WEATHER NOWCASTING FROM SATELLITE IMAGE SEQUENCES USING PICTURE FUZZY CLUSTERING AND SPATIOTEMPORAL REGRESSION

Nguyen Dinh Hoa[1], Pham Huy Thong[2], Le Hoang Son[2]

[1] VNU Information Technology Institute, Vietnam National University, 144 Xuan Thuy, Cau Giay, Hanoi, Vietnam

[2] VNU University of Science, Vietnam National University, 334 Nguyen Trai, Thanh Xuan, Hanoi, Vietnam,

Email: {hoand, thongph, sonlh}@vnu.edu.vn

ABSTRACT

Weather nowcasting comprises the detailed description of the current weather along with forecasts obtained by extrapolation for very short-range period of zero to six hours ahead. It is particularly useful when forecasting complicated processes such as rainfall, clouds, and rapidly developing or changing storms. This plays an important role for daily activities like working, travelling, daily planning, flying, etc. Weather forecast can be solved by latest radar, satellite or observational data. However, the main challenges associated with nowcasting are the flawed characterization of transitions between different meteorological structures.

In this paper, we propose an innovative method for weather nowcasting from satellite image sequences using the combination of picture fuzzy clustering and spatiotemporal regression. Firstly, the satellite image pixels are partitioned into clusters by using picture fuzzy clustering - a fuzzy clustering method based on the theory of picture fuzzy sets. Secondly, the Fast Fourier Transform method is used to filter out non-predictable scales leading to the increasing in time ranges of predictability. Finally, the spatiotemporal regression method is used to forecast the predicted sequences of images. The experiments indicate that the proposed method is better than the relevant ones in weather nowcasting.

Keywords: Picture fuzzy clustering; Satellite images; Spatiotemporal regression; Weather nowcasting.

1. INTRODUCTION

Weather nowcasting, according to Mass (2011), combines a description of the current state of the atmosphere and a short-term forecast of how the atmosphere will evolve during the next several hours. In this time range it is possible to forecast small features such as rainfall, clouds and individual storms with reasonable accuracy. A forecaster using the latest radar, satellite and observational data is able to make analysis of the small-scale features present in a small area such as a city and make an accurate forecast for the following few hours. Although nowcasting techniques have mostly been applied to radar measurements by some researchers such as Germann and Zawadzki (2002), Turner et al. (2004), etc., there are many regions that are away from radar coverage particularly in developing countries. To forecast the weather in these areas using nowcasting techniques, satellite observations are the appropriate alternative according to Rivolta et al. (2006) and Marzano et al. (2007). Based on the observations of satellite image approach, Melgani (2006), Evans (2006) and Shukla et al. (2012) developed methods to track and nowcast meteorological parameters. Melgani (2006) reconstructed the context of cloud-contaminated multi-temporal and multispectral images; Evans (2006) used multi-channel correlation-relaxation labeling to analyze cloud motion; Shukla et al. (2012) proposed an approach to study the evolution of convective cells. Recently, Shukla et al. (2014) proposed a method for predicting satellite image sequences combining spatiotemporal autoregressive (STAR) model with fuzzy clustering to increase the forecast accuracy. Although this technique resulted in better prediction accuracy than those of Melgani (2006), Evans (2006) and Shukla et al. (2012), the forecasting output was not good enough because of the limitations of fuzzy sets such as the hesitation and vagueness.

In this paper, we combine the STAR technique with the picture fuzzy clustering (FC-PFS) in a new algorithm named as *Picture Fuzzy Clustering - SpatioTemporal AutoRegressive* (PFC-STAR) in order to enhance the weather nowcasting quality. FC-PFS (Thong and Son, 2014) was constructed on the basis of picture fuzzy sets, which in essence are the generalization of the traditional fuzzy sets and are able to handle the mentioned limitations of fuzzy sets. FC-PFS was proven in the equivalent paper to be better than some fuzzy clustering algorithms. The proposed PFC-STAR method consists of three steps. *Firstly*, FC-PFS is used to partition the training sample into clusters. *Secondly*, all elements of these clusters are labeled and Discrete Fourier Transform (DFT) - based filter is embedded to clarify nonpredictable scales leading to the increase in time range of predictability. *Finally*, the STAR technique is used to predict the output image of weather nowcasting. The experimental dataset is taken from the satellite image sequences of Southeast Asia. The experimental results indicate that the proposed method is better than the relevant ones in weather nowcasting, particularly in the accuracy of rain-rate retrieval. The rest of the paper is organized as follows. In Section 2, we present the proposed PFC-STAR method combining the FC-PFS, the DFT-based filter and the STAR techniques for weather nowcasting problem. In Section 3, we apply the PFC-STAR to predict the weather from satellite image sequences of Southeast Asia. Finally, the conclusions are covered in Section 4.

