification of merged AVHRR and SMMR Arctic data with neural networks", *Photogramm. Eng. Remote Sens.*, **55**, 1989, pp. 1331-1338.

[2] Lee, J., R.C. Weger, S.K. Sengupta and R.M. Welch, "A neural network approach to cloud classification", *IEEE Trans. on Geosci. and Remote Sensing*, **28**, 1990, pp.846-855.

[3] Lee, V.C.S., S.L. Hung, A.Y.S. Cheng, C.Y. Lam and C.M.Tam, "An operational cloud classifier for satellite images", *Proceedings of the 2nd International Conference on East Asia and Western Pacific Meteorology and Climate*, 1992, pp.480-487.

[4] Welch, R.M., S.K. Sengupta, A.K.P. Rabindra, N. Rangaraj and M.S. Xvar, "Polar cloud and surface classification using AVHRR imagery: and intercomparison of methods", *J. Appl. Meteorol.*, **31**, 1992, pp.405-420.

[5] Pankiewicz, G.S., "Pattern recognition techniques for the identification of cloud and cloud systems," *Meteorol. Application*, **2**, 1995, pp.257-271.

[6] Weszka, J.S., C.R. Dyer and A. Rosenfeld, *IEEE Trans. Syst. Man. Cybern*. SMC-6, 1976, pp.269-285.

[7] Duda, R.O. and P.E. Hart, "Use of Hough transform to detect lines and curves in pictures", *Comm. of ACM*, **15**, 1972, pp.11-15.