

# **Short-Term Precipitation Forecasting Based on Radar Reflectivity Images**

**Po-Jui Lin**

## **Abstract**

This article describes the image pre-processing and feature extraction stage of a framework that predicts short-term precipitation at Hong Kong based on radar reflectivity images. This framework utilizes cloud trajectory method based on velocity vector obtained by the SIFT algorithm which matches descriptors in the adjacent time frames. Features including statistics of radar images and global descriptions of cloud pattern are selected as the input of a convolutional neural network.

## **Introduction**

High-intensity rainfall within a short period causes severe damage in populated areas. For instance, Typhoon Mangkhut, the most intense cyclone in the recorded history that battered Hong Kong in September 2018, injured nearly 500 people and resulted in an economic loss of 7.3 billion HKD in the megacity alone [1]. Therefore, an accurate weather prediction service can not only provide early warning to the public, but it can also be utilized in everyday life such as outdoor activities. Weather radar, with its advantage of short-term predictive capability and high-spatiotemporal resolution, has been one of the primary tools to forecast weather. In the article, radar maps that cover the radar reflectivity of a 41 km \* 41 km area are used. Adjacent images have a time span of 10 minutes, and the images are taken at a single height for simplicity. The dataset has been collected from the Hong Kong Meteorological Bureau.

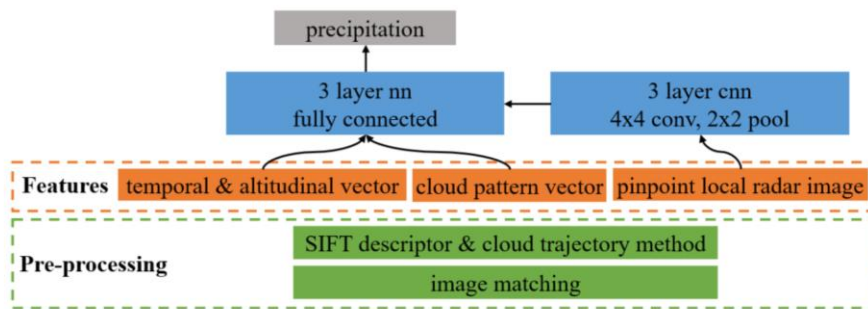
## Solution

### Framework

The framework is based on the final solution from Team Marmot, who finished in 1st place in the CIKM AnalytiCup 2017 – “Short-Term Quantitative Precipitation Forecasting Challenge”.

Figure 1 illustrates the flowchart of this solution. This article will mainly focused on the pre-processing and feature extraction stage.

**Figure 1:** flowchart of the solution



### Image Preprocessing

An accurate prediction of precipitation requires precise tracking on the trajectory of the cloud movement within a specified range around the target site.

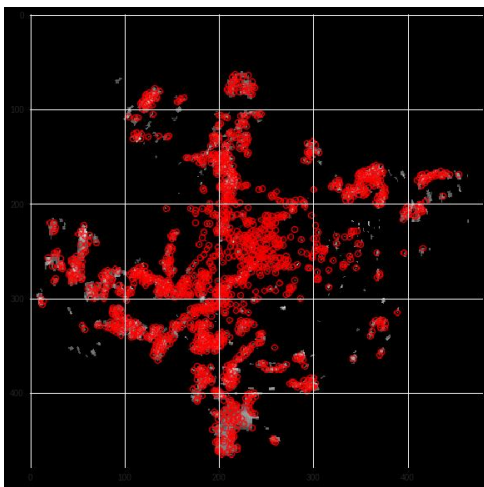
#### -Taylor Frozen Hypothesis

Taylor Hypothesis is an assumption that advection contributed by turbulent circulations themselves is small so that the advection of a field of turbulence past a fixed point can be taken to be entirely due to the mean flow [2]. It implies that the temporal variations at a fixed site can be deduced from the spatial variations.

## **-Key point detection by FAST**

To acquire the velocity field of the cloud movement, the relative displacement of key points at two adjacent timestamp images should be calculated. First of all, key points in radar images are detected using FAST (Features from Accelerated Segment Test) algorithm. FAST is a corner detection method originally developed by Edward Rosten and Tom Drummond. Compared to other corner detectors, FAST main advantage is its computational efficiency. While the difference of Gaussians (DoG) is used by the SIFT and Harris detectors, FAST algorithm applies several techniques to reduce the computational resource to identify corners:

1. Consider the surrounding circle of pixels of a possible corner pixel  $p$ , if a set of contiguous pixels in the circle exceeds a certain threshold,  $p$  is identified as a corner.
2. A high-speed test is added to exclude a large number of non-corners. This test examines only 4 pixels in the circle mentioned above, which is consisted of the pixel in the up, down, left and right direction. At least three of these pixels must exceed the certain threshold for  $p$  to be a corner.