2. THE PFC-STAR ALGORITHM

2.1 The flow of PFC-STAR

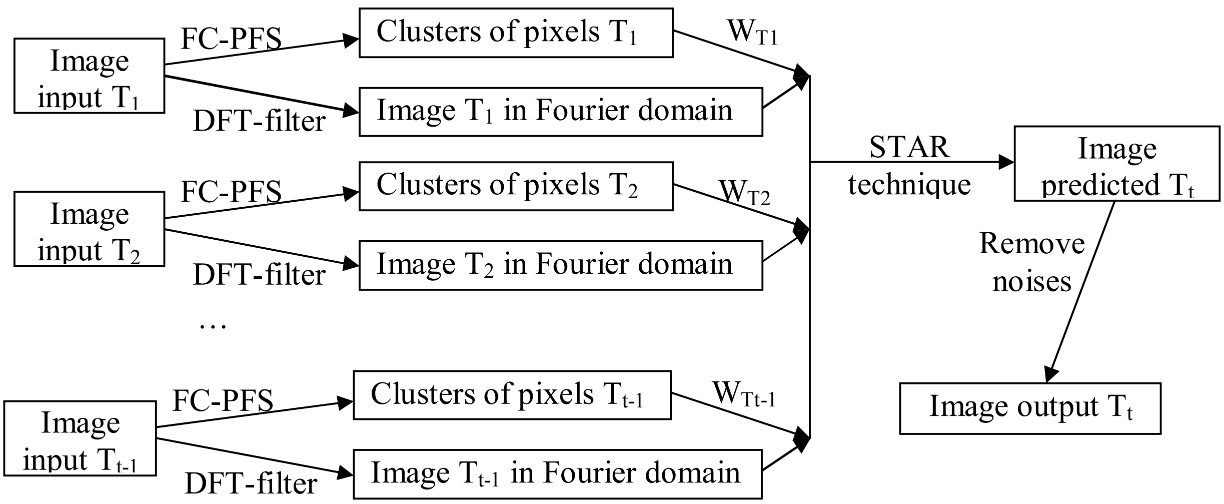


Figure 1. The PFC-STAR algorithm

The input image sequences of the PFC-STAR algorithm are firstly processed by FC-PFS algorithm and DFT-filter. FC-PFS is used to partition all pixels in these images into clusters and mark them by different colors. DFT-filter is employed to remove nonpredictable scales of images and change them to the Fourier domain. Secondly, STAR technique is applied to predict the image result from all images in the Fourier domain and the clusters of pixels. It aims to find out the weight for each cluster of pixels in each image to determine the predicted image. Finally, before resulting in the final predicted image, some noises such as the salt-and-peeper are removed by the Adaptive Median Filtering method. The proposed algorithm is described in Figure 1.

2.2 Picture fuzzy clustering

Thong and Son (2014) have proposed a picture fuzzy clustering algorithm based on PFS. Suppose that we have a dataset X consisting of N data points in d dimensions. The algorithm divided the dataset into C groups satisfying the objective function below.

$$J = \sum_{k=1}^{N}\sum_{j=1}^{C}\left(u_{kj}\left(2-\xi_{kj}\right)\right)^{m}\left\|X_k - V_j\right\|^2 + \sum_{k=1}^{N}\sum_{j=1}^{C}\eta_{kj}\left(\log\eta_{kj} + \xi_{kj}\right) \to \min, \quad (1)$$

Some constraints are defined as follows.

$$u_{kj}, \eta_{kj}, \xi_{kj} \in [0,1], \quad (2)$$

$$u_{kj} + \eta_{kj} + \xi_{kj} \le 1, \quad (3)$$

$$\sum_{j=1}^{C}\left(u_{kj}\left(2-\xi_{kj}\right)\right) = 1, \quad (4)$$

$$\sum_{j=1}^{C}\left(\eta_{kj} + \frac{\xi_{kj}}{C}\right) = 1,\ k = \overline{1,N},\ j = \overline{1,C}. \quad (5)$$

The optimal solutions of the systems (1-5) are:

$$V_j = \frac{\sum_{k=1}^{N}\left(u_{kj}\left(2-\xi_{kj}\right)\right)^{m} X_k}{\sum_{k=1}^{N}\left(u_{kj}\left(2-\xi_{kj}\right)\right)^{m}},\ j = \overline{1,C}, \quad (6)$$

$$u_{kj} = \frac{1}{\sum_{i=1}^{C}\left(2-\xi_{kj}\right)\left(\frac{\left\|X_k - V_j\right\|}{\left\|X_k - V_i\right\|}\right)^{\frac{2}{m-1}}},\ k = \overline{1,N},\ j = \overline{1,C}, \quad (7)$$

$$\eta_{kj} = \frac{e^{-\xi_{kj}}}{\sum_{i=1}^{C} e^{-\xi_{ki}}}\left(1 - \frac{1}{C}\sum_{i=1}^{C}\xi_{ki}\right),\ k = \overline{1,N},\ j = \overline{1,C}, \quad (8)$$

$$\xi_{kj} = 1 - \left(u_{kj} + \eta_{kj}\right) - \left(1 - \left(u_{kj} + \eta_{kj}\right)^{\alpha}\right)^{\frac{1}{\alpha}},\ k = \overline{1,N},\ j = \overline{1,C}. \quad (9)$$

Fuzzy Clustering Method on Picture Fuzzy Sets

I: Data X whose number of elements (N) in d dimensions; Number of clusters (C); fuzzifier m; threshold ε; maximum iteration *maxSteps > 0.*

O: Matrices u, η, ξ and centers V;

FC-PFS:

1: t = 0

2: $u_{kj}^{(t)} \leftarrow random$; $\eta_{kj}^{(t)} \leftarrow random$; $\xi_{kj}^{(t)} \leftarrow random$ ($k = \overline{1,N}$, $j = \overline{1,C}$) satisfy equations (2, 3)

3: Repeat

4: t = t + 1

5: Calculate $V_j^{(t)}$ ($j = \overline{1,C}$) by equation (6)

6: Calculate $u_{kj}^{(t)}$ ($k = \overline{1,N}$; $j = \overline{1,C}$) by equation (7)

7: Calculate $\eta_{kj}^{(t)}$ ($k = \overline{1,N}$; $j = \overline{1,C}$) by equation (8)

8: Calculate $\xi_{kj}^{(t)}$ ($k = \overline{1,N}$; $j = \overline{1,C}$) by equation (9)

9: Until $\left\|u^{(t)} - u^{(t-1)}\right\| + \left\|\eta^{(t)} - \eta^{(t-1)}\right\| + \left\|\xi^{(t)} - \xi^{(t-1)}\right\| \le \varepsilon$ or *maxSteps* has reached

2.3 Discrete Fourier Transform – based filter

Discrete Fourier Transform (DFT) is the sampled Fourier Transform and therefore does not contain all frequencies forming an image, but only a set of samples which is large enough to fully describe the spatial domain image. The number of frequencies corresponds to the number of pixels in the spatial domain image, i.e. the images in the spatial and Fourier domains are of the same size. For a square image of size $H \times K$, the two-dimensional DFT is given by equation (10).

$$P(x,y,t)=\sum_{h=0}^{H-1}\sum_{k=0}^{K-1}p(h,k,t)e^{-i2\pi\left(\frac{hi}{H}+\frac{kj}{K}\right)}, \tag{10}$$

where $p(h,k,t)$ is the image in the spatial domain at time t and the exponential term is the basis function corresponding to each point $P(x,y,t)$ in the Fourier space. The inverse DFT from Fourier space into spatial domain is described in equation (11).

$$p(x,y,t)=\frac{1}{HK}\sum_{h=0}^{H-1}\sum_{k=0}^{K-1}P(h,k,t)e^{i2\pi\left(\frac{hi}{H}+\frac{kj}{K}\right)}. \tag{11}$$

2.4 STAR

Pfeifer and Deutsch (1980) proposed a linear spatio-temporal autoregressive model which was a 3D version of the regular autoregressive model (AR) and it enabled users to forecast a spatio-temporal series based on the information of its own past in space and time. The intensity of every pixel P(x, y, t) in an image domain can be modeled as a function "ψ" of the intensity of neighboring pixels in space and time, with assumption of causality.