**Figure 2:** A sample result of applying FAST algorithm on a radar image for key point detection.

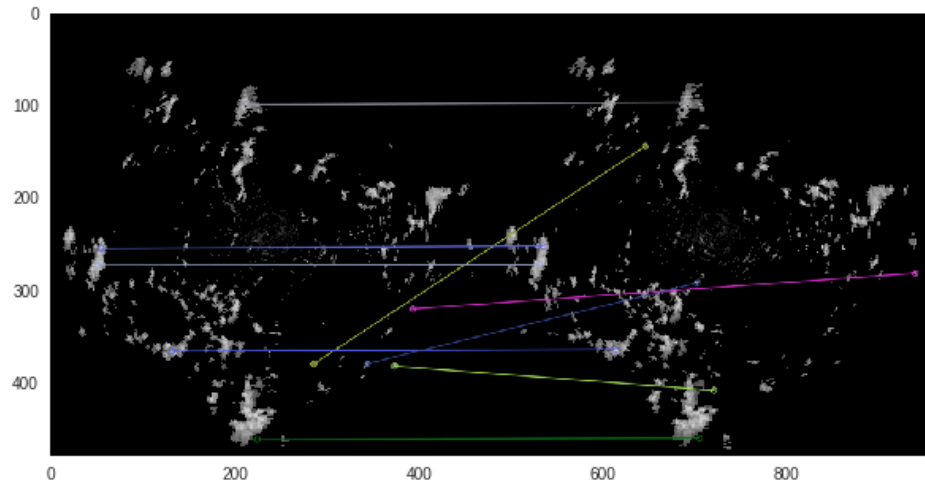
It can be observed that almost every cloud segments in the radar images are identified as corner possibly since the background is black so each cloud element can easily stand out as an interest point.

### **-SIFT descriptor matching**

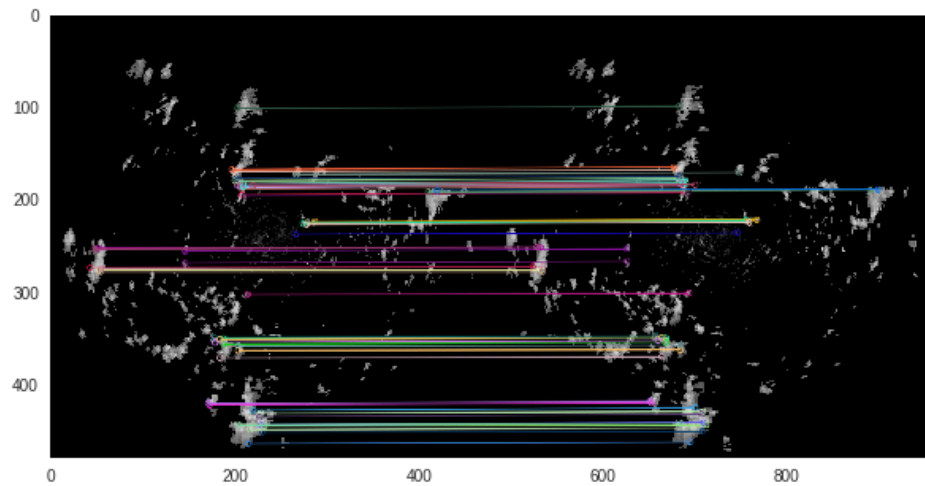
SIFT (scale-invariant feature transform) is a feature detection algorithm that can both detect and describe key points in images. SIFT key points are generated by constructing a scale space to ensure scale invariance and utilize DoG as a LoG (Laplacian of Gaussian) approximation to find key points. To generate key point descriptors, a 16x16 neighborhood around the key point is taken. It is divided into 16 sub-blocks of 4x4 size. For each sub-block, a 8 bin orientation histogram is generated. Therefore, a total of 128 bin values are available and it is represented as a vector to form key point descriptor.

However, in this article, key points are identified based on FAST rather than SIFT because SIFT descriptors on FAST key points achieves better result on the radar images. Figure 3 and 4 demonstrates this difference by matching SIFT descriptor vectors of two radar images at two adjacent timestamps.