$$P(x,y,t)=\psi\left(P(x+\Delta x_i, y+\Delta y_i, t+\Delta t_i)\right), \tag{12}$$

where $(\Delta x_i, \Delta y_i, \Delta t_i)$ is the neighborhood coverage in space and time with $\Delta t_k < 0$. Equation (12) is modified for clustering based Spatiotemporal Autoregressive (STAR) by Shukla et al. (2014) as follow.

$$P(x,y,t)=\sum_{k=t-T}^{t-1}\sum_{j=y-J}^{y+J}\sum_{i=x-I}^{x+I}W_{i,j,k}^{G}P(i,j,k), \tag{13}$$

where $2I+1$, $2J+1$ and T correspond to the total number of rows, columns and frames respectively, which are included in the predictor set (neighborhood), and $W_{i,j,k}^{G}$ is the corresponding weight with the superscript G denotes the weight for G^{th} cluster, $x=\overline{0,H}$, $y=\overline{0,K}$. The weights are calculated by minimizing the function,

$$\left\| P(x,y,t)-\sum_{k=t-T}^{t-1}\sum_{j=y-J}^{y+J}\sum_{i=x-I}^{x+I}W_{i,j,k}^{G}P(i,j,k)\right\| \to \min, \tag{14}$$

where $\|.\|$ denotes to the Euclidean distance. Equation (14) is solved by the least square method using QR factorization (Paige & Saunders, 1982). After finding the weights, all pixels in predicted image will be calculated by equation (13).

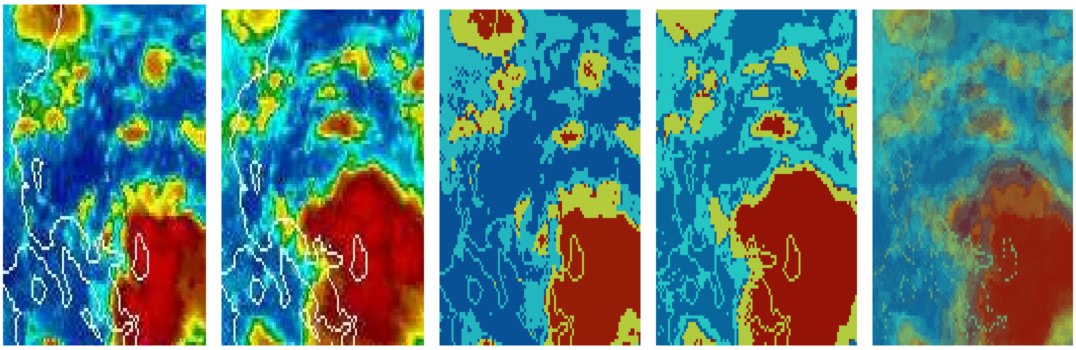
2.5 Remove noises from the predicted image

The predicted image could consist of noises and out-of-bound pixels. The noises are removed by using the Adaptive Median Filtering method (Al-amri et al., 2010). This method performs spatial processing to preserve detail and smooth non-impulsive noises. A prime benefit to this adaptive approach to Median Filtering is that the adaptive windows of Adaptive Median Filtering do not erode away edges or other small structures in the image. Additional, some out-of-bound pixels are normalized by equation (15).

$$P(x,y,t)=\sum_{k=t-T}^{t-1}\sum_{j=1}^{C}\mu_j^{(k)}\left(2-\xi_j^{(k)}\right)V_j^{(k)} \quad (15)$$

3. EVALUATION

The experiments are conducted on the computer with configuration of 2G RAM, 2.13 GHz core 2 Duo. The data includes three sets of images: Malaysian, Luzon – Philippines and Jakarta – Indonesia. Each set contains five images consecutively from 9.30 am to 13.30 pm. All images have the same size (100x100 pixels). Figure 2 shows the images of predicting process for Luzon, Philippines. Figure 2a (resp. 2b) shows the original images at 9.30 am (resp. 10.30 am). Figure 2c (resp. 2d) indicates the clustering images (resp. the forecasted image). Table 1 also shows that PFC-STAR produces better RMSE compared to the method of Shukla et al. (2014).



(a) (b) (c) (d) (e)

Figure 2. Images of Luzon, Philippines: (a) The original image at 9.30 am; (b) the original image at 10.30 am; (c) the clustered image at 9.30; (d) the clustered image at 10.30; (e) the forecasted image at 11.30.

Table 1. RMSE and computational time of PFC-STAR and Shukla et al.

Data	RMSE (%)		Computational time (sec)	
	PFC-STAR	Shukla et al. (2014)'s method	PFC-STAR	Shukla et al. (2014)'s method
Malaysian	26.77	27.11	362.745	359.88
Luzon – Philippines	33.61	33.45	345.672	343.43
Jakarta – Indonesia	30.12	32.04	342.76	339.97

4. CONCLUSIONS

This paper presented a novel method for weather nowcasting from satellite image sequences using the combination of picture fuzzy clustering and spatiotemporal regression. Experimental results indicated that the proposed method is better than that of Shukla et al. (2014). Further works will improve PFC-STAR to get higher accuracy of prediction and lower computational time.

5. ACKNOWLEDGEMENTS

This work is sponsored by a VNU project under the contract No. QG.14.60.

6. REFERENCES

Al-amri, S. S., Kalyankar, N. V., & Khamitkar, S. D. (2010). A comparative study of removal noise from remote sensing image. *arXiv preprint arXiv:1002.1148*.

Cuong, B.C., Kreinovich, V., Picture Fuzzy Sets - a new concept for computational intelligence problems, Proceeding of 2013 Third World Congress on Information and Communication Technologies (WICT 2013) 1-6.

Evans, A. N. (2006). Cloud motion analysis using multichannel correlation-relaxation labeling. *Geoscience and Remote Sensing Letters, IEEE*, *3*(3), 392-396.

Germann, U., & Zawadzki, I. (2002). Scale-dependence of the predictability of precipitation from continental radar images. Part I: Description of the methodology. *Monthly Weather Review*, *130*(12), 2859-2873.

Marzano, F. S., Rivolta, G., Coppola, E., Tomassetti, B., & Verdecchia, M. (2007). Rainfall nowcasting from multisatellite passive-sensor images using a recurrent neural network. *Geoscience and Remote Sensing, IEEE Transactions on*, *45*(11), 3800-3812.

Mass, C., & Mass, C. F. (2011). Nowcasting: The Next Revolution in Weather Prediction. *Bulletin of the American Meteorological Society*.

Meisinger, K., & Kaup, A. (2007). Spatiotemporal selective extrapolation for 3-D signals and its applications in video communications. *Image Processing, IEEE Transactions on*, *16*(9), 2348-2360.

Melgani, F. (2006). Contextual reconstruction of cloud-contaminated multitemporal multispectral images. *Geoscience and Remote Sensing, IEEE Transactions on*, *44*(2), 442-455.

Paige, C. C., Saunders, M. A. (1982). LSQR: An algorithm for sparse linear equations and sparse least squares. *ACM Transactions on Mathematical Software (TOMS)*, *8*(1), 43-71.

Pfeifer, P. E., & Deutrch, S. J. (1980). A Three-Stage Iterative Procedure for Space-Time Modeling Phillip. *Technometrics*, *22*(1), 35-47.

Rivolta, G., Marzano, F. S., Coppola, E., & Verdecchia, M. (2006). Artificial neural-network technique for precipitation nowcasting from satellite imagery. *Advances in Geosciences*, *7*(7), 97-103.

Shukla, B. P., & Pal, P. K. (2012). A source apportionment approach to study the evolution of convective cells: An application to the nowcasting of convective weather systems. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, *5*(1), 242-247.

Shukla, B. P., Kishtawal, C. M., & Pal, P. K. (2014). Prediction of Satellite Image Sequence for Weather Nowcasting Using Cluster-Based Spatiotemporal Regression. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(7), 4155 – 4160.

Thong, P.H., Son, L.H. (2014). A new approach to multi-variables fuzzy forecasting using picture fuzzy clustering and picture fuzzy rules interpolation method. *Proceeding of 6th International Conference on Knowledge and Systems Engineering (KSE 2014)*, October 9-11, 2014, Hanoi, Vietnam, 679 - 690.

Turner, B. J., Zawadzki, I., & Germann, U. (2004). "Predictability of precipitation from continental radar images—Part III: Operational nowcasting implementation (MAPLE)," *J. Appl. Meteorol.*, vol. 43, no. 2, 231–248, Feb.