We accomplish the matching by using Brute-Force matcher. It takes the descriptor of one key point in the first set matches with all other key points in the second set using some distance calculation. The closest one is returned. Parameters of `cv2.BFMatcher()` are set to be `cv2.NORM_L2` and `crossCheck = True` to best fit SIFT descriptors and provide consistent results. Again, the matching result is shown in Figure 3 and 4.



**Figure 3:** SIFT descriptor matching on SIFT key points of two radar images at two adjacent timestamp.



**Figure 4:** SIFT descriptor matching on FAST key points of two radar images at two adjacent timestamp.

## **Feature Extraction**

### **-Histogram of typical SIFT descriptors**

The typical descriptors are acquired by applying k-means clustering on all local SIFT descriptors in the dataset, and the histogram of descriptor of each cluster is obtained as the vector representation of the cloud type.

### **K-means clustering**

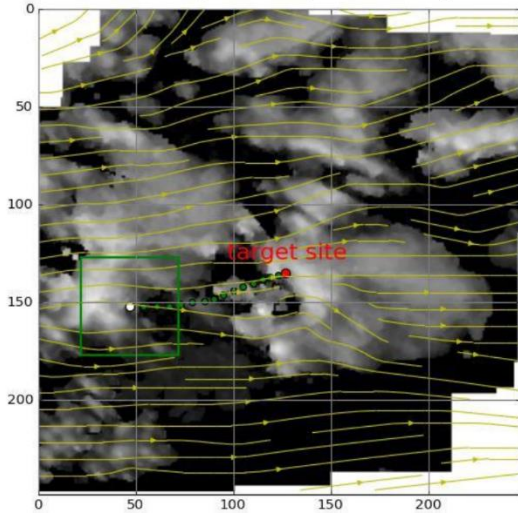
K-means clustering is categorized as unsupervised learning that best fit unlabeled data. The goal of this algorithm is to identify groups in the data, which the variable K represents the number of groups. The algorithm iteratively assign every data point to one of the K groups based on the provided features. In other words, data points are clustered based on feature similarity.

### **-Histogram of radar reflectivity intensity**

Radar reflectivity is a measure of the efficiency of a radar target in intercepting and returning radio energy [3]. By analyzing the distribution of radar reflectivity, the rainfall difference under various cloud formation can be distinguished.

### **-Cloud velocity, direction, acceleration, trace curvature.**

The velocity vector is acquired by matching SIFT descriptors in two adjacent time frames mentioned above. Figure 5 is a illustration of the trajectory lines that would shift to the target site as time goes by.



**Figure 5:** Illustration of the trajectory lines

Red dot is the target site that the cloud pattern would move to after a short period. Green dots are the trace positions with fixed temporal interval. Yellow lines are the convoluted trajectory lines along the velocity vector. The green square is the pinpoint surrounding area.

### **-Pinpoint radar reflectivity image**

Based on Taylor frozen hypothesis and the matching of SIFT descriptor mentioned above, the trajectory that passes through the target site could be extrapolated. The local images along the extrapolated trace are one of the input features of a convolutional neural network to provide linkage between radar reflectivity and precipitation.

## Conclusion

In this article, the pre-processing and feature extraction methods of a framework for short-term precipitation forecasting is provided. In the pre-processing stage, cloud trajectory lines are tracked based on the velocity vector, which is obtained by matching SIFT descriptors in two adjacent time frames. Afterward, features including pinpoint local radar images and the global descriptions of the cloud pattern would be adopted for a convolutional neural network.

## Reference

- [1] Tony Cheung and Su Xinqi, “Typhoon Mangkhut officially Hong Kong’s most intense storm since records began”, *South China Morning Post*, September 2018, [Online]. Available: <https://www.scmp.com/news/hong-kong/health-environment/article/2164523/its-official-typhoon-mangkhut-was-most-intense>
- [2] “Taylor’s hypothesis”, *American Meteorological Society*, February 2012, [Online]. Available: [http://glossary.ametsoc.org/wiki/Taylor%27s\\_hypothesis](http://glossary.ametsoc.org/wiki/Taylor%27s_hypothesis)
- [3] “Radar reflectivity”, *American Meteorological Society*, April 2012, [Online]. Available: [http://glossary.ametsoc.org/wiki/Radar\\_reflectivity](http://glossary.ametsoc.org/wiki/Radar_reflectivity)